ELECTRON TRAPPING IN GATE DIELECTRICS AND NBTI OF MOSFETs

MO HUAI CHANG

A thesis submitted in partial fulfillment of the requirements of Liverpool John Moores University for the degree of Doctor of Philosophy

November 2006

Acknowledgements

I would like to thank Prof. J. F. Zhang for his continuous guidance, advice, encouragement and fruitful discussions throughout the project. Without his investment of time and efforts, this project would never reach the fruition.

I also deeply appreciate Dr. C. Z. Zhao for his advice and support throughout the project. Many thanks are given to other members in our research group, Dr. W. D. Zhang, Mr. M. Zahid and Mr. Y. G. Wang for their extreme helpfulness.

Finally, I would like to express my deeply gratitude to my family, especially my parents for their understanding, support and encouragement during all these years.

Summary

The transistors in integrated circuits have increased in speed and reduced in cost and power consumption because of reduced transistor size in successive technology generations. However, it is not possible to reduce the voltage to the same scaling factors so the electric fields have increased. At elevated field, carriers can be easily injected into the oxide and lead to a gradual degradation of the oxide. Generated defects can trap some of the injected carriers and change the electrical behavior of the MOSFET. Eventually a sudden breakdown occurs as the generation sums up to a critical amount.

Generation of acceptor-like electron traps in gate oxide is an important source for the instability of MOS transistors. Agreements have not been reached on the dominant damaging species. When injected electrons were orders of magnitude higher than injected holes, it was proposed that hydrogen release and its subsequent transportation through the oxide dominated the generation. It was also reported that holes were more efficient in creating electron traps than electrons. However, the physical process for the hole-induced generation is not clear. The release and subsequent transportation of hydrogenous species for hole-induced electron trap generation is investigated in Chapter 3. Effects of hydrogenous species released near the two interfaces and in the bulk of oxides are examined. It is found that the release and subsequent transportation of hydrogenous species are not important for the hole-induced generation. Results support that holes can interact directly with the oxide to generate electron traps without going through hydrogen as intermediate species.

Furthermore, the capture cross sections of generated acceptor-like electron trap are not unambiguously determined and there are confusions on how many capture cross sections genuinely existing. The dependence of trap density for a given capture cross section on stress level is not clear, either. To fill to knowledge gap, the electron trapping kinetics is investigated in Chapter 4. There are a number of obstacles for

such an investigation, including the simultaneous occurrence of trapping and trap generation, stability of trapping, and effects of positive charges. Through careful selection of experimental conditions and testing samples, we have been able to overcome them. In particular, recent work at this university in this area has allowed us to develop a new method for correcting the effect of positive charges. After removing all uncertainties, a capture cross section as large as $10^{-13} \sim 10^{-14}$ cm^2 is found for the generated acceptor-like trap. It is shown that electron trapping follows the first order model and there is also a smaller capture cross section in the region of $10^{-15} \sim 10^{-16}$ cm^2. For the first time, it is shown that the density of the larger trap increases with stress, but the density of the smaller trap clearly saturates.

As the nitrogen concentration in silicon oxynitrides (SiON) increases, the negative bias temperature instability (NBTI) becomes a limiting factor for device lifetime. Despite recent efforts, there are confusion and issues remaining unsolved. One of them is how important positive charge formation in SiON is for NBTI and whether all positive charges are the same type. Positive charges formed in SiON during negative bias temperature stresses (NBTS) is investigated in Chapter 5. It is shown that NBTS can induce three different types of positive charges: as-grown hole trapping, anti-neutralization positive charges (ANPC) and cyclic positive charges. Efforts have been made to search for the features of NBTI, which cannot be explained without involving positive charges. It is unambiguously identified that the impact of measurement temperature on NBTI originates from only one types of defect: ANPC. By using 'On-The-Fly' measurement technique, the positive charge density observed in a 2.7 nm SiON can reach the up half of 10^{12} cm^{-2}, which is comparable with the positive charges reported for relatively thick SiO_2 (> 5 nm). The relative importance of positive charge formation depends on measurement interruption time. The shorter the interruption, the more important positive charges become for NBTI.

1 A Review of the Degradation of Gate Dielectrics in MOSFETs

1.1. Introduction

The advance in MOS devices has improved the quality of our life significantly in the last three decades or so. The success of MOS devices heavily relies on the excellent insulating properties of gate silicon dioxides and their near perfect interface with silicon. The gate silicon dioxide is a amorphous insulator with a high bandgap of about 9 eV and high energy barriers against electron and hole injection from silicon [1]. It can be easily grown on Si with low defect density. The excellent scaling and process integration of gate silicon dioxides are mainly due to the stability and insensitivity of the insulator to process steps following the oxidation.

In a continuous drive to increase integrated circuit performance through shrinkage of the circuit elements, the dimensions of MOSFETs have been scaled down following a trend known as Moores's law [2-4]. Moores's Law describes the exponential growth of chip complexity due to decreasing minimum feature size, accompanied by concurrent improvements in circuit speed, memory capacity, and cost per bit. To maintain the gate control and reduce the short channel effects, the gate SiO_2 has been decreased in thickness from hundred of nanometers (nm) 40 years ago to less than 2 nm today, as shown in Figure 1.

Gate silicon dioxides are not perfect and suffer from reliability problems. A very narrow region of transition occurs from the mono-crystalline silicon substrate to the amorphous silicon dioxide layer, which contains a high density of non-saturated bonds, strained bonds and broken bonds [5,6]. These imperfections or defects introduce localized energy levels in the silicon forbidden gap at the SiO_2/Si interface. Defects are also present in the oxide bulks. The density of these defects is a function

of processing conditions, such as the growth temperature/pressure and the chemical interaction of these so-called "intrinsic" defects with chemical impurities, such as H and Cl, introduced during IC processing [5]. As the oxide scaling down, its intrinsic reliability limits are approached and might become a major showstopper for future technology.

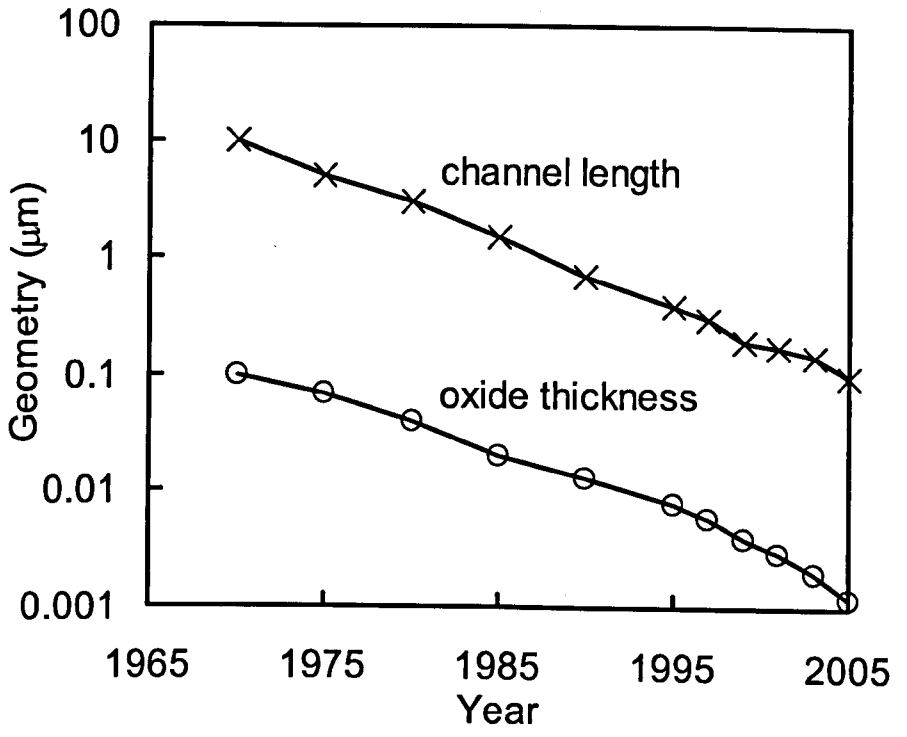


Figure 1. Trend in MOSFET scaling. The channel length and the oxide thickness have been reduced by two orders of magnitude since early seventies.

During device operation, the thin gate oxide is subjected to a high electrical field. As the CMOS technology progresses, Figure 2 shows that this field has gradually increased. Even the lowering of operation voltage for recent CMOS generations has not stopped this trend. Values as high as 6 MV/cm are reached in the current generation of CMOS technology. At this field, carriers can be easily injected into the oxide and lead to a gradual degradation of the oxide. Generated defects can trap some of the injected carriers and change the electrical behavior of the MOSFET. Eventually a sudden breakdown occurs as the generation sums up to a critical amount.

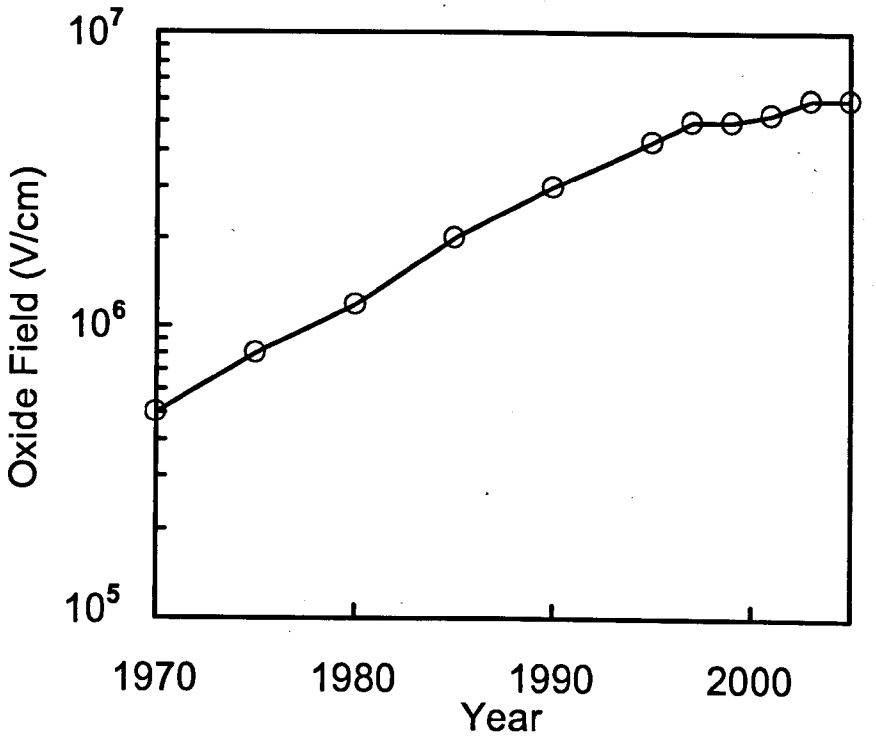


Figure 2. Oxide electric field as a function of time for CMOS circuits. Oxide field have increased from below 1 MV/cm in the early seventies to 6 MV/cm today.

According to their spatial location, generated traps are classified as either bulk oxide traps or interface traps. Interface traps are located at or very near the Si/SiO_2 interface and are able to exchange charges with the silicon bands on time scales ranging from picoseconds to hours [6]. They are distributed in energy throughout the silicon band gap. Their charge states are determined by the interface Fermi-level. Bulk oxide traps are located farther away from the Si/SiO_2 interface and are much more difficult to exchange charges directly with the silicon. Bulk defects include both electron traps and hole traps, which can be either neutral or charged [5].

Generated defects can affect device operation in three ways. First, they can capture electrons or holes and build up space charges, which shift parameters, such as threshold voltage and transconductance [7,8]. Second, electron traps can assist electron tunneling through the dielectric [9,10]. This stress induced leakage current (SILC) can considerably reduce the non-volatile memory retention time. Third, the build-up of electron traps triggers the dielectric breakdown [11,12]. This time-dependent dielectric breakdown (TDDB) and SILC and their relation to the

preceding trap generation have been numerously studied over the past three decades. But even today, a complete understanding has not yet been achieved. The knowledge of the trap generation mechanism and trapping-detrapping characteristics is the key issue for understanding TDDB and SILC.

In this chapter, the important concepts related to the degradation of gate oxides will be reviewed. These include the sources of instabilities, charge carrier injection, interface states generation, electron trap generation, trapping/detrapping, and positive charge generation.

1.2. Sources of instabilities and charge carrier injection

1.2.1. Irradiation

The generation of defects in the oxide and at the interface of MOSFETs under irradiation has been widely reported [14-21]. The radiation source can be vacuum ultraviolet light, X-ray, electron beam or any other high energy sources. The common feature of these sources is that their photon energy is large than 9 eV, the band gap of silicon dioxides [21]. The irradiation induced degradation in MOSFETs begins when ionising radiation is absorbed in the oxide, creating electron-hole pairs. Depending on the kind of irradiation and the polarity of the applied field, some fraction of these charged pairs will undergo recombination process. Electrons that have higher mobility can be swept out of the oxide and the holes of lower mobility will have a high probability of being trapped in the oxide. Degraded characteristics include changes in MOSFET threshold voltage, reduced inversion layer mobility, increased minority-carrier generation and increased low-frequency noise [20].

1.2.2. Plasma charging

Plasma-assisted processes are widely used in the manufacturing of VLSI devices. The degradation of gate oxides in MOS devices due to plasma processing was

reported [22-27]. During plasma processing, ions and electrons are collected by the metal or polysilicon electrodes connected to the gate. Due to this charge collection, a voltage appears on the electrode and causes tunnelling current to flow through the gate oxide. This results in charge build-up, the generation of new oxide traps, and the generation of interface states. H. C. Shin and C. M. Hu [22] suggested that the plasma charging could be reduced by reducing the ion density or the electron temperature. Maintaining a very uniform plasma charging current during the etching, the proper use of protection diodes and the antenna design rule will reduce the plasma damage to an acceptable level.

1.2.3. Fowler-Nordheim injection and substrate hot carrier injection

These are uniform injection techniques, and will be described in details in the section 2.5 of Chapter 2.

1.2.4. Channel hot carrier injection

When MOSFETs operating in saturation, the electric field peaks in the vicinity of the substrate-drain junction due to the pinch-off condition. If the lateral electric field is sufficiently high, strong carrier heating can result in electrons with an energy high enough to create electron-hole pairs by impact ionisation [28-31]. Those hot carriers having energy over 3.2 eV for electrons and 4.8 eV for holes can be injected into the oxide [29]. Hot carrier injection into the SiO_2 depends strongly on the biasing conditions of the device and on the mode of operation (subthreshold, linear or saturation). The device degradation is most severe if it is biased to maximise the substrate current, which occurs at $V_g \approx V_d/2$ [28]. For $V_g > V_d/2$, electron injection is dominant. However, when V_g is close to the threshold voltage ($V_g \approx V_t$), hole injection dominates [30]. In addition to the degradation of transistor performance, electron trapping and related damages also cause significant reliability problems for EPROM devices, such as the "program window closure" effect [31].

1.2.5. Avalanche injection

In avalanche injection [21,32-34], the MOS capacitor is operated in the deep depletion condition, as shown by the energy band diagram in Figure 3. Charge carriers are accelerated rapidly by the applied electric fields. When the gate voltage is high enough to produce avalanche breakdown in the silicon, carriers generated in the depletion layer will be accelerated to sufficient energy for impact ionisation to occur. This creates plasma of energetic hole-electron pairs near the interface. Some of the electrons or holes created will have enough energy to surmount the interfacial energy barrier and enter the SiO_2. The energy barrier for electron injection is 3.2 eV, whereas it is 4.8 eV for hole injection. As a result, electrons have a higher injection probability. Avalanche is a uniform injection technique and is used in MOS capacitors.

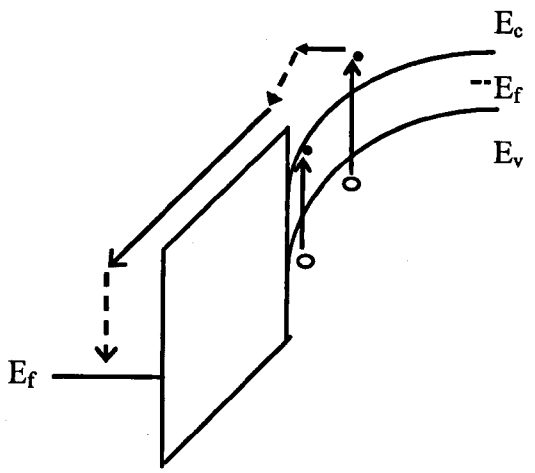


Figure 3. Energy band diagram shows a MOS capacitor operated in the deep depletion condition during avalanche injection.

1.3. Electron traps in the oxide

1.3.1. Electron trap generation

Electron trapping in silicon dioxides has been investigated for the last three decades or so. Early work [35-38] was focused on impurity-related traps. There are three classes of impurities: sodium, implants and hydrogen/water. Sodium introduces both deep and shallow level of electron traps in the oxide [36,39]. For modern semiconductor industry, the use of ultra-clean technology reduces the sodium contamination to a negligible level. The most commonly used implants are Arsenic (As), Phosphorus (P) and Boron (B). Boron does not introduce traps into the oxide, but both As and P do [36,37]. The trap density was reported to be 0.7-1 times of the ion density in the oxide [37]. With careful control of the implantation, these traps are generally not important for modern CMOS processes.

It is most difficult to eliminate hydrogen/water from the device and it is generally believed that it still plays a crucial role in modern device instability. Nicollian et al. [35] showed that diffusion of H_2O into the oxide created a trap with a capture cross section (σ) in the order of 10^{-17} cm^2. When Aluminum is used as the gate, hydrogen also introduces a trap of $\sigma = 10^{-18}$ cm^2 [38]. For a poly-si-gated device used in industry, however, these relatively large traps were absent. The electron trap in the gate oxide of a modern MOSFET has a capture cross-section in the order of 10^{-19} cm^2 or less [40,41].

Since all the 'well-known' electron traps mentioned above are not important in a modern device, previous work on them is of limited use. For a poly-si-gated device, most of electron traps are generated under high field stress. It is these generated traps that cause leakage and trigger breakdown [11-13]. Despite the past efforts, our understanding of the electron traps is still poor. For example, the microstructure of generated electron traps has not been unambiguously identified and agreement on damaging species has not been reached [42-44].

The researchers at IBM [40] believe that the damaging species are hydrogen-related for thin oxides, while the researchers at Lucent Technology [44] believe that holes are the species. For the hydrogen theory, the problem is that there is no convincing direct evidence showing that hydrogen can create new traps in the oxide bulk, although it is well accepted that hydrogen generates interface states at room temperature [45]. For the hole theory, it was proposed that the recombination of electrons and holes led to the generation [46]. It was also shown that the oxide broke down at a constant hole fluence, Q_h, under Fowler-Nordheim injection (FNI) at room temperature [11]. However, this is not strong enough to suggest that there is a causal relation between these two. Q_h is not a constant anymore under substrate hot hole injection [47] or at different temperatures [48]. The impact of hydrogen on hole-induced electron trap generation will be investigated in Chapter 3.

1.3.2. Models proposed on electron trap generation

Several models have been proposed for the electron trap generation. In summary, there are at least four trap generation models: the "hydrogen release model", the "anode hole injection model", the "electric field energy model", and the "electron-hole recombination model".

1.3.2.1. Hydrogen release model

The hydrogen release model has been suggested since 1980's [40,49-51], in which the electrons tunnel through the oxide potential barrier and reach the anode with sufficient energy to release hydrogen from the anode/oxide interface. Hydrogen is always present in sufficient amounts because of interface annealing applied to reduce the initial interface trap density. The released hydrogen diffuses through the oxide and can generate electron traps. This phenomenon is thermally activated above 200K. This model is supported by the observation that, when an oxide (without poly-Si gate) on a silicon substrate is exposed to a remote hydrogen plasma, the interface trap creation evolves similarly to the interface trap creation induced by a hot electron stress [24,52]. It is also supported by the observation of hydrogen buildup in the

dielectric after stress [53], and by the theoretical calculations and experimental observation of the interaction of hydrogen with the oxide lattice [54-56].

1.3.2.2. Anode hole injection model

In the anode hole injection model [57], it is believed that holes injected from the anode dominate the electron trap generation. However, the precise role of holes in the trap generation process and the details of the microscopic mechanism of the trap generation are still uncertain. Difficulties in studying this process arise from the inability of many techniques to separately control the hole and the electron injection. When the gate oxide of a nMOSFET is stressed with a positive gate voltage, with source and drain grounded, electrons tunneling through the oxide are injected from the transistor channel and provided from the source and drain. In this configuration, a positive current can be measured at the substrate (charge separation technique) [58,59]. The substrate current density has a similar oxide field dependence as the FN-current density. The ratio between gate and substrate current depends on the oxide field.

A well-known and widely accepted explanation for the physical origin of this substrate current is given in refs. [57,58,60]. When the injected electrons enter the anode (the poly-Si gate), they lose their energy by creating high energetic holes and the holes can then be injected back into the oxide. The hole flow reaches the cathode and is measured as a positive substrate current. Another possible explanation for the substrate current is the creation of holes in the cathode by photons generated in the anode [60]. In anode hole injection model, the probability for a tunneling electron to generate an anode hole is found to be almost constant during the whole stress under a constant-voltage-stress. This probability slightly increases under a constant-current-stress, where the increase is very small for thin oxides. At room temperature, it has been observed [11,58] that the hole fluence reaches a critical value at breakdown, which remains constant in a wide range of electrical field from 8 MV/cm to 14 MV/cm.

For lower temperature, the hole fluency at breakdown decreases with increasing oxide field [61]. For temperatures above 300K, it increases for higher oxide field [62]. These temperature effects indicate that the hole fluency is possibly not controlling the breakdown. In ref. [60], it is demonstrated that with a gate bias below 7.8 V, the dominant source of the hole current is photo-excitation of valence band electrons in the cathode by light generation in the anode. It is also shown [63] that the injected electrons relax their energy by light emission. Recent work by IMEC [64] shows that most of the holes in substrate current do not traverse through the gate, which means that the substrate current could significantly overestimate the anode hole injection.

1.3.2.3. Electric field model

Some authors have suggested that the electric field itself induces sufficient energy directly into the oxide to cause electron trap creation (electric field energy model) [65-69]. However, experimental results under both tunneling stress and hot electron injection have confirmed that the breakdown of silicon dioxide is related to the electron fluency through the dielectric and not solely due to the interaction of the electric field with the dielectric lattice [70]. Recent studies have also shown a direct correlation between the energy dependence of damage produced using tunneling electron injection and that produced using hot electron injection [71-73]. This means that electron energy could be the dominating factor, rather than the oxide field. These results suggest that the electric field model does not correctly describe the breakdown of ultra-thin oxides.

1.3.2.4. Electron-hole recombination model

It is well known that some injected holes can be trapped and then neutralized by recombining with electrons. The energy released by electron-hole recombination in the oxide could generate electron traps [74,75]. The holes are produced by either anode hole injection or by impact ionization in the oxide layer near the anode [76]. These secondary hole generation mechanism can be observed down to hot electron

energies of 5 eV and 9 eV respectively [76]. Once the holes are generated and/or introduced into the oxide layer, they are mobile and can move throughout the oxide bulk to the interfaces where some are trapped in energetically "deep" sites believed to be caused by oxygen vacancies [77]. Conduction band electrons can continuously recombine with these holes and generate electron traps [21,46,49,58]. This defect generation mechanism is process dependent and appears to be related to the hydrogen content of the oxide [21]. It reaches a steady-state value dictated by the hole generation probability and the recombination kinetics with free electrons. It has been suggested that defect generation by the recombination of an electron with a trapped hole is observed only when electrons are injected into the conduction band of the oxide [21].

1.3.3. Electron trapping kinetics

Generation of acceptor-like electron traps in gate oxides is an important source for device instability. It has been found that electrons injected into the gate oxide have a certain, small probability of being captured in electron traps present in the oxide bulk [40].

Our understanding of the generated trap is still limited, partially because of the lack of a well accepted trapping model. Developing a trapping model is highly desirable for understanding the generated electron traps, since it allows us to determine their effective physical size, namely capture cross section. It will also give us information on how many types of electron traps are generated. Information like this will be valuable when the origin of generated traps is explored and models for oxide breakdown and SILC are developed [11,78].

The most well known trapping model is the first order model [35,37,79] and it has been applied successfully for as-grown electron traps [35,37]. Although it was proposed that this model could also be applied for generated traps [38], this verdict has not been adopted generally. Only limited works [80] were carried out on electron trapping kinetics for poly-si gated devices. Non-uniform electron injection was used and consequently, uncertainty exists in the density of injected electrons [79,81].

In the first order model, it is assumed that the traps may have multiple-capture cross sections. The areal density of the trapped electrons, N, is described by,

$$N = \sum_{i=1}^{M} N_i \cdot \left[1 - \exp\left(\frac{-\sigma_i \cdot Jt}{q}\right)\right] \tag{1}$$

where M is the number of measurable electron capture cross sections, N_i is the saturation density of traps with a capture cross section of σ_i.

The minimum numbers of distinct cross sections and the corresponding trap densities can be determined by fitting with experimental data. Using this approach, the more cross sections are used, the better fitting will be. Oxide defects with trapping cross sections of approximately 10 ~ 100 atomic dimensions ($\sigma \approx 10^{-13} \sim 10^{-14}$ cm^2 for charged traps) and one atomic dimension ($\sigma \approx 10^{-16} \sim 10^{-17}$ cm^2 for neutral traps) have been identified by using this method, which is the reasonable dimensions for coulombic and neutral defect sites respectively. Introduction of more different electron trap cross sections can give better fitting with the test results. However, this may be an artefact. For example, the generation of new traps at very large electron injection level may be responsible for very small capture cross sections. Therefore, it is important to find the supporting evidence for the presence of different cross sections, so that they are physically meaningful.

In Chapter 4, the difficulties of studying the electron trapping kinetics for generated traps will be overcome by a careful selection of experimental conditions and testing samples. A new method is developed to compensate the offsetting effects of positive charge formation on electron trapping kinetics. As a result, the number of capture cross sections can be unambiguously determined. The dependence of trap density for a given capture cross section on stress level is examined.

1.4. Positive charges in the oxide

1.4.1. Positive charge formation

As the downscaling of gate oxides continues, trap density in the oxide bulk will reduce, but positive charges formed near to the SiO_2/Si interface become relatively important. For gate oxides used in industry, hole trapping is the most important process for positive charge formation.

Positive charges can be formed in silicon dioxides in a number of ways. In the early generation of metal-oxide-semiconductor devices, both mobile and immobile sodium ions were found in the oxide [82,83], which caused device instabilities. In a modern MOS device, the sodium ion is eliminated by using the ultra-clean process. However, fixed positive charges are still formed near the oxide/silicon interface during the fabrication. It was suggested that these fixed charges were partially ionized silicon atoms [84], although this was challenged recently [85,86]. Recent work [85-87] shows that positive charges can be formed in the oxide by interacting with H_2 at a temperature over 500°C. The formed charges are fixed when the device has no gate or has a metal gate [86]. However, they become mobile when the oxide is capped by a layer of polycrystalline silicon [87]. It has been proposed that these mobile protons can be used for non-volatile memories [87].

After fabrication, positive charges can also be introduced into the oxide during the device operation. In an irradiative environment, it is well-known that positive charge formation is one of the main sources for device instabilities [88]. They are also formed under a high electrical field (e.g., > 7 MV/cm [38,50]), which is used for programming non-volatile memories [89,90]. Even during the operation of a MOS field-effect transistor (MOSFET) at the relatively low oxide field, formation of positive charges was observed [90].

In addition, two different types of positive charges have been reported [91,92]. One of them is the trapped hole, which, once neutralized, can not be recharged without further hole injection. The other is referred to by various names, including

anomalous positive charges [93,94], slow states [95], border traps [96], and switching oxide traps [97]. The main difference from the trapped hole is that it can be positively charged again under a negative gate bias without switching on the hole injection [91,92]. Formation of anomalous positive charges is generally considered to be a complicated issue and our understanding of it is still poor. For example, it is not known if there is more than one type of anomalous positive charges. There is also a lack of evidences for clarifying the relation between anomalous positive charges and hole traps.

Previous works [98-101] at this university have shown that that generated hole traps behave quite differently from as-grown ones, although they have similar capture cross section in the order of $10^{-13} \sim 10^{-14}$ cm^2 [31,98,102,103]. Three different types of hole traps are identified [101,104]. One of them is as-grown, the other two are created. Generated hole traps consist of two components: anti-neutralization positive charges (ANPC) and cyclic positive charges (CPC). After neutralization, both of them can be recharged positively under a negative gate bias, $V_g < 0$ (-5 MV/cm), without hole injection. When the gate polarity is switched to positive, the CPC can be neutralized at a speed similar to its charging. The neutralization of ANPC, however, is much more difficult. Additionally, the discharging of CPC is insensitive to temperature, while the discharging of ANPC is thermally accelerated. It was proposed that the most important process for hole-trap generation is the direct interaction of injected holes with the oxide [104].

1.4.2. Negative bias temperature instability

Negative bias temperature instability (NBTI) takes place in p-channel MOS devices under negative gate voltage at elevated temperature. NBTI leads to a number of adverse effects on devices, such as absolute drain current I_{dsat}, and transconductance g_m, decrease and absolute "off" current I_{off}, and threshold voltage V_t increase. Typical stress temperatures lie in the 100°C ~ 200°C range. The increased importance of NBTI has motivated many investigations in this area. As early as 1967, Deal et al. [108,109] reported that both the interface trap density N_{it}, and oxide charge density N_{ox}, increased upon negative bias stress. The rate of increase of both

N_{it} and N_{ox} were similar. Six years later, Goetzberger et al [110] reported the same observation. Again in 1977, Jeppson and Svensson [111] confirmed the observation that there is an equal growth of the oxide charge and the interface trap density, independent of the BTI stress field and temperature. The close correlation between N_{it} and N_{ox} was also been confirmed by many other researches [111-115]. Another common observation is the generation of interface trap and positive oxide charge follow a power-law time dependence with a power factor in the range of 0.2 ~ 0.3 [105,111,116,117].

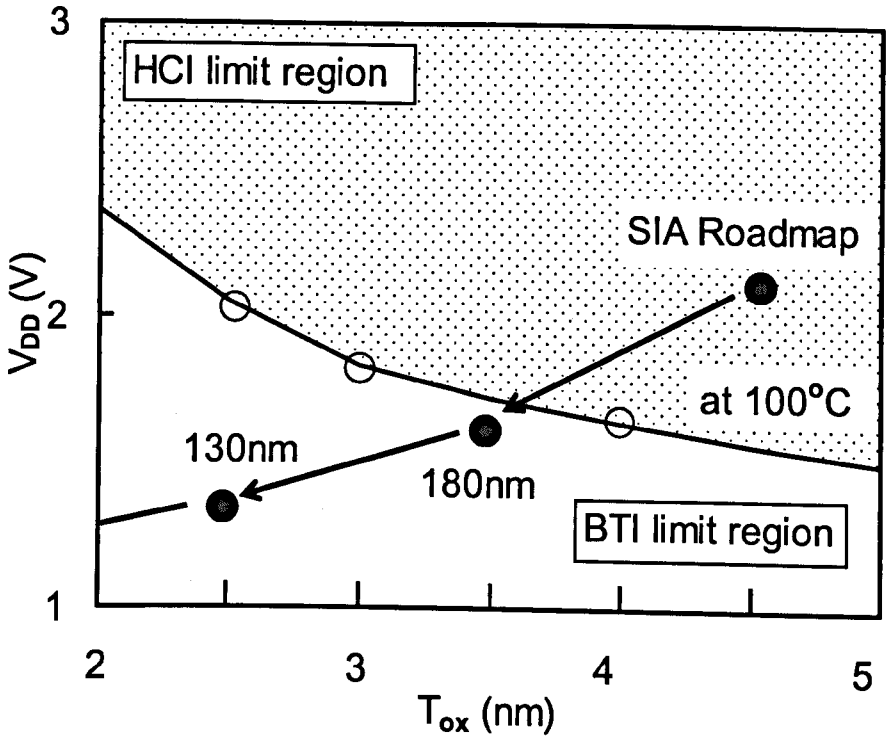


Figure 4. The transition of lifetime limitation mechanisms as a function of gate oxide thickness. When the thickness is below 3.5 nm, degradation due to BTI becomes to limit the device lifetime [106].

As the oxide becomes thinner, NBTI is becoming increasingly important, mainly because of two reasons. First, for each new generation of CMOS process, both operation temperature and electrical field increases. Second, to suppress the boron penetration and increase the dielectric constant, the nitrogen density in the gate oxide is increasing rapidly. It is reported that nitridation enhances NBTI and positive charge formation [105-107]. Figure 4 shows that for a gate oxide thinner than 3.5 nm,

the NBTI replaces hot carrier induced degradation as the limiting mechanism for device lifetime. It is therefore important to revisit the NBTI.

The most common models proposed for NBTI induced trap generation will be presented below. The first model discusses trap creation via hydrogen interaction dynamics. The second model proposed trap formation by the interaction of SiH with holes at the Si/SiO_2 interface. The last model describes trap creation via chemical species interaction and diffusion.

1.4.2.1. Hydrogen model

A hydrogen terminated interface trap precursor can be represented by

$$Si_3 \equiv SiH$$

High electric field can dissociate the silicon-hydrogen bond, leading to:

$$Si_3 \equiv SiH \rightarrow Si_3 \equiv Si\bullet + H^0,$$

where H^0 is a neutral interstitial hydrogen atom or atomic hydrogen. Atomic hydrogen is highly reactive and considered being a fast diffuser in the oxide [118]. The availability of SiH bonds for dissociation under high electric field is the rate-limiting process for this reaction.

Recent first-principles calculations show that the positively charged hydrogen or proton (H^+) is the only stable charge state of hydrogen at the interface. H^+ reacts directly with the SiH to form an interface trap, according to the reaction [119]:

$$Si_3 \equiv SiH + H^+ \rightarrow Si_3 \equiv Si\bullet + H_2.$$

The above reaction uses the fact that the SiH complex (or passivated dangling bond chemical species) is polarized such that a more positive charge resides near the Si

atom and a more negative charge resides near the hydrogen atom. Mobile H^+ migrates towards the negatively charged dipole region of the SiH molecule. The H^+ atom then reacts with the H^- to form H_2 leaving behind a positively charged Si dangling bond. In this model, H_2 can later dissociate to again act as a catalyst to disrupt additional SiH bonds. This process, in theory, can continue so long as hydrogen is available and SiH bonds are available to react. The reaction between hydrogenous species (H^+ or H_2) and SiH bonds is the rate-limiting process for this reaction model.

1.4.2.2. Hole model

This model explains NBTI induced trap formation by the interaction of SiH with "hot holes" or holes near or at the Si/SiO_2 interface [113,116,120-122]. Dissociation involving holes is given by

$$Si_3 \equiv SiH + h^+ \rightarrow Si_3 \equiv Si\bullet + H^+$$

During NBTI stress, holes are attracted to the SiO_2/Si interface. This model is consistent with the results that a positive substrate bias accelerates NBTI degradation, reduces the device lifetime [107]. Holes induced SiH bonds dissociation is the rate-limiting process for this reaction model.

1.4.2.3. Electrochemical model

Jeppson and Svensson were the first to propose a diffusion-reaction concept to explain the $t^{0.25}$ dependence of NBTI degradation [111].

$$Si_3 \equiv SiH + O_3 \equiv SiOSi \equiv O_3 \rightarrow Si_3 \equiv Si\bullet + O_3 \equiv Si^+ + O_3 \equiv SiOH + e^-$$

Y D_{it} Qf X to Si

According to this model, the silicon interface contains a large number of defects (i.e. SiH bond), which are electrically inactive, but may be electrically activated during NBTI. When the interface defect (i.e. SiH bond) is electrically activated, the diffusing species (e.g. H) leaves the defect site and reacts with the SiO_2 lattice to form an OH group bonded to an oxide atom, leaving a trivalent Si atom (Si_0^+) in the oxide and one trivalent Si_s at the Si interface. The Si_0^+ forms the oxide charge and the Si_s forms the interface trap. It has been proposed that X and Y in the reaction are H_2 and H^0, respectively. This model agrees with the observation that similar amount of interface traps and positive charges are generated. The hydrogen reaction and diffusion are the rate-limiting process for this reaction model.

References

1. S. M. Sze, "Physics of semiconductor devices," J. Willy & Sons, 1981.
2. G. E. Moore, "Cramming more components onto integrated circuits," Electronics, **38**, 114 (1965).
3. R. R. Schaller, "Moore's law: past, present, and future," IEEE Spectrum, **34**, 52 (1997).
4. R. D. Issac, "The future of CMOS technology," IBM J. Res. Dev., **44**, 369 (2000).
5. P. Balk, The Si-SiO2 system, Elsevier, Amsterdam, Chapter 1, 1998.
6. C. R. Helms, The Si-SiO2 system, Elsevier, Amsterdam, Chapter 3, 1998.
7. G. Groeseneken, R. Bellens, G. Van den bosch and H .E. Maes, "Hot-carrier degradation in submicrometre MOSFETs: from uniform injection towards the real operating conditions," Semicond. Sci. Tech., **10**, 1208 (1995).
8. J. F. Zhang and W. Eccleston, "Effects of high field injection on the hot carrier induced degradation of submicrometer pMOSFET's," IEEE Trans. Elec. Dev., **42**, 1269 (1995).
9. D. J. DiMaria and E. Cartier, "Mechanism for stress-induced leakage currents in thin silicon dioxide films," J. Appl. Phys., **78**, 3883 (1995).
10. J. De Blauwe, D. Wellekens, J. Van Houdt, R. Degraeve, L. Haspeslagh, G. Groeseneken and H. E. Maes, "Impact of tunnel-oxide nitridation on endurance and read-disturb characteristics of Flash E(2)PROM devices," Microelectronic Eng., **36**, 301 (1997).
11. R. Degraeve, G. Groeseneken, R. Bellens, J. L. Ogier, M. Depas, P. J. Roussel and H. E. Maes, "New insights in the relation between electron trap generation and the statistical properties of oxide breakdown," IEEE Trans. Elec. Dev., **45**, 904 (1998).
12. J. H. Stathis and D. J. DiMaria, "Oxide scaling limit for future logic and memory technology," Microelectronic Eng., **48**, 395 (1999).
13. R. Degraeve, B. Kaczer and G. Groeseneken, "Ultra-thin oxide reliability: searching for the thickness scaling limit," Microelectronics Reliability, **40**, 697 (2000).

14. V. V. Afanasev, A. Stesmans, A. G. Revesz and H. L. Hughes, "Trap generation in buried oxides of silicon-on-insulator structures by vacuum ultraviolet radiation," J. Electrochem. Soc., **144**, 749 (1997).
15. M. Knoll, D. Braunig and W. R. Fahrner, "Generation of oxide charge and interface states by ionizing-radiation and by tunnel injection experiments," IEEE Trans. on Nucl. Sci., **29**, 1471 (1982).
16. E. M. Verbitskaya, V. K. Eremin and A. M.Ivanov, "Formation of radiation defects in high-resistivity silicon as a result of cyclic irradiation and annealing," Semiconductors, **31**, 189 (1997).
17. P. M. Lenahan and P. V. Dressendorfer, "Effects of bias on radiation-induced paramagnetic defects at the silicon-silicon dioxide interface," Appl. Phys. Lett., **41**, 542 (1982).
18. D. L.Griscom, "Diffusion of radiolytic molecular-hyderogen as a mechanism for the post-irradiation buildup of interface states in SiO2-on-si structures," J. Appl. Phys., **58**, 2524 (1985).
19. N. S. Saks, C. M. Dozier and D. B. Brown, "Time-dependence of interface trap formation in MOSFETs following pulsed irradiation," IEEE Trans. Nucl. Sci., **35**, 1168 (1988).
20. D. B. Brown and N. S. Saks, "Time-dependence of radiation-induced interface trap formation in metal-oxide-semiconductor devices as a function of oxide thickness and applied field," J. Appl. Phys., **70**, 3734 (1991).
21. S. K. Lai, "Interface trap generation in silicon dioxide when electrons are captured by trapped holes," J. Appl. Phys., **54**, 2540 (1983).
22. H. C. Shin and C. M. Hu, "Thin gate oxide damage due to plasma processing," Semicond. Sci. Tech., **11**, 463 (1996).
23. T. Brozek and C. R. Viswanathan, "Increased hole trapping in gate oxides as latent damage from plasma charging," Semicond. Sci. Tech., **12**, 1551 (1997).
24. E. Cartier and J. H. Stathis, "Atomic hydrogen-induced degradation of the Si/SiO2 structure," Microelectronic Eng., **28**, 3 (1995).
25. T. Brozek, L. H. Peng and C. R. Viswanathan, "Role of test stress levels in detection of process-induced latent charging damage in MOS transistors," in Proc 1st Int. Symp. on Plasma Process-Induced Damage, Santa Clara, p.81, USA (1996).

26. J. C. King and C. M. Hu, "Effect of low and high-temperature anneal on process induced damage of gate oxide," IEEE Elec. Dev. Lett., **15**, 475 (1994).
27. M. Creusen, H. C. Lee, S. Vanhaelemeersch and G. Groeseneken, "The effect of plasma damage and different annealing ambients on the generation of latent interface states," in Proc 3rd Int. Symp. on Plasma Process-Induced Damage, Honolulu, p.217, USA (1998).
28. C. T. Wang, "Hot carrier design consideration for MOS devices and circuits," Van Nostrand Reinhold, New York, 1992.
29. P. E. Cottrell, R. R. Troutman and T. H. Ning "Hot-electron emission in n-channel IGFETS," IEEE Trans. Elec. Dev., **26**, 520 (1979).
30. W. L. Chen and T. P. Ma, "Channel hot-carrier induced oxide charge trapping in NMOSFET's," in Tech. Digest IEEE 37th Int. Electron Devices Meeting, Washington, p.731, USA (1991).
31. J. M. Aitken and D. R. Young, "Avalanche injection of holes into SiO2," IEEE Trans. Nucl. Sci., **24**, 2128 (1977).
32. S. Ogawa, N. Shiono and M. Shimaya, "Neutral electron trap generation in SiO2 by hot holes," Appl. Phys. Lett., **56**, 1329 (1990).
33. C. T. Sah, J. Y. C. Sun and J. J. T. Tzou, "Generation-annealing kinetics if the interface donor states at 0.25 eV above the midgap and the turn-around phenomena on oxidized silicon during avalanche electron injection," J. Appl. Phys., **54**, 2547 (1983).
34. J. F. Zhang and W. Eccleston, "Positive bias temperature instability in MOSFET's," IEEE Trans. Elec. Dev., **45**, 116 (1998).
35. E. H. Nicollian, C. N. Berglund, P. F. Schmidt and J. M. Andrews, "Electron trapping/detrapping within thin SiO_2 films in the high field tunneling regime," J. Appl. Phys., **42**, 5654 (1971).
36. D. J. DiMaria, The properties of electron and hole traps in thermal silicon dioxide layers grown on silicon. In S. T. Pantelides (ed.), *The Physics of SiO2 and its Interfaces*, New York: Pergamon, 160, 1978.
37. R. F. De Keersmaecker and D .J. DiMaria, "Electron trapping and detrapping characteristics of arsenic-implanted SiO_2 layers," J. Appl. Phys., **51**, 1085 (1980).
38. J. F. Zhang, S. Taylor, and W. Eccleston, "Electron trap generation in thermally grown SiO_2 under Fowler-Nordheim stress," J. Appl. Phys., **71**, 725 (1992).

39. J. F. Zhang, Wiley Encyclopedia of Electrical and Electronics Eng., **22**, 540 (1999).

40. D. J. DiMaria and J. W. Stasiak, "Trap creation in silicon dioxide produced by hot-electrons," J. Appl. Phys., **65**, 2342 (1989).

41. D. R. Wolters and J. J.van der Schoot, "Kinetics of charge trapping in dielectrics," J. Appl. Phys., **58**, 831 (1985).

