SWITCHING OXIDE TRAPS

Timothy R. Oldham
NASA Goddard Space Flight Center/QSS Group, Inc.
Greenbelt, MD 20771, USA

I. Introduction

We consider radiation-induced charge trapping in SiO₂ dielectric layers, primarily from the point of view of CMOS devices. However, SiO₂ insulators are used in many other ways, and the same defects occur in other contexts. The key studies, which determined the nature of the oxide charge traps, were done primarily on gate oxides in CMOS devices, because that was the main radiation problem in CMOS at one time. There are two major reviews of radiation-induced oxide charge trapping already in the literature, which discuss the subject in far greater detail than is possible here. The first of these was by McLean et al.¹ in 1989, and the second, ten years later, was intended as an update, because of additional, new work that had been reported.²

Basically, the picture that has emerged is that ionizing radiation creates electron-hole pairs in the oxide, and the electrons have much higher mobility than the holes. Therefore, the electrons are swept out of the oxide very rapidly by any field that is present, leaving behind any holes that escape the initial recombination process.³ These holes then undergo a polaron hopping transport toward the Si/SiO₂ interface (under positive bias). Near the interface, some fraction of them fall into deep, relatively stable, long-lived hole traps. See reference 1 and its bibliography for more details on the charge generation, recombination and transport processes. The nature and annealing behavior of these hole traps is the main focus of this paper, and of reference 2, Chapter 2.

II. The Interfacial Transition Layer

The presence of hole traps in the oxide is a fundamental consequence of the oxidation process by which the oxide is grown. There is a nonstoichiometric transition layer in the oxide near the Si/SiO₂ interface, due to the incomplete oxidation of the Si. An early discussion of this layer was provided by Deal et al., in a 1967 review article, which was based on work done even earlier. Basically, Si right at the interface cannot be oxidized completely because of the lattice mismatch, so there is always some excess Si. As oxide grows, the interface moves, so the excess Si, in effect, moves into the oxide, where it is eventually consumed. But new excess Si is always being formed at the interface, so a steady state concentration profile of excess Si is eventually established in the interfacial region, with stoichiometric SiO₂ in the bulk oxide. They concluded that there were positively charged defects associated with the excess Si, with energy levels apparently outside the Si bandgap. Etchback experiments showed these defects to be within 20 nm of the interface (and the total oxide thickness was much greater in those days). Of course, in present-day samples, the total oxide thickness is typically much less than 20 nm, so this picture has to be modified somewhat. But the process-induced positive charge described by Deal et al. is very consistent with our present understanding of radiation-induced trapped positive charge, except that it was not induced by radiation. Deal et al. also noted that low temperature oxidation increased the concentration of excess Si at the interface, and high temperature annealing tended to reduce it. We will discuss process dependences later, but we note that radiation hardened oxides often use low
temperature oxidation and low temperature annealing, both conditions which increase the excess Si concentration.

Later, Grunthaner et al performed a series of XPS (x-ray photoemission spectroscopy) studies to determine the physical and chemical properties of this interface layer, which they summarized in a review article. They also summarized dozens of experiments by others which also supported the idea that there is a unique, nonstoichiometric transition layer with excess Si (or oxygen deficiency, depending on one's point of view) with a high degree of local strain (see bibliography of reference 4).

III. Early Spectroscopic Studies

Spectroscopic studies on quartz and bulk glasses have identified a large number of paramagnetic defects, which could be studied by electron spin resonance measurements. Of these, the one that correlates with radiation-induced oxide positive charge trapping, is the so-called E' center. This correlation was first shown by Lenahan and Dressendorfer, and later confirmed by others. The E' center was first detected by Weeks in α-quartz, and associated with a single oxygen vacancy by others. At first, the idea that the center was due to a single oxygen vacancy was controversial, until Feigl et al. showed that it was correct. The basic idea is illustrated in Fig 1. Two Si atoms are joined by a weak, strained Si-Si bond, where an oxygen atom should be, and each is back-bonded to three oxygen atoms. When a positive charge is captured, the Si-Si bond is broken, and the lattice relaxes. The new idea introduced by Feigl et al. was that the lattice relaxation is asymmetrical. The positively charged Si relaxes back into the plane of its oxygens, or even past it, so that it puckers in the opposite direction. The neutral Si, with an unpaired electron in one orbital, relaxes toward the vacancy. Then Lenahan and Dressendorfer showed that oxide trapped charge, measured as $\Delta V_{MG}$, the voltage shift at midgap, was removed at the same rate as the E' signal in a series of isochronal annealing measurements, shown in Fig. 2. The correlation between these two measurements is strong evidence that the E' signal and positive oxide charge are due to the same defect, and the correlation has been observed repeatedly on different samples by now.

IV. Annealing Studies

Although oxide trapped charge is sometimes referred to as "fixed" charge, this description is correct only in a relative sense. That is, radiation-induced oxide trapped charge is stable on the time scale of many experiments, but it undergoes a long-term annealing process, which has a complex dependence on time, temperature and applied field. The annealing process can proceed by either of two mechanisms, tunneling or thermal excitation. The tunneling process
Figure 2. Correlation between $E'$ signal and trapped positive charge in annealing experiments.

can be analyzed in terms of a tunneling front, as illustrated in Fig. 3. The probability of tunneling to a particular trap site is an exponential function of the barrier height and the depth of the trap in the oxide, there will be position in the oxide at any given time, where the holes to the right will have been neutralized, and the holes to the left will remain. This position, called the tunneling front, is actually about 0.2 nm wide, and moves into the oxide with a velocity of about 0.2 nm per decade of time, according to model predictions. This $\ln t$ dependence provides an analytical basis for the $\ln t$ annealing behavior observed empirically in many MOS oxides.

Figure 3. Tunneling front analysis; at any time there is a position where trapped positive charge to the right has been neutralized by electron tunneling, and trapped charge to the left remains.

Thermally activated annealing occurs when a carrier is thermally excited from the SiO$_2$ valence band to the trap level, neutralizing the trap. Both the tunneling front model, and the thermal emission model were combined in an elegant fashion by McWhorter et al., which is illustrated in Fig. 4. Basically, they show both a tunneling front and a thermal activation front, where the position of both fronts varies as $\ln t$. Clearly, both mechanisms can lead to $\ln t$ annealing behavior, but they require different trap energy distributions to do so. For a trap
with a single, discrete energy level, the tunneling process will predict a straight-line $\ln t$ annealing response, if the distribution of traps is uniform with increasing depth into the oxide. This result is often observed in unhardened, commercial oxides. If the spatial distribution of traps is peaked at the interface, and drops off with increasing depth in the oxide, the slope of the annealing curve will decrease with increasing time, which is often observed in hardened gate oxides. To predict $\ln t$ annealing by the thermal process, one has to assume a uniform distribution of trap states across the bandgap (as McWhorter has done in Fig. 4), rather than a discrete energy level. The distribution of trap states as a function of energy level in the bandgap has been measured for many oxides, and there are well-defined peaks rather than a uniform distribution. But the peaks have a finite width, so the assumption of a uniform distribution may be justified in a limited range, but not in general. The response of all oxides is temperature independent at room temperature and over a reasonable range of higher temperatures. If one raises the temperature enough, the thermally activated process will become important sooner or later in all oxides. The temperature at which thermal activation becomes important varies from one oxide to another. In some oxides thermal activation is important even before 100 C. In others, the response remains independent of temperature well above 100 C. But by 200 C, all oxides will undergo significant thermal annealing in a short time.

IV. Negative Bias Annealing and the Oxygen Vacancy Dipole Model

Although the annealing process for trapped holes had been studied, and understood to a degree, the literature contained little discussion of what happened at the atomic level during annealing. Many researchers seem to have assumed, without saying so specifically, that an electron tunneled to a trapped hole, neutralized it, and reformed the broken Si-Si bond, so that the damage to the oxide was completely removed. A critical result, shown in Fig. 5, forced many new studies, and eventually led to many new insights. Schwank et al. annealed an irradiated sample at +10V and 100 C for about a week, until all the trapped positive charge (as determined by $\Delta V_{BD}$) had annealed. Then they reversed the bias, to negative 10V, and continued annealing at 100 C. Within about a day, nearly half of the "annealed" positive charge was restored. This result indicated that the annealing process involved some kind of compensation process, because the defects were neutralized without being removed (and unirradiated controls showed no such instability). Eventually, Lelis et al. carried out a careful study of this effect.
Figure 5. Trapped hole annealing results, showing that annealed holes are not really removed.