42. D. J. DiMaria and J. H. Stathis, "Anode hole injection, defect generation, and breakdown in ultrathin silicon dioxide films," J. Appl. Phys., **89**, 5015 (2001).

43. D. Esseni, J. D. Bude and L. Selmi, "On interface and oxide degradation in VLSI MOSFETs - Part I: Deuterium effect in CHE stress regime," IEEE Trans. Elec. Dev., **49**, 247 (2002).

44. J. D. Bude, B. E. Weir and P. J. Silverman, "Explanation of stress-induced damage in thin oxides," in Tech. Digest IEEE 44th Int. Electron Devices Meeting, San Francisco, p.179, USA (1998).

45. L. Do Thanh and P. Balk, "Elimination and generation of Si-SiO_2 interface traps by low-temperature hydrogen annealing," J. Electrochem. Soc., **135**, 1797 (1988).

46. I. C. Chen, S. Holland, and C. Hu, "Electron-trap generation by recombination of electrons and holes in SiO_2," J. Appl. Phys., **61**, 4544 (1987).

47. H. Satake, S. Takagi and A. Toriumi, "Evidence of electron-hole cooperation in SiO_2 dielectric breakdown," International Reliability Physics Symposium Proceedings, 156 (1997).

48. B. Kaczer, R. Degraeve, N. Pangon, T. Nigam and G. Groeseneken, "Investigation of temperature acceleration of thin oxide time-to-breakdown," Microelectronic Eng., **48**, 47 (1999).

49. D. J. DiMaria, E. Cartier and D. Arnold, "Impact ionization, trap creation, degradation, and breakdown in silicon dioxide films on silicon," J. Appl. Phys., **73**, 3367 (1993).

50. D. J. DiMaria, D. J. Buchaman, D. A. Stathis and R. E. Stahlbush, "Interface states induced by the presence of trapped holes near the silicon-silicon dioxide interface," J. Appl. Phys., **77**, 2032 (1995).

51. D. J. DiMaria, "Temperature-dependence of trap creation in silicon dioxide," J. Appl. Phys., **68**, 5234 (1990).

52. E. Cartier, D. A. Buchanan and J. G. Dunn, "Atomic hydrogen-induced interface degradation of reoxidized-nitrided silicon dioxide on silicon," Appl. Phys. Lett., **64**, 901 (1994).

53. D. A. Buchanan, A. D. Marwick, D. J. DiMaria and L. Dori, "Hot-electron-induced hydrogen redistribution and defect generation in metal-oxide-semiconductor capacitors," J. Appl. Phys., **76**, 3595 (1994).

54. P. E. Blochl and J. H. Stathis, "Hydrogen electrochemistry and stress-induced leakage current in silica," Phys. Rev. Lett., **83**, 372 (1999).

55. A. Yokozawa and Y. Miyamoto, "First-principles calculations for charged states of hydrogen atoms in SiO2," Phys. Rev. B, **55**, 13783 (1997).

56. V. V. Afanas'ev and A. Stesmans, "H-complexed oxygen vacancy in SiO2: Energy level of a negatively charged state," Appl. Phys. Lett., **71**, 3844 (1997).

57. K. F. Schuegraf and C. M. Hu, "Metal-Oxide-Semiconductor field-effect-transistor substrate current during Fowler-Nordheim tunneling stress and silicon dioxide reliability," J. Appl. Phys., **76**, 3695 (1994).

58. I. C. Chen, S. Holland, K. K. Young, C. Chang and C. Hu, "Substrate hole current and oxide breakdown," Appl. Phys. Lett., **49**, 669 (1986).

59. Z. A. Weinberg, M. V. Fischetti and Y. Nissancohen, "SiO2-induced substrate current and its relation to positive charge in field-effect transistors," J. Appl. Phys., **59**, 824 (1986).

60. D. J. DiMaria, E. Cartier and D. A. Buchanan, "Anode hole injection and trapping in silicon dioxide," J. Appl. Phys., **80**, 304 (1996).

61. H. Satake and A. Toriumi, "Substrate hole current generation and oxide breakdown in Si MOSFETs under Fowler-Nordheim electron tunneling injection," in Tech. Digest IEEE 39th Int. Electron Devices Meeting, Washington, p.337, USA (1993).

62. B. Kaczer, R. Degraeve, N. Pangon and G. Groeseneken, "The influence of elevated temperature on degradation and lifetime prediction of thin silicon-dioxide films," IEEE Trans. Elec. Dev., **47**, 1514 (2000).

63. E. Cartier, J. C. Tsang, M. V. Fischetti and D. A. Buchanan, "Light emission during direct and Fowler-Nordheim tunneling in ultra thin MOS tunnel junctions," Microelectronic Eng., **36**, 103 (1997).

64. M. Rasras, D. de Wolf, G. Groeseneken, B. Kaczer, R. Degraeve and H. E. Maes, "Photo-carrier generation as the origin of Fowler-Nordheim-induced substrate hole current in thin oxides," in Tech. Digest IEEE 45th Int. Electron Devices Meeting, Washington, p.465, USA (1999).

65. D. J. Dumin, J. R. Maddux, R. S. Scott and R. Subramoniam, "A model relating wearout to breakdown in thin oxides," IEEE Trans. Elec. Dev., **41**, 1570 (1994).

66. R. S. Scott, N. A. Dumin, T. W. Hughes, D. J. Dumin and B. T. Moore, "Properties of high voltage stress generated traps in thin silicon oxides," in Proc IEEE 33rd Int. Reliability Phys. Symp., Las Vegas, p.131, USA (1995).

67. B. Schlund, C. Messick, J. S. Suehle and P. Chaparala, "A new physics-based model for time-dependent-dielectric-breakdown," in Proc IEEE 34th Int. Reliability Phys. Symp., Dallas, p.84, USA (1996).

68. J. W. McPherson and H. C. Mogul, "Disturbed bonding states in SiO2 thin-films and their impact on time-dependent dielectric breakdown," in Proc IEEE 36th Int. Reliability Phys. Symp., Reno, p.47, USA (1998).

69. M. Kimura, "Field and temperature acceleration model for time-dependent dielectric breakdown," IEEE Trans Elec. Dev., **46**, 220 (1999).

70. E. M. Vogel, J. S. Suehle, M. D. Edelstein, B. Wang, Y. Chen and J. B. Bernstein, "Reliability of ultrathin silicon dioxide under combined substrate hot-electron and constant voltage tunneling stress," IEEE Trans. Elec. Dev., **47**, 1183 (2000).

71. D. J. DiMaria, "Defect generation under substrate-hot-electron injection into ultrathin silicon dioxide layers," J. Appl. Phys., **86**, 2100 (1999).

72. D. J. DiMaria, "Defect generation in field-effect transistors under channel-hot-electron stress," J. Appl. Phys., **87**, 8707 (2000).

73. D. J. DiMaria, "Electron energy dependence of metal-oxide-semiconductor degradation," Appl. Phys. Lett., **75**, 2427 (1999).

74. H. Uchida and T. Ajioka, "Electron trap center generation due to hole trapping in SiO2 under Fowler-Nordheim tunneling stress," Appl. Phys. Lett., **51**, 433 (1987).

75. H. Satake, S. Takagi, A. Toriumi, "Evidence of electron-hole cooperation in SiO2 dielectric breakdown," in Proc IEEE 35th Int. Reliability Phys. Symp., Denver, p.156, USA (1997).

76. D. J. DiMaria, "Hole trapping, substrate currents, and breakdown in thin silicon dioxide films," IEEE Elec. Dev. Lett., **16**, 184 (1995).

77. A. Yankova, L. DoThanh and P. Balk, "Effects of thermal nitridation on the trapping characteristics of SiO2-films," Solid-State Elec., **30**, 939 (1987).

78. D. Ielmini, A. S. Spinelli, M. A. Rigamonti and A. L. Lacaita, "Modeling of SILC based on electron and hole tunneling - Part I: Transient effects," IEEE Trans. Elec. Dev., **47**, 1258 (2000).

79. J. F. Zhang, S. Taylor and W. Eccleston, "A quantitative investigation of electron trapping in SiO_2 under Fowler-Nordheim stress," J. Appl. Phys., **71**, 5989 (1992).

80. M. Brox and W. Weber, "Dynamic degradation in MOSFET's – Part I: The physical effects," IEEE Trans. Elec. Dev., **38**, 1852 (1991).

81. W. L. Chen, A. Balasinski and T. P. Ma, "Lateral profiling of oxide charge and interface traps near MOSFET junctions," IEEE Trans. Elec. Dev., **40**, 187 (1993).

82. R. Williams, "Photoemission of electrons from silicon into silicon dioxide," Physical Rev., **140**, A569 (1965).

83. D. J. DiMaria, "Capture and release of electrons on NA+-Related trapping sites in the SiO2 layer of metal-oxide-semiconductor structures at temperatures between 77-degrees-K and 296-degrees-K," J. Appl. Phys., **52**, 7251 (1981).

84. B. E. Deal, "Current understanding of charges in thermally oxidized silicon structure," J. Electrochem. Soc., **121**, 198c (1974).

85. W. L. Warren, K. Vanheusden, J. R. Schwank, D. M. Fleetwood, P. S. Winokur, and R. A. B. Devine, "Mechanism for anneal-induced interfacial charging in SiO2 thin films on Si," Appl. Phys. Lett., **68**, 2993 (1996).

86. V. V. Afanas'ev and A. Stesmans, "Positive charging of thermal SiO2/(100)Si interface by hydrogen annealing," Appl. Phys. Lett., **72**, 79 (1998).

87. K. Vanheusden, W. L. Warren, R. A. B. Devine, D. M. Fleetwood, J. R. Schwank, M. R. Shaneyfelt, P. S. Winokur and Z. J. Lemnios, "Non-volatile memory device based on mobile protons in SiO2 thin films," Nature, **386**, 587 (1997).

88. D. M. Fleetwood, M. J. Johnson, T. L. Meisenheimer, P. S. Winokur, W. L. Warren and S. C. Witczak, "1/f noise, hydrogen transport, and latent interface-

trap buildup in irradiated MOS devices," IEEE Trans. Nucl. Sci., **44**, 1810 (1997).

89. R. Bellens, E. De Schrijver, G. Van den bosch, G. Groeseneken, P. Heremans and H. E. Maes, "On the hot-carrier-induced poststress interface-trap generation in n-channel MOS-transistors," IEEE Trans. Elec. Dev., **41**, 413 (1994).
90. V. H. Chan and J. E. Chung, "The impact of nMOSFET hot-carrier degradation on CMOS analog subcircuit performance", IEEE J. Solid-St. Circ., **30**, 644 (1995).
91. L. P. Trombetta, F. J. Feigl and R. J. Zeto, "Positive charge generation in metal–oxide–semiconductor capacitors," J. Appl. Phys., **69**, 2512 (1991).
92. R. K. Freitag, D. B. Brown and C. M. Dozier, "Evidence for two types of radiation-induced trapped positive charge," IEEE Trans. Nucl. Sci., **41**, 1828 (1994).
93. D. R. Young, E. A. Irene, D. J. DiMaria, R. F. De Keersmaecker and H. Z. Massoud, "Electron trapping in SiO2 at 295 and 77 K," J. Appl. Phys., **50**, 6366 (1979).
94. R. E. Stahlbush, E. Cartier and D. A. Buchanan, "Anomalous positive charge formation by atomic hydrogen exposure," Microelectronic Eng., **28**, 15 (1995).
95. S. K. Lai and D. R. Young, "Effects of avalanche injection of electrons into silicon dioxide – generation of fast and slow interface states," J. Appl. Phys., **52**, 6231 (1981).
96. D. M. Fleetwood, "Effects of hydrogen transport and reactions on microelectronics radiation response and reliability," Microelectronics Reliability, **42**, 523 (2002).
97. A. J. Lelis and T. R. Oldham, "Time dependence of switching oxide traps," IEEE Trans. Nucl. Sci., **41**, 1835 (1994).
98. J. F. Zhang, H. K. Sii, G. Groeseneken and R. Degraeve, "Hole trapping and trap generation in the gate silicon dioxide," IEEE Trans. Elec. Dev., **48**, 1127 (2001).
99. J. F. Zhang, C. Z. Zhao, H. K. Sii, G. Groeseneken, R. Degraeve, J. N. Ellis and C. D. Beech, "Relation between hole traps and hydrogenous species in silicon dioxides," Solid-State Elec., **46**, 1839 (2002).

100. J. F. Zhang, C. Z. Zhao, G. Groeseneken, R. Degraeve, J. N. Ellis and C. D. Beech, "Hydrogen induced positive charge generation in gate oxides," J. Appl. Phys., **90**, 1911 (2001).

101. J. F. Zhang, C. Z. Zhao, A. H. Chen, G. Groeseneken and R. Degraeve, "Hole-traps in silifcon dioxides – Part I: Properties," IEEE Trans. Elec. Dev., **51**, 1267 (2004).

102. T. H. Ning, "Capture cross section and trap concentration of holes in silicon dioxide," J. Appl. Phys., **47**, 1079 (1976).

103. G. Van den bosch, G. Groeseneken, H. E. Maes, R. B. Klein and N. S. Saks, "Oxide and interface degradation resulting from substrate hot-hole injection in metal-oxide-semiconductor field-effect transistors at 295 and 77 K," J. Appl. Phys., **75**, 2073 (1994).

104. C. Z. Zhao, J. F. Zhang, G. Groeseneken and R. Degraeve, "Hole-traps in silicon dioxides – Part II: Generation mechanism," IEEE Trans. Elec. Dev., **51**, 1274 (2004).

105. S. S. Tan, T. P. Chen, J. M. Soon, K. P. Loh, C. H. Ang and L. Chen, "Nitrogen-enhanced negative bias temperature instability: an insight by experiment and first-principle calculations," Appl. Phys. Lett., **82**, 1881 (2003).

106. N. Kimizuka, T. Yamamoto, T. Mogami, K. Yamaguchi, K. Imai and T. Horiuchi, "The impact of bias temperature instability for direct-tunneling ultra-thin gate oxide on MOSFET scaling," in Tech. Digest Symp. on VLSI Technology, Kyoto, p73, Japan (1999).

107. N. Kimizuka, K. Yamaguchi, K. Imai, T. Iizuka, C. T. Liu, R. C. Keller and T. Horiuchi, in "NBTI enhancement by nitrogen incorporation into ultrathin gate oxide for 0.10 μm gate CMOS generation," in Tech. Digest Symp. on VLSI Technology, Honolulu, p.92, USA (2000).

108. B. E. Deal, M. Skalar, A. S. Grove and E. H. Snow, "Characteristics of the surface-state charge (Qss) of thermally oxidized silicon," J. Electrochem. Soc., **114**, 266 (1967).

109. B. E. Deal, "The current understanding of Qss in the thermally oxidized silicon structure," J. Electrochem. Soc., **121**, 198C (1974).

110. A. Goetzberger, A. D. Lopez and R. J. Strain, "On the formation of surface states during stress aging of thermal Si-SiO interfaces," J. Electrochem. Soc., **120**, 90 (1973).

111. K. O. Jeppson and C. M. Svensson, "Negative bias stress of MOS devices at high electric fields and degradation of MNOS devices," J. Appl. Phys., **48**, 2004 (1977).

112. G. Chen, M. F. Li, C. H. Ang, J. Z. Zheng and D. L. Kwong, "Dynamic NBTI of p-MOS transistors and its impact on MSFET scaling," IEEE Elec. Dev. Lett., **23**, 734 (2002).

113. C. E. Blat, E. H. Nicollian, E. H. Poindexter, "Mechanism of negative-bias-temperature instability," J. Appl. Phys., **69**, 1712 (1991).

114. S. S. Tan, T. P. Chen, J. M. Soon, K. P. Loh, C. H. Ang, W. Y. Teo and L. Chen, "Linear relationship between H+-trapping reaction energy and defect generation: insight into nitrogen-enhanced negative bias temperature instability," Appl. Phys. Lett., **83**, 530 (2003).

115. X. J. Zhou, L. Tsetseris, S. N. Rashkeev, D. M. Fleetwood, R. D. Schrimpf, S. T. Pantelides, J. A. Felix, E. P. Gusev and C. D'Emic, "Negative bias-temperature instabilities in metal-oxide-silicon devices with SiO2 and SiOxNy/HfO2 gate dielectrics," Appl. Phys. Lett., **84**, 4394 (2004).

116. D. K. Schroder and J. A. Babcock, "Negative bias temperature instability: road to cross in deep submicron silicon semiconductor manufacturing," J. Appl. Phys., **94**, 1 (2003).

117. S. Ogawa, M. Shimaya and N. Shiono, "Interface-trap generation at ultrathin SiO2 (4-6nm)-Si interface during negative-bias temperature aging," J. Appl. Phys., **77**, 1137 (1995).

118. D. L. Griscom, "Diffusion of radiolytic molecular-hydrogen as a mechanism for the post-irradiation buildup of interface states in SiO_2-on-Si structures," J. Appl. Phys., **58**, 2524 (1985).

119. S. N. Rashkeev, D. M. Fleetwood, R. D. Schrimpf and S. T. Pantelides, "Proton-induced defect generation at the Si-SiO2 interface," IEEE Trans. Nucl. Sci., **48**, 2086 (2001).

120. R. B. Flair and R. C. Sun, "Threshold-voltage instability in MOSFETs due to channel hot-hole emission," IEEE Trans. Elec. Dev., **28**, 83 (1981).

121. F. B. McLean, "A framework for understanding radiation-induced interface states in SiO MOS structures," IEEE Trans. Nucl. Sci., **27**, 1651 (1980).

122. S. Ogawa and N. Shiono, "Generalized diffusion-reaction model for the low-field charge-buildup instability at the Si-SiO_2 interface," Phys. Rev. B, **51**, 4218 (1995).

2 Experimental Facilities and Techniques

2.1. Introduction

In this chapter, the system and equipment used are presented. The details of the testing samples are given. The techniques used for characterizing the degradation are described. The techniques used for stressing the devices, including the substrate-hot-carrier injection and Fowler-Nordheim injection, are reviewed. Parasitic effects encountered in these techniques and the methods used to evaluate them are also given.

New testing techniques and their associated software have been developed, including 'On-The-Fly' and subthreshold swing measurements. In this project, test system was also developed to carry out tests at elevated temperature.

2.2. System and equipment

The equipment used includes a personal computer with an IEEE 488 port, a probe station with micro-positioners, a pulse generator (Agilent 8110A), a parametric measurement mainframe with 4 medium power source/monitor units (MPSMUs) (Agilent E5270A). A schematic diagram of the system is given in Figure 1.

The system is fully controlled by a computer through an IEEE 488 port. The control program was specifically designed for this system and was written in turbo C language. The probe station with micro-positioners is used to connect the test device with the measurement equipment. The pulse generator (Agilent 8110A) is used to generate pulses required by the charge pumping measurements. The parametric

measurement mainframe (Agilent E5270A) has 4 MPSMU, which can be used to supply/measure the gate, drain, well and substrate voltage/current.

A furnace with a ceramic and a quartz tube wound with resistance wire (Carbolite, model CTF 12/100/900) is used in the annealing experiments. The maximum temperature is 1200°C and the uniform zone length is 375 mm. Either high purity nitrogen or forming gas (10% H_2 and 90% N_2) is used during the annealing. The upstream pressure is 1.2 ~ 1.4 bars and the flow rate is 2 litres per minute. The typical temperature used for annealing is 400°C. The maximum annealing temperature is 450°C.

The instruments used in the low temperature experiments include a cryogenerator (compressor unit Leybold RW2 and cool head), a refrigerator-cooled cryostat (Leybold ROK 10-300), an intelligent temperature controller (Oxford ITC4), and a rotary vacuum pump (Metrovac GDR1). A schematic diagram of the low temperature system is given in Figure 2. The packaged device is placed in a sample holder in the cryostat, which is vacuumed and radiation-shielded, and is connected to the testing instruments. The temperature in the cryostats can be set between 70 ~ 300K.

For experiments at elevated temperature, a hotplate (Bibby Sterilin HC500) is used, and the temperature is measured by a thermocouple (RS 206-3750). The temperature of the hotplate can be set as high as 250°C. The testing device is placed on the hotplate, which is isolated from the ground by a thin layer of vacuum grease. The probes are used to connect the device to the rest of testing instruments.

The typical experimental sequence is summarised in the flow-chart of Figure 3. The MOSFET is placed on the probe station and its terminals are connected to the equipment through the micro-positioners. Transfer characteristics and charge pumping measurements are first carried out in order to assess the initial properties of the testing device. The typical interface state density for a fresh device is $1 \sim 3\times10^{10}$ $cm^{-2}eV^{-1}$. The device is then stressed, either using the substrate hot carrier injection or Fowler-Nordheim injection techniques for a pre-set time. During the stress, the

interface state density, the oxide charge density, and the level of charge injection are monitored against time. The stress is terminated when the pre-specified stress time is reached.

2.3. Samples used in experiments

The testing samples used in this study were manufactured by 0.13 μm and 0.35 μm CMOS processes at Interuniversity Microelectronics Research Centre (IMEC), Belgium. The surface channel pMOSFETs (nMOSFETs) have a p+ (n+) poly-Si gate. For the device fabricated by the 0.13 μm CMOS technology, the gate dielectric is a silicon oxynitride of 2.7 nm, oxidized at 850°C and then nitrided in NO at 1050°C for 10 sec. The typical channel length used is 0.15 μm and the channel width is 10 μm. For the device fabricated by the 0.35 μm CMOS technology, the gate oxide was grown in dry O_2 at 900°C to a thickness of 7.1 nm. The n well was doped to a level of 2×10^{17} cm^{-3} approximately and no threshold voltage adjustment implantation was carried out. The channel length and width used are 10 μm and 200 μm, respectively. The device is passivated by a 1 μm layer of oxide. A top view of the MOSFET layout is given in Figure 4.

2.4. Techniques for characterizing the degradation

The Si/SiO_2 interface state density and the oxide trap density are two key parameters for characterizing the oxide degradation of MOSFETs. The density of interface states can be measured by using the charge pumping technique [1-3], the capacitor-voltage method (C-V), and the slope of transfer characteristics in the subthreshold region [4-6]. The density of trapped charges in the SiO_2 layer can be obtained from the shift of the drain current versus gate voltage characteristics in the subthreshold region [7] or the midgap point [6,8-9].

2.4.1. Measurement of interface states

The standard, two-level charge pumping technique [1] is used to measure interface state density for oxide thicker than 5.5 nm. For thin gate oxides ($t_{ox} \leq 2.7$ nm), the interface state density is extracted from the slope of transfer characteristics in the subthreshold region [6].

2.4.1.1. Basic principle of the charge pumping technique

The charge pumping technique was first introduced by Brugler and Jespers in 1969 [10]. It is now a widely accepted technique for measuring interface states in MOS devices. Their successes are mainly due to the easiness to implement and have excellent accuracy and resolution. Moreover, the charge pumping technique is applicable to small geometry MOS transistors rather than requiring large area capacitors, which are needed for the capacitance-voltage method.

The basic experimental set up for charge pumping measurement is shown in Figure 5. An AC voltage is applied to the transistor gate, and the charge pumping current, I_{cp}, is measured at the substrate. The source and drain of the transistor are connected together to a certain reverse bias voltage with respect to the substrate or grounded.

The waveform of gate voltage is shown in Figure 6. We start from $V_g = V_{top}$ and use nMOSFET as an example. When the gate voltage is larger than the V_t, inversion is obtained, all interface traps up to the Fermi level, E_f, are filled with electrons. As represented by the shaded area in Figure 7(a), these electrons are drawn into the device from the source and drain. When the gate voltage drops below V_t, the Fermi level drops accordingly in the bandgap, and the interface electron concentration is reduced exponentially with the decrease of the Fermi level, as shown in Figure 7(b). The energy level for the emission of electrons from the interface traps may, however, not be able to follow the decrease of the Fermi level. When the gate voltage reaches

the flat-band voltage, V_{fb}, the electron emission reaches an energy level, $E_{em,e}$, given by [11]

$$E_{em,e}=E_i-k\cdot T\cdot \ln\left(V_{th}\cdot\sigma_n\cdot n_i\cdot t_{em,e}+e^{\frac{E_i-E_{f,inv}}{k\cdot T}}\right) \quad (1)$$

where V_{th} is the thermal velocity of carriers,

n_i is the intrinsic concentration,

σ_n is the capture cross section for electrons,

E_i is the intrinsic energy level, and

$t_{em,e}$ is the time available for the emission of electrons from the interface traps during the fall time of the gate pulse.

The electrons trapped in the energy level below $E_{em,e}$ will be recombined with the majority carriers from the substrate. At the moment the gate voltage reaches V_{fb}, the interface is flooded by holes from the substrate, and the thermal equilibrium is established suddenly in the interface traps by recombination with these holes, as shown in Figure 7(c). This recombination gives rise to a transfer of electrons from the source and drain to the substrate, which forms the charge pumping current.

Analogously, as shown in Figure 7(d) ~ (f), when the gate voltage increases from V_{base} to the threshold voltage, V_t, the emission of holes from interface traps only reaches the energy level, $E_{em,h}$, by [11]

$$E_{em,h}=E_i+k\cdot T\cdot \ln\left(V_{th}\cdot\sigma_p\cdot n_i\cdot t_{em,h}+e^{\frac{E_{f,acc}-E_i}{k\cdot T}}\right) \quad (2)$$

where σ_p is the capture cross section for holes,

$E_{f,acc}$ is the Fermi level in accumulation,

The holes still trapped in the interface states between $E_{em,h}$ and $E_{f,inv}$ will recombine with electrons coming from the source and drain. The repetitive occurrence of both recombination processes gives rise to a net current, I_{cp}, which can be measured either at the substrate or at the source and drain. I_{cp} is determined by [11]

$$I_{cp}=2\cdot q\cdot f\cdot A_g\cdot k\cdot T\cdot D_{it}\cdot \ln\left(v_{th}\cdot n_i\cdot\sqrt{\sigma_n\cdot\sigma_p}\cdot\sqrt{t_{em,e}\cdot t_{em,h}}\right) \tag{3}$$

where D_{it} is the average interface state density between $E_{em,e}$ and $E_{em,h}$. If all other parameters are known, D_{it} can then be determined by measuring I_{cp}. The typical values for the parameters in the above expression are as follow: f = 100 KHz, T = 300K, $v_{th} = 10^7$ cm/sec, $n_i = 1.5\times10^{10}$ cm^{-3}, $\sigma_n = \sigma_p = 10^{-15}$ cm^2. The $t_{em,e}$ and $t_{em,h}$ is determined by the falling time, t_f, and rising time, t_r, respectively.

$$t_{em,e}=\frac{|V_{FB}-V_T|}{|\Delta V_G|}\cdot t_f \tag{4}$$

$$t_{em,h}=\frac{|V_{FB}-V_T|}{|\Delta V_G|}\cdot t_r \tag{5}$$

Proper selection of charge pumping pulse amplitude, rising and falling time is very important, in order to obtain reliable results. Experiments were carried out to select the pulse amplitude, rising and falling time in previous work [12]. The pulse amplitude used in this project was fixed at 1.5 V for both pMOSFETs and nMOSFETs, and $t_r = t_f = 0.02$ μs is used for pMOSFETs, $t_r = t_f = 1$ μs is used for nMOSFETs. Using rise and fall time shorter than these criteria will cause parasitic effects, such geometric effect [13], which will induce significant error to the results.

Figure 8 shows the evolution of the charge pumping current, I_{cp}, versus bias voltage curve with different amount of electron injection during a FN experiment. The increase of I_{cp} means the generation of interface states density throughout the stress.

2.4.1.2. Extraction of interface state density from subthreshold swing

A useful parameter in the subthreshold regime of I_d-V_g characteristics is the gate-voltage swing S, which can be used to determine the generated interface trap density. It is defined as the gate voltage required to change the current by one decade [4-6], and is given by,

$$S = 2.3 \cdot \frac{k \cdot T}{q} \cdot \frac{1 + \frac{C_D}{C_{ox}} + \frac{C_{it}}{C_{ox}}}{1 - \left(\frac{C_D}{C_{FB}}\right)^2}, \tag{6}$$

where C_D, C_{ox}, C_{it}, and C_{FB} are the depletion layer, the oxide, the interface trap, and the flat-band capacitances, respectively. $C_{it} = q \cdot D_{it}$, and D_{it} is interface trap density. Considering the generation of interface traps due to stress, the subthreshold swing change ΔS can be expressed as

$$\Delta S = S(\text{after stress}) - S(\text{before stress})$$

$$\Delta S = 2.3 \cdot \frac{k \cdot T}{q} \cdot \frac{\frac{\Delta C_{it}}{C_{ox}}}{1 - \left(\frac{C_D}{C_{FB}}\right)^2} \tag{7}$$

where ΔC_{it} is the interface trap capacitance change due to stress. Therefore, the change in interface state density due to stress, ΔD_{it}, is directly related to ΔS as

$$\Delta D_{it} = \frac{C_{ox} \cdot \Delta S}{2.3 \cdot k \cdot T} \cdot \left[1 - \left(\frac{C_D}{C_{FB}}\right)^2\right] \tag{8}$$

Figure 9(a) shows the subthreshold I_d-V_g characteristics under SHI (Substrate Hole Injection) stress. The subthreshold I_d-V_g curve shifts along the V_g axis and its slope is degraded with stress. This is an indication of both oxide charge generation and interface state generation. A comparison of different interface state measurement

techniques is shown in Figure 9(b). A 5.5 nm pMOSFET was stressed by SHI, the results of the charge pumping technique and subthreshold swing technique agrees well.

2.4.2. Measurements of oxide charges

Previous works in this university [7,14] used the shift of I_d-V_g characteristics in the subthreshold region to measure the density of trapped charges in the SiO_2 layer. This method is again employed on 0.35 μm CMOS test samples. For 0.13 μm CMOS test samples, the shift of I_d-V_g characteristics in the midgap point is used.

2.4.2.1. I_d-V_g shift in the subthreshold region

The density of oxide charges is calculated from the shift of the gate voltage at a fixed drain current of 10^{-10} × W / L, where W is the channel width and L is the channel length. Our 0.35 μm CMOS test samples have a channel width and channel length of 200 μm and 10 μm, respectively. Therefore, this current level is 2 nA, which is in the deep subthreshold region. The drain voltage is fixed at -0.05 V for pMOSFETs and +0.05 V for nMOSFETs. The basic experimental set up for transfer characteristics measurement is shown in Figure 10. As the charge centroid is not known, we assume that it is at the Si/SiO_2 interface to comply with previous work [15, 16]. The effective density of trapped charges, ΔN_{ot}, is given by

$$\Delta N_{ot} = -\frac{C_{ox} \cdot \Delta V_g}{q} \tag{9}$$

where C_{ox} is the oxide capacitance, and q is the one electronic charge.

In the above expression, the effects of interface state generation is not taking into account. This can be acceptable for relatively low stress level, as ΔN_{ot} is generally over one order of magnitude higher than the density of interface states. However, at

severe stress level, a large amount of interface states can be generated, which can affect the slope in subthreshold region. As shown in Figure 11, interface states generated by SHI can be considerable when $Q_h > 10^{17}$ holes/cm^2. The effects of the generated interface states must be taken into account when estimating the trapped hole density now. For generated electron traps, their density is generally in the same order of the generated interface states at severe stress level. The effect of ΔD_{it} on the electron traps at severe stress level has to be evaluated.

To evaluate the effect of interface states on ΔN_{ot}, the following experiment was carried out. A pMOSFET was stressed by SHI. When Q_h reached a certain level, SHI was interrupted and an electron injection at E_{ox} = +8 MV/cm was carried out to neutralize the trapped holes and partially fill the generated electron traps, so that the effect of ΔN_{it} on the generated electron traps can be evaluated. When the amount of generated interface states is relatively small, as shown in Figure 12(a), it is apparent that the main effect is the shift of the curve towards positive gate voltage due to the generation of electron traps. The dotted curve is obtained from the shift of the I_d-V_g curve of the fresh device along the x-axis for the comparison with the curves after stress. Almost no distortion of the curve slope is present. However, when the generated interface states increase further, as shown in Figure 12(b), significant distortion in the slope can be observed, which will affect the accuracy of measured ΔN_{ot}. The following methods were used to remove the effect of interface states on ΔN_{ot}.

As shown in Figure 13(a), ΔN_{ot} induced by interface states can be corrected from the ΔV_g at I_d = 2 nA, which is caused by the distortion of the curve slope. To determine ΔV_g, we have to calculate the midgap current, I_{d_midgap}, where the generated interface states are essentially neutral. For 0.35 μm CMOS sample, I_{d_midgap} is found to be 7.3×10^{-14} A. The contribution of ΔD_{it} to ΔN_{ot} is given in Figure 13(b). ΔN_{ot} (ΔD_{it}) increased linearly against the generated interface states. Hereafter in this project, the measured trapped charges on 0.35 μm CMOS sample are systematically corrected at I_{d_midgap}.

$\Delta N_{ot} = \Delta N_{ot}$ (measured) – ΔN_{ot} (ΔD_{it}) for trapped holes

$\Delta N_{ot} = \Delta N_{ot}$ (measured) + ΔN_{ot} (ΔD_{it}) for generated electron traps

An example of the above correction is given in Figure 14, where electron traps were generated by SHI. As expected, the effect of generated interface states is negligible at relatively low injection level, but increases for high Q_h.

2.4.2.2. I_d-V_g shift at the midgap

The threshold voltage shift under high field stress reflects the generation of both interface traps and trapped oxide charges, i.e., $\Delta V_{th} = \Delta V_{it} + \Delta V_{ot}$, where ΔV_{it} and ΔV_{ot} are the voltage shifts due to interface traps and trapped oxide charges, respectively. It is often that ΔV_{th} alone cannot provide enough information on oxide degradation due to stress, and the separation of ΔV_{th} into ΔV_{it} and ΔV_{ot} is needed [8].

Based on the general assumption that interface traps in the upper half of the band gap are acceptors and those in lower half of the band gap are donors, interface traps are uncharged at midgap $\phi_S = \phi_F$. The midgap current, I_{mg}, can be obtained from the subthreshold current equation [17],

$$I_d\Big|_{Vg=Vmg} = \mu_{eff} \cdot \frac{W}{2 \cdot L} \cdot \frac{q \cdot N_A}{\phi_T} \cdot L_D \cdot \left(\frac{n_i}{N_A}\right)^2 \cdot \exp(\phi_T \cdot \phi_S) \cdot (\phi_T \cdot \phi_S)^{-\frac{1}{2}} \cdot (1 - \exp(-\phi_T \cdot V_d)) \quad , \tag{10}$$

where L_D is the Debye length given by

$$L_D = \left(\frac{2 \cdot \varepsilon_r \cdot \varepsilon_o}{\phi_T \cdot q \cdot N_A}\right)^{\frac{1}{2}} \tag{11}$$

$\phi_T = q / (K_B \cdot T)$, n_i is the intrinsic carrier concentration, ε_r is the relative dielectric constant of Si, and ϕs is the band bending at the surface, which can be defined as ϕ_S

$= (K_B \cdot T / q) \cdot \ln (N_A / n_i)$. When pMOSFETs is used, μ_{eff} is the effective hole mobility. Since the midgap current $I_d|_{Vg=Vmg}$ is in the range of 0.001 ~ 0.1 pA, the linear extrapolation of the subthreshold curve ($\log I_d$-V_g) to this low current level yields the midgap voltage V_{mg}. The midgap voltage shift can be used to evaluate charge trapping in the oxide.

2.4.3. On-The-Fly measurement technique

In order to avoid the recovery during the characterization steps, a new technique was proposed [18] to characterize the Negative Bias Temperature Instability (NBTI). 'On-The-Fly' measurement keeps a quasi-constant gate voltage during the characterization, and measuring the linear drain current all along the stress, as shown in Figure 15. To estimate the trans-conductance, $g_m(n)$, V_g was perturbed by a small ±DV to give:

$$g_m(n) = \frac{I_{d(V_g+DV)} - I_{d(V_g-DV)}}{2 \cdot DV} \tag{12}$$

The degradation of drain current between two measurement points 'n' and 'n-1' is,

$$\Delta I_d(n) = I_d(n) - I_d(n-1) \tag{13}$$

The shift of threshold voltage between these two points can be evaluated by,

$$\Delta V_t(n) = -\frac{\Delta I_d(n)}{g_m(n)} \tag{14}$$

The accumulative shift of threshold voltage is,

$$\Delta V_t = -\sum_{n=1}^{M} \frac{I_d(n) - I_d(n-1)}{g_m(n)} \tag{15}$$

where M+1 is the number of I_d measurements and $g_m(n)$ is the mean value of the trans-conductance between the n^{th} and $n\text{-}1^{th}$ I_d measurements, as shown in Figure 16. Hence, periodical three point I_d measurements are enough to monitor ΔI_d, g_m, ΔV_t during stress.

In equation (12), a small perturbation, DV, was applied on V_g to obtain the trans-conductance, g_m. DV has a direct influence on the assessment of g_m. Figure 17 shows the measured drain currents during NBTS when DV = 250 mV. ΔI_d(down) is only half of the ΔI_d(up). This is because at relatively high V_g, I_d tends to saturate, as shown in Figure 18. In order to assess the g_m accurately, it is important to have ΔI_d(down) $\approx$ ΔI_d(up). Figure 19 shows the ΔI_d(down) / ΔI_d(up) ratio under different DV. DV = 25mV will be used hereafter in this project.

In equation (15), g_m is assumed being a constant between the two measurement points and the average $\overline{g_m}$ is used, $\overline{g_m(n)} = (g_m(n) + g_m(n-1))/2$. To estimate the error caused by this assumption, Figure 20 compares the ΔV_t calculated by using $g_m(n)$, $g_m(n-1)$, and $\overline{g_m(n)}$. It is clear that the error is insignificant.

Since 'On-The-Fly' measurement is based on I_d degradation at stress voltage, NBTS still occurs during the measurement. When I_d was measured for the first time, it typically took 0.15 sec and some degradation occurs during this period, as shown in Figure 21(a), where I_d degradation is monitored on a fresh device. The 20 ms measurement time is the fastest time the current system can achieve. Figure 21(b) shows if I_d measured at 20 ms is used as the reference value, ΔV_t shifts upward. This uncertainty in the reference I_d leads to an underestimation of ΔV_t, which cannot be completely corrected at this stage. Care has been exercised to take this uncertainty into account when analyzing the results.

2.5. Techniques for stressing devices

There are a number of carrier injection techniques, which are used to investigate degradations. The most commonly used are avalanche carrier injection, channel hot carrier injection (CHC), optically induced hot-electron injection, Fowler-Nordheim electron injection (FN), substrate hot electron injection (SHE) and substrate hole injection (SHI). Among them, the CHC is the only non-uniform injection technique.

In the following paragraphs, the principles of high field injection technique (FN) and substrate hot carrier injection techniques (SHE and SHI) are explained.

2.5.1. High field injection technique

The Fowler-Nordheim Injection [19] was first studied by Fowler and Nordheim and is widely used in the breakdown test under high oxide electric field. The basic experimental set up of FN electron injection is shown in Figure 22. The source and drain of MOS transistor are connected to zero volts. A large voltage is applied to gate to create a high electrical field across the gate oxide.

When this electrical field is above 6 ~ 7 MV/cm; Figure 23 shows that the physical distance between the conduction band of the gate oxide and that of the substrate becomes so thin that electrons can tunnel through. This is called as Fowler-Nordheim tunneling. The electrons gain kinetic energy in the oxide and can create electron-hole pairs by impact ionization. When electrons enter the anode, the energy released can be transferred to holes and causes hole injection. The FN current density can be calculated by

$$J_{FN} = A \cdot E_{ox}^{\ 2} \cdot \exp(\frac{-B}{E_{ox}}) \qquad (16)$$

where A and B are two constants related to the electron effective mass in the oxide conduction band and the Si/SiO_2 barrier height.

The MOSFETs used in this project have polysilicon gate. Gate depletion occurs under positive bias for a nMOSFET with n+ poly-gate or negative bias for a pMOSFET with p+ poly-gate. This causes an extra voltage drop on the polysilicon due to the polysilicon band bending. The voltage drop on the polysilicon can be calculated by [20],

$$V_{poly} = \frac{\varepsilon_{ox}^{\ 2} \cdot E_{ox}^{\ 2}}{2 \cdot q \cdot \varepsilon_{si} \cdot N_{poly}} \tag{17}$$

until V_{poly} is pinned at 1.12 V due to the saturation of band bending in strong inversion. V_{poly} is dependent on the doping of polysilicon, N_{poly}, and the oxide field. Attention must be paid to take the effect of V_{poly} into account, when calculating the oxide field E_{ox} [20],

$$E_{ox} = \frac{V_g - V_{poly} - \phi_s - \phi_{ms}}{t_{ox}} \tag{18}$$

where ϕ_s is the substrate surface potential and ϕ_{ms} is the work function. As shown in Figure 24, the difference between the gate current under positive and negative bias in a nMOSFET is caused by polysilicon depletion.

The oxide field under FN condition on 0.35 μm CMOS test sample is calculated using this method with a gate doping density of 10^{20} cm^{-3}. The doping of a p+ poly gate is approximately the same as the n+ poly gate.

2.5.2. Substrate hot carrier injection techniques

The basic experimental set up of SHE [21] is given in Figure 25. The source and drain of a MOS transistor are connected to zero volts. The gate voltage should be larger than the threshold voltage to produce an inverted channel underneath the gate. The underlying pn junction is forward biased. The electrons supplied from the forward biased underlying pn junction diffuse into the space charge layer, then drift

towards the Si/SiO_2 interface, as illustrate in Figure 26. They gain energy on their way from the high field set up by the p-well bias. When they arrive at the interface, they can be injected into the gate oxide [15].

The advantage of this technique is that the oxide field and injection current can be varied independently. Using this technique, it becomes possible to distinguish the dependence of the charge trapping on oxide field, injection current density and energy of injected electrons at the moment of injection.

In the SHI [22] technique, holes can be injected into the n-well by forward biasing the well-substrate pn junction. The basic experimental set up is given in Figure 27. The schematic energy band diagram is shown in Figure 28. The operation principle of SHI is similar to that of SHE. Holes are injected from the p-substrate into the n-well. They then diffuse towards the space charge region in which they are accelerated towards the Si/SiO_2 interface. A small fraction of holes can gain enough energy to be injected into the gate oxide, giving rise to a gate current, I_g. The others are collected at the source and drain terminals. By adjusting the n-well bias, $V_{n\text{-well}}$, and the gate bias, V_g, the silicon and oxide field can be independently controlled. The injectable carrier supply is controlled by the substrate, and n-well biases.

References

1. G. Groeseneken and H. E. Maes, "A reliable approach to charge pumping measurements in MOS transistors", IEEE Trans. Elec. Dev., **31**, 42 (1984).
2. P. Heremans and J. Witters, "Analysis of the charge pumping technique and its application for the evaluation of MOSFET degradation", IEEE Trans. Elec. Dev., **36**, 1318 (1989).
3. N. S. Saks and M. G. Ancona, "Evaluation of the 3-level charge pumping technique for characterizing interface traps", Appl. Phys. Lett., **60**, 2261 (1992).
4. J. R. Brews, "Subthreshold behavior of uniformly and nonuniformly doped long-channel MOSFET", IEEE Trans. Elec. Dev., **26**, 1282 (1979).
5. S. Horiguchi, T. Kobayashi and K. Saito, "Interface-trap generation modeling of Fowler-Nordheim tunnel injection into ultra-thin gate oxide", J. Appl. Phys., **58**, 387 (1985).
6. Y. Park and D. K. Schroder, "Degradation of thin tunnel gate oxide under constant Fowler-Nordheim current stress for a flash EEPROM", IEEE Trans. Elec. Dev., **45**, 1361 (1998).
7. I. S. Al-Kofahi, J. F. Zhang and G. Groeseneken, "The enhanced degradation of MOSFETs damaged by hot holes", Physics and chemistry of SiO_2 and the Si-SiO_2 interface-3, The Electrochemical Society Series, **96**, 711 (1996).
8. P. J. McWhorter and P. S. Winokur, "Simple technique for separating the effects of interface traps and trapped-oxide charge in metal-oxide-semiconductor transistors", Appl. Phys. Lett., **48**, 133 (1986).
9. D. C. Sekhar, P. P. Ray, M. M. De Souza P. Chaparala, "Edge effect under temperature bias stress of 0.18 μm PMOS technology", in Proc 24th Int. Conference on Microelectronics, Nis, p.16, Yugoslavia (2004).
10. J. S. Brugler and G. A. Jespers. "Charge pumping in MOS devices", IEEE Trans. Elec. Dev., **16**, 297 (1969).
11. C. T. Wang, PhD, "Hot carrier design considerations for MOS devices and circuits", Chapter 1, The mechanisms of hot-carrier degradation, pp. 30-31.
12. H. K. Sii, PhD, "Degradation of gate dielectric under hot hole stress", Chapter 2, Experimental facilities and Techniques, pp. 27-28.