Figure 6. Alternate positive and negative bias annealing, showing reversible annealing over multiple cycles.

One of their key results is shown in Fig. 6, which was obtained by irradiating a hardened oxide with a short Linac pulse, and annealing with the applied field alternately ±1 MV/cm. The negative bias reverse annealing amounts to about 0.5V, which is significant, but it is clearly only a fraction of the total annealing. They interpreted these results as indicating two different processes—a permanent "true" annealing, and a reversible compensation process, where trapped charge was neutralized without being removed. The charge transfer seemed to be more or less constant over several cycles, and roughly logarithmic with time. These observations were taken to mean that charge was tunneling in and out of a hole trap defect in a more or less reversible way. The model proposed by Lelis et al. to explain these results is illustrated in Fig. 7. They propose that the electron tunneling into the hole trap goes to the neutral side of the E' center, creating a dipole structure, a positively charged Si adjacent to a negatively charged Si. In the interface region where the E' centers are found, there is a high degree of local strain because the lattice mismatch between the Si substrate and the oxide is being accommodated. Therefore, one would expect a distribution of separation distances between the Si atoms in the E' centers. Lelis et al. suggested that if the positive and negative Si atoms were close together, the coulomb attraction would be sufficient to pull them together,
reforming the broken bond, resulting in true annealing. If the relaxation were greater, so that the Si atoms were far apart, then the dipole would be stable, or at least metastable, and it could exchange charge with the substrate in a reversible manner, as in Fig. 6. Other results, not shown here, also suggested that if one cycled charge long enough, some of the dipoles would undergo true annealing—that is, even if the Si atoms were relatively far apart, random thermal vibration might eventually bring them close enough together to reform the broken Si-Si bond. Lelis et al. also noted that there was significant variation from one oxide to another in the amount of negative bias reverse annealing. The only unhardened, commercial oxide in their initial set of samples had almost no reversible annealing, and the hardened oxides all had a significant amount. This pattern held when other oxides were tested later. We will return to this point later, in Section X.

The dipole hypothesis proposed by Lelis et al. was attractive because it explained complex annealing behavior, true annealing and reversible annealing, in terms of a single defect that was already well known. It was consistent with the existing spectroscopic data, which showed only a single defect contributing to the hole trapping. The ESR technique detects unpaired electron spins, and the unpaired spins are eliminated in the dipole structure as the positive charge is compensated. In fact, the spectroscopic data was one thing that led Lelis et al. to propose dipoles instead of having the electron tunnel to the positive Si, without reforming the broken bond. Then the number of unpaired spins would increase, not decrease, as the positive charge was compensated. Almost immediately, independent experiments by other groups began to provide experimental results that could be interpreted as supporting the dipole hypothesis, and we will discuss some of these in the next several sections. But we note that it was also controversial for several years. Three different groups presented papers which, at least implicitly, criticized it. We cannot discuss these papers fully here, but reference 2 contains a relatively complete discussion. Freitag et al. and Edwards concluded that the electrical switching behavior was due to some other (unknown) defect, not to E centers. And Warren et al. concluded it was too early to say where the electron tunneling into the hole trap went (in 1994, six years and many confirming experiments later), that the topic needed further study. The experiment that settled these controversies was reported by Conley et al. in 1995, seven years after the first dipole model paper by Lelis et al. It is discussed in Section IX, below.

V. TSC Studies
There have been numerous TSC (thermally stimulated current) measurements on MOS devices, because the technique is a good way to determine the energy levels of trapped charges. Two of these studies are particularly relevant here. The first of these was done by Shanfield et al.\textsuperscript{22} in 1988, about the same time as the first paper by Lelis et al., but not published in the open literature until later. Shanfield et al. used capacitor samples from the same source as some of the transistors used by Lelis et al., and they obtained results very consistent with those of Lelis et al. That is, they found significant compensation in the hardened oxide, but not in the unhardened oxide. The key point is that in a TSC measurement, one measures the current flow of charge freed by thermal excitation. If both electrons and holes are trapped, the current reflects the sum of these two components. In a C-V measurement, the voltage shift reflects the net trapped charge—that is, the difference between the trapped electrons and holes. If $Q_{\text{TSC}} > Q_{\text{CV}}$, then it is evidence for significant electron trapping. In the hardened oxide, where Lelis et al. had seen significant charge switching, the TSC results also indicated significant electron trapping. In the unhardened oxide, where Lelis et al. had seen no switching, and no evidence of compensation, Shanfield et al. saw no evidence for compensation in the TSC results, either. (Actually, the TSC results on the unhardened oxide were difficult to interpret because space charge effects had not been accounted for, but the results did not indicate the presence of electron traps with or without correction.) For this reason, the Shanfield et al. experiments were the first independent confirmation of the Lelis et al. model, in the sense that the model could easily explain the results, but they were difficult to explain with any other model.

Later, Fleetwood et al.\textsuperscript{21} did another TSC measurement, intended to repeat and extend the Shanfield et al. work. They measured TSC as a function of applied bias, and confirmed that space charge effects had been significant in the work on unhardened oxides by Shanfield et al. They used samples from a different source than Shanfield et al. or Lelis et al., but also had a hardened gate oxide, and a very soft, thick field oxide. They found significant compensation in the hardened oxide ($Q_{\text{TSC}} > Q_{\text{CV}}$), but not in the soft oxide ($Q_{\text{TSC}}= Q_{\text{CV}}$, within experimental uncertainty). The Fleetwood et al. experiment was the third time a different group had reported the same basic result, compensation in hard oxides but not in soft oxides, and for the second source of samples. We will discuss this further in Section X.

VI. 1/f Noise Measurements

Many authors have used 1/f noise to characterize radiation damage in MOS devices, starting with Schofield et al.\textsuperscript{24} in 1989. They concluded that 1/f noise correlated with oxide-trapped charge, not with interface trapped charge. This conclusion may have been surprising to some initially, but it has since been confirmed by others. Schofield et al. cited the first paper by Lelis et al.\textsuperscript{13} to support the idea that there is an electron trap associated with the hole trap that can exchange charge with the substrate. Fleetwood et al.\textsuperscript{25} have recently presented further analysis, DFT (Density Functional Theory) modeling, which suggests several possible mechanisms which may account for different experimental observations. They note that capture cross sections inferred from 1/f noise, hole trapping, and trapped hole neutralization are all about the same, and the energy levels for the defects are also about the same; all of which suggests that the defects actually are the same. They argue that the $E_{\text{g}}'$ and the $E_{\text{g}}''$ can both contribute to 1/f noise. They distinguish between the $E_{\text{g}}$ and the $E_{\text{g}}'$, where the $E_{\text{g}}'$ has the neutral silicon atom puckered, and back-bonded to a fourth Si. The $E_{\text{g}}'$ is similar, except that a nearest neighbor Si is also close enough to interact, which shifts the energy levels so that the complex no longer forms a stable dipole. They suggest that the noise is mediated by lattice relaxation, as the structure shifts back and forth between these two configurations, it
gains and loses the dipole electron. The $E_{a'}$, or dimer, may account for an unexplained observation, namely, that pre-irradiation 1/f noise correlated with post-irradiation hole trapping. If the oxygen vacancy precursor loses an electron to the substrate, it would create an $E_{a'}$ center initially, which could alternately gain and lose an electron, which would also be mediated by lattice relaxation. These ideas from the modeling have not been tested experimentally, but they provide plausible explanations for a number of observations.

![Figure 8. Positive trapped charge and neutral centers as a function of dose in rad (SiO$_2$).](image)