13. G. Van den bosch, G Groeseneken and H. E. Maes, "On the geometric component of charge-pumping current in MOSFET's", IEEE Trans. Elec. Dev., **14**, 107 (1993).
14. I. S. Al-Kofahi, J. F. Zhang and G. Groeseneken, "Continuing degradation of the Si/SiO_2 interface after hot hole stress", J. Appl. Phys., **81**, 2686 (1997).
15. G. Van den bosch, G. Groeseneken and H. E. Maes, "Oxide and interface degradation resulting from substrate hot-hole injection in metal-oxide-semiconductor field-effect transistors at 295 and 77K", J. Appl. Phys., **75**, 2073 (1994).
16. A. V. Schwerin, M. M. Heyns and W. Weber, "Investigation on the oxide field dependence of hole trapping and interface state generation in SiO_2 layers using homogeneous nonavalanche injection of holes", J. Appl. Phys., **67**, 7595 (1990).
17. S. M. Sze, Physics of semiconductor devices. New York: Wiley, 446, 1990.
18. M. Denais, A. Bravaix, V. Huard, C. Parthasarathy, G. Ribes, F. Perrier, Y. Rey-Tauriac and N. Revil, "On-the-fly characterization of NBTI in ultra-thin gate oxide pMOSFET's," in Tech. Digest IEEE 50^{th} Int. Electron Devices Meeting, San Francisco, p.109, USA (2004).
19. R. H. Fowler and L. Nordheim, "Electron emission in intense electric field", in Proc. R. Soc. London, Ser. A, p.173 (1928).
20. K. F. Schuegraf, C. C. King and C. M. Hu, "Impact of polysilicon depletion in thin oxide MOS technology", in Proc IEEE Int. Symp. on VLSI Technology, Systems and Applications, Taipei, p.86, Taiwan (1993).
21. A. V. Schwerin and M. M. Heyns, "Oxide field dependence of bulk and interface trap generation in SiO_2 due to electron injection", in Proc 7^{th} Biennial Conference on Insulating Films On Semiconductors, Liverpool, p.263, United Kingdom (1991).
22. G. Van den bosch, G. Groeseneken and H. E. Maes, "Critical analysis of the substrate hot-hole injection technique", Solid-State Electronics, **37**, 393 (1994).

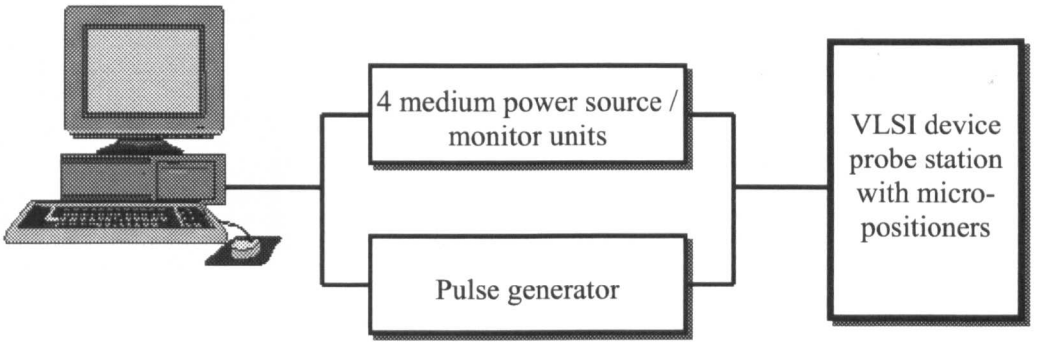


Figure 1. Schematic diagram of the testing system.

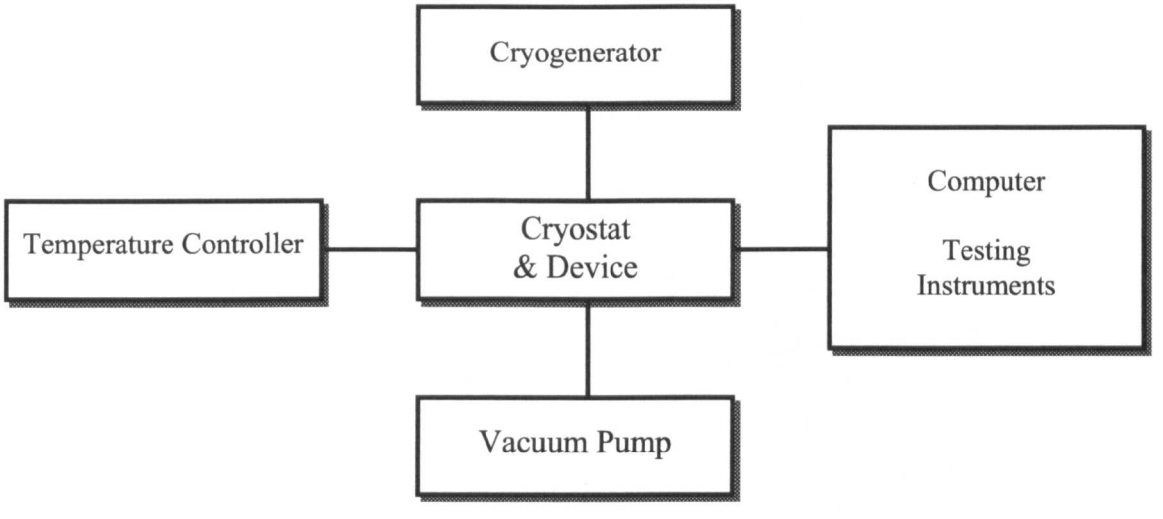


Figure 2. Schematic diagram of the low temperature system.

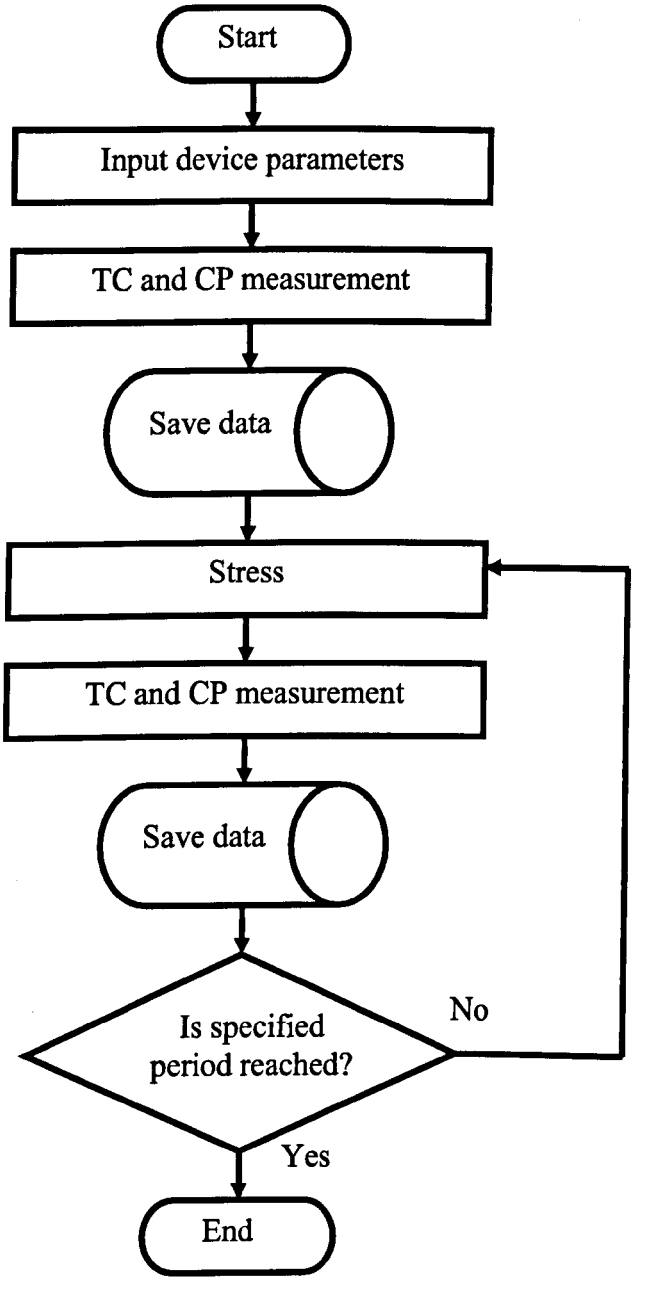
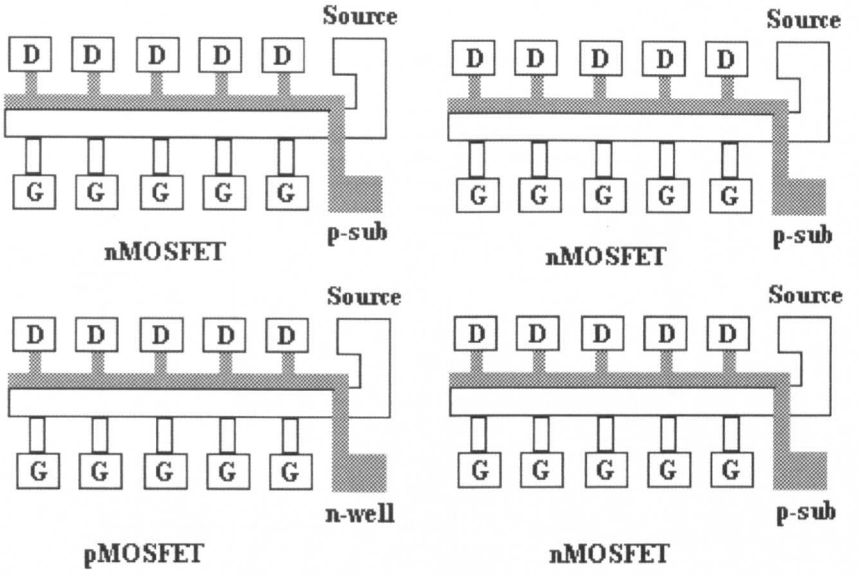
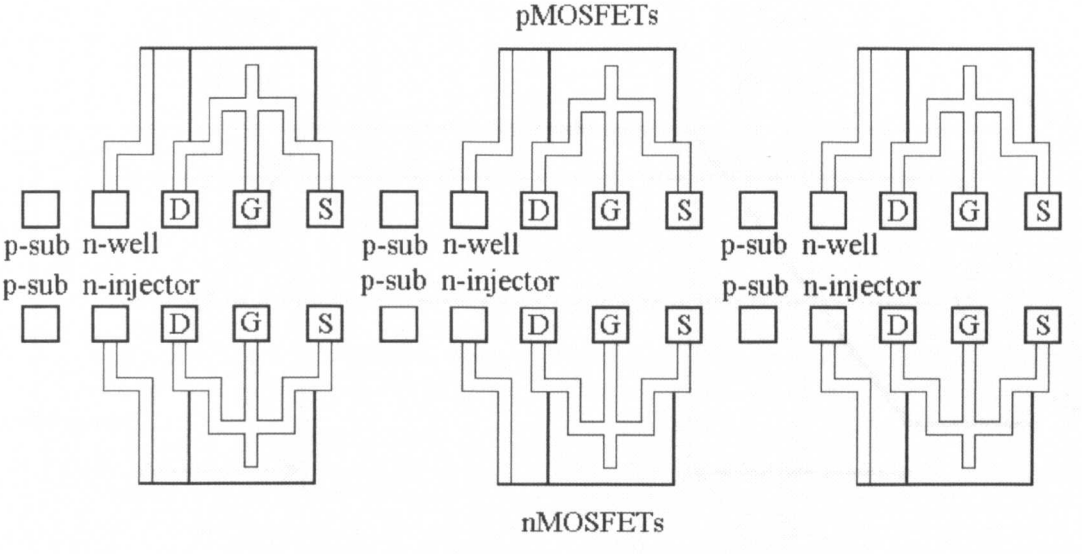


Figure 3. Flow chart of a typical experiment sequence.



(a)



(b)

Figure 4. Layout of MOSFETs fabricated by (a) 0.13 μm CMOS technology, and (b) 0.35 μm CMOS technology.

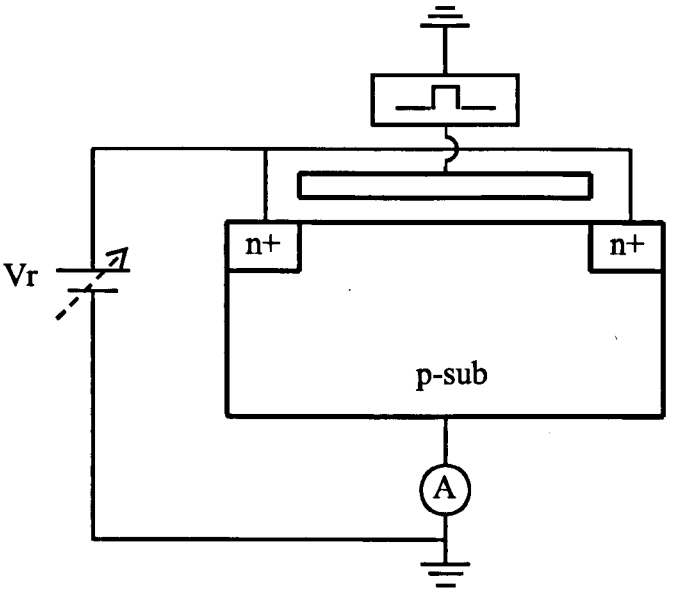


Figure 5. Experimental set up of charge pumping measurement.

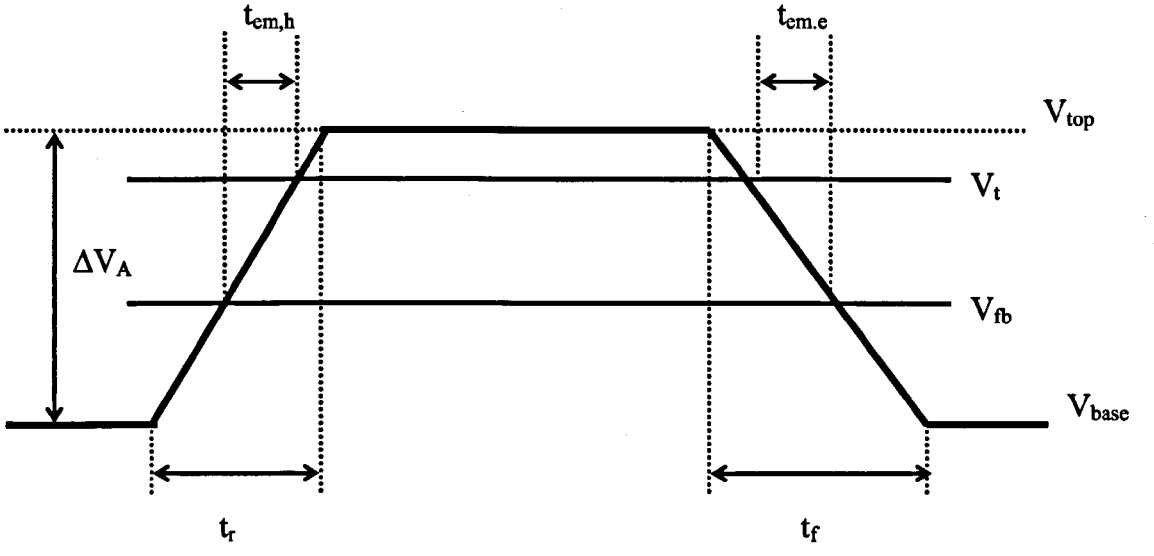


Figure 6. Waveform applied to the gate during the charge pumping measurement.

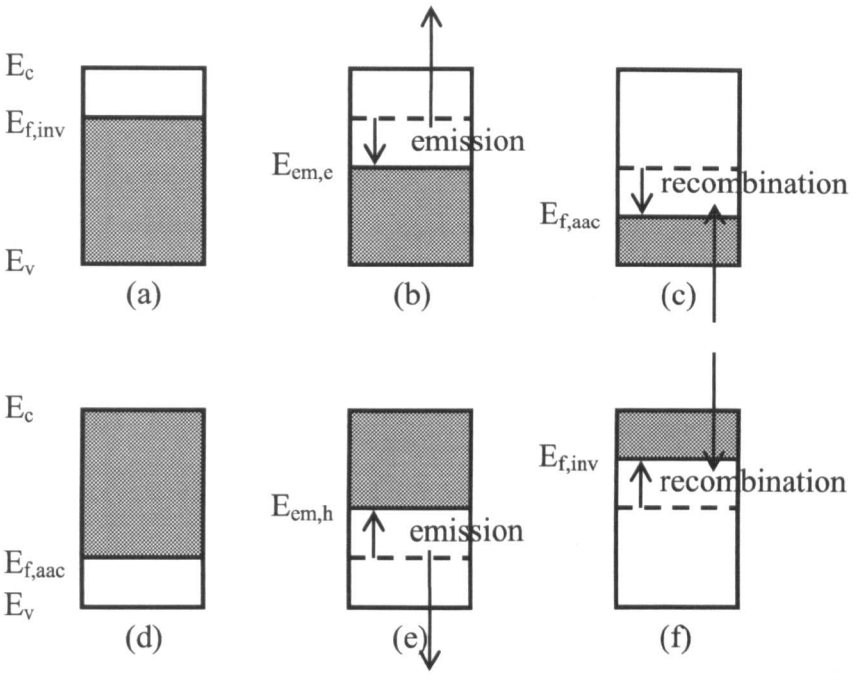


Figure 7. Schematic energy band representation of the various emission and recombination processes that are occurring during one charge pumping cycle.

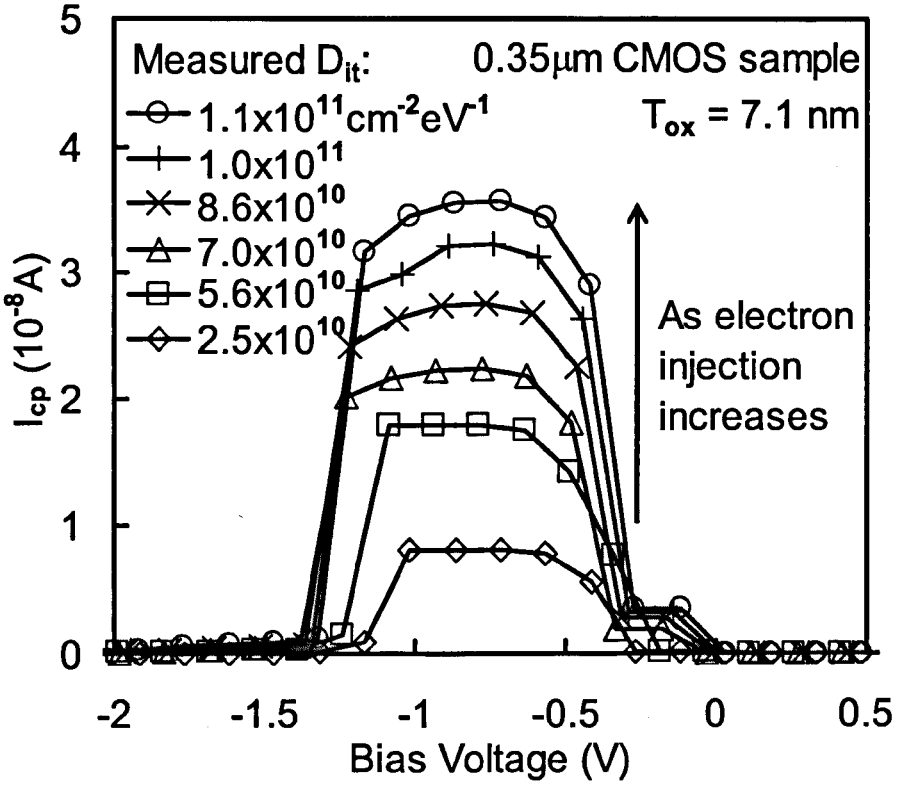
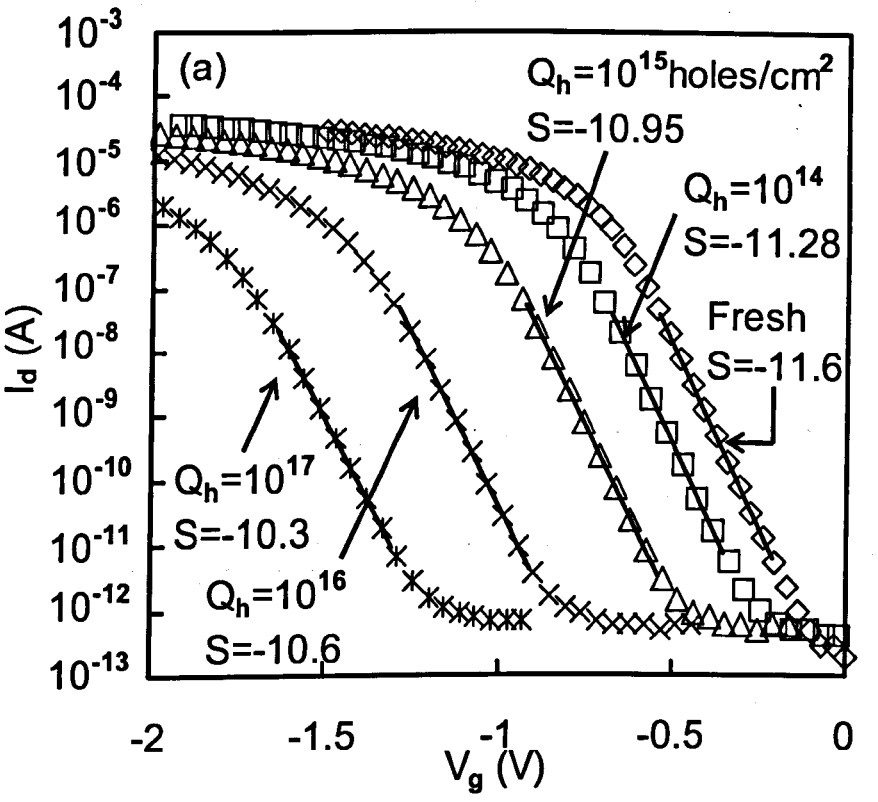


Figure 8. Charge pumping curves at six injection levels during the FNI under E_{ox} = +11 MV/cm. Symbols '◇', '□', '△', 'x', '+', and '○' correspond to Q_{inj} = 0, 1×10^{18}, 2×10^{18}, 5×10^{18}, 1×10^{19}, and 2×10^{19} carriers/cm^2, respectively. The increase of I_{cp} means the generation of interface states.


(a)
Q_h=10^{15}holes/cm^2
S=-10.95
Q_h=10^{14}
S=-11.28
Fresh
S=-11.6
Q_h=10^{17}
S=-10.3
Q_h=10^{16}
S=-10.6
I_d (A)
V_g (V)

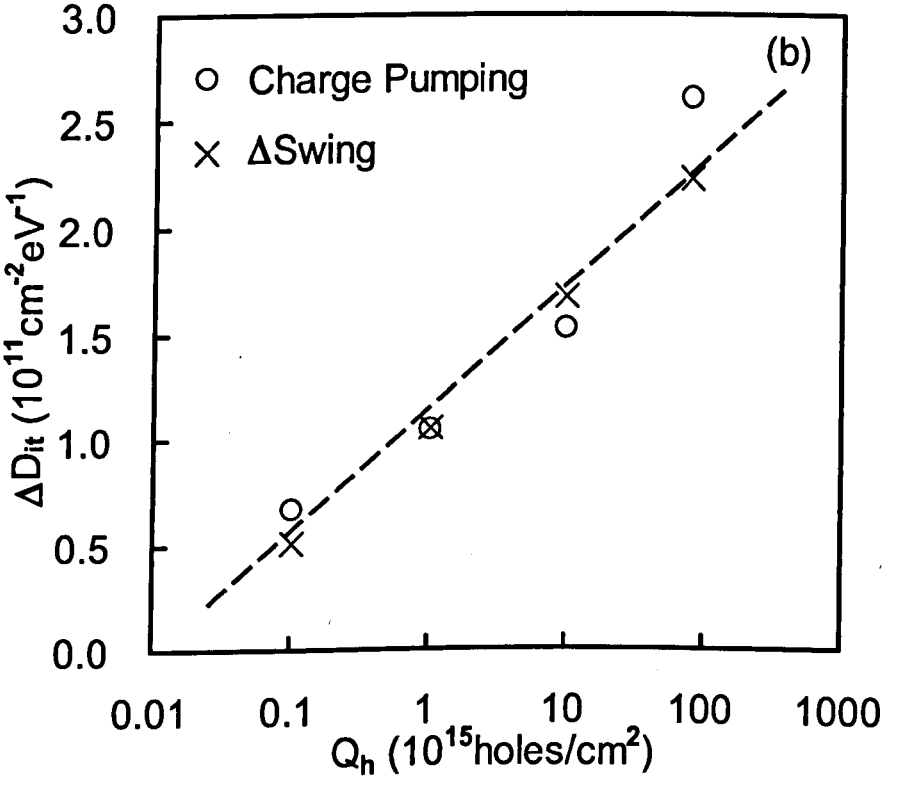

(b)
Charge Pumping
ΔSwing
ΔD_{it} (10^{11}cm^{-2}eV^{-1})
Q_h (10^{15}holes/cm^2)

Figure 9. 0.35 μm CMOS sample (T_{ox} = 5.5 nm) was stressed by SHI (E_{ox} = -5 MV/cm, V_w = 8.8 V, V_s = 9.8 V). (a) I_d-V_g characteristics during the stress. Interface states generation can be seen by the decrease of slope in subthreshold region. (b) Interface states were measured by charge pumping and subthreshold swing techniques, with $C_{ox} = 6.3\times10^{-7}$ Fcm^{-2}, $C_D = 4.2\times10^{-7}$ Fcm^{-2}, and $C_{FB} = 5.0\times10^{-7}$ Fcm^{-2}.

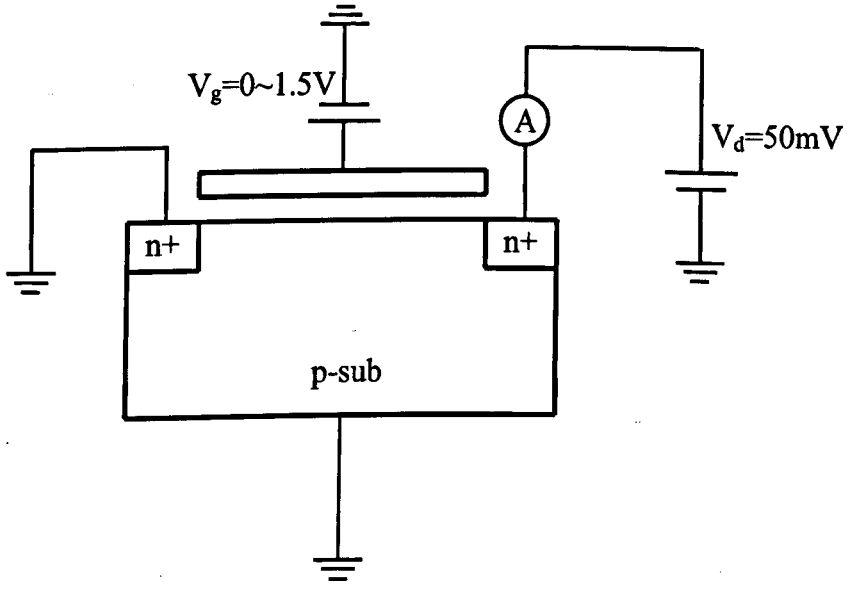


Figure 10. Experimental set up of transfer characteristics measurement.

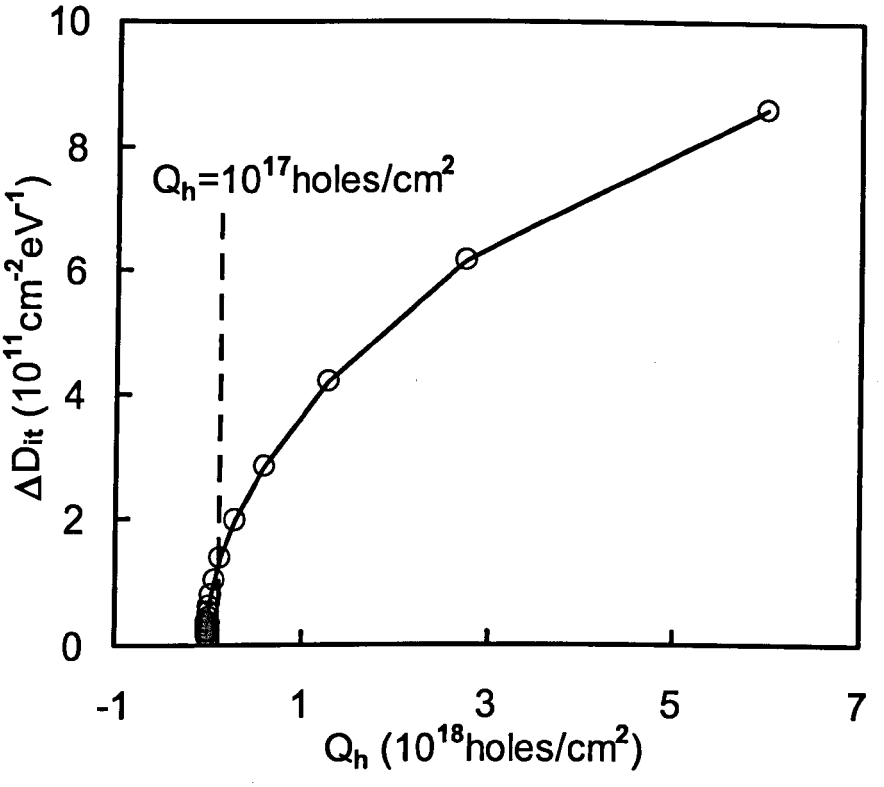


Figure 11. Interface states generation under SHI (t_{ox} = 7.1 nm, E_{ox} = -5 MV/cm, V_{well} = 8.8 V, V_{sub} = 10 V). The generation is substantial when $Q_h > 10^{17}$ holes/cm^2 (dashed line).

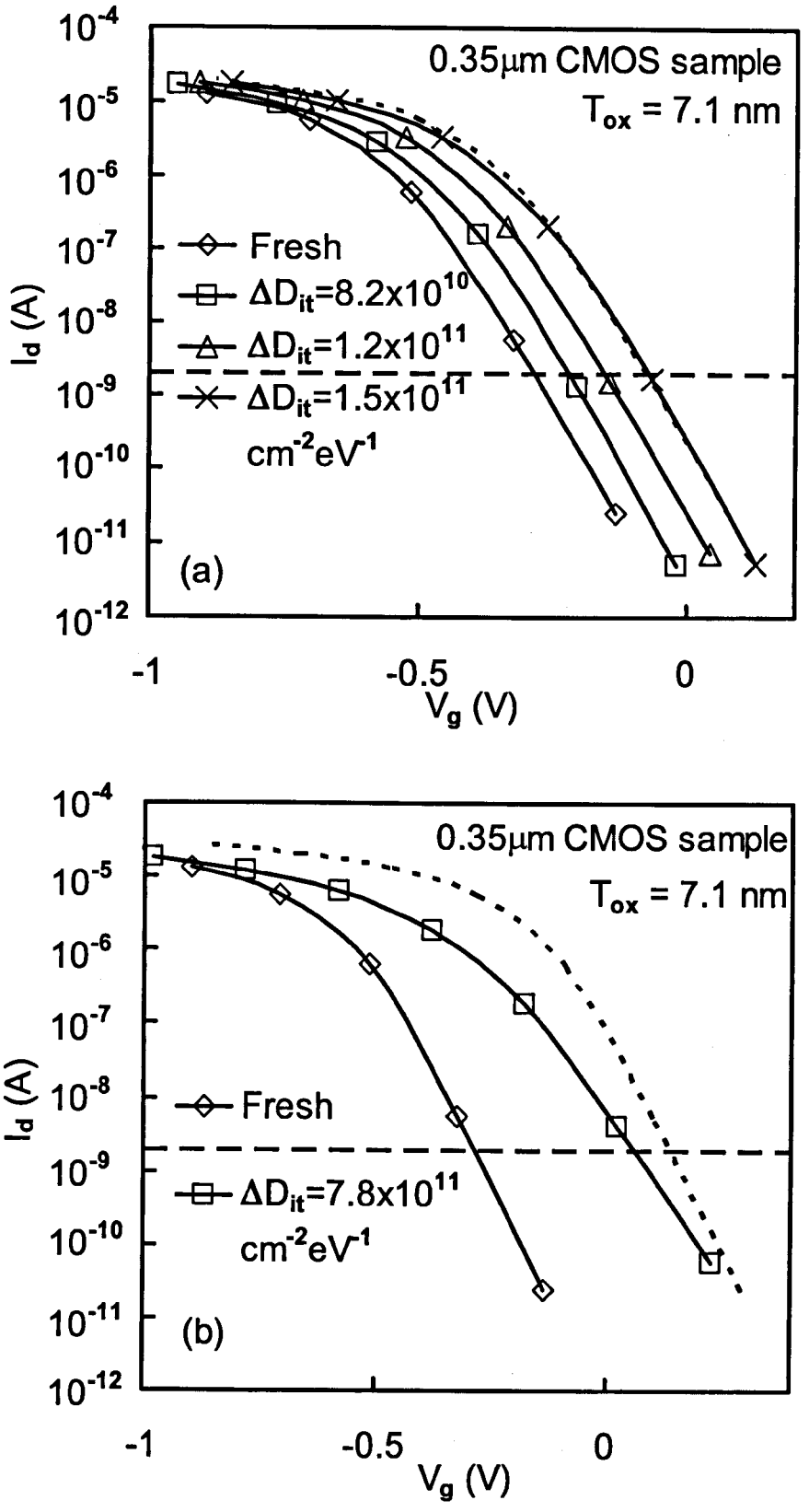


Figure 12. I_d-V_g characteristics. (a) Interface states generation is small, no distortion is noticeable. (b) Interface states generation is large, a clear distortion is observed. The dotted curve is a parallel shift of the fresh curve.

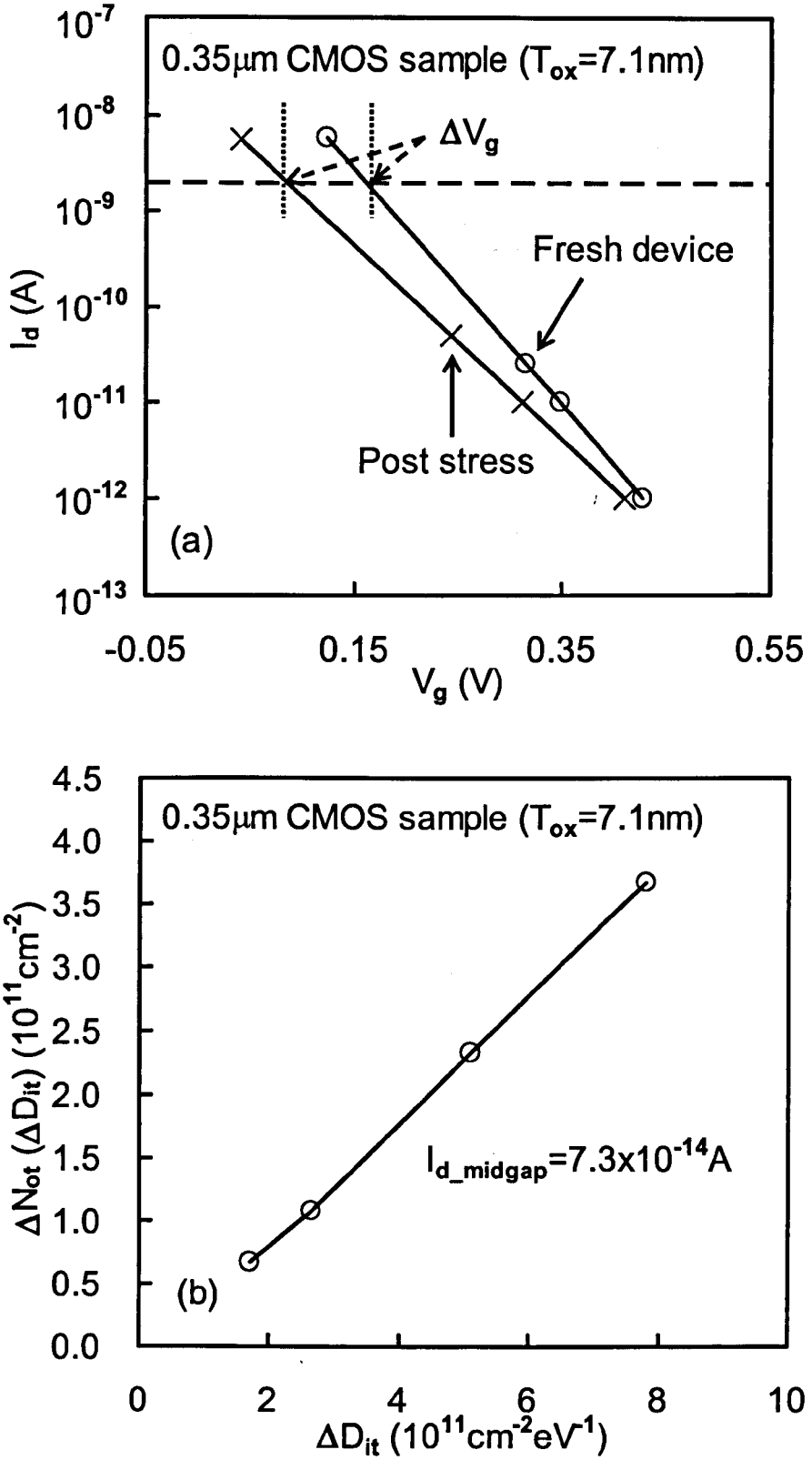


Figure 13. Evaluation of the effect of generated interface states on the measurement of trapped charges. (a) ΔV_g at I_d = 2 nA is caused by the distortion of the curve slope. (b) Correction shows ΔN_{ot} is proportional to the generated interface states.

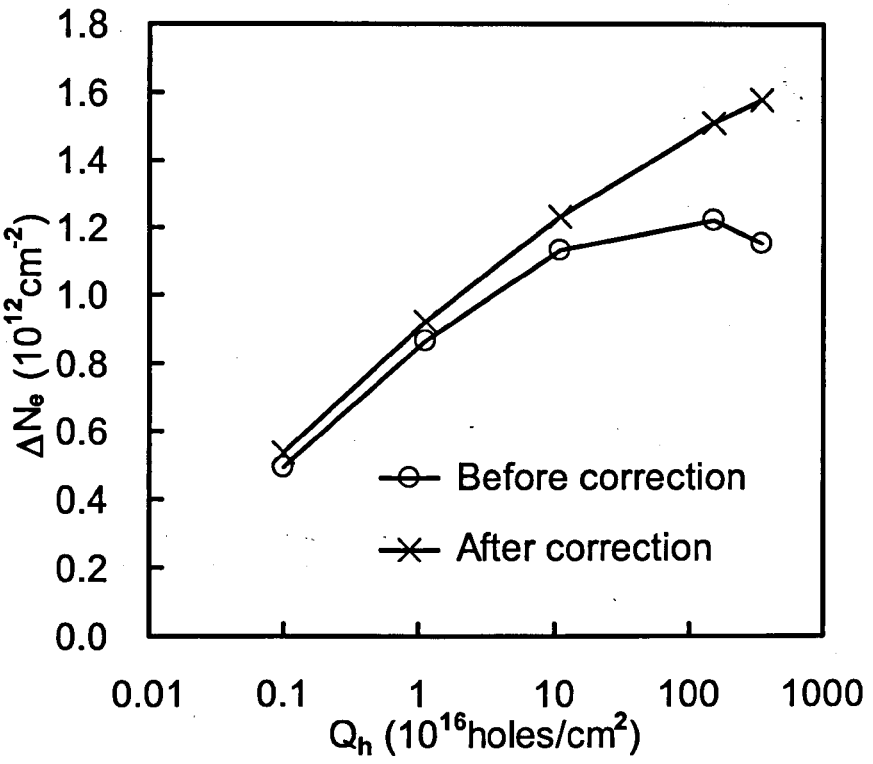


Figure 14. Comparison of the electron trap generation with and without the correction of the effect of generated interface states. Electron traps were generated by SHI then filled filled by FNI at 8 MV/cm.

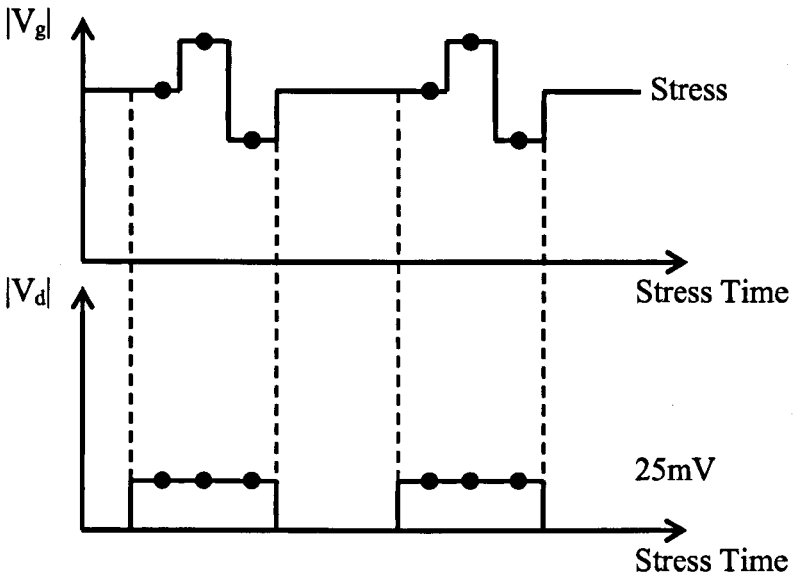


Figure 15. On-The-Fly characterization. A periodic pulse around the stress voltage, and a small drain voltage are applied simultaneously during the linear drain current measurement. Three drain currents are measured in each periodic pulse.

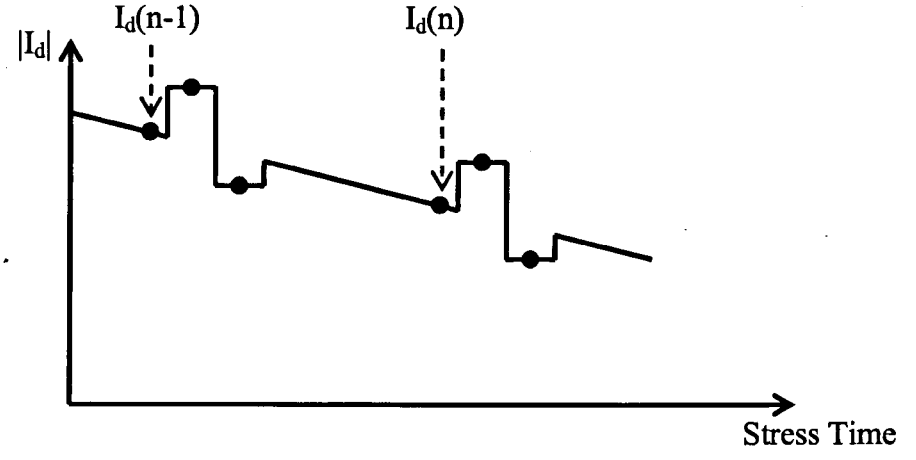


Figure 16. The n^{th} and $n\text{-}1^{th}$ I_d measurements, together with the transconductance $g_m(n)$, can gives the threshold voltage shift, ΔV_t, between the n^{th} and $n\text{-}1^{th}$ measurement points

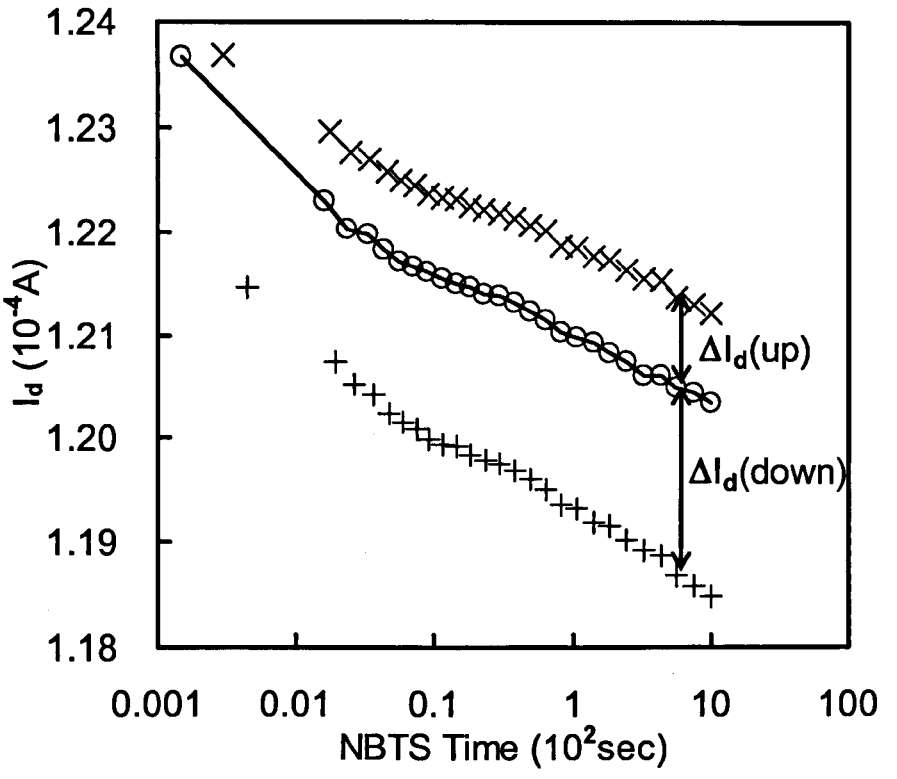


Figure 17. Measured drain currents during NBTS. DV = 250 mV. Symbol 'o', 'x', and '+' were measured at V_g, V_g - DV, and V_g + DV, respectively.

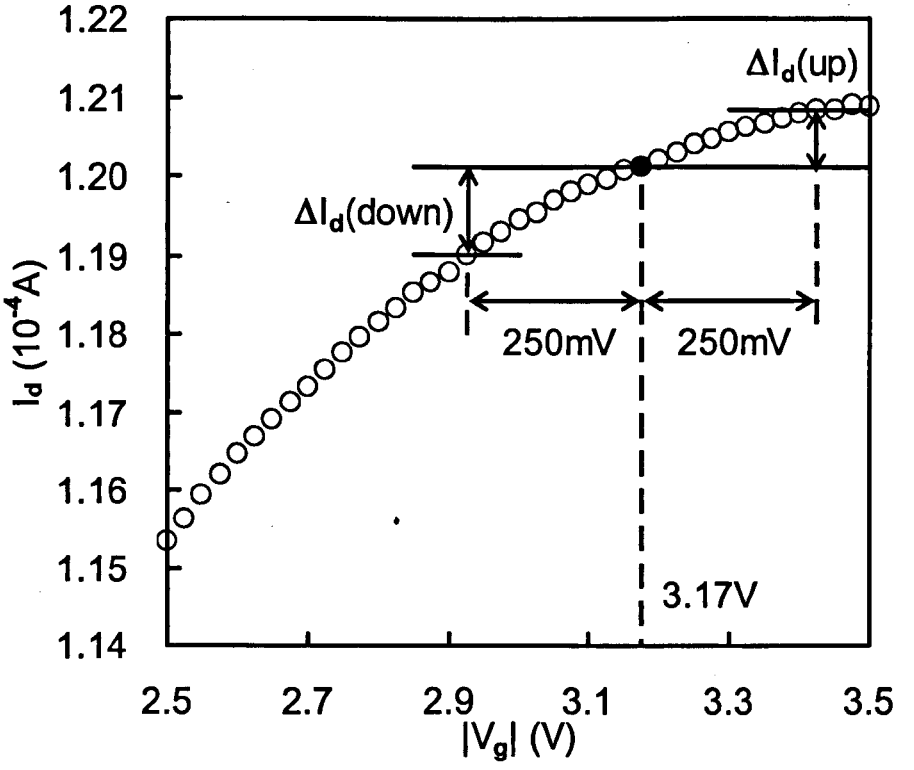


Figure 18. At relatively high V_g, I_d tends to saturate. A large DV will induce larger ΔI_d(down) than ΔI_d(up), which will affect the accuracy of g_m assessment.

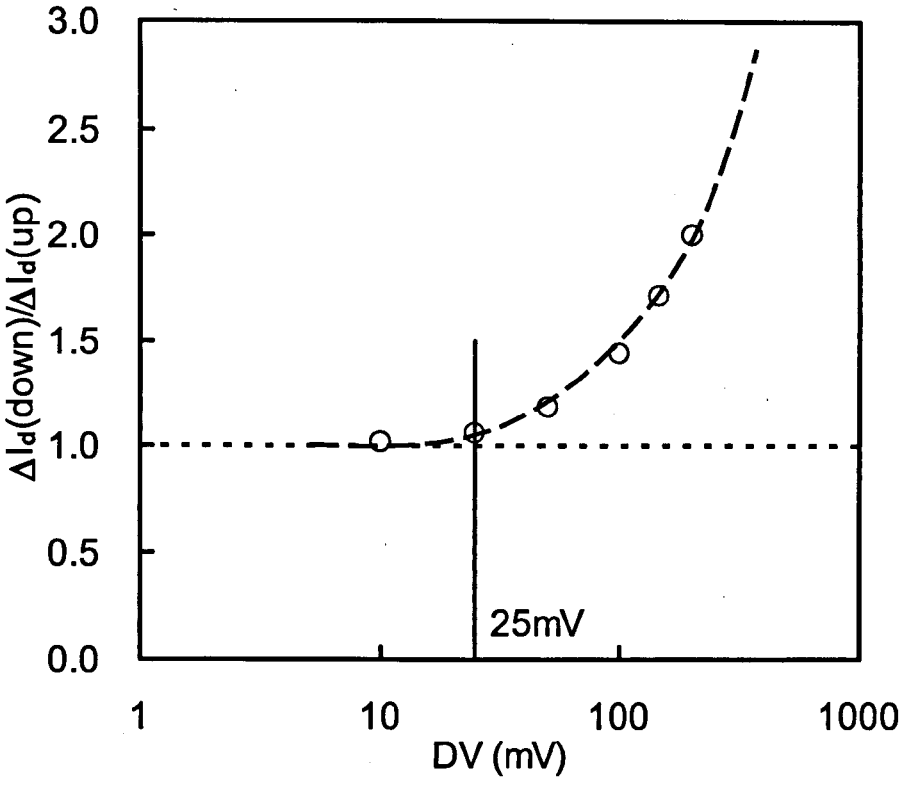


Figure 19. The ratio of $\Delta I_d(down)/\Delta I_d(up)$ under different DV. The ratio approaches to one when DV = 25 mV.

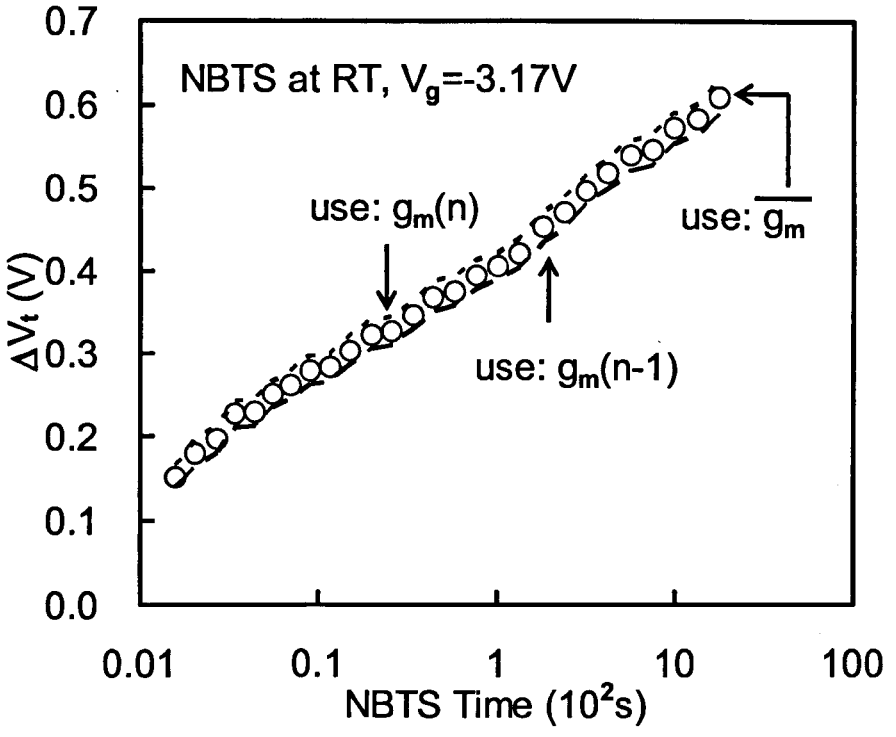


Figure 20. The effect of g_m on ΔV_t. Symbol 'o' is calculated using the average g_m value between two measurement points, while the dashed line uses the previous g_m value, and the dot line uses the current g_m value. It is clear that ΔV_t is weekly dependent on g_m value used.

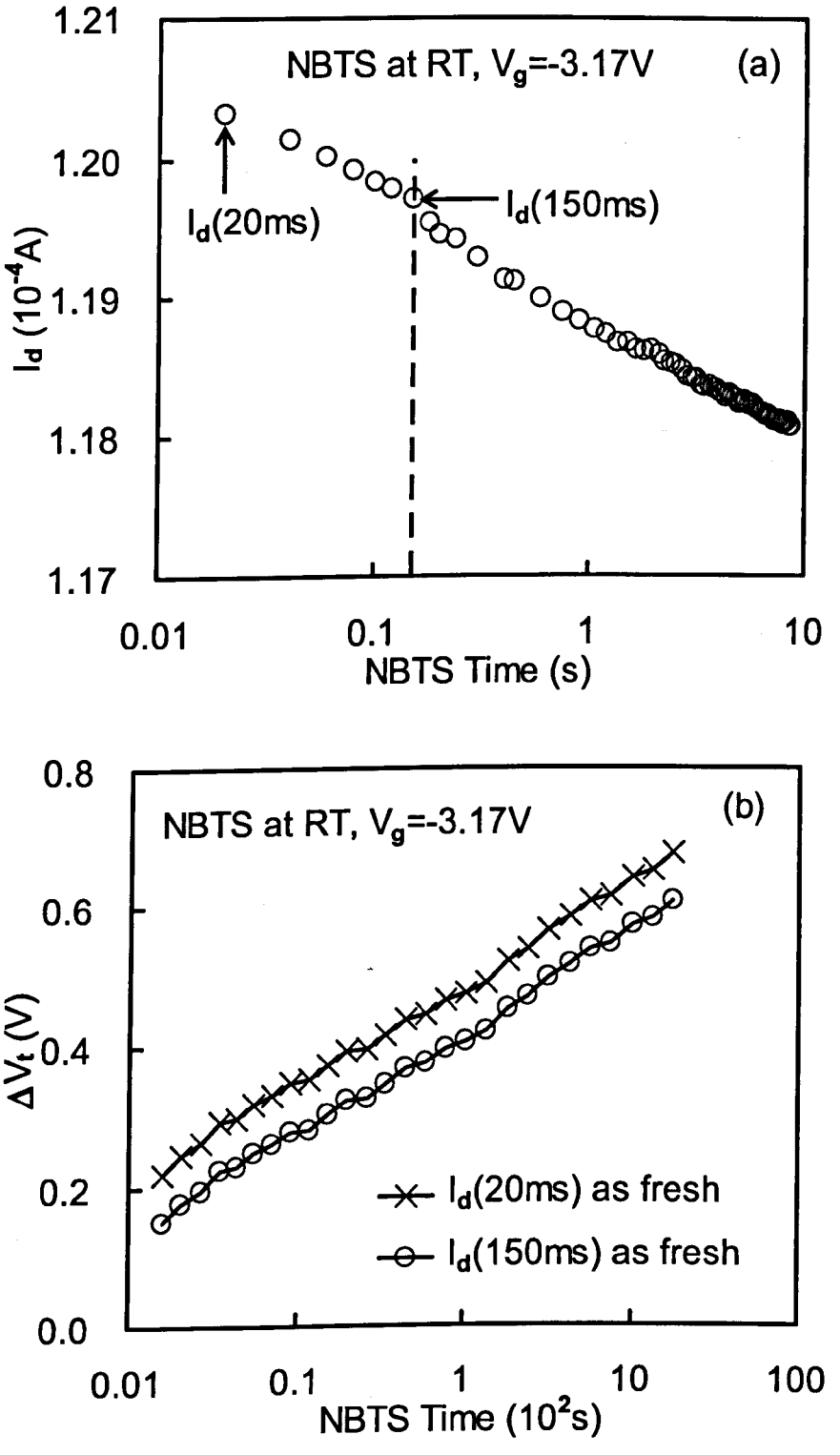


Figure 21. When I_d was measured for the first time, it took 0.15 sec and some degradation occurs during this period. (a) I_d degradation is monitored on a fresh device. (b) This uncertainty in the reference I_d leads to an underestimation of ΔV_t, but does not change the trend of generation.

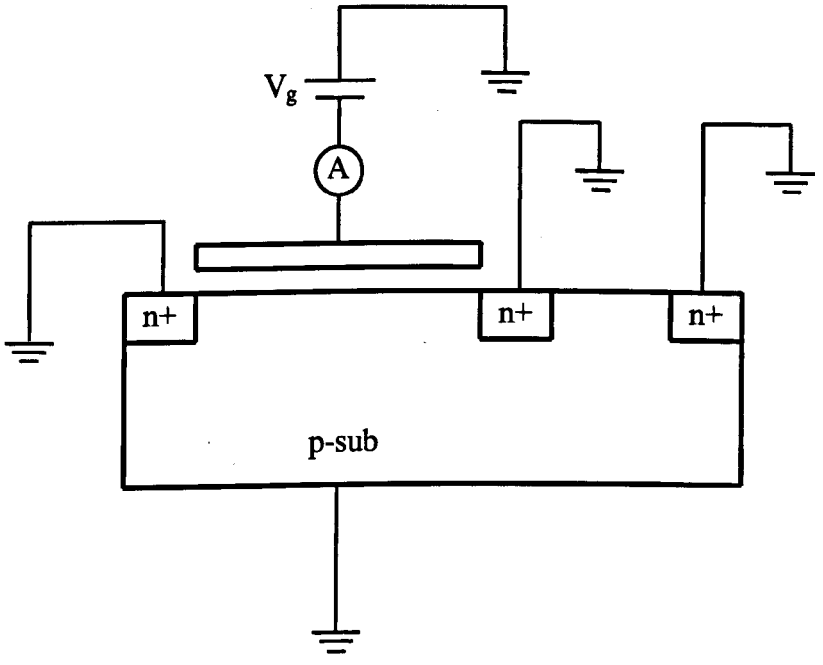


Figure 22. Experimental set up of Fowler-Nordheim Injection (FNI).

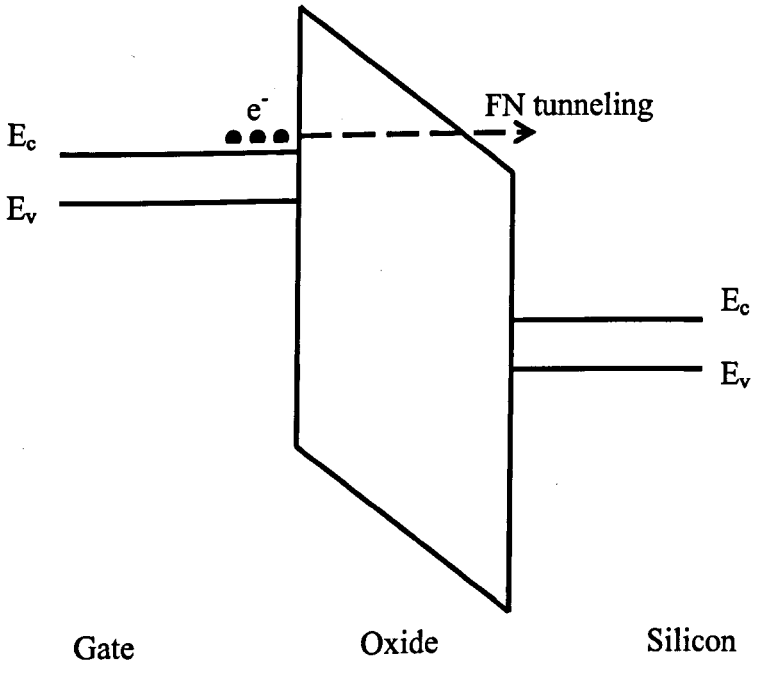


Figure 23. Schematic energy band diagram of FNI. When the oxide field is above 6 ~ 7 MV/cm, the physical distance between the conduction band of the gate and that of the substrate becomes so thin that electrons can tunnel through.

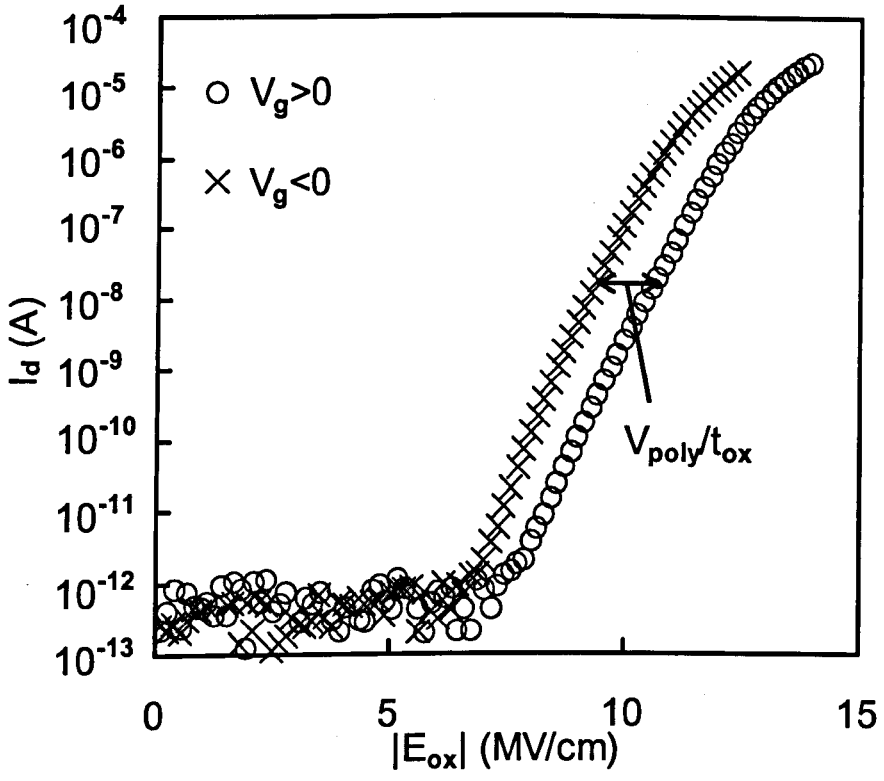


Figure 24. Evaluation of the effect of gate depletion in a 0.35 μm nMOSFET on the calculation of oxide field.

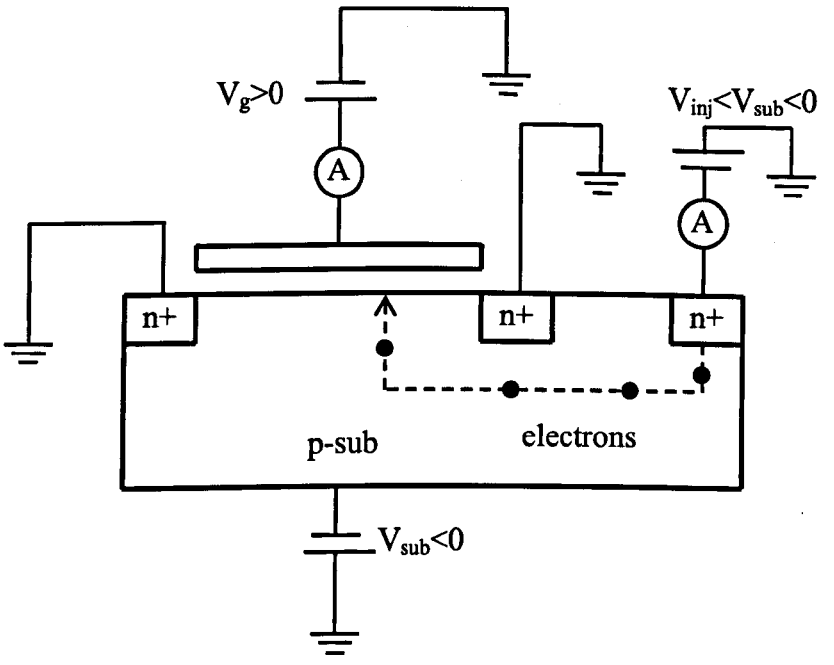


Figure 25. Experimental set up of Substrate Hot Electron injection (SHE). Electrons are supplied from the forward biased underlying pn junction diffuse into the space charge layer, then drift towards the Si/SiO_2 interface.

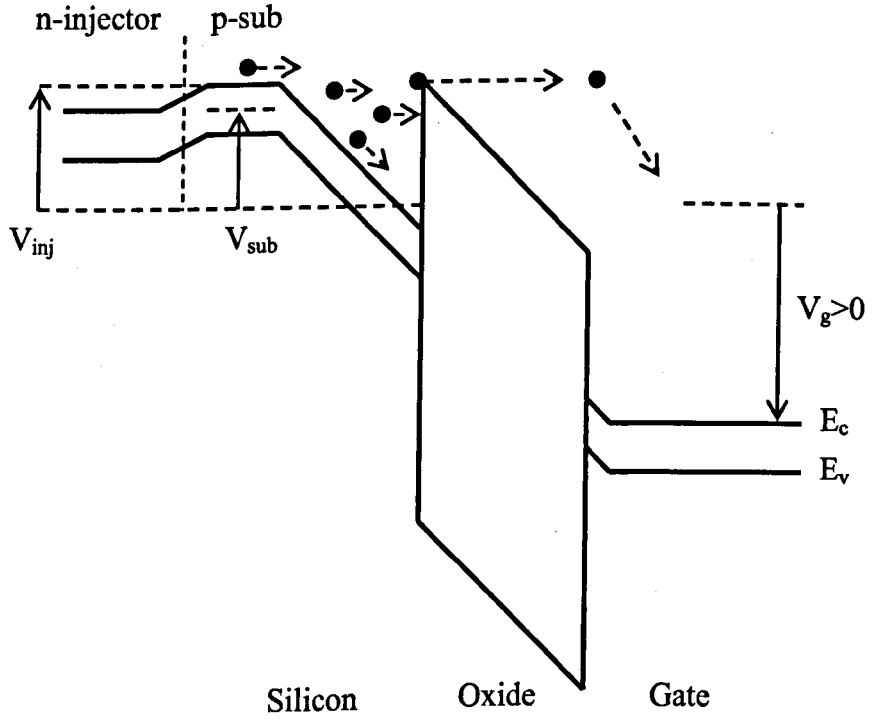


Figure 26. Schematic energy band diagram of SHE injection.

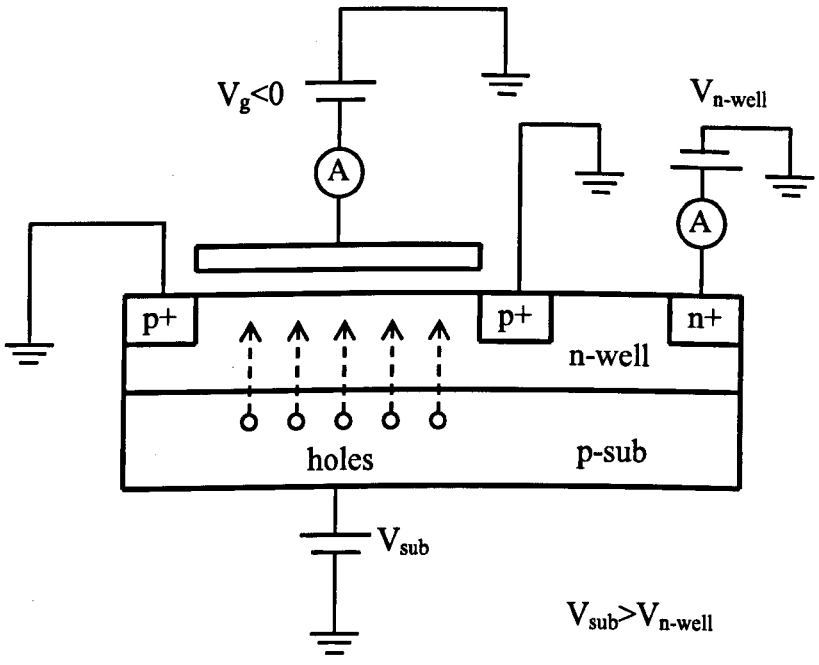


Figure 27. Experimental set up of Substrate Hole Injection (SHI). Holes are injected from the p-substrate into the n-well. They then diffuse towards the space charge region in which they are accelerated towards the Si/SiO_2 interface.

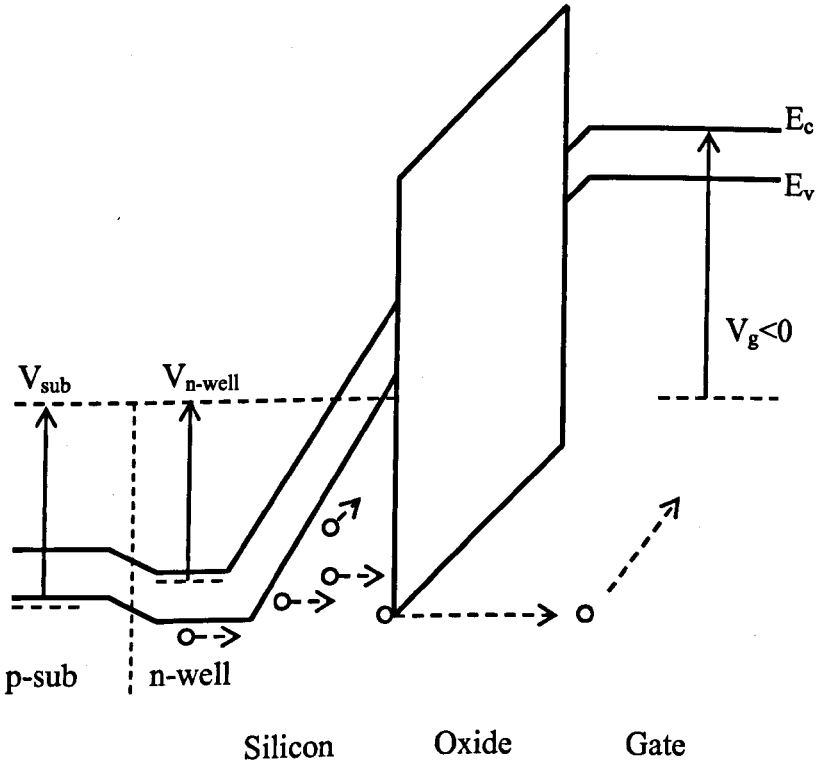


Figure 28. Schematic energy band diagram of SHI injection.

3 Role of Hydrogen in Hole-Induced Electron Trap Creation

3.1. Introduction

Defect generation is one important reliability issue for metal-oxide-semiconductor field-effect-transistors (MOSFETs). The defect generation has received much attention recently [1-3]. The generated defects can have a number of adverse effects on devices, including a shift in device parameters [4,5], the stress-induced-leakage current [2,6], and eventually, the oxide breakdown [2,7-9]. Defects can be formed in the bulk of gate dielectric [2,8-12], in the interfacial region [13,14], at the dielectric/Si interface [4,5,15-17] and in Si [18,19]. The gate dielectric used in the current industry is silicon dioxides (SiO_2) or silicon oxynitrides (SiO_xN_y). Both donor-like and acceptor-like traps can be created in these dielectrics [14,20,21]. The acceptor-like trap is commonly referred to as 'electron trap'. Despite past efforts, our understanding of the electron trap is still poor. For example, the microstructure of generated electron trap has not been unambiguously identified and agreement on damaging species has not been reached [1,2,22].

It is well known that hydrogenous species play an active and complex role both in the oxide/Si interfacial region [17,23-27] and in the bulk of silicon [18,19]. For example, although atomic hydrogen passivates interface states at 400°C, there is compelling evidence that they actually create interface states at room temperature [23,28]. Some researchers proposed that hydrogenous species also dominated electron trap creation [2]. During electrical stress, electrons injected into the oxide gain energy as they travel through the oxide. When they exit the oxide, the released energy can free hydrogenous species. These hydrogenous species then travel through the oxide and create electron traps [2]. The evidence presented for hydrogen-induced electron trap creation, however, is not as convincing as that for interface state

generation and agreement has not been reached among different groups of researchers [1,2,22].

Some researchers [1,22] proposed that holes injected during the stress could dominate electron trap creation. Here, holes can interact with the oxide directly [1,21] and hydrogen is not involved. This proposal is supported by two observations. First, it was shown that there was a good correlation between the electron trap generation and the fluency of holes through the oxide [9]. Second, it was found that holes were much more efficient than electrons in generating electron traps [22]. These observations, however, are not unambiguous enough to convince everyone [2]. For instance, it is fair to say that a correlation between hole injection and trap generation does not necessarily mean that there is a causal relationship between them. Holes, like electrons, may generate traps also through the release and subsequent transportation of hydrogenous species [2]. The higher generation efficiency of holes does not rule out the involvement of hydrogen. The greater generation can be caused by the higher energy delivered to electrode. This means that the supply of hydrogenous species is essential for creating electron traps.

In this chapter, we focus on whether holes can generate electron traps without going through hydrogen as intermediate species. To simplify testing conditions, the substrate hole injection (SHI) is chosen to stress devices. During the SHI, there are three potential regions for releasing hydrogen: Si/oxide interfacial region, oxide bulk and gate/oxide interfacial region. Tests will be carried out to assess their impact on hole-induced electron trap generation.

3.2. Experimental conditions

In previous works [2,6,8,9], stresses were often carried out under a condition where electron injection dominated in numbers. This is not suitable to study the hole-induced generation, since electrons could simultaneously release other damaging species. To concentrate on hole-induced electron trap creation, electron injection

during the stress should be suppressed. This can be achieved by using the substrate hole injection for creating electron traps.

The SHI technique was extensively used in previous work at this university [17,27,29]. Its principle and set up are explained early in section 2.5.2. In brief, during the SHI, gate was typically biased with an oxide field strength, E_{ox}, of -5 MV/cm. Under this E_{ox}, electron injection from the p+ poly-si gate was negligible for a 7.1 nm oxide. The source and drain was grounded. The n-well/p-substrate junction underneath the MOSFET was forward biased, which supplied holes from the substrate. These holes were then accelerated by the electrical field in the space charge region of n-well. Some holes became sufficiently energetic to overcome the potential energy barrier at the interface and be injected into the oxide.

The test follows a 'stress-then-sense' procedure [12,29] and a typical result is given in Figure 1. During the stress by hole injection, there was a gradual build up of trapped holes (symbol 'o'). After hole fluency reached a preset level, it was interrupted and 10^{17} cm^{-2} electrons were used to fill generated traps at an oxide field of +8 MV/cm (symbol 'x') [12]. Trapped electron density, N_e, was monitored from gate voltage shift in subthreshold region. Since the centroid of trapped charge is not known, the effective density is used by assuming the centroid being at the oxide/Si (n-well) interface, to comply with previous work in this area [12,20].

Care has been exercised to ensure that additional trap generation by the trap filling step is negligible, as is shown by the symbol '+' in Figure 2(a). Figure 2(a) also shows that higher hole fuency, Q_h, results in higher generation and hole injection is effective in creating electron traps. Attention was also paid to simultaneous interface state generation during stress. Interface state density was determined by standard charge pumping technique. Figure 2(b) shows that generated interface states are typically much less than created electron traps, where N_{es} represents the saturation level of N_e in Figure 1. Effects of these interface states on gate voltage shift were taken into account when determining electron trap density.

As mentioned earlier, electron injection was suppressed when stressing a device by substrate hole injection. There is little doubt that this leads to a considerable reduction of the amount of released hydrogen. However, some hydrogenous species can still be released by holes. Figure 3 shows that there are three important regions for releasing hydrogen: the Si (n-well)/oxide interfacial region, the bulk of oxide, and the gate/oxide interfacial region. In the following, their impact on hole-injection-induced electron trap generation will be examined one-by-one.

3.3. Hydrogen species released near the Si/SiO_2 interface

During the substrate hole injection, holes are accelerated in the space charge region as they travel towards the interface. It is well known that there are hydrogenous species at the Si/SiO_2 interface. For example, the precursor of interface states can appear in the form of Si-H [17,23,27]. When energetic holes bombard the Si/SiO_2 interface, hydrogenous species are released and new interface states are formed [27]. The released hydrogenous species could travel through the oxide and generate traps. The importance of this hydrogen release for electron trap generation during SHI is assessed by the following experiments.

3.3.1. Bombardment effects on electron trap generation

If hydrogen released at the Si/SiO_2 interface played an important role in electron trap generation, one would expect that the generation is closely related to the interface bombardment. In this section, we intend to examine whether this expectation can be experimentally observed.

Bombardment of the Si/SiO_2 is controlled by biases applied to the n-well, V_w, and p-substrate, V_s. (V_s - V_w) determines the number of holes, and V_w determines the energy of holes.

Experiment 1: The same interface bombardment, different generation

For the same V_w and V_s, the bombardment should be the same for a given time. Consequently, the same generation is expected. Figure 4 shows the result obtained on two devices under E_{ox} = -1 and -5 MV/cm, respectively, for the same V_w and V_s. Against the expectation, Figure 4(a) clearly shows that the generation is different for a given time. As a result, it does not appear that hydrogen released at the Si/SiO_2 interface plays an important role in the generation.

To explore why the generation in these two devices is different, we plot gate current due to hole injection in Figure 4(b). The different oxide field resulted in a considerable difference in the hole injection current. For a given bombardment time, lower generation under E_{ox} = -1 MV/cm could be caused by the smaller number of injected holes. When the generated traps were plotted against hole fluency, Figure 4(c) shows that there is little difference between these two devices, and there is a good correlation between electron trap generation and hole injection.

Experiment 2: Higher carrier energy and bombardment, the same generation

The above experimental results show that the amount of holes bombarding the Si/SiO_2 interface is not the dominant factor controlling the generation. However, it is possible that their energy at which they arrive at the interface could control the generation. One would expect that the more energetic the holes, the more hydrogenous species would be released. If this process is important, it should lead to higher generation.

Two devices were stressed separately by SHI. V_s - V_w is the same for both devices, so the same amount of holes will arrive at the interface. A higher V_w was applied on one device to make the holes 'hotter', thus the energy effects on generation can be observed. The oxide field is then adjusted to obtain similar gate current during SHI as shown in Figure 5(a). This is important to eliminate trap generation difference caused by different hole injection current. Figure 5(b) shows the electron trap generation is clearly insensitive to V_w. Figure 5(c) indicates that an increase of the

hot carrier energy at the interface is not significant enough to have an observable effect on the generation.

Experiment 3: Larger number of bombarding holes, the same generation

Two devices were stressed under E_{ox} = -3.5 MV/cm, V_w = 9.3 V, V_s = 9.88 V and E_{ox} = -3.5 MV/cm, V_w = 6.8 V, V_s = 7.8 V, respectively. The injection current during the stress is plotted in Figure 6(a), which is similar for the two stress conditions. However, an increase of V_s - V_w from 0.53 to 1 V should substantially enhance the number of holes supplied to the interface. Figure 6(b) shows that there is little difference in the generation.

The observations made from the above experiments are that electron trap generation is not controlled by the energy and the number of holes bombarding the Si/SiO_2 interface. The same bombardment does not necessarily leads to the same electron trap generation. A good correlation between hole injection and electron trap generation is found. As mentioned in the introduction, a good correlation, however, does not necessarily mean that hole injection is responsible for electron trap generation. The relation between these two will be further investigated.

3.3.2. Threshold for electron trap generation and hydrogen release

In this section, we study whether threshold voltage for creating electron traps and releasing hydrogen is different. To start, V_w = 3 V and V_s - V_w = 0.7 V was applied for 4600 s under E_{ox} = -1 MV/cm. The V_w was then increased by a series of voltage step of 0.5 V. For each level of V_w, V_s - V_w was maintained at 0.7 V and the stress time was kept at 4600 s. The gate current, I_g, generated electron traps, N_{es}, and trapped hole, N_h, were recorded and plotted in Figure 7(a) and (b).

It is clear that there is a threshold V_w of 5 V approximately for I_g, N_{es} and N_h. This is in good agreement with the potential barrier height for holes at the interface (~ 4.8 eV). The same threshold value for both hole injection and electron trap generation strongly supports that hole injection is responsible for creating electron traps under

our experimental conditions. In contrast, Figure 7(c) shows no clear threshold value for interface state generation, D_{it}. For the lowest V_w applied here (3 V), some interface states can be created already through bombardment, although hole injection is negligible. Since rupturing hydrogen bond is a main source for creating interface states [23,27], the generation of interface states can be used as a detector for hydrogen release at the interface. The different behaviour between electron trap generation and hydrogen release at Si/SiO_2 interface further supports that the latter is not the main source for the former during SHI.

3.4. Hydrogen species released in the SiO_2

Since hole mobility in the oxide is six orders of magnitude lower than electron mobility [30], we do not expect that acceleration of holes in the oxide will be important for releasing hydrogen and the subsequent electron trap generation. This is confirmed by Figure 4(c), which shows that the generation at a given Q_h is clearly insensitive to the oxide field strength. It, however, does not mean that holes cannot free hydrogenous species in the oxide. Previous works at this university [17,31,32] showed that hydrogenous species could be released by both hole trapping and detrapping. If we use D–H representing a hydrogen-related hole trap, the process for releasing hydrogen can be schematically illustrated below.

During hole trapping:

$$\text{Hole} + \text{D–H} \rightarrow \text{D}\cdots\text{H}^+ \rightarrow \text{D}\bullet + \text{H}^+ \lessgtr \text{Released}$$

During detrapping:

$$\text{electron} + \text{D}\cdots\text{H}^+ \rightarrow \text{D} + \text{H}^0 \lessgtr \text{Released}$$

where the dotted line represents a weakened bond. Although hydrogenous species can be emitted both in the form of H^+ and H^0, the emission of H^0 is a more efficient process [31].

To study the importance of hydrogen released through the above processes on electron trap generation, it is desirable to be able to vary hydrogen concentration in a device. If the hydrogen released in the oxide is important for electron trap creation, an increase of its concentration should substantially enhance the electron trap generation. Previous work at this university [17,29,31] unambiguously showed that hydrogen concentration in a device could be significantly increased, by exposing a stressed device with holes trapped in the oxide to H_2 at 400°C. Hydrogen molecules were cracked by trapped holes [29,33], which resulted in the increase of reactive hydrogenous species. Figure 8 shows a typical example of H_2 exposure effects. The interface states post H_2 exposure was significantly enhanced, compared with a normal device without H_2 exposure.

In Figure 9(a) and (b), a device was first stressed by hole injection to a level of 10^{15} holes/cm^2 (symbol 'o'). It was then exposed to forming gas (10% H_2) for 40 min at 400°C. After the exposure, hole injection was resumed and the generated electron traps and interface states are represented by symbol 'x' in Figure 9(a) and (b), respectively. For comparison, typical results (symbol '△') are also given when a device was not exposed to forming gas. As expected, the significant enhancement in interface state generation implies that the hydrogen concentration was substantially increased after the forming gas exposure. The generation of electron traps, however, has not been substantially increased. This indicates that hydrogen does not play a significant role in hole-induced electron trap creation during SHI.

3.5. Hydrogen species released near the gate/SiO_2 interface

Apart from oxide/Si interface and oxide bulk, Figure 3 shows that hydrogen can also be released at the poly-Si/oxide interface and/or within the poly-Si gate. When holes arrive at the gate, they lose their energy, which could lead to release of hydrogen. In the presence of holes, it is likely that majority of released hydrogenous species can capture holes and become positively charged [16,34]. These positive species will not travel through the oxide under a negative bias during the SHI. Consequently, they will not contribute to the generation in the oxide. However, we cannot rule out that

there are some neutral hydrogenous species, which can pass through the oxide and create traps. To find how important this is, the following tests were carried out.

Experiment 1: Nidridation

Firstly, it was reported that nitridation blocked not only boron, but also hydrogenous species [35], as illustrated in Figure 10. One can expect that, for the same amount of holes arrived at the gate, there will be less hydrogen reaching the dielectric/substrate interface after nitridation. This in turn should leads to less interface states generation for oxynitrides. Figure 11, however, shows that the generated interface states are actually more for the oxynitride. This indicates that, during the substrate hole injection, the hydrogen movement from the gate to the dielectric/substrate interface is insignificant.

Experiment 2: High temperature:

Secondly, it is well known that hydrogen release and the transportation in the oxide are thermally activated [34,36,37]. If hydrogen plays an important role in electron trap generation, one would expect that the generation should also be thermally activated. This activation was observed when electrons were injected into the oxide during stress [8]. During SHI, electron injection was suppressed.

Four devices were stressed by SHI (E_{ox} = -5 MV/cm, V_w = 8.8 V, V_s = 9.8 V) at room temperature, 65°C, 100°C and 150°C, respectively. After the hole injection reached the preset level, the stress was interrupted, and the devices were cooled down to room temperature before the filling of generated electron traps.

Figure 12 shows that hole-induced generation is insensitive to temperature up to 150°C. This is clearly in contrast with the thermally enhanced generation, when electron is injected into the oxide. This insensitivity of hole-induced electron trap generation supports that hydrogen released from the other two regions, the Si/oxide interface and the oxide bulk, are not important, either.

Experiment 3: Low temperature:

Thirdly, the hydrogen release and the transportation in the oxide were examined at low temperature. At 77K, the hydrogen movement should be frozen [34]. If hydrogen plays an important role in electron trap generation, one would expect that the generation is significantly reduced.

A fresh device was stressed by SHI (E_{ox} = -5 MV/cm, V_w = 8.8 V, V_s = 9.8 V) at 77K. After the hole injection reached the preset level, the stress was interrupted and followed by trap filling at 77K. Figure 13 shows that up to $Q_h = 10^{16}$ holes/cm^2, both the electron trapping kinetics and electron trap generation are insensitive to temperature. However, with increasing hole injection, a drop in generation is observed for 77K device. This drop is being caused by the generation of so-called 'anti-neutralization positive charges, (ANPC)' whose neutralization is thermally activated [38].