VII. Neutral Electron Traps

Electron trapping in oxides exposed to intense radiation during processing, either by electron beam or x-ray lithography, or by plasma etching, ion implantation, reactive ion etch (RIE), or other processes, has been studied for more than 25 years. Early studies showed that radiation in processing created large amounts of positive charge, which could be annealed fairly easily. But after the positive charge was removed, large numbers of neutral centers remained. These neutral centers could trap electrons during normal device operation, and they were very difficult to remove. An important early insight was provided by Aitken, who suggested the traps were dipolar in nature—the positive end of the dipole capturing electrons, and the negative end capturing holes. He also suggested that a distribution of separation distances between the ends of the dipole would lead to a wide range of field strengths, explaining why electron capture cross sections vary by orders of magnitude, even when measured by the same experimenter on the same sample. (The idea that electron traps and hole traps have a common origin was later supported by Aslam, who showed that process steps that increase the density of one also increase the density of the other.) In 1980, when Aitken suggested a dipole, there was no dipolar structure that could be suggested, that seemed to be relevant. However, Lelis et al. did eventually propose one. An experimental study was eventually done by Walters et al., to test the idea that the dipole structure proposed by Lelis et al. was the neutral trapping center proposed by Aitken. One of their key results is shown in Fig. 8. He measured positive charge trapping, and electron trapping as a function of dose, using optically assisted electron injection. Positive charge is observed to buildup first, but the number of positive centers starts to saturate around $10^8$ rads. In terms of the dipole model, this means that electron-hole pairs are being created in a region with a significant density of
trapped positive charges, which recombine with the electrons to form dipoles. The relative saturation means that the positive charges are being consumed almost as fast as they are created, but then the number of neutral centers is growing rapidly. Then, as they continued to inject electrons, the amount of electron trapping increased very rapidly, becoming the dominant effect by $10^7$ rads. Walters performed the experiment in Fig. 8 with no bias applied during irradiation, but he also irradiated with bias. And he found that positive bias, which increases hole trapping, also increased the formation of neutral centers. He also found that the spatial distribution of neutral centers was the same as the distribution of hole traps, within his experimental resolution. For these reasons, he concluded that the dipole structure proposed by Lelis et al. was the neutral electron trap. Under appropriate experimental conditions, the positive end of the dipole could capture a second electron, making the whole complex net negatively charged. These results were a significant confirmation of the dipole model, and also a significant extension of it. Certainly, the fact that another group had used the model to make a prediction, and that was tested and confirmed experimentally was important independent support for the dipole model. And by showing the connection of the dipole model to electron trapping, it was shown to be relevant to processing and reliability problems outside the traditional radiation effects area. Walters was the first to propose any specific structure for an electron trap, so his work was a breakthrough in that area.

Figure 9. Annealing of neutral electron traps by exposure to hydrogen.

Efforts to anneal neutral electron traps have mostly involved attempts to passivate the centers by annealing in hydrogen rich ambients. Results of some of these experiments are summarized in Fig. 9. For the oxides in these experiments ($t_{ox} = 35$ nm), the initial NET density was about $1.3 \times 10^{12}$/cm$^2$. The samples were then exposed to a series of annealing experiments, where each test involved more hydrogen than the one before, the number of unpassivated defects was reduced in a consistent manner. But even after extreme measures (annealing in 50 atm forming gas, which is equivalent to 5 atm pure hydrogen), more than $10^{11}$/cm$^2$ NET remained. The simplest view of how these results fit together is illustrated in Fig. 10. Parts a, b, and c are the same as Fig. 7; part d is the electron trap, as described by Walters; and part e is the neutral electron trap (dipole) passivated by reaction with hydrogen. The experimental difficulty in getting the hydrogen passivation reaction to go to completion suggests that the passivation and depassivation reactions are in equilibrium, and the depassivation reaction is strong enough that the defects are never fully passivated.

In a wafer that has been through a CMOS process, one would expect a large number of trapped holes (part b in Fig. 10) at some point, because radiation is so common in CMOS processing. (In fact, DeLaus made the comment that the most severe radiation environment
most parts ever see, is in the process line.) High temperature processing will neutralize these by exciting carriers from the valence band, leaving a high concentration of dipole states (part c). Many of these will be passivated, forming Si-H bonds (part e), because hydrogen-rich ambients are commonly used in processing. But the reaction is not efficient enough to passivate them all, so many NETs will remain, which can subsequently trap electrons (part d). The arrows between part c of Fig. 10, and parts d and e, run in both directions because other work by Thompson and Nishida and by Hsu et al. Thompson and Nishida studied the detrapping of trapped electrons as a function of temperature and applied field, and showed that the traps are shallow, and relatively easy to empty. Hsu et al. also showed that in hot electron injection experiments there is a trap filling process, and a trap creation process. That is, there are pre-existing traps which can be filled, consistent with the transition from part c to part d in Fig. 10. And if the injected electrons are energetic enough, they can create “new” traps, by breaking a Si-H bond, for example (10(e) to 10(c) or 10(d)). The dipole model, together with follow-up experiments by others, provide a framework, then, which can be very useful for interpreting a variety of reliability physics experiments. We will discuss this further in Section XI.