To confirm the above explanation, a second test was carried out using an already heavily stressed device (previous stress level $Q_h = 2\times10^{18}$ holes/cm^2). Figure 14(a) shows the experimental sequence. The heavily stressed device was filled at room temperature first (symbol '●'), then cooled down to 77K and filled again (symbol '◇'). The cooled down process does not affect the level of generated electron traps, as similar amount is again observed at 77K. A comparatively short SHI ($Q_h = 10^{17}$ holes/cm^2, symbol 'x') is applied to induce positive charges with little further generation of electron traps. This time, the electron trapping at 77K (symbol '△') appears lower than that before SHI (symbol '◇'), because the ANPC are not neutralized at 77K. The positive charge offsets the true level of electron trapping. By warming up to room temperature, the neutralization of ANPC become easier and the level of electron trapping recovers, as shown in Figure 14(b).

To remove the impact of temperature on trap filling, Figure 15 shows a fresh device stressed by SHI (E_{ox} = -5 MV/cm, V_w = 8.8 V, V_s = 9.8 V) at 77K, but filled at room temperature. The electron trap generation at 77K (symbol '◇') is similar compare to the one at room temperature (symbol 'o'). This temperature insensitivity further

supports that hydrogen released from the other two regions, the Si/oxide interface and the oxide bulk, are not important, either.

3.6. Conclusions

In this chapter, we focused on whether holes can generate electron traps without going through hydrogen as intermediate species. Under substrate hole injection, hydrogenous species can potentially be released from three regions: Si/oxide interfacial region, oxide bulk and gate/oxide interfacial region. The impact of hydrogenous species released from these three regions on the generation was examined one-by-one.

Near the Si/oxide interface, it was shown that electron trap generation could be different for the same bombardment of interface and, consequently, the same hydrogen release. The threshold n-well voltage for electron trap generation is also different from that for interface states generation resultant from breaking the hydrogen bond. These results do not support that hydrogenous species released from the Si/oxide interface are important for hole-induced electron trap generation during SHI.

Under our experimental conditions, electron trap generation is strongly correlated with hole injection. These two also have the same threshold n-well voltage, indicating that hole injection controls the generation. After hole injection, its acceleration in the oxide is not responsible for creating electron traps. An increase of hydrogen density in a device has little impact on the hole-induced generation.

Finally, near the gate/oxide interface, any released H^+ would not travel through the oxide under negative gate bias. Since hydrogen-release and the subsequent transportation are thermally accelerated processes, the insensitivity of hole-induced electron trap generation to temperature in the range of 77K to 150°C suggests that these processes are not important for hole-induced electron trap creation. Our results

support that holes can interact directly with the oxide to generate electron traps without going through hydrogen as intermediate species.

References

1. D. Esseni, J. D. Bude and L. Selmi, "On interface and oxide degradation in VLSI MOSFETs - Part I: Deuterium effect in CHE stress regime," IEEE Trans. Elec. Dev., **49**, 247 (2002).
2. D. J. DiMaria and J. H. Stathis, "Anode hole injection, defect generation, and breakdown in ultrathin silicon dioxide films," J. Appl. Phys., **89**, 5015 (2001).
3. D. Ielmini, A. S. Spinelli, A. L. Lacaita and A. Modelli, "A new two-trap tunneling model for the anomalous stress-induced leakage current (SILC) in Flash memories," Microelectronic Eng., **59**, 189 (2001).
4. W. L. Chen, A. Balasinski and T. P. Ma, "Lateral profiling of oxide charge and interface traps near MOSFET junctions," IEEE Trans. Elec. Dev., **40**, 187 (1993).
5. J. F. Zhang and W. Eccleston, "Effects of high field injection on the hot carrier induced degradation of submicrometer pMOSFET's," IEEE Trans. Elec. Dev., **42**, 1269 (1995).
6. D. Ielmini, A. S. Spinelli, M. A. Rigamonti and A. L. Lacaita, "Modeling of SILC based on electron and hole tunneling - Part I: Transient effects," IEEE Trans. Elec. Dev., **47**, 1258 (2000).
7. M. J. Kirton and M. J. Uren, "Noise in solid-state microstructures – A new perspective on individual defects, interface states and low-frequency (1/F) noise," Adv. Phys., **38**, 367 (1989).
8. B. Kaczer, R. Degraeve, N. Pangon and G. Groeseneken, "The influence of elevated temperature on degradation and lifetime prediction of thin silicon-dioxide films," IEEE Trans. Elec. Dev., **47**, 1514 (2000).
9. R. Degraeve, G. Groeseneken, R. Bellens, J. L. Ogier, M. Depas, P. J. Roussel and H. E. Maes, "New insights in the relation between electron trap generation and the statistical properties of oxide breakdown," IEEE Trans. Elec. Dev., **45**, 904 (1998).
10. J. F. Zhang, S. Taylor and W. Eccleston, "A quantitative investigation of electron trapping in SiO_2 under Fowler-Nordheim stress," J. Appl. Phys., **71**, 5989 (1992).

11. L. Vishnubhotla, T. P. Ma, H. H. Tseng and P. J. Tobin"Interface trap generation and electron trapping in fluorinated SIO2," Appl. Phys. Lett., **59**, 3595 (1991).
12. W. D. Zhang, J. F. Zhang, M. Lalor, B. Burton, G. Groeseneken and R. Degraeve, "Two types of neutral electron traps generated in the gate silicon dioxide," IEEE Trans. Elec. Dev., **49**, 1868 (2002).
13. D. M. Fleetwood and N. S. Saks, "Oxide, interface, and border traps in thermal, N2O, and N2O-nitrided oxides," J. Appl. Phys., **79**, 1583 (1996).
14. D. M. Fleetwood, P. S. Winokur, O. Flament and J. L. Leray, "Stability of trapped electrons in SiO2," Appl. Phys. Lett., **74**, 2969 (1999).
15. N. Haneji, L. Vishnubhotla and T. P. Ma, "Possible observation of PB0 and PB1 centers at irradiated (100)Si/SiO2 interface from electrical measurements," Appl. Phys. Lett., **59**, 3416 (1991).
16. D. B. Brown and N. S. Saks, "Time-dependence of radiation-induced interface trap formation in metal-oxide-semiconductor devices as a function of oxide thickness and applied field," J. Appl. Phys., **70**, 3734 (1991).
17. C. Z. Zhao, J. F. Zhang, G. Groeseneken, R. Degraeve, J. N. Ellis and C. D. Beech, "Interface state generation after hole injection," J. Appl. Phys., **90**, 328 (2001).
18. C. T. Sah, J. Y. C. Sun and J. J. T. Tzou, "Deactivation of the boron acceptor in silicon by hydrogen," Appl. Phys. Lett., **43**, 204 (1983).
19. J. Coutinho, O. Andersen, L. Dobaczewski, K. B. Nielsen, A. R. Peaker, R. Jones, S. Oberg and P. R. Briddon, "Effect of stress on the energy levels of the vacancy-oxygen-hydrogen complex in Si," Phys. Rev. B, **68**, 184106 (2003).
20. D. J. DiMaria, The properties of electron and hole traps in thermal silicon dioxide layers grown on silicon. In S. T. Pantelides (ed.), *The Physics of SiO2 and its Interfaces*, New York: Pergamon, 160, 1978.
21. J. F. Zhang, S. Taylor, and W. Eccleston, "Electron trap generation in thermally grown SiO_2 under Fowler-Nordheim stress," J. Appl. Phys., **71**, 725 (1992).
22. J. D. Bude, B. E. Weir and P. J. Silverman, "Explanation of stress-induced damage in thin oxides," in Tech. Digest IEEE 44^{th} Int. Electron Devices Meeting, San Francisco, p.179, USA (1998).

23. L. Do Thanh and P. Balk, "Elimination and generation of Si-SiO_2 interface traps by low-temperature hydrogen annealing," J. Electrochem. Soc., **135**, 1797 (1988).
24. R. E. Stahlbush, B. J. Mrstik and R. K. Lawrence, "Postirradiation behavior of the interface state density and the trapped positive charge," IEEE Trans. Nucl. Sci., **37**, 1641 (1990).
25. W. B. Fowler and A. H. Edwards, "Defects and defect processes in silicon dioxide," Radiat. Eff. Defects Solids, **146**, 11 (1998).
26. D. M. Fleetwood, "Effects of hydrogen transport and reactions on microelectronics radiation response and reliability," Microelectronics Reliability, **42**, 523 (2002).
27. W. D. Zhang, J. F. Zhang, M. J. Uren, G. Groeseneken, R. Degraeve, M. Lalor and B. Burton, "On the interface states generated under different stress conditions," Appl. Phys. Lett., **79**, 3092 (2001).
28. R. E. Stahlbush and E. Cartier, "Interface defect formation in MOSFETs by atomic-hydrogen exposure," IEEE Trans. Nucl. Sci., **41**, 1844 (1994).
29. J. F. Zhang, H. K. Sii, R. Degraeve and G. Groeseneken, "Mechanism for the generation of interface state precursors," J. Appl. Phys., **87**, 2967 (2000).
30. R. C. Hughes, "High-Field Electronic Properties of SiO2," Solid-State Electron., **21**, 251 (1978).
31. J. F. Zhang, C. Z. Zhao, G. Groeseneken and R. Degraeve, "Analysis of the kinetics for interface state generation following hole injection," J. Appl. Phys., **93**, 6107 (2003).
32. I. S. Al-kofani, J. F. Zhang and G. Groeseneken, "Continuing degradation of the SiO2/Si interface after hot hole stress," J. Appl. Phys., **81**, 2686 (1997).
33. R. E. Stahlbush, A. H. Edwards, D. L. Griscom and B. J. Mrstik, "Postirradiation cracking of H2 and formation of interface states in irradiated metal-oxide-semiconductor field-effect transistors," J. Appl. Phys., **73**, 658 (1993).
34. N. S. Saks, R. B. Klein and D. L. Griscom, "Formation of interface traps in MOSFETs during annealing following low-temperature irradiation," IEEE Trans. Nucl. Sci., **35**, 1234 (1988).

35. E. Cartier, D. A. Buchanan and G. J. Dunn, "Atomic hydrogen-induced interface degradation of reoxidized-nitrided silicon dioxide on silicon," Appl. Phys. Lett., **64**, 901 (1994).
36. G. Van den bosch, G. Groeseneken, H. E. Maes, R. B. Klein and N. S. Saks, "Oxide and interface degradation resulting from substrate hot-hole injection in metal-oxide-semiconductor field-effect transistors at 295 and 77 K," J. Appl. Phys., **75**, 2073 (1994).
37. N. S. Saks, R. B. Klein, R. E. Stahlbush, B. J. Mrstik and R. W. Rendell, "Effects of past-stress hydrogen annealing on MOS oxides after CO-60 irradiation or Fowler-Nordheim injection," IEEE Trans. Nucl. Sci., **40**, 1341 (1993).
38. J. F. Zhang, C. Z. Zhao, A. H. Chen, G. Groeseneken and R. Degraeve, "Hole-traps in silicon dioxides - Part I: Properties," IEEE Trans. Electron Dev., **51**, 1267 (2004).

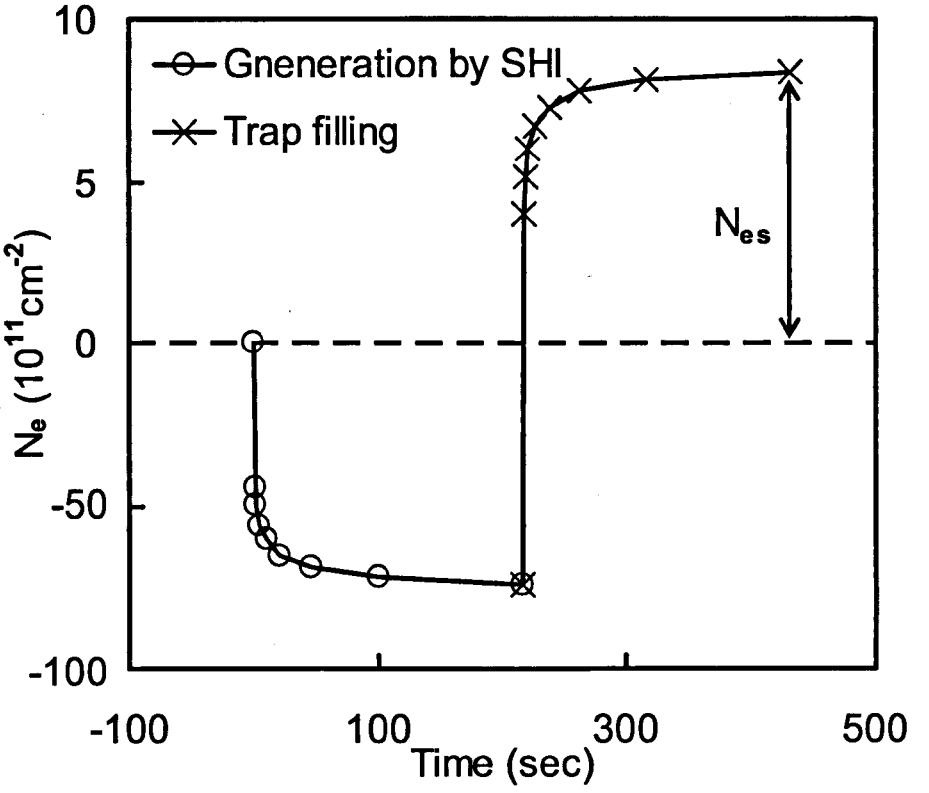
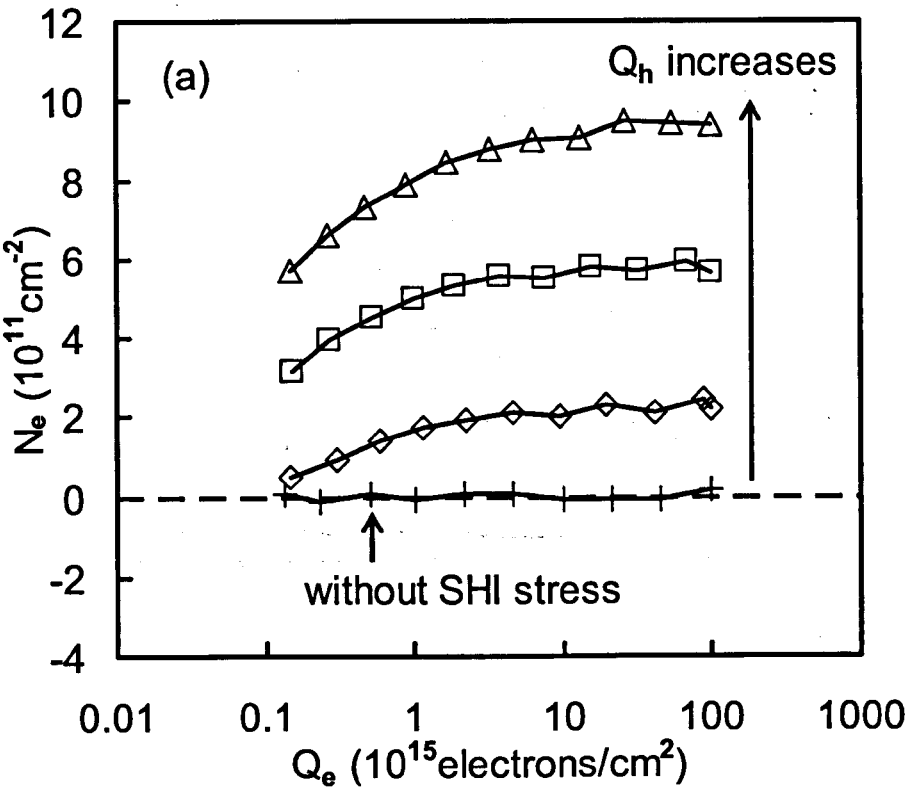


Figure 1. Hole-induced electron trap creation. N_e is the effective density of electron trapping. Device was stressed by SHI with an oxide field strength of E_{ox} = -5 MV/cm, a n-well bias of V_w = 9 V and a p-substrate of V_s = 10 V, which leads to the build-up of trapped holes (symbol 'o'). Generated electron traps were then filled by an electron injection of 10^{17} cm^{-2} under E_{ox} = +8 MV/cm (symbol 'x'). N_{es} is the saturation value of electron trapping.


(a)
Q_h increases
without SHI stress
N_e (10^{11}cm^{-2})
Q_e (10^{15}electrons/cm^2)
12
10
8
6
4
2
0
-2
-4
0.01
0.1
1
10
100
1000

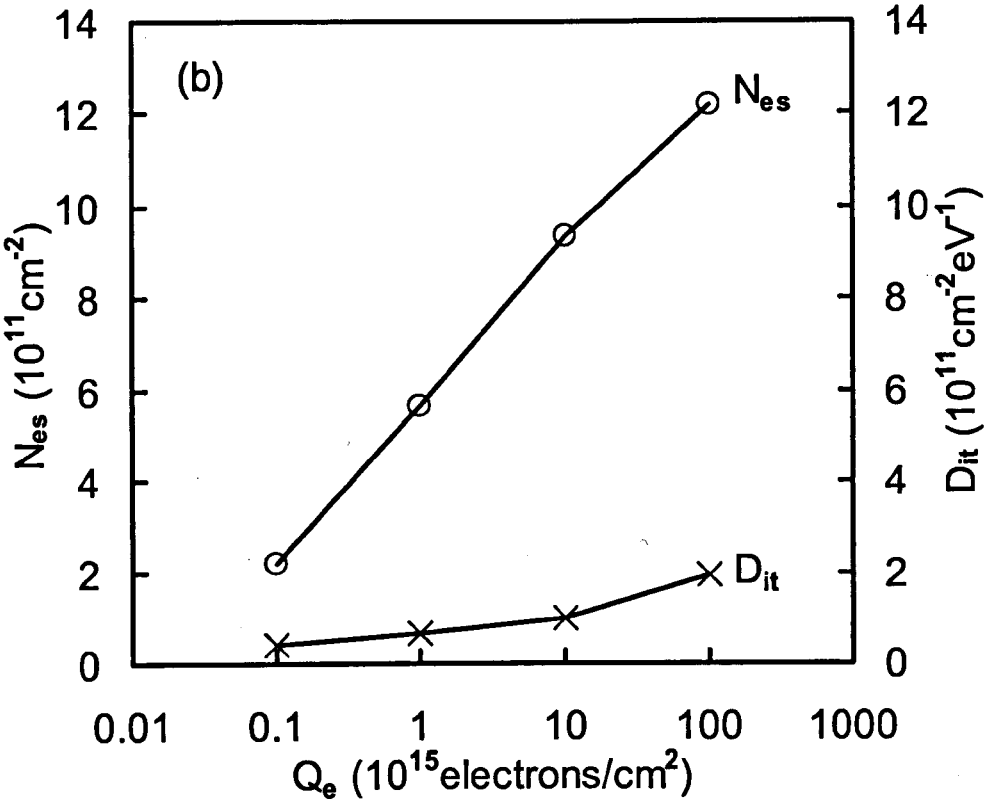

(b)
N_{es}
D_{it}
N_{es} (10^{11}cm^{-2})
D_{it} (10^{11}cm^{-2}eV^{-1})
Q_e (10^{15}electrons/cm^2)
14
12
10
8
6
4
2
0
0.01
0.1
1
10
100
1000

Figure 2. (a) Shows the higher the hole fluency, Q_h, the more traps were created. Without hole injection (symbol '+'), electron trapping is negligible. (b) Compares N_{es} with the density of generated interface states D_{it}. Symbols '+', '◇', 'o' and '△' correspond to Q_h = 0, 10^{14}, 10^{15} and 10^{16} holes/cm^2, respectively.

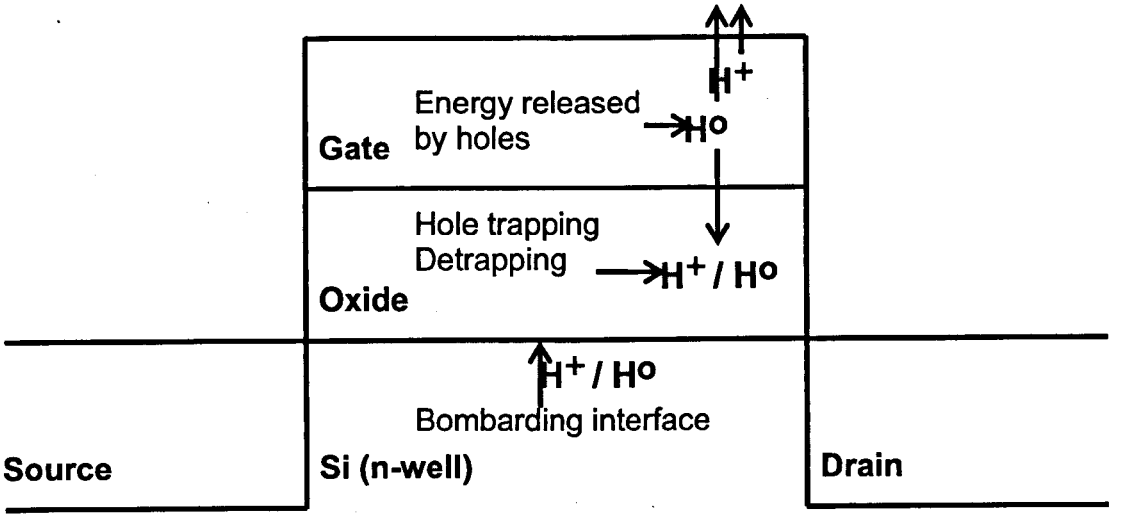
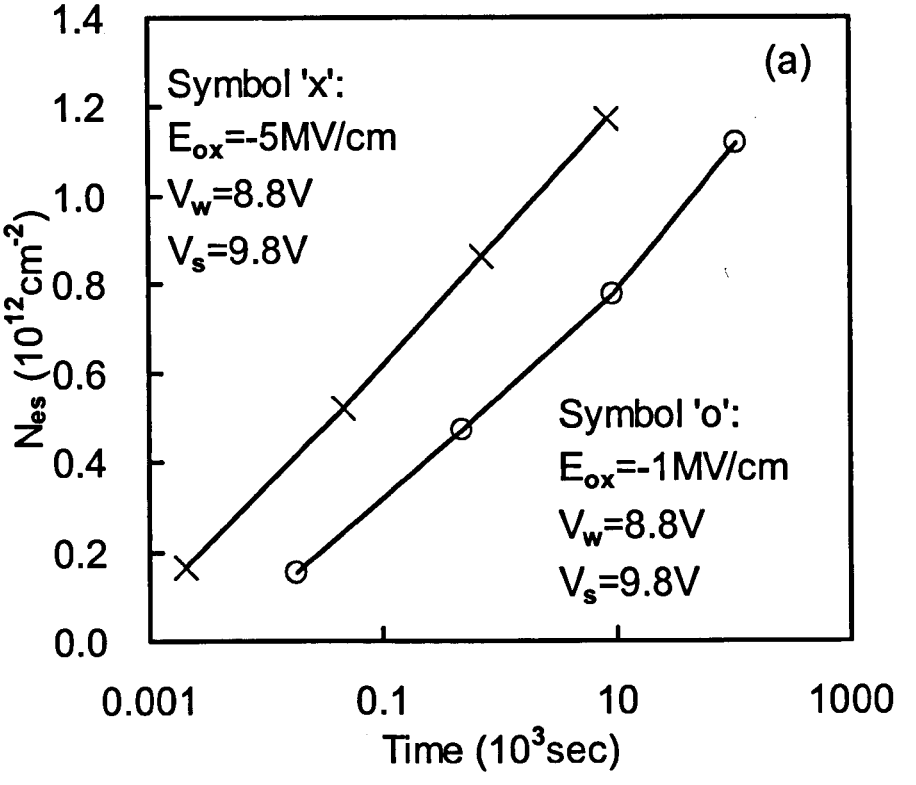
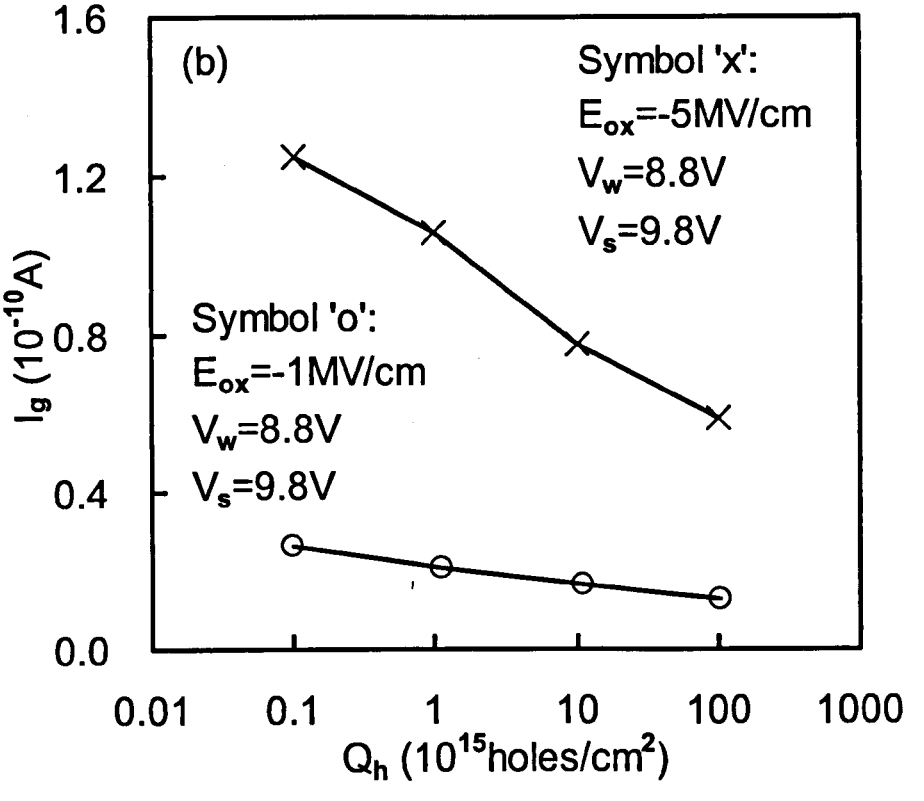


Figure 3. A schematic illustration of hydrogen release during substrate hole injection from three region: Si/oxide interfacial region, oxide bulk and gate/oxide interfacial region.


(a)
Symbol 'x':
E_{ox}=-5MV/cm
V_w=8.8V
V_s=9.8V
Symbol 'o':
E_{ox}=-1MV/cm
V_w=8.8V
V_s=9.8V
N_{es} (10^{12}cm^{-2})
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0.001
0.1
10
1000
Time (10^3sec)


(b)
Symbol 'x':
E_{ox}=-5MV/cm
V_w=8.8V
V_s=9.8V
Symbol 'o':
E_{ox}=-1MV/cm
V_w=8.8V
V_s=9.8V
I_g (10^{-10}A)
1.6
1.2
0.8
0.4
0.0
0.01
0.1
1
10
100
1000
Q_h (10^{15}holes/cm^2)

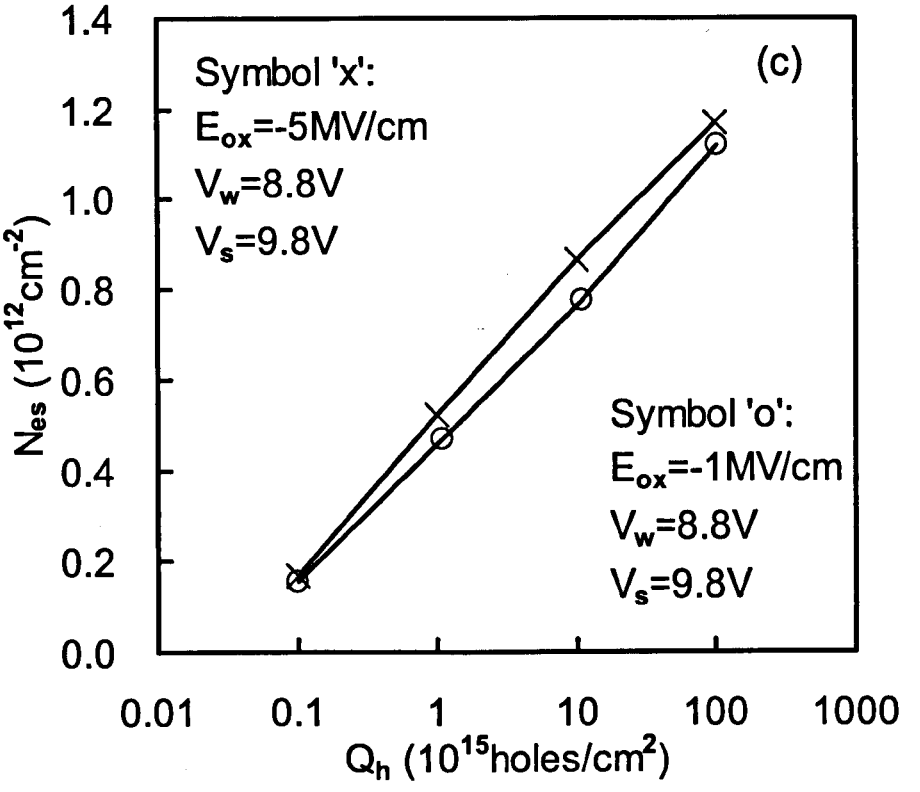
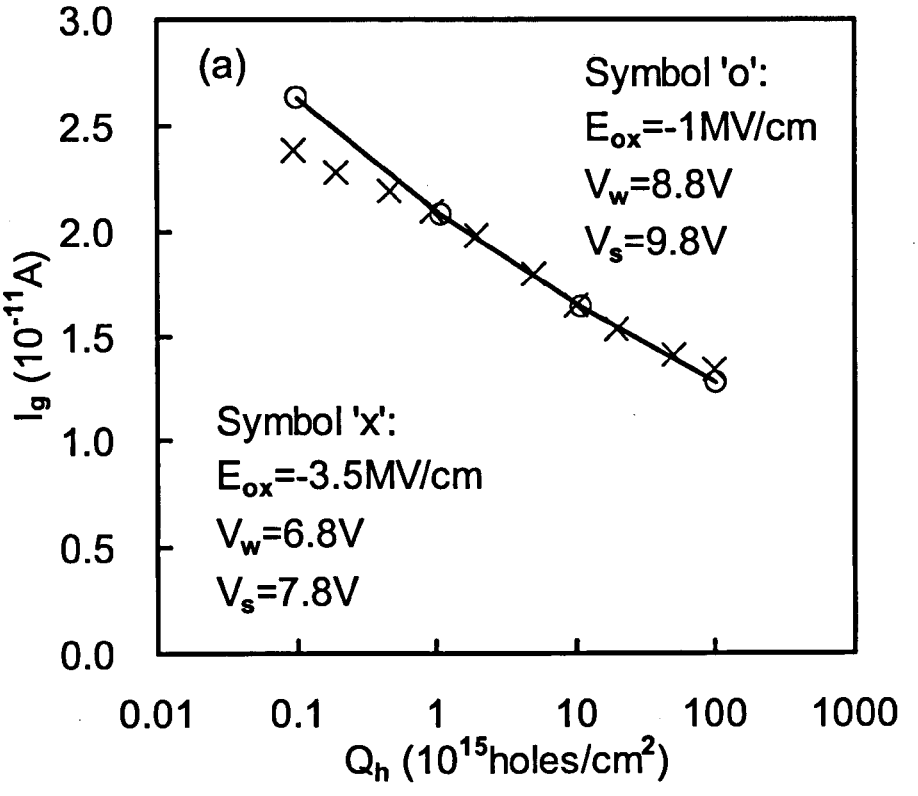


Figure 4. Effects of bombarding the Si/oxide interface on electron trap generation. For a given time, the bombardment of the Si/oxide interface is the same for the two devices. (a) Electron trap creation for a given time is clearly different. (b) The difference originates from the different injection current. (c) When plotted against hole fluency, the generation in these two devices agrees well.


(a)
Symbol 'o':
E_{ox}=-1MV/cm
V_w=8.8V
V_s=9.8V
Symbol 'x':
E_{ox}=-3.5MV/cm
V_w=6.8V
V_s=7.8V
I_g (10^{-11}A)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0.01
0.1
1
10
100
1000
Q_h (10^{15}holes/cm^2)

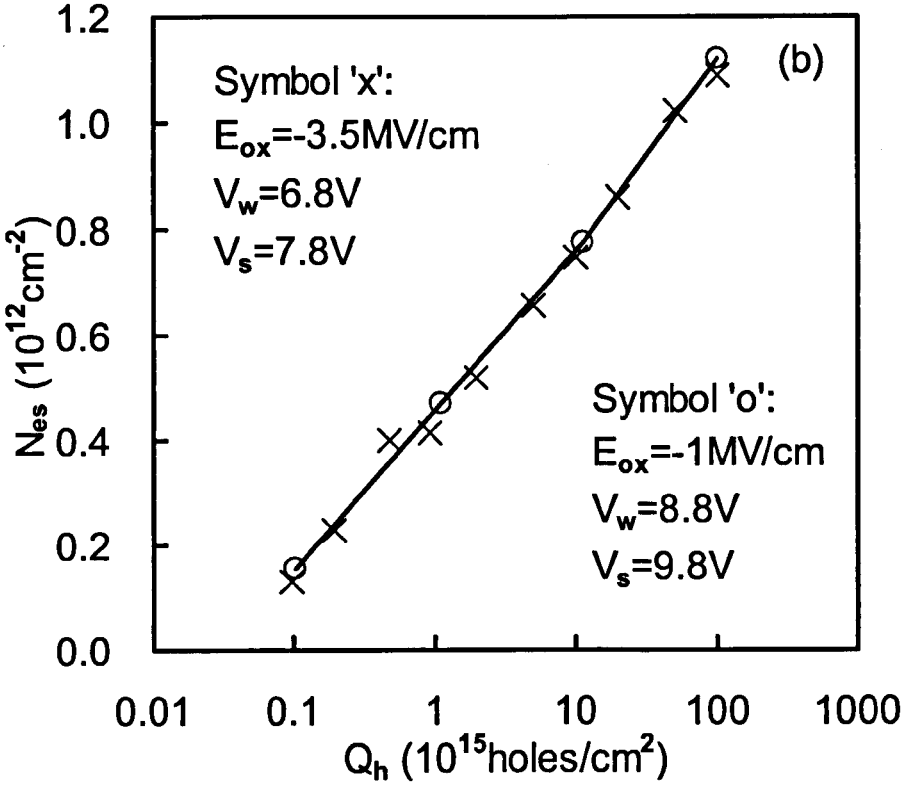
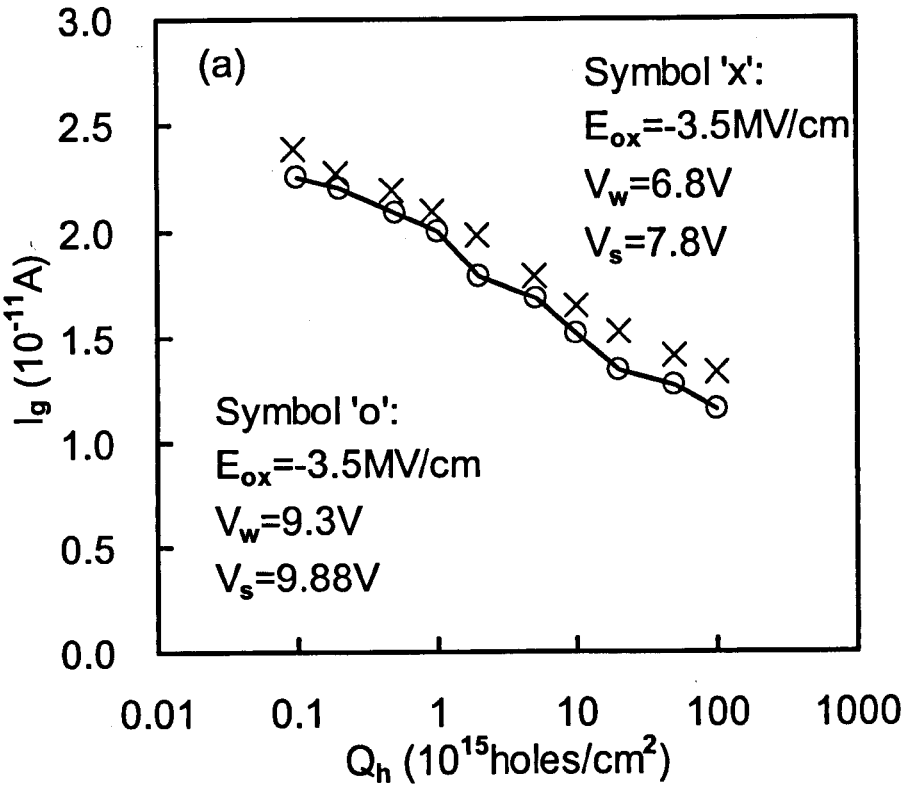

(b)
Symbol 'x':
E_{ox}=-3.5MV/cm
V_w=6.8V
V_s=7.8V
Symbol 'o':
E_{ox}=-1MV/cm
V_w=8.8V
V_s=9.8V
N_{es} (10^{12}cm^{-2})
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0.01
0.1
1
10
100
1000
Q_h (10^{15}holes/cm^2)

Figure 5. Effects of energy of holes bombarding the Si/oxide interface on electron trap generation. V_s - V_w is the same for both devices, so the same amount of holes will arrive at the interface. A higher V_w was applied on one device to make the holes more energetic. (a) The oxide field were adjusted to produce similar injection current during SHI. (b) When plotted against hole fluency, electron trap generation is insensitive to the n-well bias, V_w.



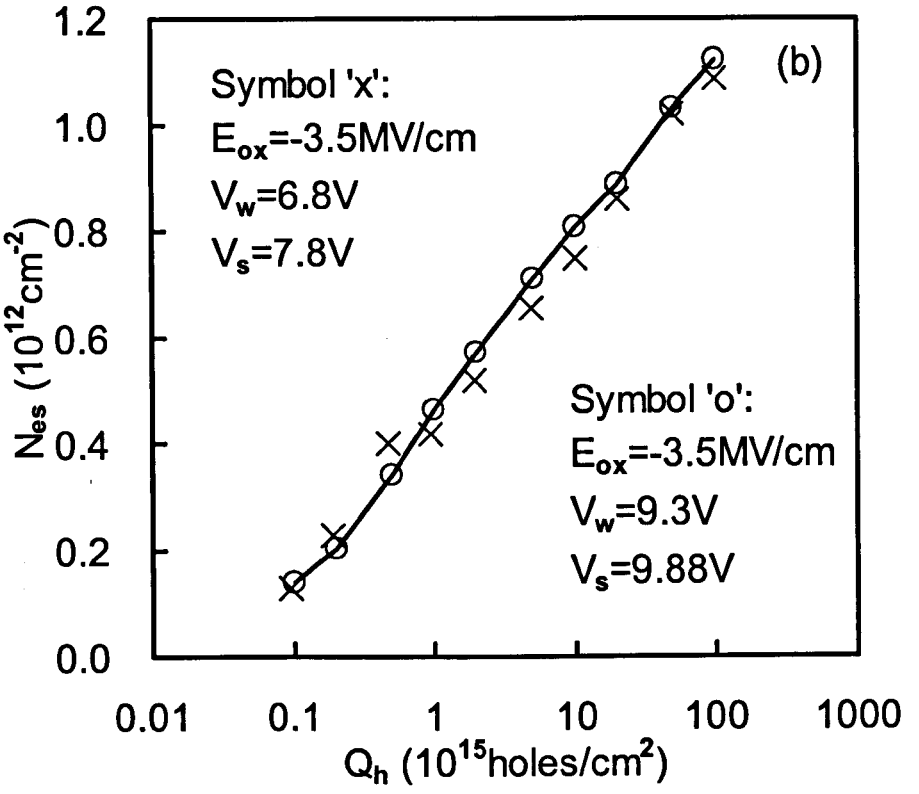
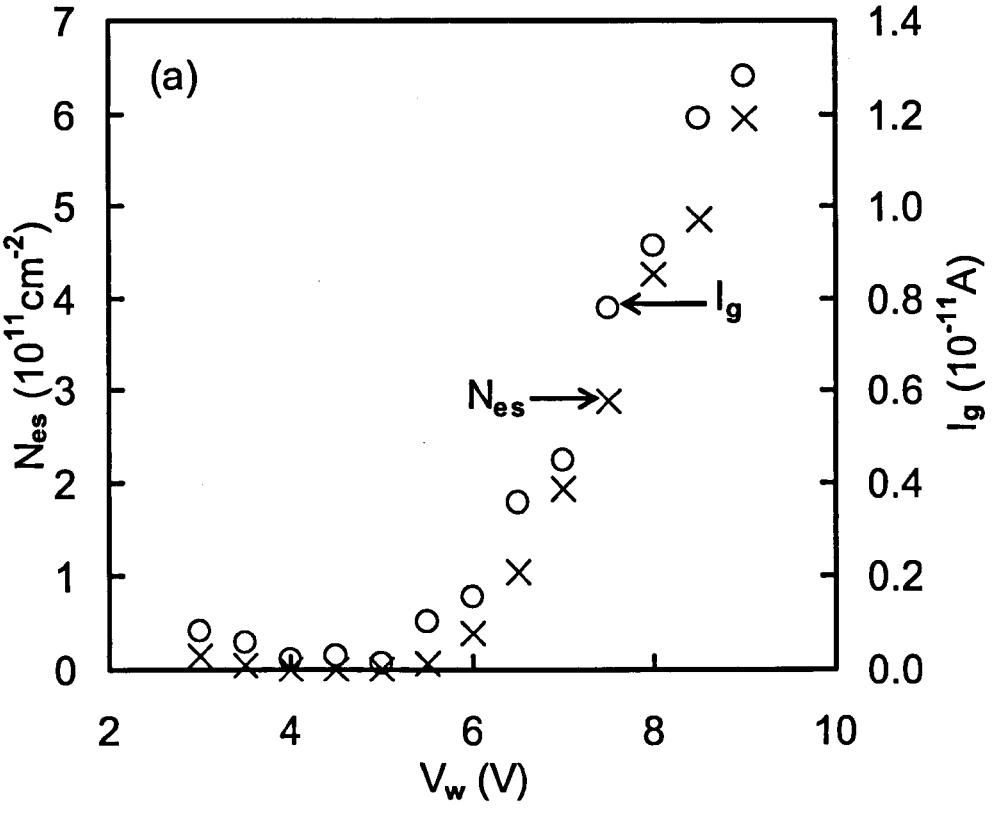
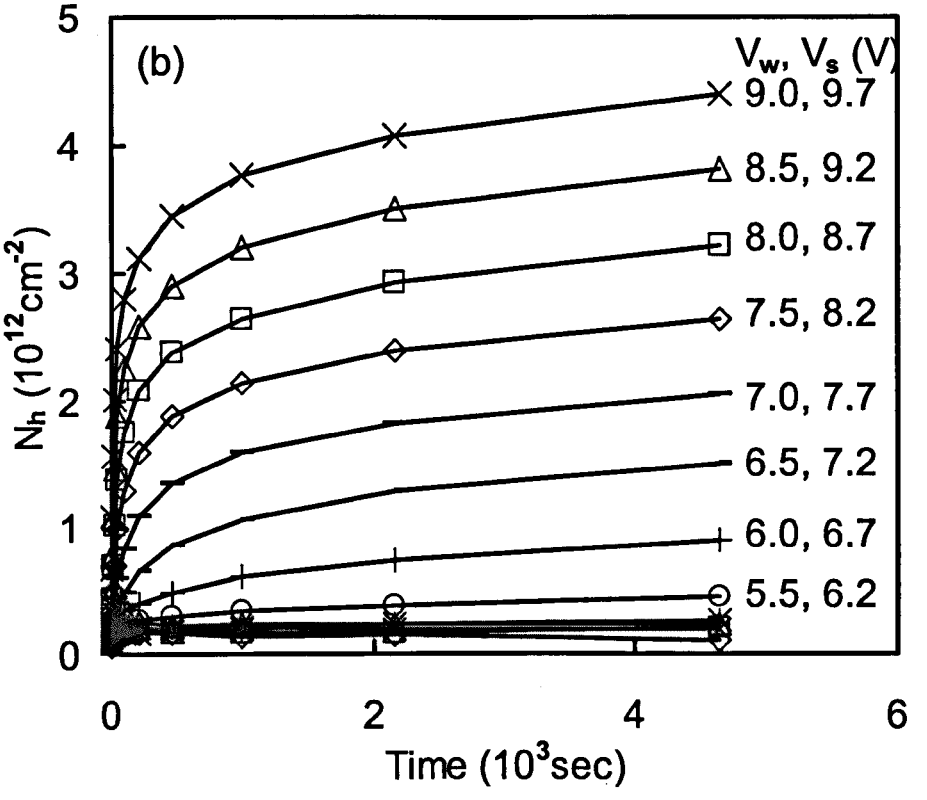


Figure 6. Effects holes bombarding the Si/oxide interface on electron trap generation. V_s - V_w was adjusted so both devices will have similar injection current during the SHI. (a) Injection current during SHI is similar for the two stress conditions. (b) Electron trap generation agrees well on both devices.


(a)
N_{es} ($10^{11}cm^{-2}$)
I_g ($10^{-11}A$)
N_{es}
I_g
V_w (V)
0
1
2
3
4
5
6
7
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
2
4
6
8
10


(b)
N_h ($10^{12}cm^{-2}$)
Time (10^3sec)
V_w, V_s (V)
9.0, 9.7
8.5, 9.2
8.0, 8.7
7.5, 8.2
7.0, 7.7
6.5, 7.2
6.0, 6.7
5.5, 6.2
0
1
2
3
4
5
0
2
4
6

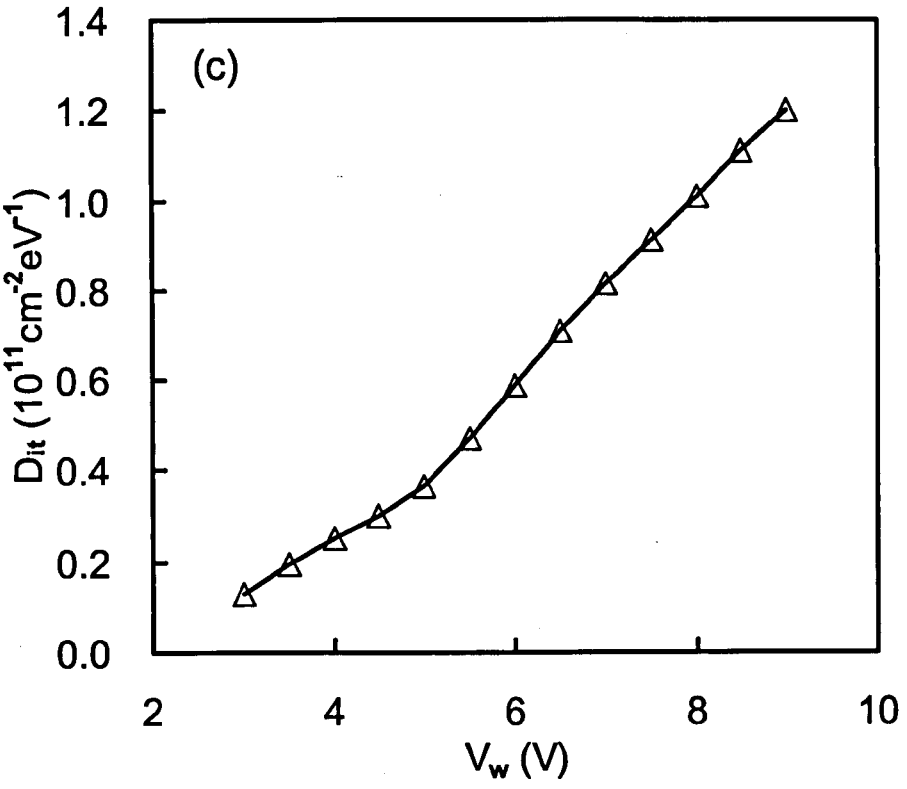


Figure 7. (a) Dependence of gate current, generated electron traps, (b) trapped hole, (c) and interface states on n-well biases, V_w. V_w was increased with a step size of 0.5 V. For each V_w, the stress time was 4600 s at E_{ox} = -1 MV/cm and (V_s - V_w) was kept at 0.7 V. Gate current, generated electron traps, and trapped hole have a threshold V_w around 5 V, while no clear threshold can be observed for interface state generation.

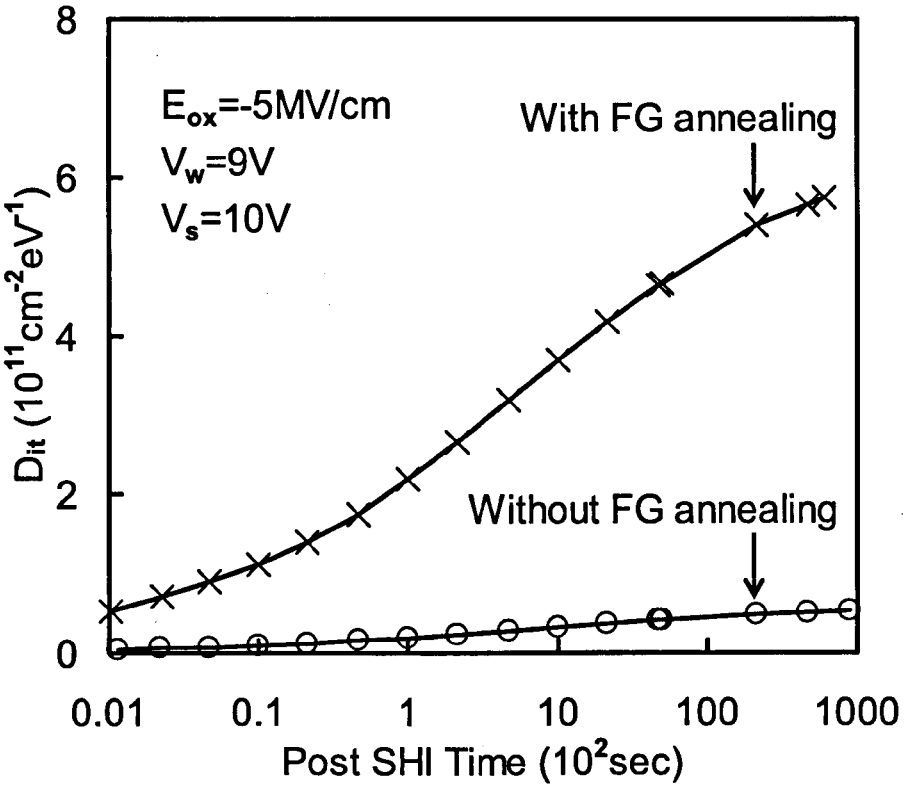
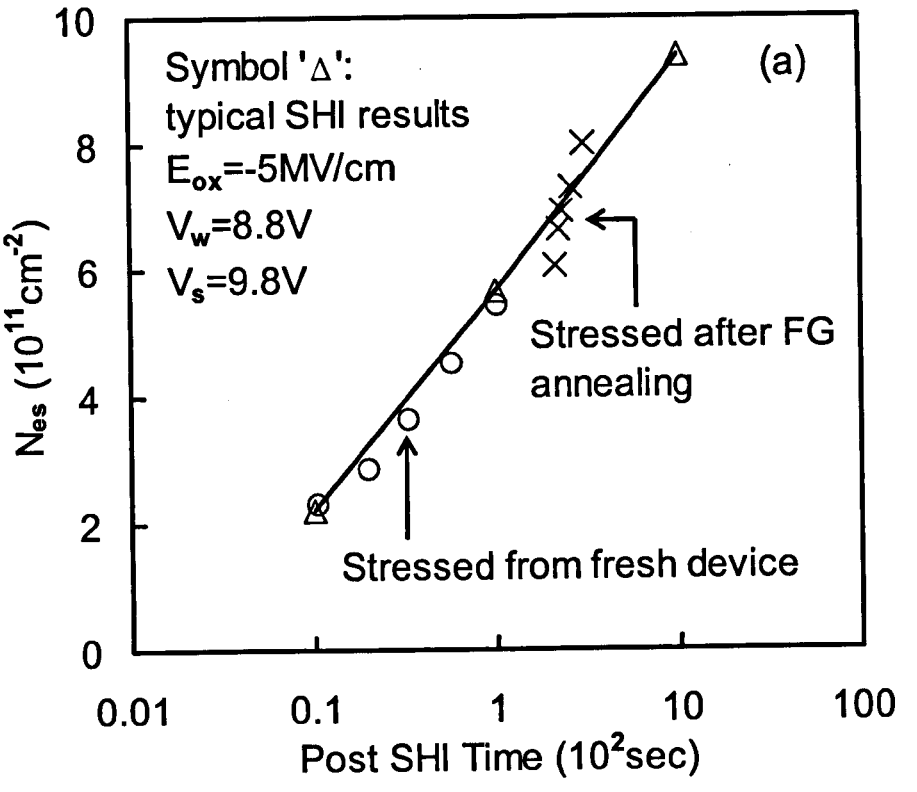


Figure 8. The forming gas anneal was carried out at 450°C for 40 min after hole injection. The annealing increase the amount of hydrogen in the device, which results in the substantial increase in the interface states generation post hole injection (E_{ox} = -5 MV/cm, V_w = 9 V, V_s = 10 V).


Symbol 'Δ':
typical SHI results
E_{ox}=-5MV/cm
V_w=8.8V
V_s=9.8V
(a)
Stressed after FG annealing
Stressed from fresh device
N_{es} ($10^{11}cm^{-2}$)
Post SHI Time (10^2sec)
0
2
4
6
8
10
0.01
0.1
1
10
100

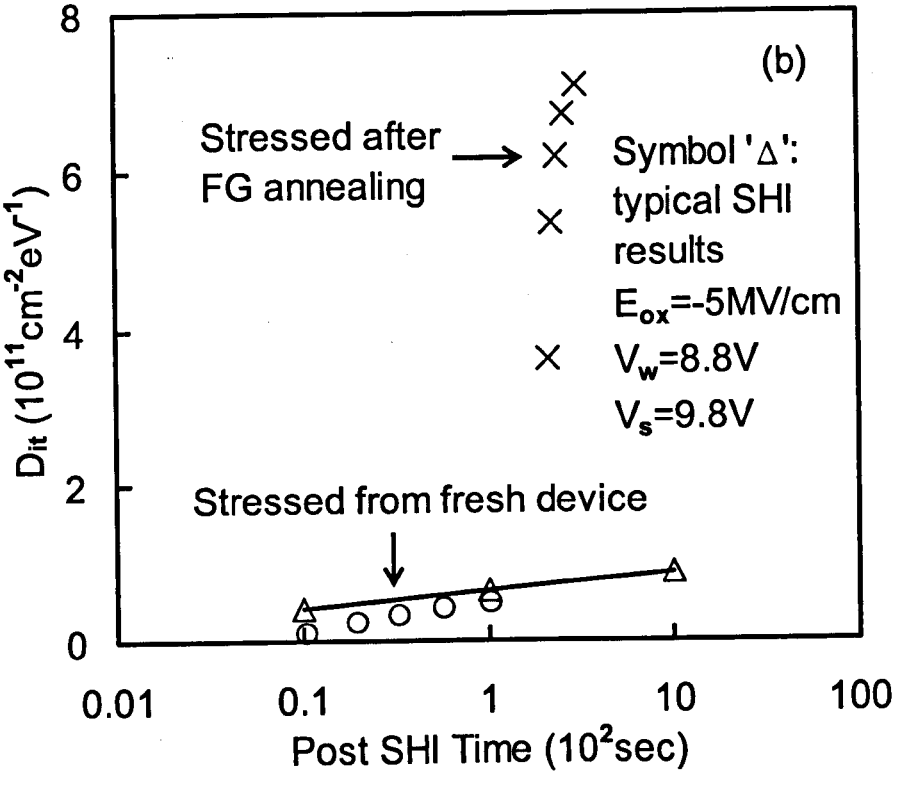

(b)
Stressed after FG annealing
Symbol 'Δ':
typical SHI results
E_{ox}=-5MV/cm
V_w=8.8V
V_s=9.8V
Stressed from fresh device
D_{it} ($10^{11}cm^{-2}eV^{-1}$)
Post SHI Time (10^2sec)
0
2
4
6
8
0.01
0.1
1
10
100

Figure 9. Impact of hydrogen density in a device on (a) the generation of electron traps and (b) interface states. A device was first stressed by SHI up to $Q_h = 10^{15}$ holes/cm^2 (symbol '○'). The SHI was then interrupted and the device exposed to forming gas (10% H_2) at 400°C for 40 min. After the exposure, the SHI was resumed and the results were represented by symbol 'x'. The exposure significantly enhanced the generation of interface states, but not the creation of electron traps. The solid line and symbol '△' represents results when a device was not given the exposure.

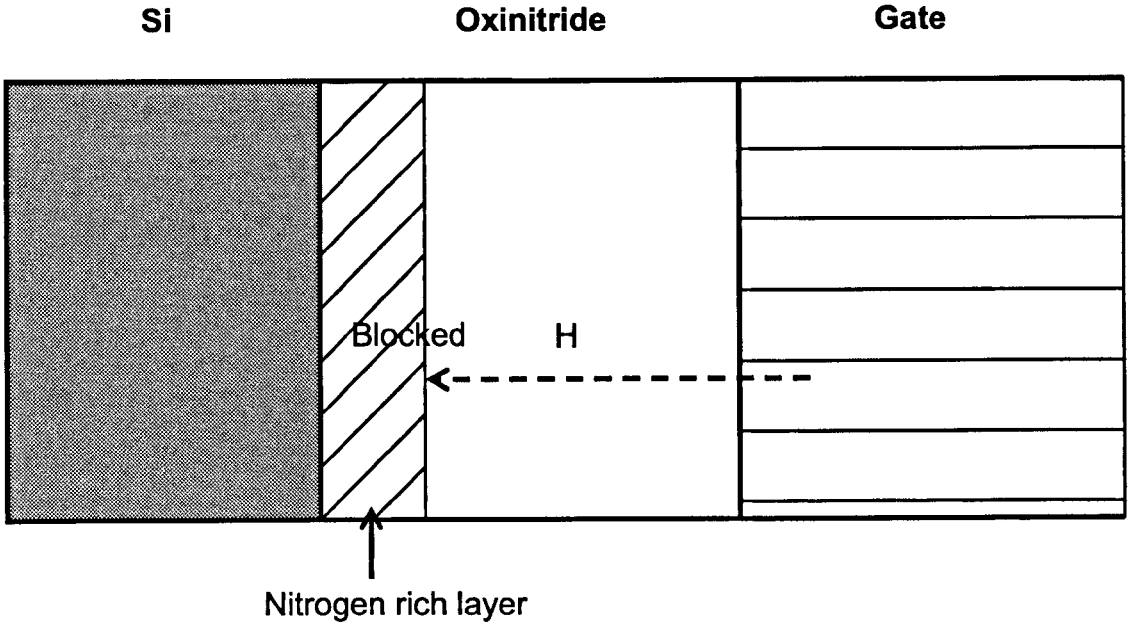


Figure 10. On the oxynitride, the nitrogen rich layer should block hydrogenous species transportation. Therefore, less hydrogen reaches the dielectric/substrate interface. This in turn should leads to less interface states generation.

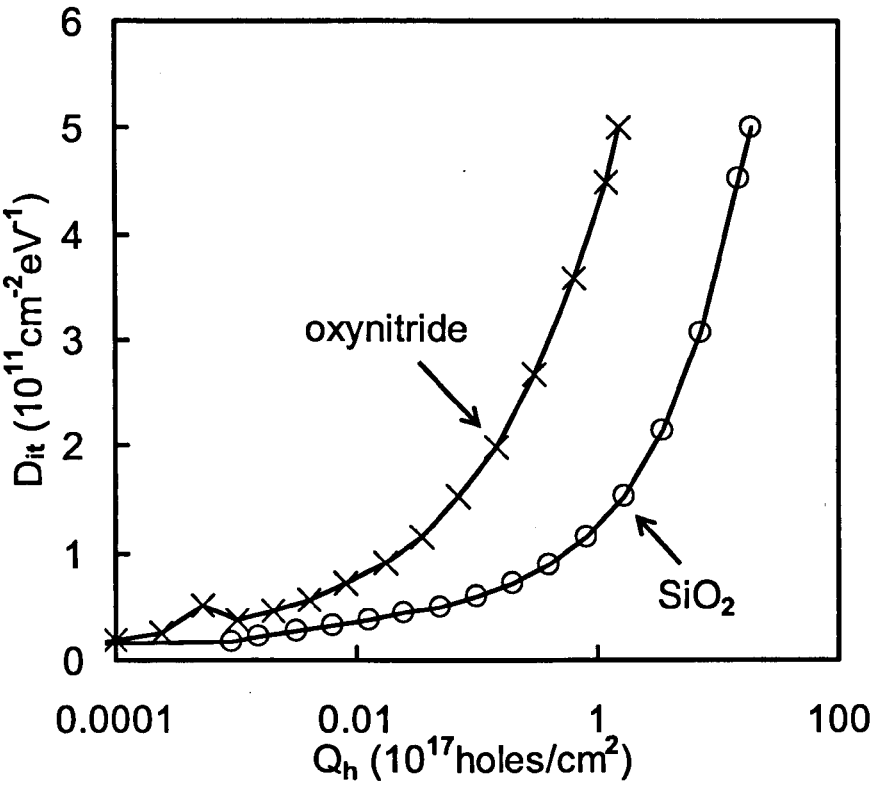


Figure 11. Interface state generation during hole injection (E_{ox} = -5 MV/cm, V_w = 8.8 V, V_s = 9 V). For the same number of injected holes, D_{it} is clearly higher for the oxynitride.

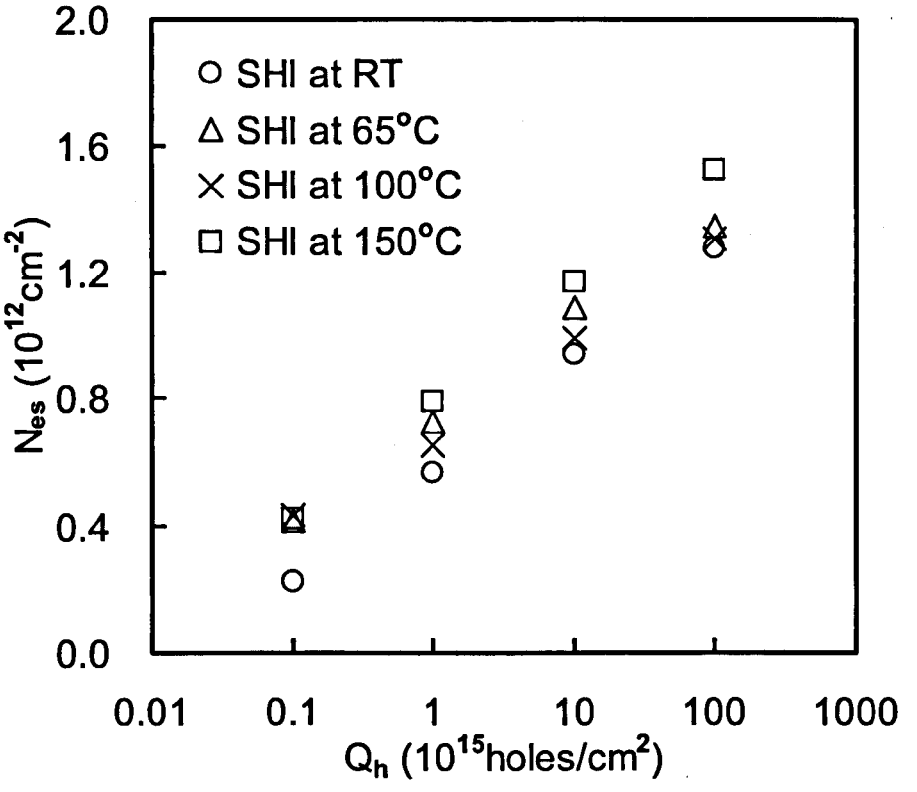
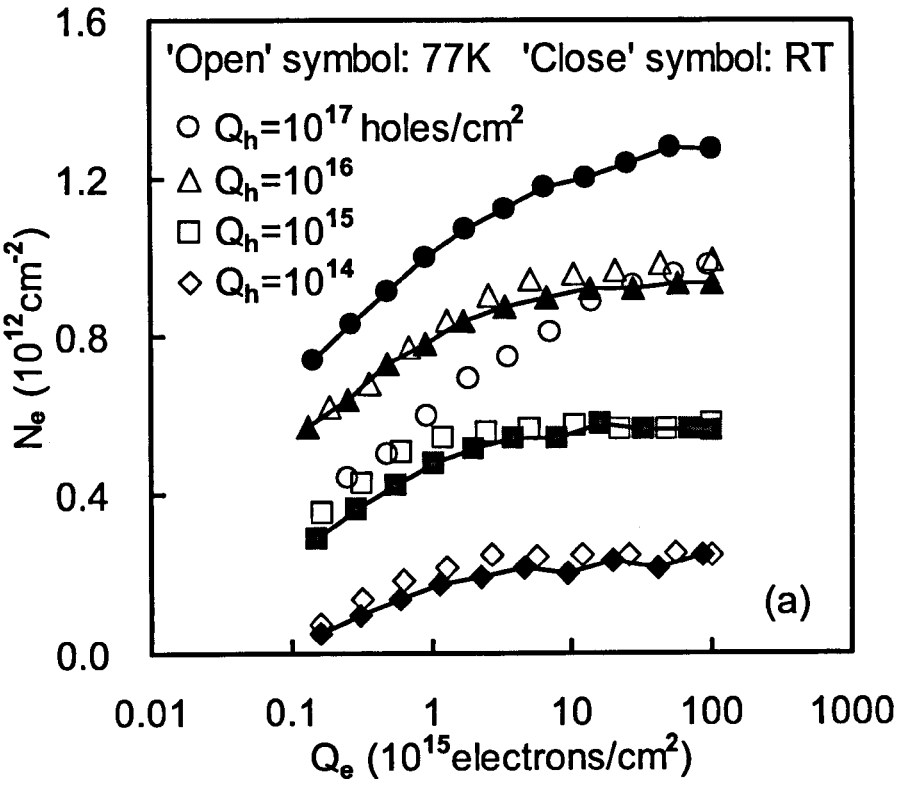


Figure 12. Impact of elevated temperature on hole-induced electron trap generation. Devices were streesed by SHI under E_{ox} = -5 MV/cm, V_w = 8.8 V, V_s = 9 V. It is clear that hole-induced electron trap creation is insensitive to temperature up to 150°C.


'Open' symbol: 77K 'Close' symbol: RT
○ $Q_h=10^{17}$ holes/cm^2
△ $Q_h=10^{16}$
□ $Q_h=10^{15}$
◇ $Q_h=10^{14}$
(a)
N_e (10^{12}cm^{-2})
Q_e (10^{15}electrons/cm^2)
1.6
1.2
0.8
0.4
0.0
0.01
0.1
1
10
100
1000

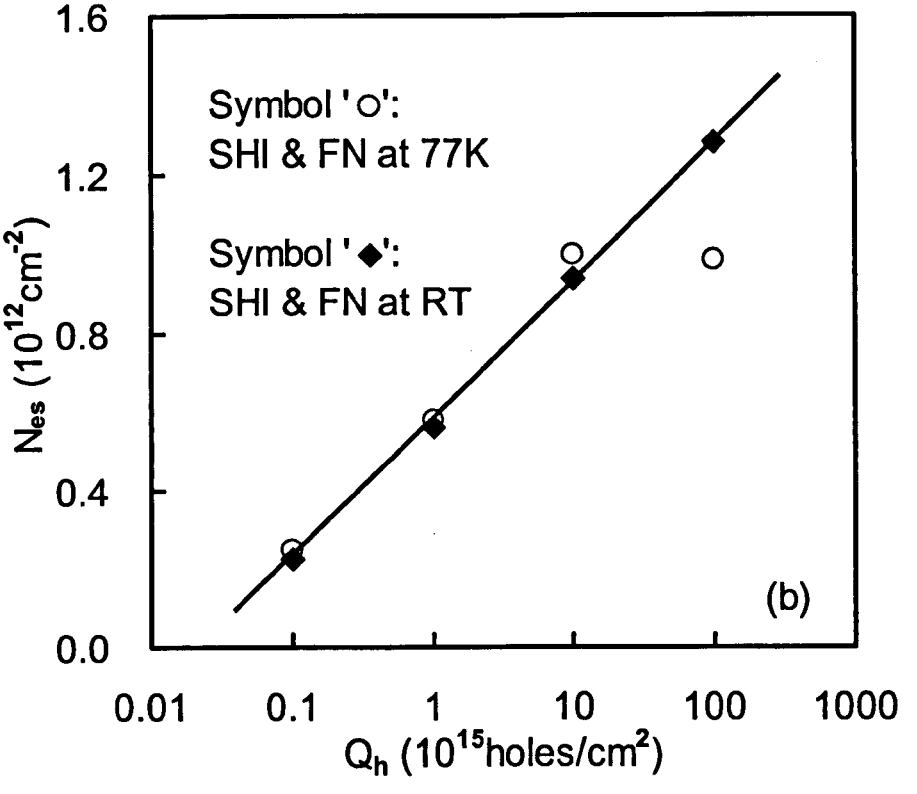
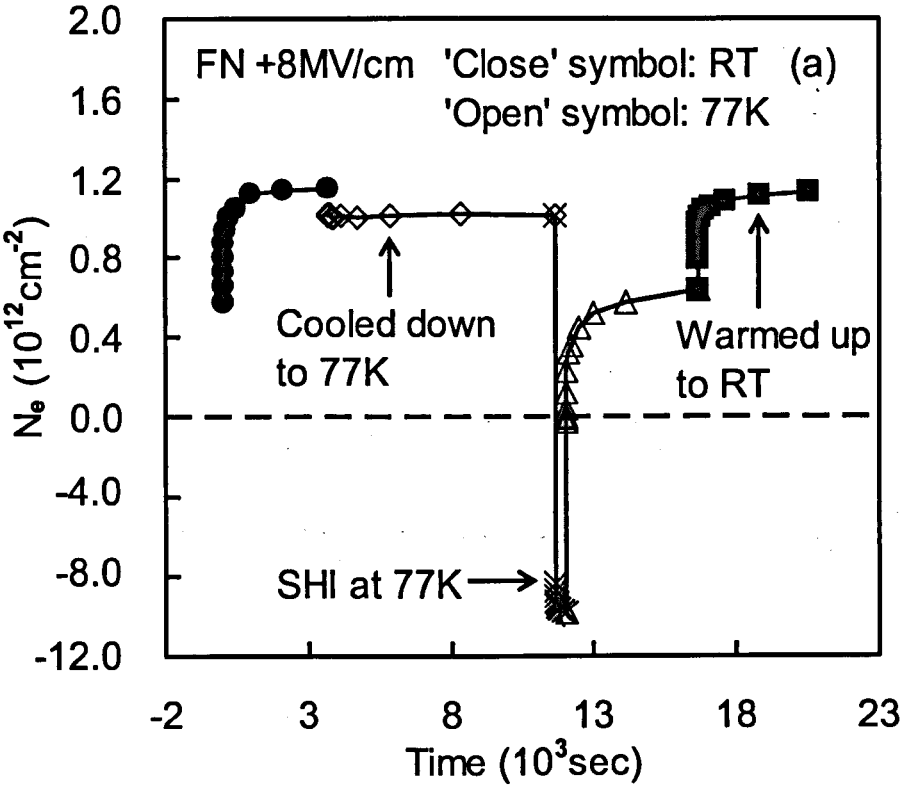

Symbol '○':
SHI & FN at 77K
Symbol '◆':
SHI & FN at RT
(b)
N_{es} (10^{12}cm^{-2})
Q_h (10^{15}holes/cm^2)
1.6
1.2
0.8
0.4
0.0
0.01
0.1
1
10
100
1000

Figure 13. Impact of low temperature on hole-induced electron trap generation. One device was streesed and filled at 77K (open symbol) compare to other one stressed and filled at RT (close symbol). (a) Trapping kinetics at different hole injection level. Symbols '◇', '□', '△' and '○' correspond to Q_h = 10^{14}, 10^{15}, 10^{16}, and 10^{17} holes/cm^2, respectively. (b) Generation of electron traps against Q_h. For Q_h up to 10^{16} holes/cm^2, both the trapping kinetics and electron trap generation are insensitive to temperature. However, with increasing hole injection, a drop in generation is observed at 77K. This drop is likely to be caused by an increase of positive charges formation under SHI, which becomes difficult to neutralize at low temperature. The solid line is a guide for the eye.


FN +8MV/cm 'Close' symbol: RT (a)
'Open' symbol: 77K
Cooled down to 77K
Warmed up to RT
SHI at 77K
N_e (10^{12}cm^{-2})
Time (10^3sec)
2.0
1.6
1.2
0.8
0.4
0.0
-4.0
-8.0
-12.0
-2
3
8
13
18
23

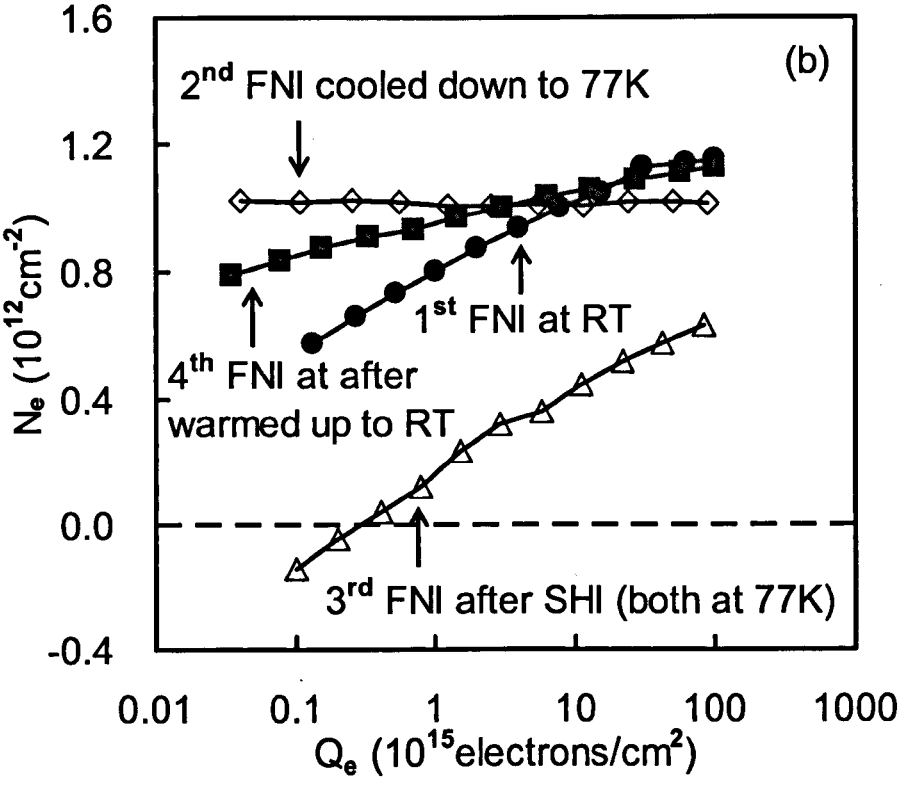

(b)
2nd FNI cooled down to 77K
1st FNI at RT
4th FNI at after warmed up to RT
3rd FNI after SHI (both at 77K)
N_e (10^{12}cm^{-2})
Q_e (10^{15}electrons/cm^2)
1.6
1.2
0.8
0.4
0.0
-0.4
0.01
0.1
1
10
100
1000

Figure 14. A heavily stressed device (previous stress level Q_h = $2x10^{18}$ holes/cm^2) was used here to investigate the positive charges neutralization at 77K. (a) The entire experiment sequence. (b) The trapping kinetics of each filling. The fillings were carried out at RT (symbol '●'), 77K (symbol '◇'), 77K (symbol '△') and RT (symbol '■'). A short SHI (Q_h = 10^{17} holes/cm^2) was applied between the two fillings at 77K to induce positive charges. It is clear that the amount filled after SHI at 77K (symbol '△') is significantly lower to the one before SHI at 77K (symbol '◇'). By warming up to room temperature, electrons injected under E_{ox} = +8 MV/cm will be more energetic, which makes the positive charges neutralization easier (symbol '■'), thus, the true amount of generation can be obtained (symbol '■' ≈ '●').

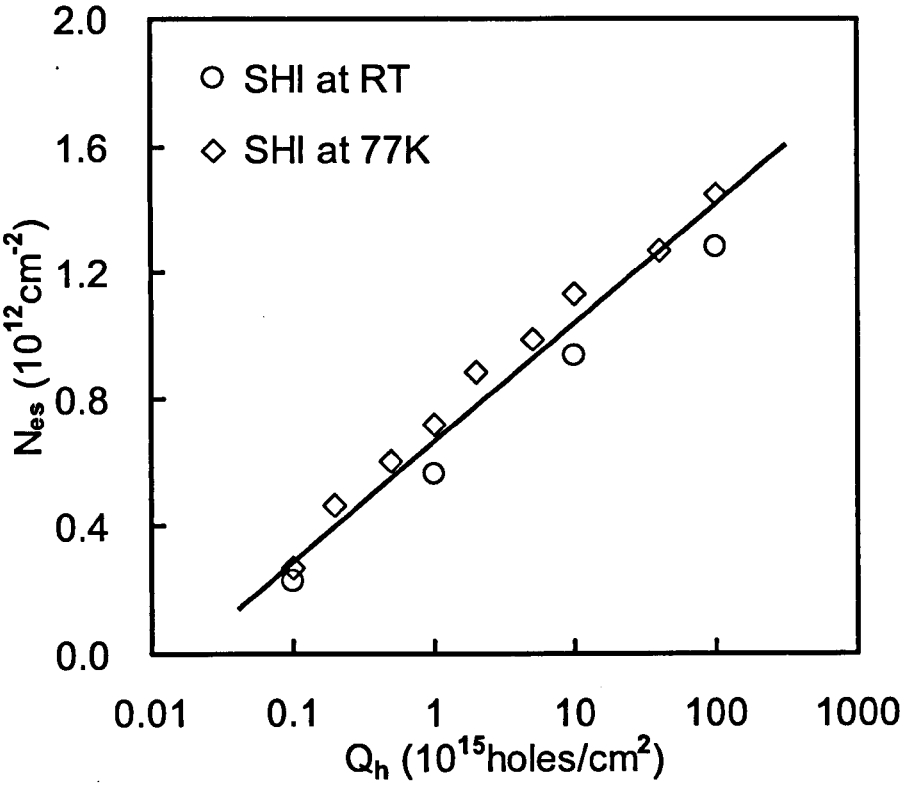


Figure 15. Impact of low temperature on hole-induced electron trap generation. A device was stressed by SHI (E_{ox} = -5 MV/cm, V_w = 8.8 V, V_s = 9.8 V) at 77K, then filled at room temperature, in order to fully neutralize SHI induced positive charges. The electron trap generation is again insensitive to temperature, which further supports that hydrogen released from the other two regions, the Si/oxide interface and the oxide bulk, are not important, either. The solid line is a guide for the eye.

4 Trapping Kinetics of Electron Traps Generated in Silicon Dioxides

4.1. Introduction

Electron traps can be both 'as-grown' or created after device fabrication. Early efforts were focused on the as-grown traps [1-3]. It was found that as-grown traps originated from impurities in the oxide, such as water molecules [1] and implants [2]. For a modern MOSFET, however, the use of poly-si gates and a careful control of impurities have effectively eliminated these as-grown traps [4]. As a result, we will focus on generated electron traps here.

Generated electron traps play an important role in device reliability. They can have several adverse effects on device performance, such as a shift in device parameters [5-7] and stress-induced leakage current (SILC) [8]. As their number accumulates, they can form a conduction path between gate and substrate, which leads to oxide breakdown [8-10]. For a high-k/SiO_2 stack, it has been suggested that its breakdown can be controlled by the interfacial SiO_2 layer [10]. Consequently, generated traps are important for the stack as well.

Despite of their importance, our knowledge on them is still limited. For example, their origin and microstructure is not known. There is no well-accepted trapping model, either. Developing a trapping model is highly desirable for understanding these traps, since it allows us to determine their effective physical size, namely capture cross section. It will also give us information on how many types of electron traps are generated. Information like this will be valuable when the origin of generated traps is explored and models for oxide breakdown and SILC are developed [8,11].

Only limited works [12] were carried out on electron trapping kinetics for poly-si gated devices. Non-uniform electron injection was used and consequently, uncertainty exists in the density of injected electrons [5,12]. Most of previous works on generated electron traps did not address the trapping kinetics, because a number of difficulties, as summarized below:

First, for thin oxides (< 3 nm), traps can be located sufficiently close to the electrodes that trapped electrons can tunnel away. In this case, trapping will not be stable and it is difficult to study the kinetics. The SILC has been widely used as a measure of generated defects in thin oxides [8]. However, electron traps are not the only source for SILC [11]. The important trapping parameters, such as capture cross sections and density, were rarely reported from the work based on SILC.

Second, if relatively thick oxide is used, stable electron trapping can be achieved, but the difficulty is in separating the trap filling from the trap creation. In most of previous works, electrons were injected into the oxide during trap generation. Once traps are generated, these electrons can fill the traps. Since the filling is typically much faster than the creation, the trapping level is determined by the generation, rather than filling. One example is given in Figure 1. Here, one MOSFET was subjected to Fowler-Nordheim injection (FNI) under an oxide field strength of +11 MV/cm. The electron trapping density, N_e, is negative initially, indicating net positive charges. As the stress increases, N_e becomes positive, representing net negative charges, and builds up continuously without saturation. The non-saturation of N_e observed here results from the continuous creation of new traps, rather than filling the existing traps. As a result, it cannot be used for studying the kinetics of trap filling.

Third, recent works [13-16] show that electrical stresses generate not only electron traps, but also hole traps. Hole trapping leads to positive charge formation in the oxide, which offsets electron trapping. This complicates the study and care must be exercised to take it into account.

To investigate the trapping kinetics of generated electron traps, the difficulties mentioned above must be overcome. We believe that this can be achieved now, since recent works at this university have improved our understanding of both electron traps [4,17] and hole traps [13-16]. The objective of this work is to investigate the trapping kinetics of generated electron traps by selecting test conditions and developing methods for overcoming these difficulties. Based on the kinetics, the properties and types of generated traps will be explored.

4.2. Trapping kinetics of electron traps generated under SHI

When selecting testing samples and stressing techniques, priority is given to overcome the difficulties mentioned earlier in section 4.1 and to simplify the testing conditions as far as possible. To avoid the uncertainty caused by the use of SILC for measuring electron traps, we chose relatively thick oxide (7.1 nm) here, since it gives stable electron trapping in large quantities (~ 10^{12} cm^{-2}). The uncertainty in the lateral distribution of generated traps was removed by stressing devices uniformly. Uniform stress can be achieved by either Fowler-Nordheim injection (FNI) [6] or substrate hole/electron injection [7,18].

During the substrate hole injection (SHI), electron injection was negligible and positive charges were built up through hole trapping, as is shown by symbol '◇' in Figure 2. Generated electron traps are neutral during the stress, since there were no electrons available for filling them. A subsequent electron injection was applied to fill these traps (symbol 'x'). As a result, the trap filling process is separated from the generation phase, allowing the investigation of trapping kinetics. We chose SHI for creating electron traps in this section. Unless otherwise specified, SHI will be carried out under an oxide field of E_{ox} = -5 MV/cm, an n-well bias of 8.8 V and a p-substrate bias of 9.8 V. The trap filling is carried out under E_{ox} = +8 MV/cm with all other terminals grounded.

4.2.1. Effects of positive charges

After obtaining stable electron trapping and separating the trap filling from the generation, a remaining obstacle for studying electron trapping kinetics is the offsetting effects of positive charges. There are two ways for forming positive charges in oxides: exposure to H_2 at elevated temperature above 500°C [19] and hole trapping [13-15,20]. For this section, the device was not exposed to H_2 at such temperature and only hole trapping is relevant.

4.2.1.1. Low stress level

When hole fluency during the SHI stress, Q_h, is relatively low (e.g. $Q_h < 10^{16}$ holes/cm^2), hole trapping is dominated by as-grown hole traps. These hole traps are pre-existed in the device after fabrication. After stopping hole injection and starting electron injection, electron traps are filled and positive charges are neutralized, as shown in Figure 2 & 3. When the electron fluency for filling, Q_e, is below 10^{14} electrons/cm^2 approximately, Figure 3 shows that as-grown hole traps are not fully neutralized and the measured N_e is the 'net' charge. This means that electron trapping kinetics for acceptor-like traps cannot be unambiguously investigated in the region of $Q_e < 10^{14}$ electrons/cm^2.

Earlier works [14,17,21] showed that the neutralization of as-grown hole traps was completed when electron fluency reached the level of 10^{14} electrons/cm^2. Consequently, the electron trapping for $Q_e > 10^{14}$ electrons/cm^2 is not affected by as-grown hole trapping. As a result, only N_e measured at $Q_e > 10^{14}$ electrons/cm^2 will be used for studying electron trapping kinetics. The impact of a lack of data for $Q_e < 10^{14}$ electrons/cm^2 will be discussed later in section 4.2.2.2.

4.2.1.2. High stress level

Typical test results are given in Figure 4. As the stress level increases, more traps were generated and the whole trapping curve moves up, when hole fluency, Q_h, is less than 10^{16} holes/cm^2. When Q_h rises further, however, the increase of trapping disappears at relatively low Q_e, although trapping is still clearly increased at higher Q_e. This is because considerable amount of new hole traps were created at higher Q_h [13-15,20]. Although as-grown hole traps have an energy level below the bottom edge of silicon conduction band, some of the generated hole traps have energy levels above it, as illustrated in Figure 5 [14,15,20]. Since the number of electrons in silicon conduction band drops rapidly for higher energy level, the number of electrons available for neutralizing these generated hole traps is less than that for as-grown hole traps. As a result, it is more difficult of neutralizing these generated hole traps and they are called as 'anti-neutralization positive charges (ANPC)' [14,15,20]. When Q_e in Figure 4 is relatively low, the neutralization of ANPC is not completed and the remaining positive ANPC brings down N_e through offsetting. As Q_e increases further, ANPC is gradually neutralized, which leads to higher N_e for higher Q_h. Since the ANPC has a substantial effect on the dynamic behavior of N_e, it is essential to correct this effect, before the trapping kinetics of generated electron traps can be investigated. A correction method is proposed below.

In Figure 6(a), a device was heavily stressed by SHI (symbol '□') with $Q_h = 10^{18}$ holes/cm^2, which created ANPC. After the stress, electrons were injected into the oxide to fill the generated electron traps and neutralize positive charges and symbol '△' represents the net charge density. This is followed by applying a negative gate bias (symbol '○'), during which both filled electron traps and neutralized ANPC can lose their electrons through electron tunneling to silicon, resulting in the drop of N_e. Electron injection was then switched on again in an attempt to refill them (symbol 'x'). In Figure 6(b), the trapping density during the 1st and 2nd filling is compared. It is obvious that the trapping during the 2nd filling failed to reach the same level as that for the 1st filling. Recent work at this university [17] shows that this is because the electron trap cannot recapture an electron, after detrapping through tunneling, although the ANPC can [14,15,20]. To further illustrate the difference of generated

electron traps and ANPC in refilling, the test procedure in Figure 6(a) was used again, but the hole fluency was chosen to be sufficiently low this time that the creation of ANPC is insignificant. With little ANPC, Figure 6(c) shows that the rise of N_e during the 2nd filling is also insignificant. The nearly flat N_e (see the dashed line) confirms that, after tunneling-induced detrapping, an electron trap cannot capture an electron again. Consequently, the rise of N_e during the 2nd filling in Figure 6(b) can only originate from the re-neutralization of ANPC.