VIII. Recent Spectroscopic and Theoretical Studies

Two ESR spectroscopic studies and two theoretical modeling efforts bearing on the Lelis et al. dipole model have been published, several years after the model was originally published. The first of these was by Warren et al., which concluded that the “fixed” oxide charge and the switching oxide charge were probably both $E_{1/2}$ centers. Of course, this conclusion is completely consistent with the dipole model, but we should note that Warren et al. did not accept this idea. They pointed out that Edwards had argued that the dipole configuration was energetically unfavorable, because of the electron-electron repulsive energy. And that Fleetwood had studied the charge exchange process, without reaching a conclusion on exactly where electrons tunnel to, when they tunnel into a hole trap. For these reasons, they concluded it was premature to specify the exact location of the tunneling electron, and that the question should be studied further.

The second ESR study was by Conley et al., who monitored the $E_{1/2}$ signal during a series of alternating bias anneals. The samples were irradiated under positive bias, then annealed unbiased for a few days, then annealed again under alternating bias, negative bias first.
Figure 11. E' density during alternating positive and negative bias annealing.

Typical results are shown in Fig. 11. The E' signal increases under negative bias, and decreases under positive bias, the same as positive oxide trapped charge in electrical measurements by Lelis et al., and others. This result is very strong evidence for the dipole model, because the unpaired spin detected by ESR and the positive charge detected electrically are increasing or decreasing together, even though they reside on different Si atoms. The dipole model is the only model that has been proposed, which is consistent with this result. (We note that Fleetwood et al. has also proposed hemi-E' centers as possible trap structures—in effect, half an E' center. The trouble with these structures is that they have an unpaired spin in the neutral state, not in the charged state. In Fig. 11, they would have valleys where the peaks are, and peaks where the valleys are. Fleetwood later agreed these centers are not consistent with the results in Fig. 11.37) From this result, it is clear that the dipole structure of Lelis et al. is the dominant switching oxide trap, or border trap. Of course, it is always possible that other structures will later be shown to produce small, second order effects.

Two recent theoretical studies have also supported the Lelis et al. model. The first of these was by Kama et al., who concluded, based on first principles quantum mechanical calculations, that the E' center can take two forms. The first of these is the E' center, which is a dimer structure where the hole is delocalized—spread over both Si atoms. This structure is symmetrical, with the Si atoms not relaxing far away from each other. When an electron tunnels to this structure, the broken bond reforms, and the trap is permanently annealed. The second form is the E', where the Si atoms relax asymmetrically, and one of them ends up puckered away from the other, back-bonded to another oxygen atom. The electron-electron repulsion energy means this state has a higher energy than the ground state, but it forms metastable dipoles. This calculation was the first to provide a theoretical explanation for stable dipoles.

The second theoretical study was by Nicklaw et al., who used density functional theory (DFT), a different mathematical method, but reached very similar conclusions. The main difference is that, depending on the positions of nearby Si neighbors, the E' can assume two configurations, and only one of them produces a metastable dipole.

Taken together, these four follow-up studies by different groups have provided strong additional support for the dipole model of Lelis et al.
IX. Process Dependencies