We are now in a position to correct the offsetting effect of positive charges on electron trapping. The real level of electron trapping, N_{ec}, can be found from,

$$N_{ec} = N_e + \Delta N_p \tag{1}$$

where N_e is the net trapping and is directly measured. The key question is how to estimate ΔN_p. As explained earlier, the rise of N_e during the 2nd filling in Figure 6(b) originates from the re-neutralization of positive charges. As a result, the amount of positive charges neutralized between a given Q_e and the end of filling is,

$$\Delta N_p(Q_e) = N_e(End\ of\ 2^{nd}\ filling) - N_e(Q_e\ of\ 2^{nd}\ filling) \tag{2}$$

In another word, the un-neutralized positive charge at Q_e is ΔN_p, as shown in Figure 6(b). Once ΔN_p is known, the real level of trapping can be found from equation (1) and is plotted in Figure 6(b).

4.2.1.3. Support for the correction method

Recent work at this university shows that the ANPC is thermally unstable and can be annealed at 150°C [14]. In Figure 7, two devices were subjected to the same stresses with $Q_h = 10^{17}$ holes/cm^2. After filling the generated electron traps, one device was exposed to 150°C for 160 min to remove ANPC. A short hole injection ($Q_h = 10^{15}$ holes/cm^2) was then used to empty electron traps still filled at the end of the thermal exposure. This short hole injection is too low to recreate the ANPC [14,20]. When

the traps were refilled (symbol 'x'), N_e was obviously enhanced at low Q_e, but remain essentially the same at the end of the filling. This supports our assumption that positive ANPC lowers the measured N_e at low Q_e. As Q_e increases, ANPC is neutralized and N_e approaches the true level of electron trapping. The other devices went through the test sequence shown in Figure 6(a) and the effect of positive charges was corrected by the proposed method. Figure 7 compares the corrected trapping density (symbol '△') with the measured N_e after annealing. The trapping after correction for positive charges agrees well with the trapping after ANPC is thermally annealed. This strongly supports the newly proposed correction method.

4.2.2. Trapping kinetics and trap properties

4.2.2.1. The first order model

After correcting the effect of positive charges on electron trapping, we are ready to study the trapping kinetics of generated electron traps and assess trap properties. The simplest model is the first order model with a single capture cross section, σ. It can be expressed as [6],

$$N_e = N_S \cdot [1 - \exp(-\sigma \cdot Q_e)] \tag{3}$$

As the electron fluency, Q_e, increases, this model predicts that trapping will saturate at a level of N_S. Such saturation is indeed observed in Figure 8. However, the correlation between the experimental data and equation (3) (dashed line) is poor. The correlation can be improved (solid line) by assuming that there are two capture cross sections for generated traps, namely,

$$N_e = N_{S1} \cdot [1 - \exp(-\sigma_1 \cdot Q_e)] + N_{S2} \cdot [1 - \exp(-\sigma_2 \cdot Q_e)] \tag{4}$$

The extracted capture cross sections are $\sigma_1 = 1.1\times10^{-14}$ cm^2 and $\sigma_2 = 4.2\times10^{-16}$ cm^2. Here, the adjustable parameters have been increased from two (N_S and σ) in equation

(3) to four (N_{S1}, σ_1, N_{S2}, and σ_2) in equation (4). As a result, the improved correlation may simply result from this increase in adjustable parameters. Indeed, one may say that an excellent agreement between experimental data and a model can always be achieved if the model contains enough adjustable parameters. In fact, this is where many researchers have their reservations about the validation of the first order model. The use of two capture cross sections in equation (4) must be further supported.

4.2.2.2. Support for the presence of two capture cross sections

After correcting the offset effect by positive charges, electron trapping post different stress levels is plotted in Figure 9(a). When compared with the data before correction in Figure 4, the trapping is higher over the whole range of Q_e now for larger Q_h, because a larger Q_h generated more traps. Supports for the existence of two genuine capture cross sections can be obtained by examining their dependence on stress levels, as detailed below:

(i) Different dependence of trap densities on stress levels

On one hand, if the two capture cross sections originated from the same defect, the improved correlation would be an artifact, since there were no evidences for the same defect possessing two well separated capture cross sections. On the other hand, if we can show that these two capture cross sections originate from two different defects, it will support the existence of two capture cross sections strongly.

When there are two different electron traps, it is possible that their densities can change independently as stresses increase. Figure 9(b) shows that this is the case. As stresses increase, the effective density of the larger trap (σ_1: $10^{-13} \sim 10^{-14}$ cm^2), N_{S1}, increases continuously. In contrast, the effective density of the smaller trap (σ_2: $10^{-15} \sim 10^{-16}$ cm^2), N_{S2}, clearly saturates. If they were from the same defect, they should increase simultaneously, which is against Figure 9(b). Consequently, the results

strongly support the presence of two different electron traps with well separated capture cross sections.

(ii) Insensitivity of extracted capture cross sections to stress level

If there exist two capture cross sections, the same values should be observed at all stress levels. Figure 9(c) shows that the two extracted capture cross sections are indeed insensitive to stress levels, further supporting our claim.

(iii) Agreement with direct observations

Figure 9(a) shows that, after correcting the effects of positive charges, the curves are nearly a parallel upward shift as stresses increase. This indicates that the increase in trapping mainly occurs at low Q_e, while the increase is insignificant for further trapping as Q_e becomes higher. To show this more clearly and to avoid curve fitting, the N_{ec} at $Q_e = 4\times10^{14}$ electrons/cm^2 and $\Delta N_{ec} = N_{ec}$ ($Q_e = 10^{17}$ cm^{-2}) - N_{ec} (4×10^{14} cm^{-2}) are plotted in Figure 9(b). The former rises continuously and the latter clearly saturates, indicating the presence of two different trapping processes. This behavior agrees well with that of extracted trap densities.

Finally, it should be point out that the lack of data for $Q_e < 10^{14}$ electrons/cm^2 will affect the accuracy of extracted larger capture cross (σ_1). This is because that its filling partially occurs in the region of $Q_e < 10^{14}$ electrons/cm^2, where no reliable data are available. However, Figure 9(a) shows that electron trapping at $Q_e = 10^{14}$ electrons/cm^2 is substantial and there is no doubt that some acceptor-like traps have an effective physical size in the order of $10^{-13} \sim 10^{-14}$ cm^2. The order of magnitude of σ_1 will not be changed by the lack of data for $Q_e < 10^{14}$ electrons/cm^2. The identification of such a large capture cross section for an acceptor-like electron trap is an achievement of this work, since it can be often masked by the offsetting effect of positive charges. Its large effective physical size and non-saturating nature

indicate that it will play an important role in oxide breakdown and stress-induced leakage current, which need further investigation.

4.3. Trapping kinetics of electron traps generated under FNI

Up to now, only electron traps created by substrate hole injection have been investigated. During the real operation of MOSFETs and the time dependent dielectric breakdown (TDDB) test, Fowler-Nordheim/Direct Tunneling injection of electrons occurs. It is widely believed that hydrogenous species are released by electrons and hydrogen can create electron traps [22-25]. It is interesting and important to investigate whether hydrogen and hole-induced electron traps have the same properties.

4.3.1. Difficulties

There are two main difficulties for characterizing electron trapping after Fowler-Nordheim injection. One of them is that, during Fowler-Nordheim injection, electrons are continuously injected into the oxide. Generated traps can be filled during the stress. Consequently, the generation phase is not clearly separated from the filling process, which complicates the investigation of trapping kinetics. The other difficulty is that the incomplete neutralization of positive charges under E_{ox} = +8 MV/cm. These problems will be addressed.

4.3.1.1. Positive charges formed during FNI

Figure 10 shows that, after a stress of $Q_{inj} = 10^{18}$ carriers/cm^2 under E_{ox} = +11 MV/cm, N_e remains negative at the end of trap-filling period. This means that the neutralization of positive charges under E_{ox} = +8 MV/cm is incomplete when Q_e reached 10^{17} electrons/cm^2. For longer stresses, Figure 11 shows $N_e > 0$. This positive N_e results from the increased creation of electron traps. However, it is not

certain whether N_e is the real level of electron trapping or a net balance level with the simultaneous presence of positive and negative charges. This is investigated next.

In Figure 12(a), a device was first heavily stressed with $Q_{inj} = 1.6\times10^{20}$ carriers/cm^2 under E_{ox} = +11 MV/cm. The subsequent '1st FNI filling' has little effect on the trapping level, which will be addressed further in section 4.3.1.2. To study if there are positive charges, the substrate hole injection is used and a hole fluency of 10^{16} cm^{-2} is sufficiently high to recharge neutralized positive charges and empty trapped electrons, leading to $N_e < 0$ at the end of the SHI. When the trap-filling was restarted (2nd FNI filling), it is clear that the trapping level does not reach that of '1st FNI filling'. This indicates that some positive charges can be neutralized under E_{ox} = +11 MV/cm, but cannot become neutral under E_{ox} = +8 MV/cm at the end of filling period. This amount of positive charges are shown as $\Delta N_p(11)$ in Figure 12(b).

To support the claim that $\Delta N_p(11)$ can be fully neutralized under E_{ox} = +11 MV/cm, the FNI at +11MV/cm was switched on again. The symbol '+' in Figure 12(a) shows that the positive charges are 'rapidly' neutralized and N_e now can reach its level of '1st FNI filling'. Figure 12(b) shows that an electron injection of 10^{19} cm^{-2} under +11 MV/cm is needed to fully neutralize the positive charges. The increase of N_e represented by the symbol '+' in Figure 12(a) & (b) is not caused by further electron trap creation, since the injection level of 10^{19} carriers/cm^2 is insignificant after the device was stressed with 1.6×10^{20} carriers/cm^2.

In summary, it is found that the FNI induced positive charges are not fully neutralized at the end of trap-filling period with an electron fluency of 10^{17} cm^{-2} under +8 MV/cm. These positive charges were only neutralized when electron fluency reached 10^{19} cm^{-2} under +11 MV/cm. The N_e measured with $Q_{inj} \geq 10^{19}$ carriers/cm^2 under +11 MV/cm represents the true level of electron trapping.

4.3.1.2. Separation of trap creation from filling

After stressing under +11 MV/cm, Figure 11 shows that the filling curves under +8 MV/cm are almost flat, indicating that traps are filled already during the stress. This is obviously not suitable for investigating the trapping kinetics. If hole injection is used for emptying the filled traps, Figure 12 shows that the positive charges may not be fully neutralized, complicating the analysis.

To overcome this difficulty, one can explore the difference in the rate for emptying trapped electrons and recharging the neutralized positive charges. Since the former is a Coulombic attractive process and the latter is not, there may be an hole fluency that is high enough to empty the trapped electrons, but not enough for the recharge.

In Figure 12, the hole fluency used is 10^{16} cm^{-2}. When it is reduced to 10^{13} holes/cm^2, Figure 13(a) & (b) show that this hole fluency is high enough to empty trapped electrons, so that $N_e < 0$ is reached. During the '2nd FNI filling', the level of N_e is approximately the same as that of '1st FNI filling', in contrast with Figure 12(a). It confirms that the 'un-neutralizable' defects are not recharged when hole fluency is limited to 10^{13} cm^{-2}. This allows the separation of trap generation phase from the trap filling.

4.3.2. Trapping kinetics

After the hole injection in Figure 13(a), the trapping kinetics can be analyzed by using the method developed in the section 4.2.1.2. The same correction method is used and Figure 13(c) shows the trapping level before and after the correction.

Figure 14(a) shows that the trapping after different level of stresses. When applying the first order model with two capture cross sections, the extracted trapping densities and capture cross sections are shown in Figure 14(b) & (c), respectively. Both trap densities and capture cross sections generated by FNI behave similarly to those created by SHI. The two capture cross sections are again in the order of 10^{-13} ~ 10^{-14}

cm^2 and $10^{-15} \sim 10^{-16}$ cm^2. The generation of the smaller traps saturates, while the creation of the larger traps does not.

To further compare the trapping kinetics of traps generated by SHI and FNI, Figure 15 plots them together. The trapping kinetics is essentially the same. These results strongly indicate that the same types of traps are generated by FNI and SHI. Traps generated by hydrogenous species and holes have the same electrical signature.

4.4. Trapping kinetics of electron traps generated in thick oxides

Up to now, only electron traps created in a 7.1 nm oxide were investigated. It will be interesting to find out if the extracted capture cross sections are sensitive to oxide thickness. If the properties of generated electron traps are insensitive to oxide thickness, then above results should also be applicable for today's state of the art oxides.

As pointed out earlier in section 4.1, for thin oxides (< 3 nm), traps can be located sufficiently close to the electrodes that trapped electrons can tunnel away, which makes the trapping unstable, and adds difficulties to study trapping kinetics. As a result, thinner oxides are unsuitable and thicker oxides have to be used.

In the following sections, the properties of generated electron traps are investigated on 10.8 nm and 13.8 nm oxides.

4.4.1. Trapping kinetics

Figure 16 shows the generated electron traps after stressing a 13.8 nm oxide under -11 MV/cm. The results are similar to those given in Figure 4 for a 7.1 nm oxide: when the stress, Q_{inj}, increases, the trapping level increases more at the end of trap-filling than at the start of the filling. This indicates that positive charge correction is

needed again. Figure 17(a) & (b) show the correction process and Figure 17(c) gives the corrected results.

By applying the first order model to Figure 17(c), the extracted trap densities and capture cross sections are given in Figure 18(a) & (b), respectively.

4.4.2. A comparison of traps generated in oxides of different thicknesses

The similarity of traps created in oxides of different thicknesses can be seen:

(i) The extracted capture cross sections are insensitive to the thickness. Figure 18(b) shows that the two capture cross sections are again in the order of $10^{-13} \sim 10^{-14}$ cm^2 and $10^{-15} \sim 10^{-16}$ cm^2.

(ii) When stresses increase, the density of the smaller traps saturates, but the density of the larger traps continuously build up.

Figure 19 compares the trapping kinetics of traps created in 7.1 nm and 13.8 nm. It is clear that the two sets of data can be fitted with the same capture cross sections. This strongly supports that the same traps were generated in SiO_2 of different thickness.

For the oxide breakdown, the most successful model is the percolation model [9]. It assumes that electron traps are generated statistically randomly in the oxide layer. To test this assumption, the volume density is plotted against stress levels for 7.1 nm, 10.8 nm and 13.8 nm oxides in Figure 20. The volume density is clearly insensitive to the thickness, indicating a 'uniform' distribution of generated defects. A statistically random generation at a microscopic scale will lead to a 'uniform' distribution at a macroscopic scale. This result supports the assumption.

4.5. Conclusions

To improve our understanding of generated electron traps, trapping kinetics is investigated in this work. By stressing the oxide with substrate hole injection, the generation phase is separated from the trap filling phase. The use of uniform stress and filling removes the uncertainty in lateral distribution. The selection of relatively thick SiO_2 layer allows the direct measurement of trapped electron density and the uncertainty caused by the use of SILC is avoided. An achievement of this work is the development of a new method, which successfully corrected the effect of anti-neutralization positive charges (ANPC).

After the correction, electron capture cross section as large as $10^{-13} \sim 10^{-14}$ cm^2 is observed for generated acceptor-like electron traps. The filling kinetics follows the first order model and there are two genuine and well separated capture cross sections. The smaller capture cross section is in the region of $10^{-15} \sim 10^{-16}$ cm^2. For the first time, it is clearly shown that the density of the smaller trap saturates, while the density of the larger trap does not.

It is found that the traps created by FNI are similar to those by SHI in terms of both capture cross sections and the dependence of trap density on stress levels. This result supports that the same types of traps were created under different stresses, and the electrical signatures of traps were the same with or without hydrogen during the generation.

It is also shown that the electrical signature of generated electron traps does not depend on oxide thickness. The volume density is also insensitive to oxide thickness. This is in agreement with early assumption that the generation is a statistically random process. This work should provide useful information for modeling the SILC and oxide breakdown in the future. For example, the continuous generation of the larger traps implies that they will dominate the breakdown, while the saturation of the smaller traps makes them less important.

References

1. E. H. Nicollian, C. N. Berglund, P. F. Schmidt and J. M. Andrews, "Electrochemical charging of thermal SiO_2 films by injected electron currents," J. Appl. Phys., **42**, 5654 (1971).
2. R. F. DeKeersmaecker and D. J. DiMaria, "Electron trapping and detrapping characteristics of arsenic-implanted SiO_2 layers," J. Appl. Phys., **51**, 1085 (1980).
3. J. F. Zhang, S. Taylor and W. Eccleston, "A quantitative investigation of electron trapping in SiO_2 under Fowler-Nordheim," J. Appl. Phys., **71**, 5989 (1992).
4. W. D. Zhang, J. F. Zhang, M. Lalor, D. Burton, G. Groeseneken and R. Degraeve, "Two types of neutral electron traps generated in the gate silicon dioxide," IEEE Trans. Electron Dev., **49**, 1868 (2002).
5. W. Chen, A. Balasinski and T. P. Ma, "Lateral profiling of oxide charge and interface traps near MOSFET junctions," IEEE Trans. Electron Dev., **40**, 187 (1993).
6. J. F. Zhang, S. Taylor and W. Eccleston, "Electron trap generation in thermally grown SiO_2 under Fowler-Nordheim stress," J. Appl. Phys., **71**, 725 (1992).
7. M. H. Chang and J. F. Zhang, "On the role of hydrogen in hole-induced electron trap creation," Semicond. Sci. Technol., **19**, 1333 (2004).
8. D. J. DiMaria and J. H. Stathis, "Anode hole injection, defect generation, and breakdown in ultrathin silicon dioxide films," J. Appl. Phys., **89**, 5015 (2001).
9. R. Degraeve, G. Groeseneken, R. Bellens, J. L. Ogier, M. Depas, P. J. Roussel and H. E. Maes, "New insights in the relation between electron trap generation and the statistical properties of oxide breakdown," IEEE Trans. Electron Dev., **45**, 904 (1998).
10. R. Degraeve, T. Kauerauf, A. Kerber, E. Cartier, B. Govoreanu, Ph. Roussel, L. Pantisano, P. Blomme, B. Kaczer and G. Groeseneken, "Stress polarity dependence of degradation and breakdown of SiO_2/high-k stacks," in Proc. IEEE 41st Int. Reliability Phys. Symp., Dallas, p.23, USA (2003).

11. D. Ielmini, A. S. Spinelli, M. A. Rigamonti and A. L. Lacaita, "Modeling of SILC based on electron and hole tunneling. Part I: Transient effects," IEEE Trans. Electron Devices, **47**, 1258 (2000).
12. M. Brox and W. Weber, "Dynamic degradation in MOSFET's. Part I: The physical effects," IEEE Trans. Electron Devices, **38**, 1852 (1991).
13. J. F. Zhang, H. K. Sii, G. Groeseneken and R. Degraeve, "Hole trapping and trap generation in the gate silicon dioxide," IEEE Trans. Electron Dev., **48**, 1127 (2001).
14. J. F. Zhang, C. Z. Zhao, A. H. Chen, G. Groeseneken and R. Degraeve, "Hole traps in silicon dioxides. Part I: Properties," IEEE Trans. Electron Dev., **51**, 1267 (2004).
15. C. Z. Zhao, J. F. Zhang, G. Groeseneken and R. Degraeve, "Hole traps in silicon dioxides. Part II: Generation mechanism," IEEE Trans. Electron Dev., **51**, 1274 (2004).
16. C. Z. Zhao and J. F. Zhang, "Effects of hydrogen on positive charges in gate oxides," J. Appl. Phys., **97**, 073703 (2005).
17. W. D. Zhang, J. F. Zhang, M. J. Lalor, D. R. Burton, G. Groeseneken and R. Degraeve, "Effects of detrapping on electron traps generated in gate oxides," Semicond. Sci. Technol., **18**, 174 (2003).
18. D. J. DiMaria, "Defect generation under substrate-hot-electron injection into ultrathin silicon dioxide layers," J. Appl. Phys., **86**, 2100 (1999).
19. J. F. Zhang, C. Z. Zhao, G. Groeseneken, R. Degraeve, J. N. Ellis and C. D. Beech, "Hydrogen induced positive charge generation in gate oxides," J. Appl. Phys., **90**, 1911 (2001).
20. C. Z. Zhao and J. F. Zhang, "Effects of hydrogen on positive charges in gate oxides," J. Appl. Phys., **97**, 073703 (2005).
21. A. R. Stives and C. T. Sah, "A study of oxide traps and interface states of the silicon-silicon dioxide interface," J. Appl. Phys., **51**, 6296 (1980).
22. D. J. DiMaria and J. W. Stasiak, "Trap creation in silicon dioxide produced by hot-electrons," J. Appl. Phys., **65**, 2342 (1989).
23. D. J. DiMaria, E. Cartier and D. Arnold, "Impact ionization, trap creation, degradation, and breakdown in silicon dioxide films on silicon," J. Appl. Phys., **73**, 3367 (1993).

24. D. J. DiMaria, D. J. Buchaman, D. A. Stathis and R. E. Stahlbush, "Interface states induced by the presence of trapped holes near the silicon-silicon dioxide interface," J. Appl. Phys., **77**, 2032 (1995).
25. D. J. DiMaria, "Temperature-dependence of trap creation in silicon dioxide," J. Appl. Phys., **68**, 5234 (1990).

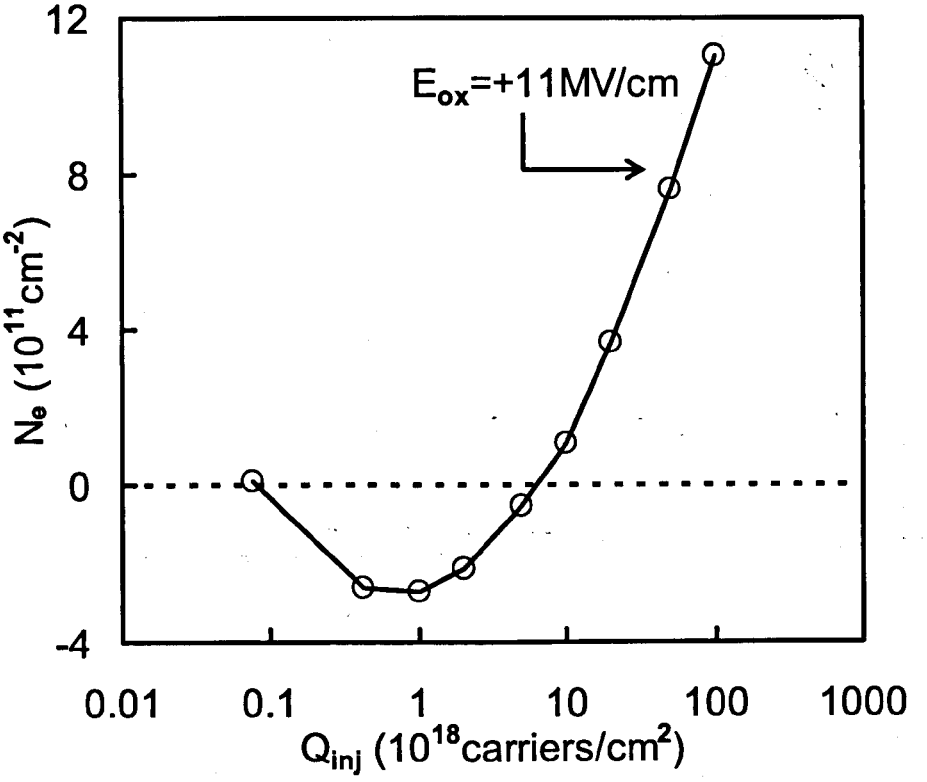


Figure 1. Electron trapping density, N_e, during stresses under an oxide field of +11 MV/cm. N_e is negative initially due to positive charge formation. As electron traps are generated and filled, N_e becomes positive, representing net negative charges. The trap generation and filling processes cannot be separated here and the data cannot be used for studying the kinetics of trap filling.

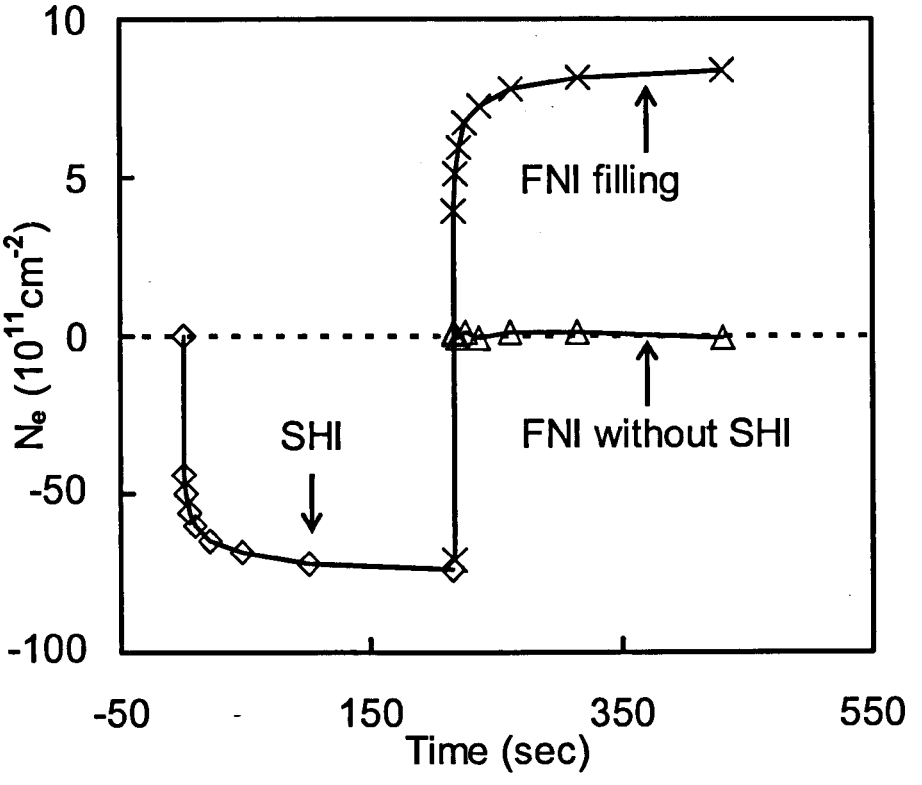


Figure 2. Typical test procedures. A device was stressed by substrate hole injection (SHI, Oxide filed: E_{ox} = -5 MV/cm, n-well bias: 8.8 V; p-substrate bias: 9.8 V), during which positive charges were formed through hole trapping. After SHI, electrons were injected into the oxide under E_{ox} = +8 MV/cm for filling the generated electron traps and neutralizing positive charges (symbol symbol 'x'). If E_{ox} = +8 MV/cm is applied to a fresh device (symbol '△'), there is little trapping, indicating the trap generation by the filling step itself is negligible.

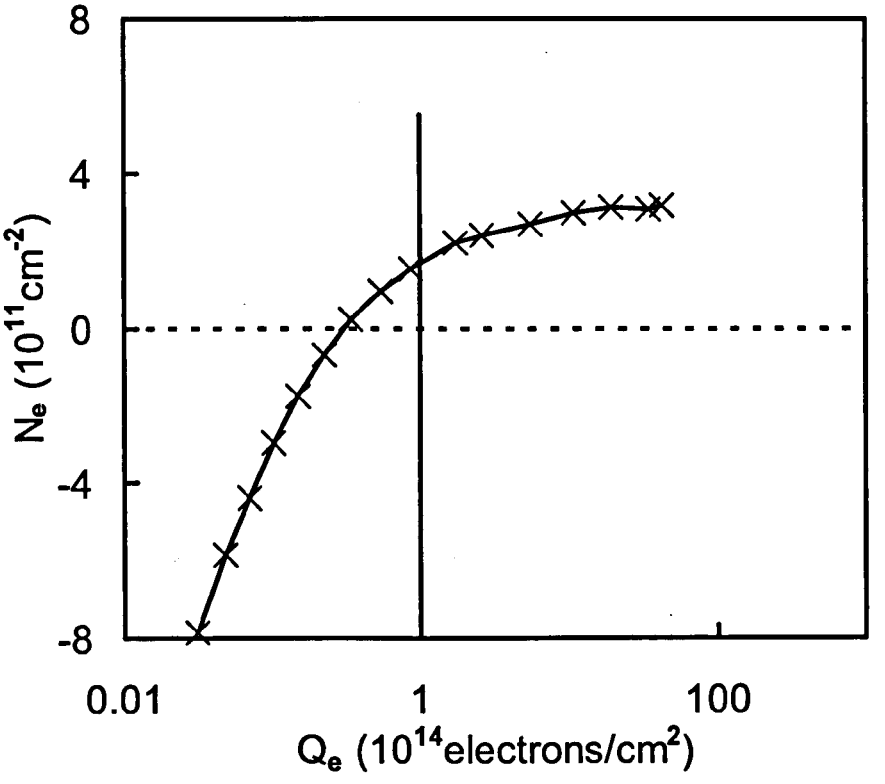


Figure 3. A device was stressed by SHI with a hole fluency of 10^{15} cm^{-2}. At low level of electron injection, Q_e, during the filling, N_e was initially negative, representing net positive charges. This is because neutralization of as-grown hole traps was not completed until Q_e reached 10^{14} cm^{-2}, approximately. As a result, data for $Q_e < 10^{14}$ cm^{-2} will not be used for studying trapping kinetics in this work. To facilitate the measurement at low Q_e, E_{ox} = +6.5 MV/cm is used here.

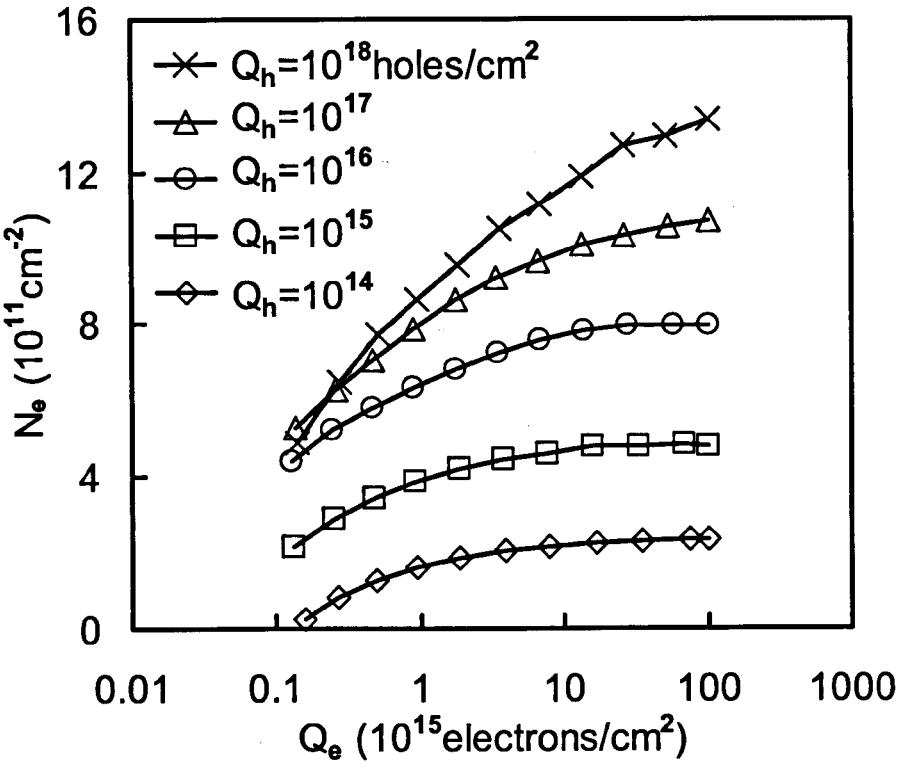


Figure 4. Electron trapping after different stress levels by SHI. When hole fluency, Q_h, is less than 10^{16} cm^{-2}, an increase in Q_h leads to higher N_e over the whole range of Q_e. For $Q_h > 10^{16}$ cm^{-2}, however, N_e increases at high Q_e, but this increase is suppressed at low Q_e.

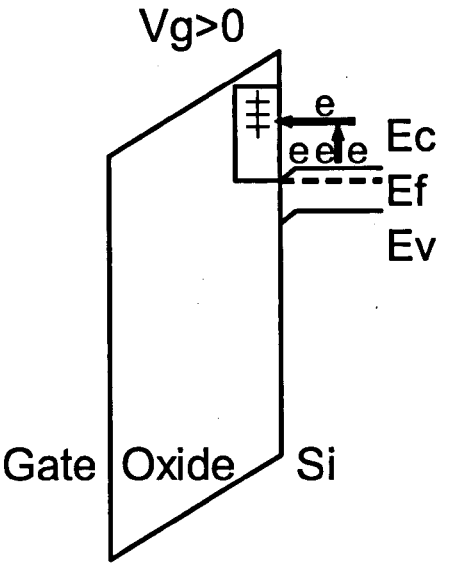
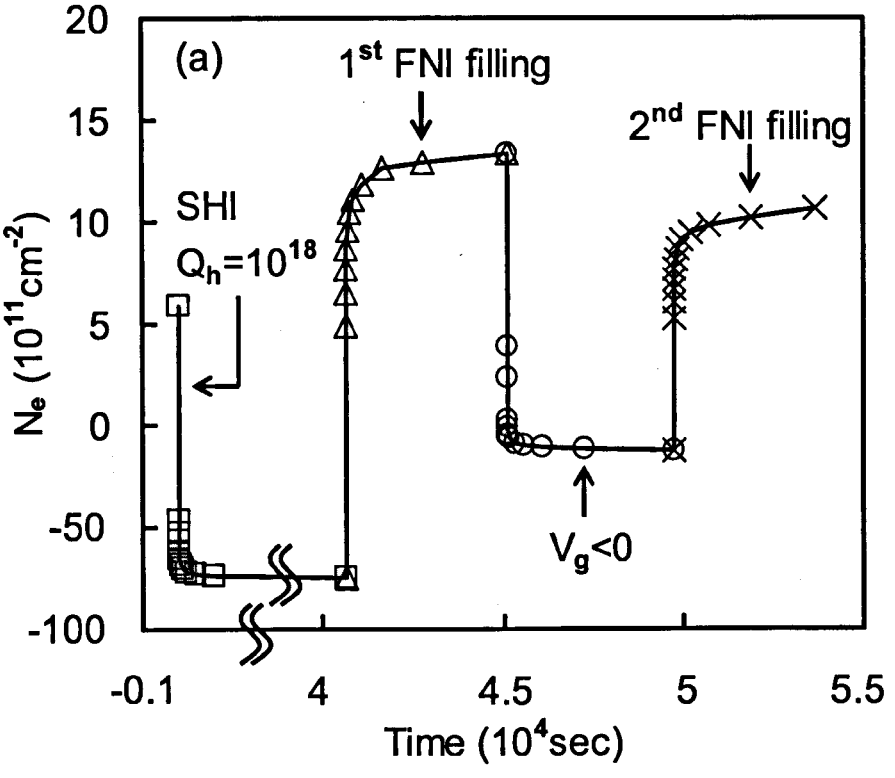
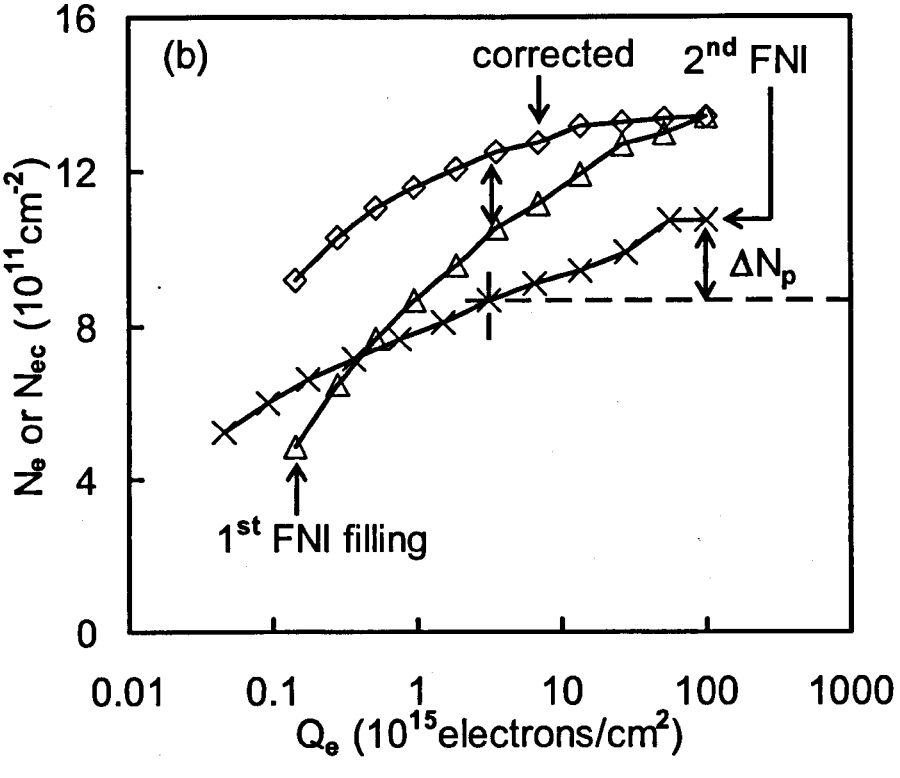


Figure 5. Energy band diagram for anti-neutralization positive charges (ANPC). ANPC has energy levels above the bottom edge of silicon conduction band. The higher the energy level, the less electrons are available in silicon and the more difficult for neutralizing a positive charge.


(a)
1st FNI filling
2nd FNI filling
SHI
Qh=10^18
Vg<0
Ne (10^11cm^-2)
Time (10^4sec)
20
15
10
5
0
-50
-100
-0.1
4
4.5
5
5.5


(b)
corrected
2nd FNI
ΔNp
1st FNI filling
Ne or Nec (10^11cm^-2)
Qe (10^15electrons/cm^2)
16
12
8
4
0
0.01
0.1
1
10
100
1000

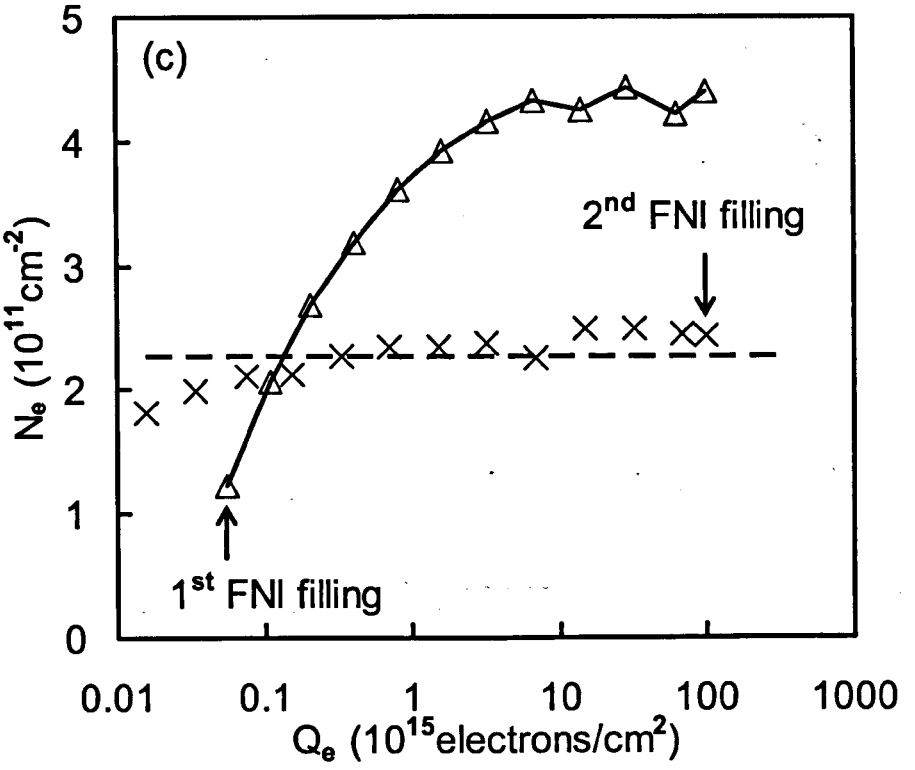


Figure 6. A method for correcting the offset effect of anti-neutralization positive charges (ANPC) on N_e. In (a), a device was heavily stressed by SHI with $Q_h = 10^{18}$ hole/cm^2, which generated not only electron traps, but also ANPC. During the following '1st FNI filling', electron traps captured electrons and became negative and positive charges were neutralized. This led to the rise of the net electron trapping (symbol '△'). A negative gate bias ($V_g < 0$, $E_{ox} = -5$ MV/cm) was then applied. Under $V_g < 0$, both filled electron traps and neutralized ANPC can lose their electrons through tunneling, which resulted in the fall of N_e (symbol 'o'). When electron injection was resumed (the '2nd FNI filling'), the detrapped electron traps cannot recapture electrons and the rise represented by symbol 'x' is entirely from re-neutralizing ANPC. In (b), the data represented by the symbols '△' and 'x' in (a) are re-plotted against Q_e. At a given level of Q_e, the remaining positive ANPC is represented by ΔN_p. The correct electron trapping level is the net trapping (symbol '△') plus ΔN_p and is shown as the symbol '◇'. In (c), the test procedure shown in (a) was used again, but Q_h was limited at 10^{15} hole/cm^2, which is sufficiently low that creation of ANPC is not significant. (c) shows that the lack of ANPC leads to a nearly flat N_e during the '2nd FNI filling'. The dashed line is a guide for the eye.

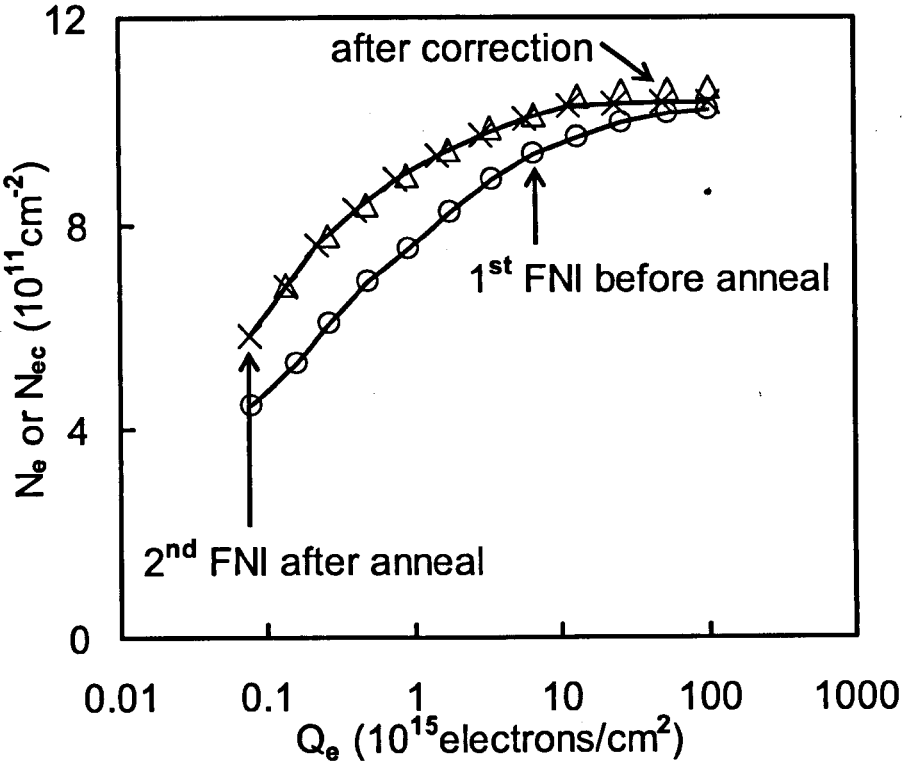


Figure 7. Two devices were subjected to the same stress of $Q_h = 10^{17}$ cm^{-2}. After filling the traps (symbol 'o'), one of them was annealed at 150°C for 160 min to remove the ANPC. After annealing ANPC, a short SHI ($Q_h = 10^{15}$ cm^{-2}) was used to empty all electron traps, before they were refilled (symbol 'x'). The other device went through the test sequence shown in Figure 6(a) and the effect of positive charges was corrected by following the method shown in Figure 6(b). The corrected data are represented by symbol '△'. The good agreement between the trapping after annealing ANPC and that after the correction supports the correction method strongly.