As we have noted already, some soft, commercial oxides seem to have no significant charge switching, or compensation, while hardened oxides generally do. In fact, Conley et al.\(^{21}\) reported that test chips from three high volume commercial process lines (Texas Instruments (TI), National Semiconductor, and Micron) had been evaluated, and none showed any significant radiation-induced switching oxide traps. This is despite the fact these are all relatively soft oxides, with a high density of other oxide traps. On the other hand, three hardened oxides ((TI), Sandia National Laboratories, and IBM/Loral/BAE) all did show switching behavior by a significant fraction of their oxide traps. The TI example is particularly instructive, since the hardened and unhardened oxides were from the same split lot, and differed in only one process step (the post-oxidation anneal, or POA). See reference 2 for a full discussion, but the main difference was that the hardened oxide received a POA at a lower temperature. The different annealing responses of these two oxides are shown in Fig. 12, where we have plotted $\Delta V / V_{ox}$ because the oxide thicknesses are different. If one compares the response at a hundred seconds, or at a thousand seconds, the hardened oxide has about an order of magnitude smaller shift. But at earlier times the difference between the oxides is smaller, and if one extrapolates back to the end of the radiation pulse (4 ps), it is possible there is no difference at all in the trapping of the two oxides. This conclusion depends on how one extrapolates, or whether one believes the extrapolation, but there is clearly less difference earlier than there is later, because the hard oxide recovers faster. These results suggest there is actually little or no difference in the number of oxide traps in the hard and soft oxides, but they are closer to the interface in the hard oxide, and so are neutralized by tunneling electrons more quickly. Recall that the hard oxide in Fig. 12 has significant negative bias reverse annealing, or charge switching, and the soft oxide has essentially none; and that these oxides differ in only one process step. So the process change that makes the hard oxide hard also introduces the radiation-induced negative bias instability, and increases the annealing rate.

![Figure 12. Normalized post-irradiation responses of “hard” and “soft” oxides, which differ only in one process step (POA).](image)

We note that Deal\(^{16}\) discussed a process-induced negative bias instability, which he attributed to excess Si near the interface. This was originally an important commercial reliability problem, but Deal showed that the instability could be eliminated by a high temperature (at the oxidation temperature) POA in an inert ambient, as was done on the unhardened oxide here. He inferred that the effect of the high temperature anneal was to remove the excess Si from the interface region. The reason high temperature anneals have been a fixture in
commercial processes since Deal's early work is that he (and others later) showed it did remove excess Si from the interface region, eliminating an important instability. It is important to realize that excess Si in the interface region is an important commercial reliability problem, even in the absence of radiation. But in the results in Fig. 12, the hard oxide is hard because the instability has been built back in by eliminating the normal POA. It has been argued that high temperature annealing steps increase the excess Si concentration at the interface by reducing the oxide. While this is certainly true if the temperature is high enough, it seems not to be a significant effect at the temperatures used in these oxides. Instead, the effect of the high temperature POA seems to be to cause the oxygen vacancies to diffuse into the oxide, away from the interface, making them less active. But they still contribute to radiation damage for a long time in the soft oxide, because they are so stable. For a more complete discussion, see Chapter 2 of reference 2, and its bibliography.

X. Reliability Physics and Electron Trap Creation

We have already introduced the subject of electron trap creation, citing Hsu et al. Their work was consistent with earlier work by DiMaria, who showed that when electrons are heated above about 2 eV by the field, Si-H bonds start to be broken. Although the average electron energy never reaches 9 eV, there is a high-energy tail to the distribution above 9 eV, which gives rise to impact ionization. The important application of these results is in breakdown studies. Generally, the oxide is considered to go through a wear-out phase, during which trapped charge builds up, followed by breakdown. At one time, there was some controversy over the details of the wear-out process, with some authors arguing that impact ionization led to positive charge buildup as the dominant mechanism, others argued against impact ionization, and for a buildup of stable electron traps as the main cause of breakdown. More recently, it has become clear that both positive and negative charge buildups can occur, and both can cause breakdown. The work by DiMaria shows that both processes can occur at the same time. The basic idea is that charge builds up near the injecting electrode, usually the Si substrate. Then a very high field develops between the oxide trapped charge and the image charges at the interface. When the field becomes large enough, avalanche multiplication starts, which leads to local heating, which destroys the oxide. Clearly, both polarities of charge can develop large space charge fields, and both can be explained is different aspects of the same oxygen vacancy defect complex we have already described. The oxygen vacancy model presented here may not explain all the breakdown results in the literature, but it seems to provide a framework, which can account for many of the known results.

XI. Oxide Traps as Interface Traps

A discussion of interface traps is beyond the scope of this paper, but there is one area where the dipole model helped to resolve a controversy about radiation-induced interface traps. There was a school of thought in the 1980s and early 1990s, which supported the idea that interface traps arose from a defect conversion process, that oxide traps somehow were converted to interface traps. One weakness of this argument is that no one was ever able to propose a plausible mechanism for the conversion process. In a review paper, Oldham et al. summarized other arguments against this idea, concluding that hole removal and interface trap formation had different time dependences, different field dependences, and different temperature dependences. Except for the fact that both processes increased with increasing dose, they appeared to be completely independent. Oldham et al. did, however, offer a new explanation of experimental results, which had been offered to support the defect conversion idea. They pointed out that no defect conversion was really necessary—oxide traps that
remained oxide traps could account for the observations. The Lelis et al. dipole model, and the charge transfer mechanisms built into it, could account for the results. The fundamental question was how to define an interface trap. Most researchers took an interface trap to be a state in the Si band-gap in equilibrium with the substrate, meaning that it responded to small voltage changes, such as the 15 mV AC probe voltage in a high frequency CV measurement. Oxide traps (dipoles) do not respond to small voltage changes because they are not in equilibrium with the substrate, but they do respond to large voltage changes (e.g., from accumulation to inversion, or the reverse). For this reason, oxide traps were being interpreted as interface traps in some experiments.

A few years after Oldham et al. presented this discussion, Fleetwood proposed that oxide traps with these electrical properties should be called “border traps.” At that time, the first Lelis et al. paper proposing the dipole model had been in the literature four years, and was already well-known. Now ten more years have passed, and the dipole model is still the only experimentally confirmed model for border traps (or for switching oxide traps, as we have usually preferred to call them here).

XII. Oxide Thickness and Scaling

In the history of the semiconductor industry, scaling to smaller feature sizes has been a pervasive theme—almost everything has depended on it. One consequence of scaling has been that the gate oxide has become thinner every year, and thinner oxides are less sensitive to radiation damage. From Q=CV, one would predict that $\Delta V_T$ is proportional to $t_{ox}^{-2}$. McGarrity estimated the hardening that could be achieved by thinning the oxide, without special processing. The most important deviation from the thickness squared dependence is that below about 10 nm, $\Delta V_T$ falls off even faster than $t_{ox}^{-2}$ would predict. Almost all the trapped positive charge is neutralized by tunneling electrons, because the tunneling distance is a large fraction of the total thickness, and tunneling from both electrodes is significant. Present day gate oxides are so thin, that gate oxide hardening is no longer a practical problem. The main radiation problem now is in the isolation structures. In oxide isolation structures, the same defects and the same physical mechanisms determine the response, but there are differences because all the mechanisms have complex time dependences, field dependences, and temperature dependences. For example, the fields are lower, which means the yield of charge from recombination is less, the transport is slower, thicker oxide also means slower transport and so on. Generally, the response of LOCOS (LOCal Oxidation of Silicon) field oxides varies widely, probably because the processing varies widely. STI (shallow trench isolation) structures are replacing LOCOS, but have not been studied as much.

One area where the dipole model has been used effectively is in understanding the ELDRS (enhanced low dose rate sensitivity) of some bipolar isolation oxides. ELDRS is at least partly due to space charge effects, and to different degrees of compensation (dipole formation) at different dose rates.

One radiation effect in the gate oxide has been reported, which is a consequence of gate oxide thinning—RILC (radiation induced leakage current). This leakage current is thought to be a form of trap-assisted tunneling, where an electron tunnels from the substrate to a trap in the oxide. Then it tunnels from the trap to the other electrode, contributing to gate oxide leakage. Of course, the oxygen vacancy dipole model is an obvious candidate for the trap state. The same effect is observed when electrical stresses, rather than radiation, generate a trapped hole or a dipole state. Then it is called SILC (stress-induced leakage current).
XIII. Conclusions

The oxygen vacancy dipole model, proposed by Lelis et al., was developed in response to radiation experiments showing what was first called negative bias reverse annealing. But it has been confirmed in numerous other experiments, and extended by other groups. The later work by Walters, showing the connection with neutral centers and electron trapping is especially significant. The dipole model is the basis for a comprehensive model of oxide charge trapping (of both polarities), applicable to many oxide reliability problems. Although the model grew out of radiation effects studies, it is now of interest to a much wider community, for this reason.

References