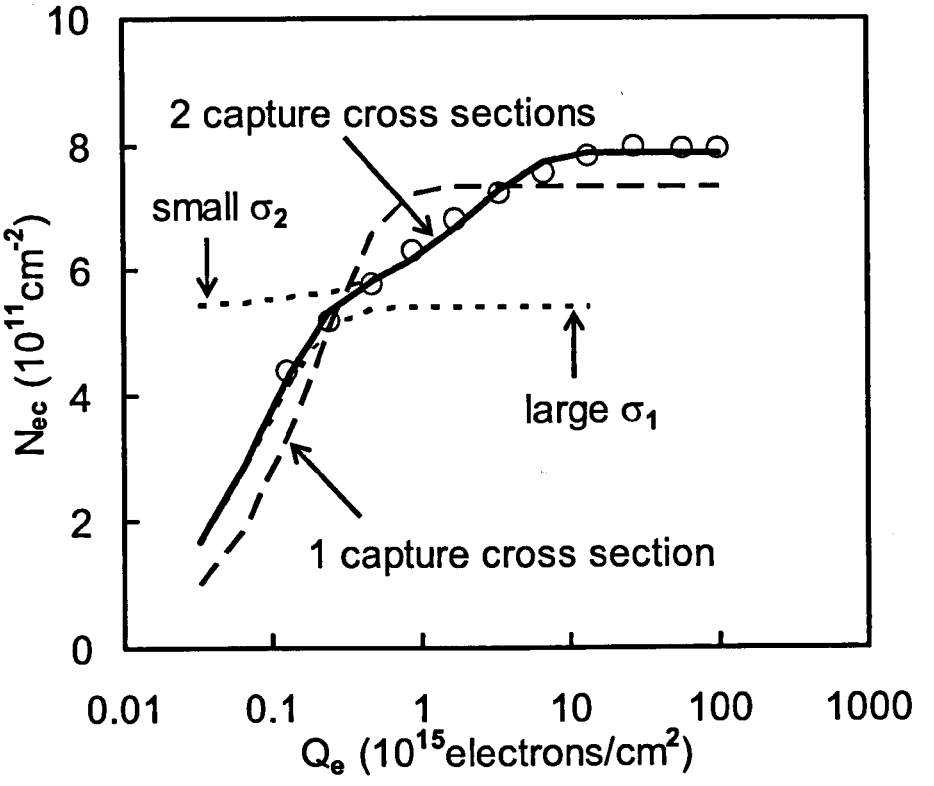
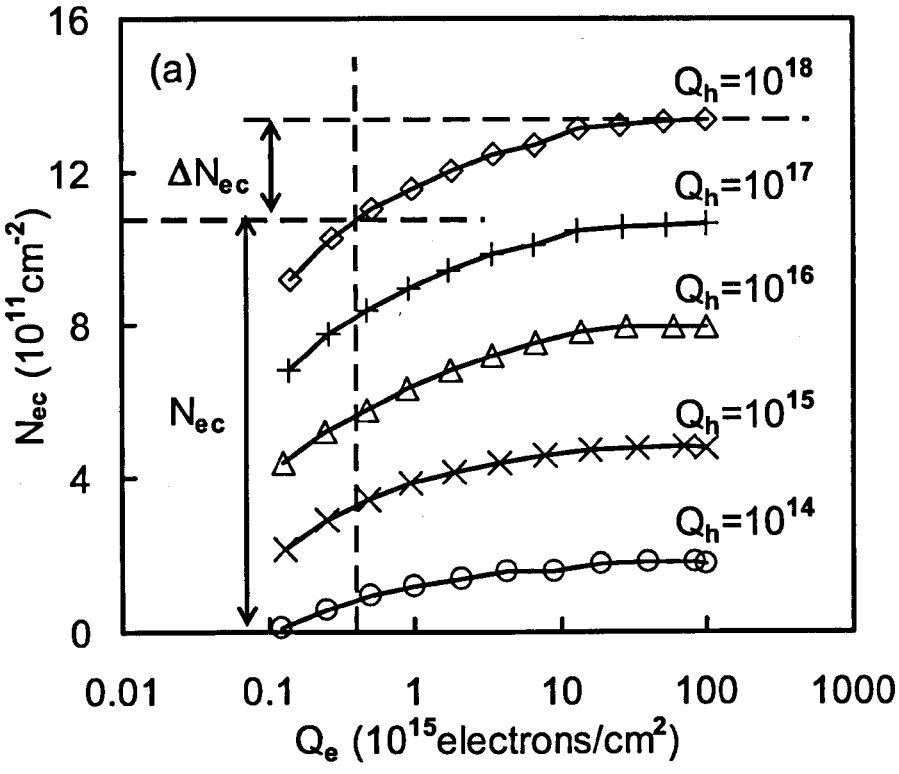
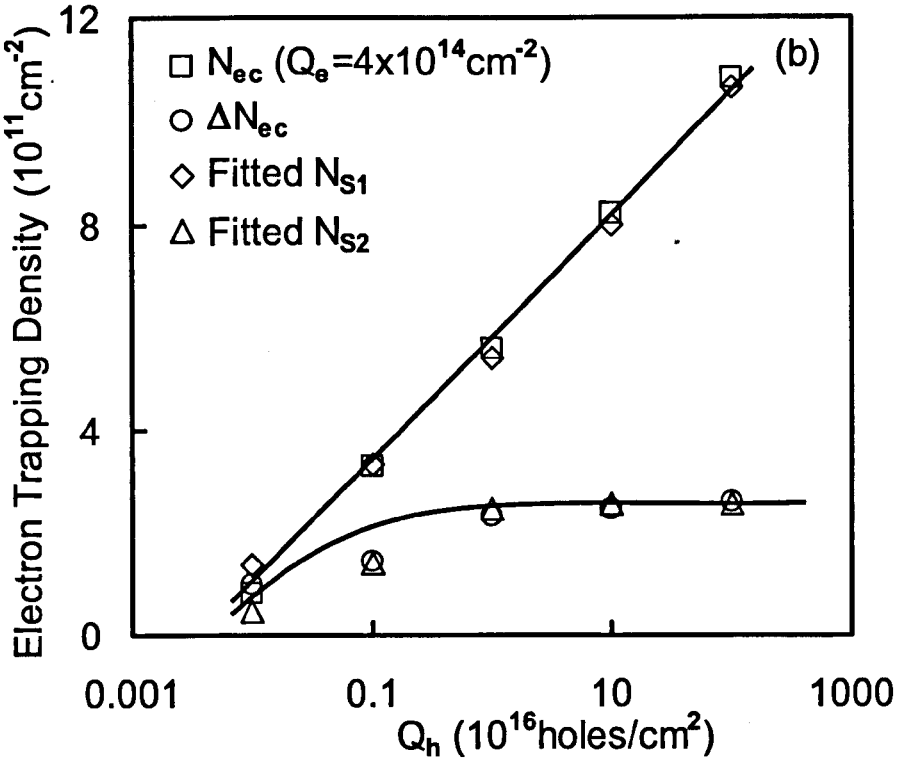


Figure 8. Trapping kinetics of electron traps generated by SHI with $Q_h = 10^{16}$ cm^{-2}. The dashed and solid line was obtained by fitting the data with equation (3) and (4), respectively. The two dotted lines show the contribution of traps with large ($\sigma_1 = 1.1\times10^{-14}$ cm^2) and small ($\sigma_2 = 4.2\times10^{-16}$ cm^2) capture cross sections.


(a)
$Q_h=10^{18}$
$Q_h=10^{17}$
$Q_h=10^{16}$
$Q_h=10^{15}$
$Q_h=10^{14}$
ΔN_{ec}
N_{ec}
N_{ec} (10^{11}cm^{-2})
Q_e (10^{15}electrons/cm^2)
16
12
8
4
0
0.01
0.1
1
10
100
1000


(b)
N_{ec} (Q_e=4x10^{14}cm^{-2})
ΔN_{ec}
Fitted N_{S1}
Fitted N_{S2}
Electron Trapping Density (10^{11}cm^{-2})
Q_h (10^{16}holes/cm^2)
12
8
4
0
0.001
0.1
10
1000

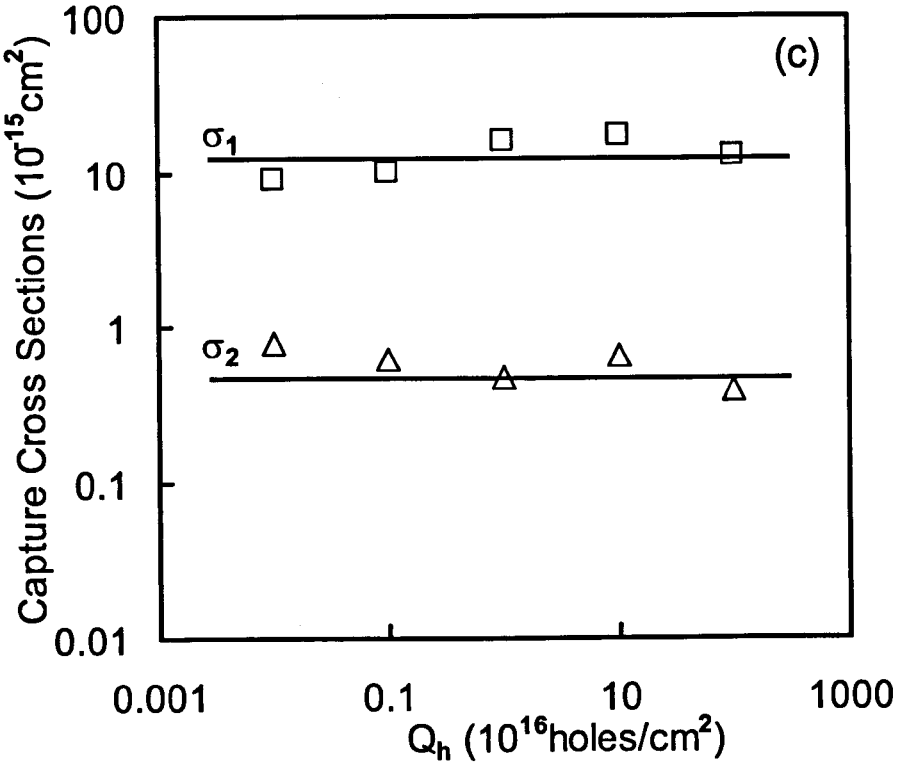


Figure 9. Supports for the presence of two capture cross sections. (a) The corrected trapping, N_{ec}, after different stress levels. The unit of Q_h in the figure is holes/cm^2. The N_{ec} at $Q_e = 4\times10^{14}$ cm^{-2} and the subsequent trapping, ΔN_{ec}, is marked out for $Q_h = 10^{18}$ holes/cm^2. (b) The extracted N_{S1} and N_{S2}, together with N_{ec} ($Q_e = 4\times10^{14}$ cm^{-2}) and ΔN_{ec}. The different dependence of N_{S1} and N_{S2} on stresses indicates that they originate from two different defects. This supports the presence of two capture cross sections strongly. (c) Shows that the extracted two capture cross sections are insensitive to stress levels.

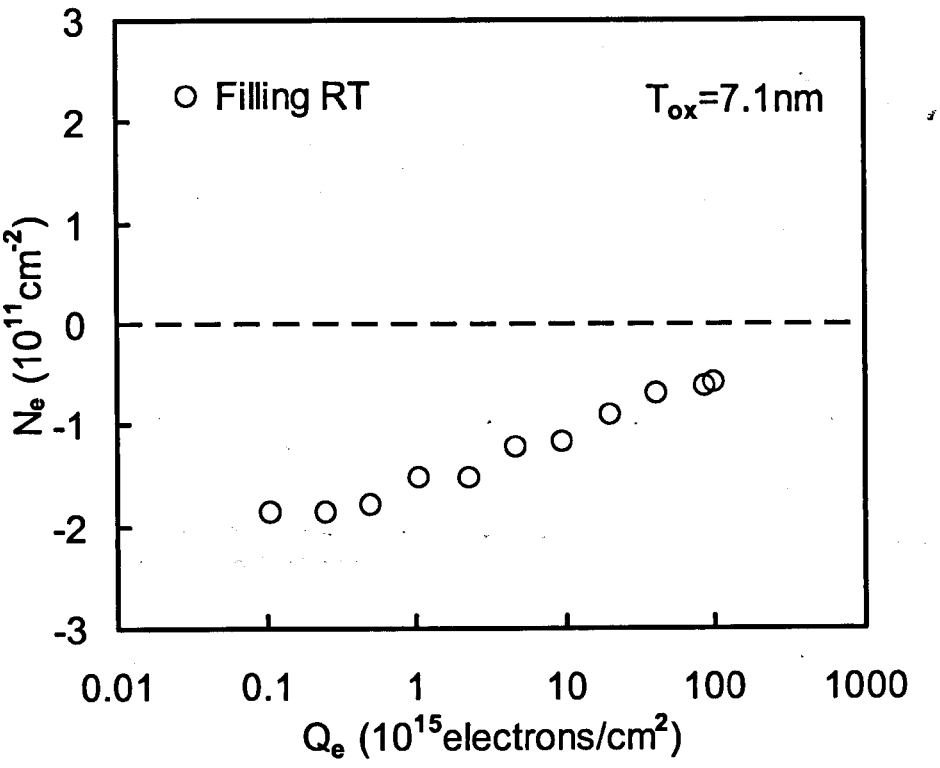


Figure 10. A device was stressed under +11 MV/cm for Q_{inj} = 1×10^{18} carriers/cm^2, then filled by +8 MV/cm at RT. Positive charges induced by +11 MV/cm, $\Delta N_p(11)$, cannot be fully neutralized. N_e remains negative after 10^{17} of electron injection, which may suggest that $\Delta N_p(11)$ have potentially a higher energy level.

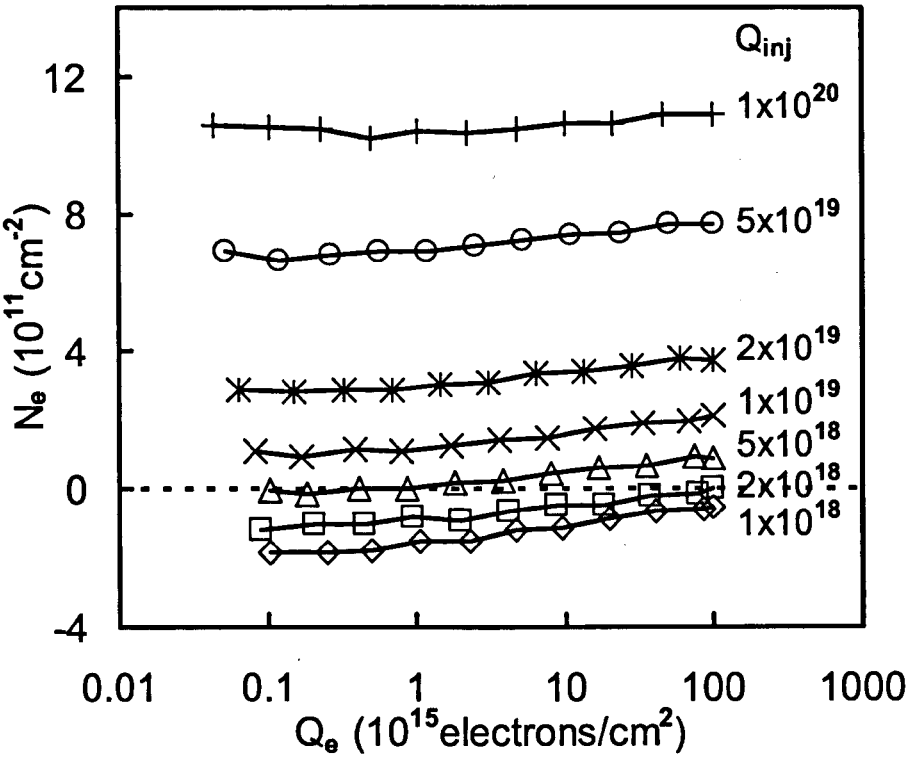
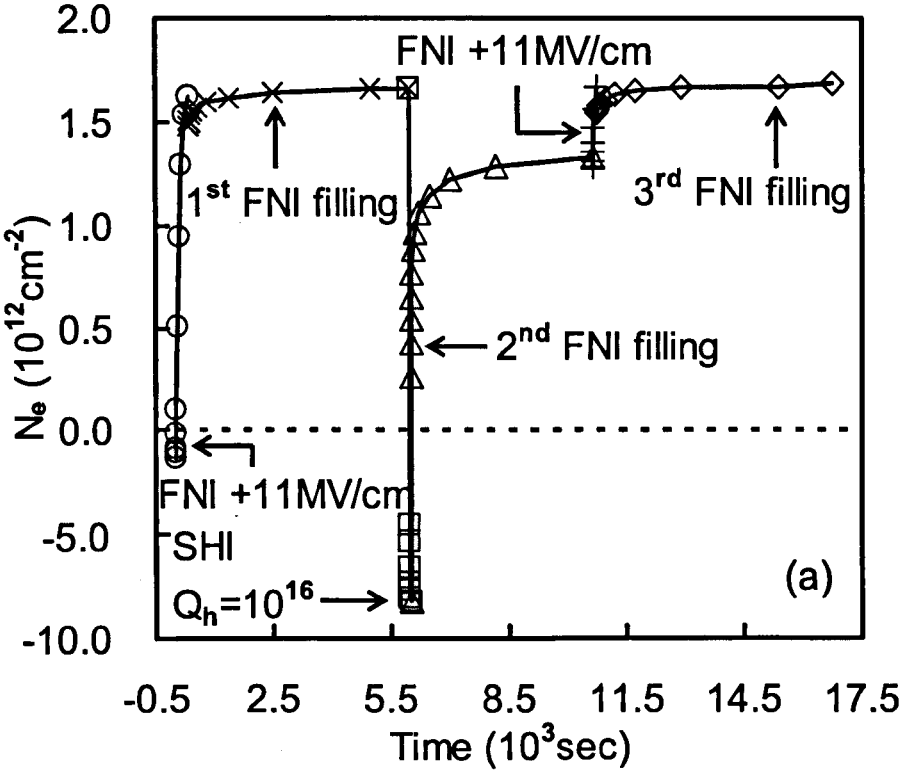


Figure 11. After longer stresses under E_{ox} = +11 MV/cm, N_e becomes positive (symbol 'x', '∗', 'o' and '+'). This positive N_e results from the increased creation of electron traps. However, it is not certain whether N_e is the real level of electron trapping or a net balance level with the simultaneous presence of positive and negative charges.


FNI +11MV/cm
1st FNI filling
3rd FNI filling
2nd FNI filling
FNI +11MV/cm
SHI
Qh=10^16
(a)
Ne (10^12cm^-2)
Time (10^3sec)
2.0
1.5
1.0
0.5
0.0
-5.0
-10.0
-0.5
2.5
5.5
8.5
11.5
14.5
17.5

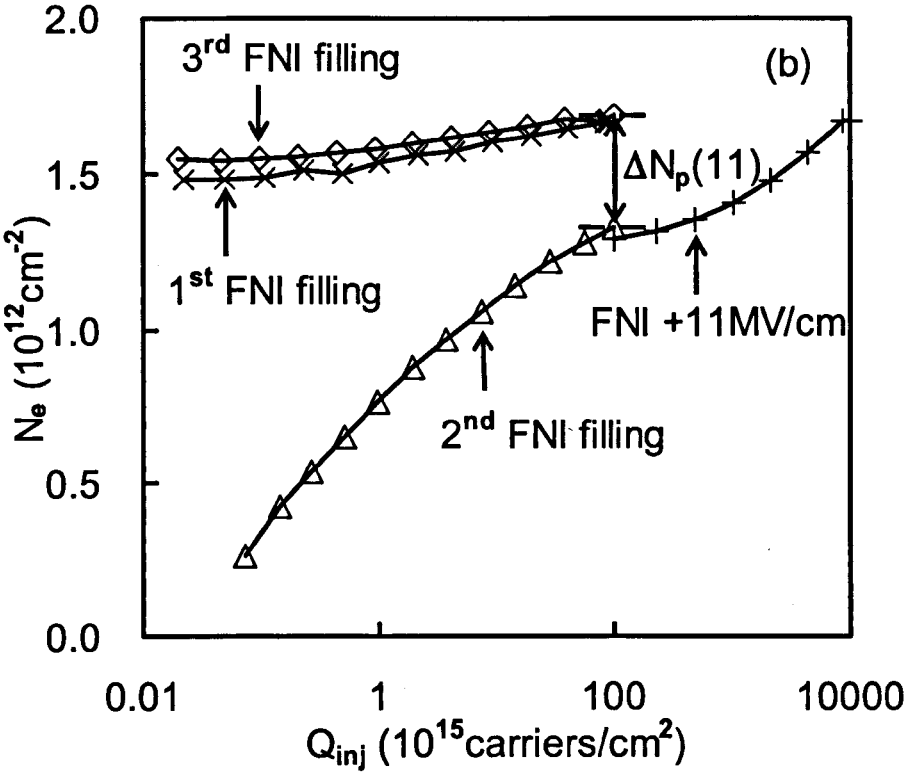
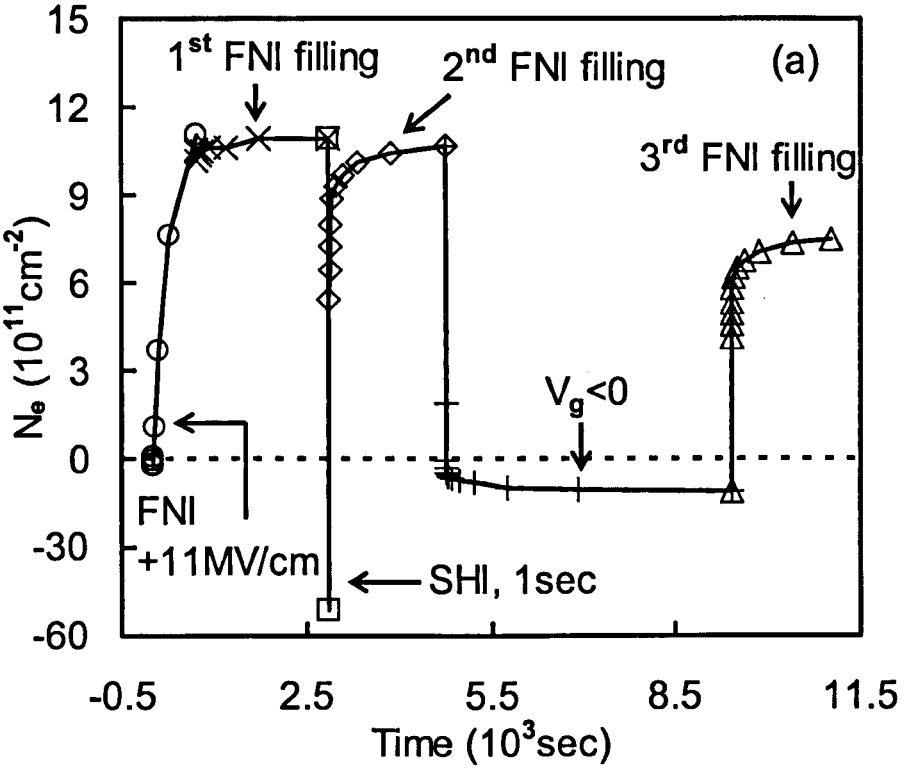
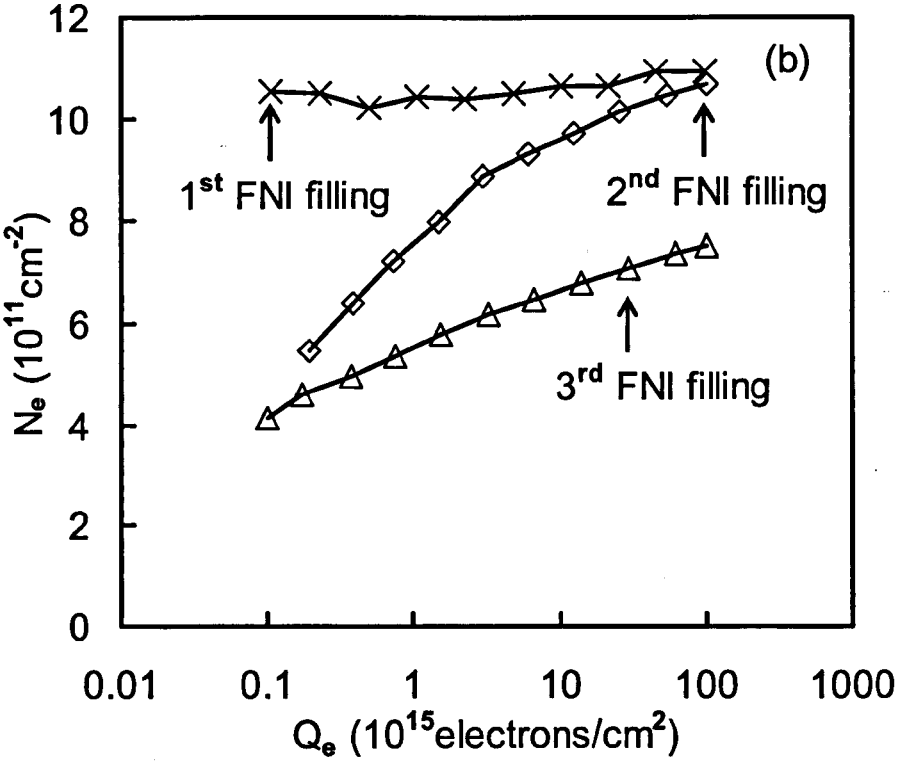

3rd FNI filling
(b)
ΔNp(11)
1st FNI filling
FNI +11MV/cm
2nd FNI filling
Ne (10^12cm^-2)
Qinj (10^15carriers/cm^2)
2.0
1.5
1.0
0.5
0.0
0.01
1
100
10000

Figure 12. (a) A device was heavily stressed under +11 MV/cm (Q_{inj} = 1.6×10^{20} carriers/cm^2, symbol '○'), then switched to +8 MV/cm (symbol 'x'). To study the positive charges formed during FNI +11 MV/cm, a SHI followed to recharge neutralized positive charges and empty trapped electrons (Q_h = 1×10^{16} holes/cm^2, symbol '□'). The trapping level of subsequent filling (symbol '△') does not reach that of 1st FNI filling, due to recharged positive charges cannot be neutralized under +8 MV/cm. FNI at +11 MV/cm was switched on again to neutralize $\Delta N_p(11)$ (Q_{inj} = 1×10^{19} carriers/cm^2, symbol '+'), then switched to +8 MV/cm (symbol '◇'). (b) The good agreement between symbol '◇' and symbol 'x' confirms that $\Delta N_p(11)$ can be effectively neutralized after under 10^{19} of carriers injection under +11 MV/cm.


15
12
9
6
3
0
-30
-60
N_e (10^{11}cm^{-2})
1st FNI filling
2nd FNI filling
(a)
3rd FNI filling
V_g<0
FNI
+11MV/cm
SHI, 1sec
-0.5
2.5
5.5
8.5
11.5
Time (10^3sec)


12
10
8
6
4
2
0
N_e (10^{11}cm^{-2})
(b)
1st FNI filling
2nd FNI filling
3rd FNI filling
0.01
0.1
1
10
100
1000
Q_e (10^{15}electrons/cm^2)

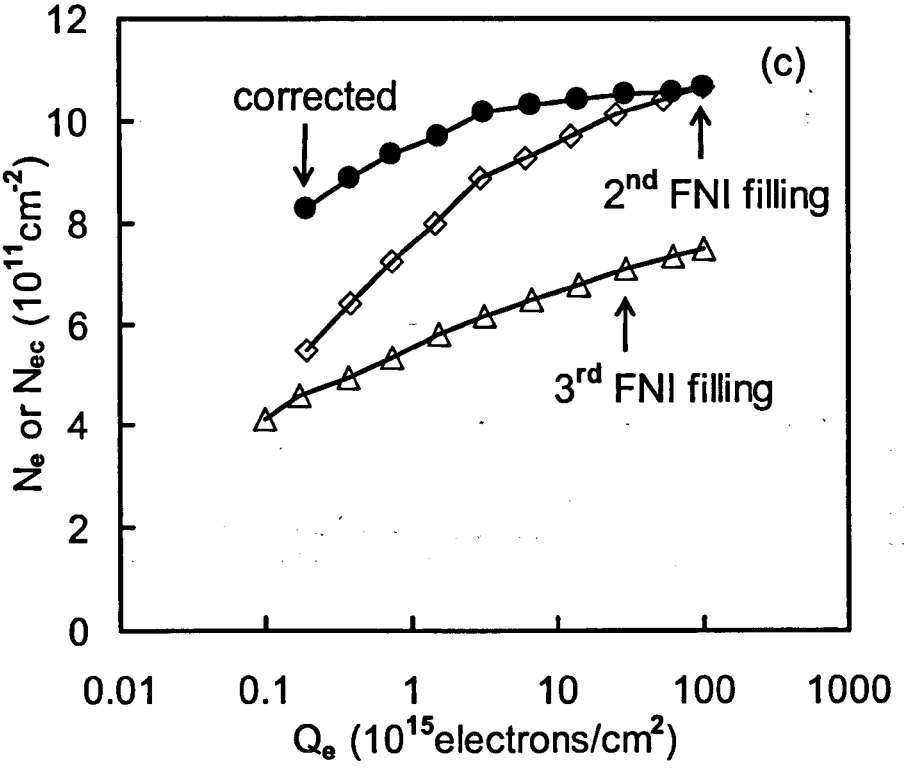
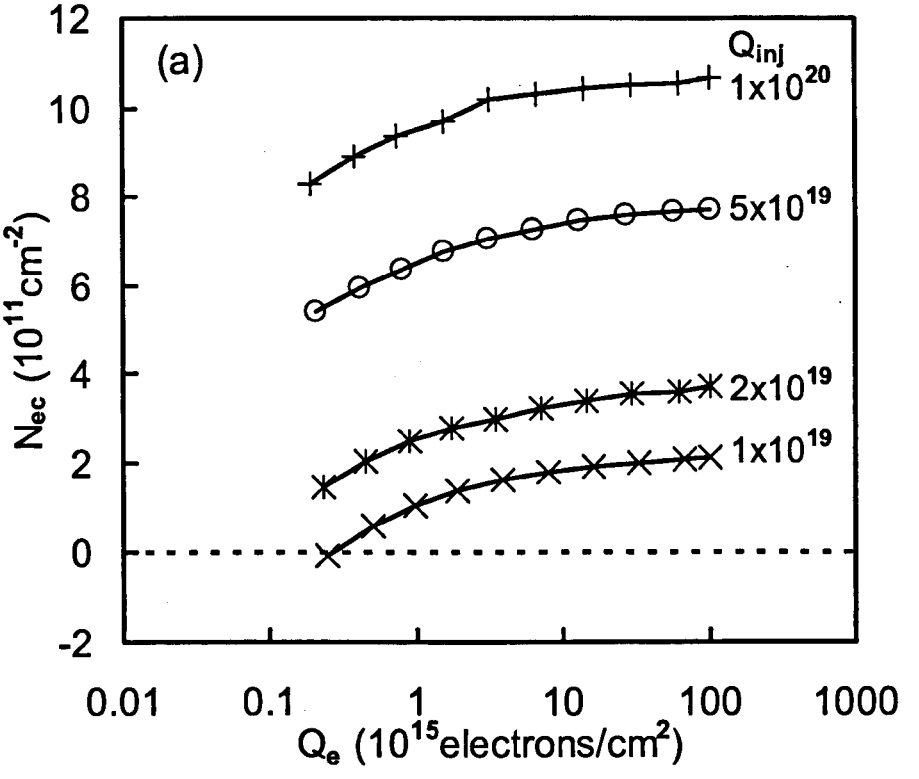
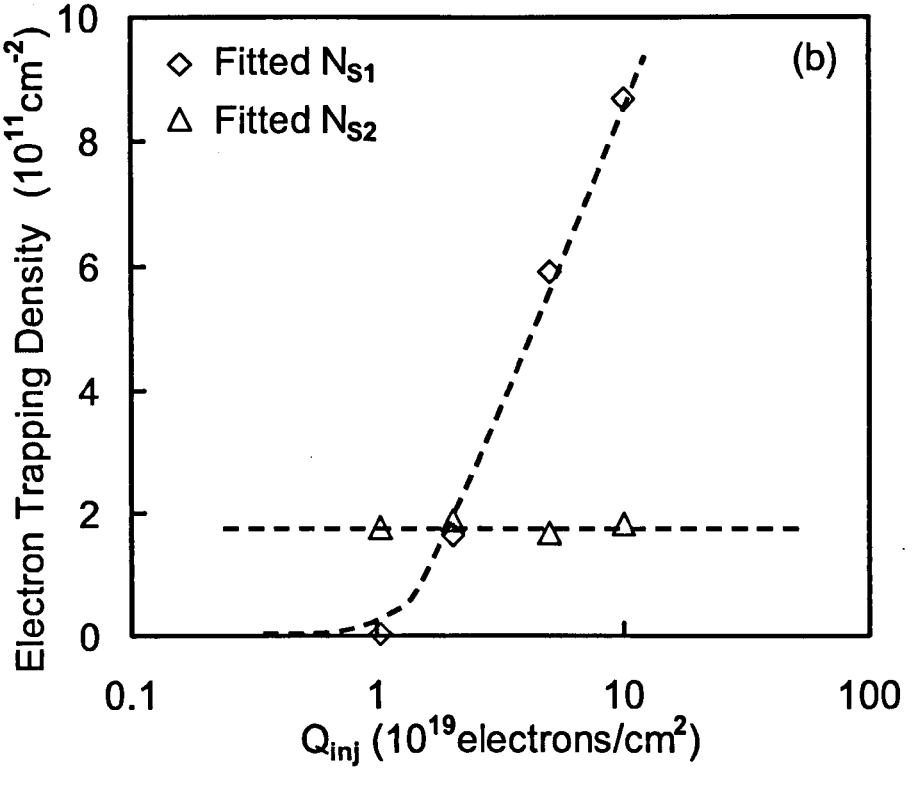


Figure 13. (a) Typical experiment sequence of trapping kinetics analyses. A device was stressed under +11 MV/cm (Q_{inj} = 1×10^{20} carriers/cm^2, symbol 'o'), then switched to +8 MV/cm (symbol 'x'). Followed by 1 sec of SHI to empty the filled electron traps (Q_h = 7×10^{13} holes/cm^2, symbol '□') then refilled again (symbol '◇'). The negative gate bias (symbol '+') and subsequence electron injection (symbol '△') will allow us to correct ANPC offsetting effects on trapping kinetics. (b) The 3rd filling (symbol '△') reveals a non-negligible amount of ANPC were created during the FNI stress. As a result, the 2nd filling (symbol '◇') cannot be used to study the electron trapping directly, the effects of ANPC on trappings kinetics need to be corrected using equation (1) & (2). (c) Symbol '●' is the trappings kinetics after ANPC correction.


(a)
Q_{inj}
1x10^20
5x10^19
2x10^19
1x10^19
N_{ec} (10^{11}cm^{-2})
Q_e (10^{15}electrons/cm^2)


(b)
◇ Fitted N_{S1}
△ Fitted N_{S2}
Electron Trapping Density (10^{11}cm^{-2})
Q_{inj} (10^{19}electrons/cm^2)

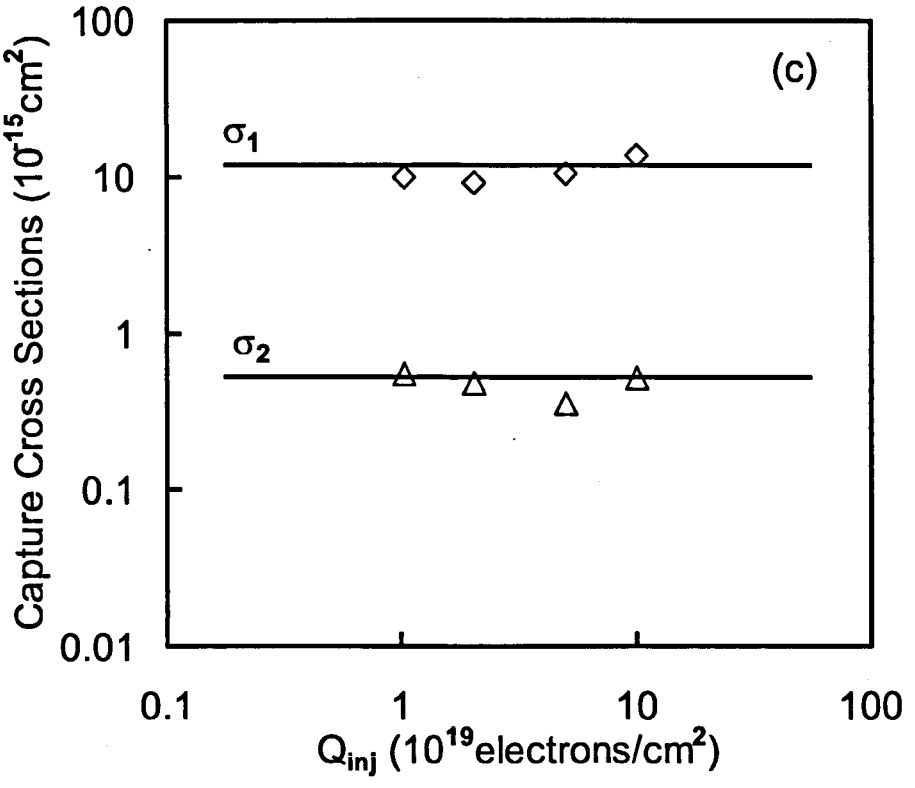


Figure 14. (a) Effects of ANPC on trappings kinetics were systematically corrected using equation (1) & (2) for each trapping curve from $Q_{inj} = 1\times10^{19}$ to 1×10^{20} carriers/cm^2. The curves are nearly a parallel upward shift as stresses increase. (b) Similarly to SHI induced electron traps, FNI induce two different electron traps with distinctively different density behavior as stresses increase. The effective density of the smaller trap, N_{S2}, saturates, while the larger trap, N_{S1}, does not. (c) The two extracted capture cross sections are insensitive to all stress levels, with the larger trap (σ_1: $10^{-13} \sim 10^{-14}$ cm^2) and the smaller trap (σ_2: $10^{-15} \sim 10^{-16}$ cm^2).

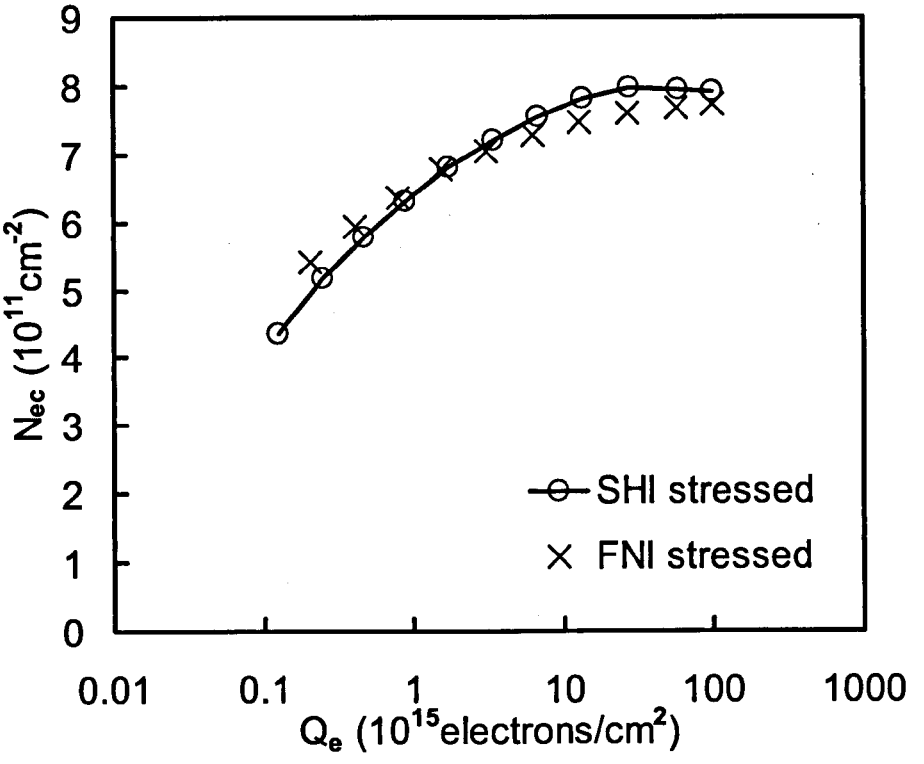


Figure 15. A comparison of the dynamic filling behaviour. Symbol 'x' represents traps created by FNI, symbol 'o' are traps created by SHI. It is apparent that the trapping kinetics is similar in these two cases. This strongly suggest that the same types of traps are generated by FNI and SHI. Traps generated by hydrogenous species and holes have the same electrical signature.

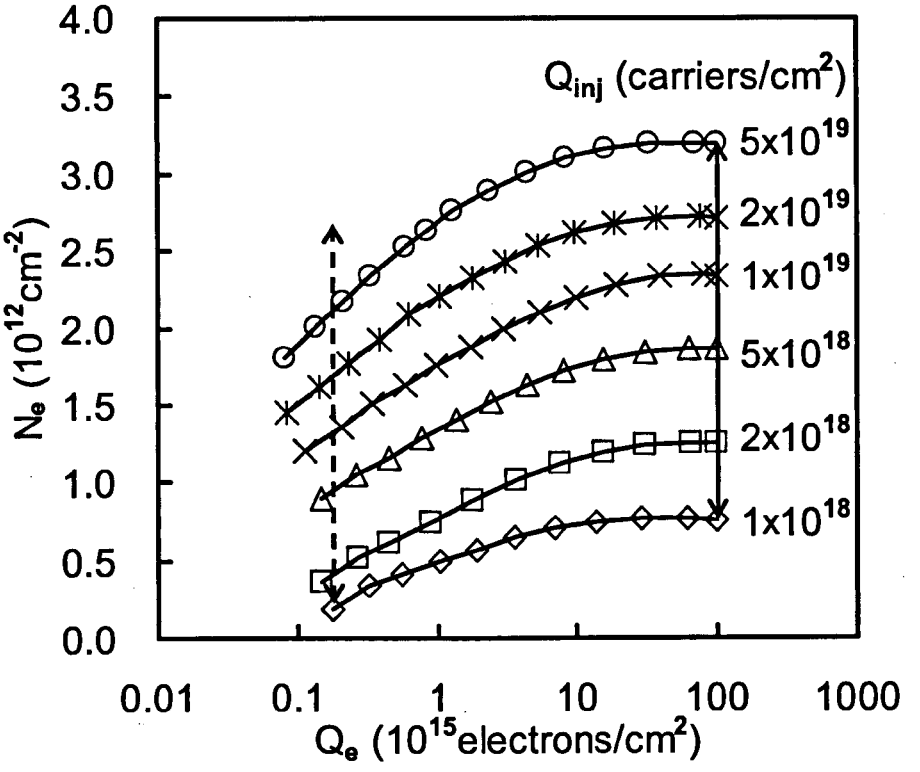
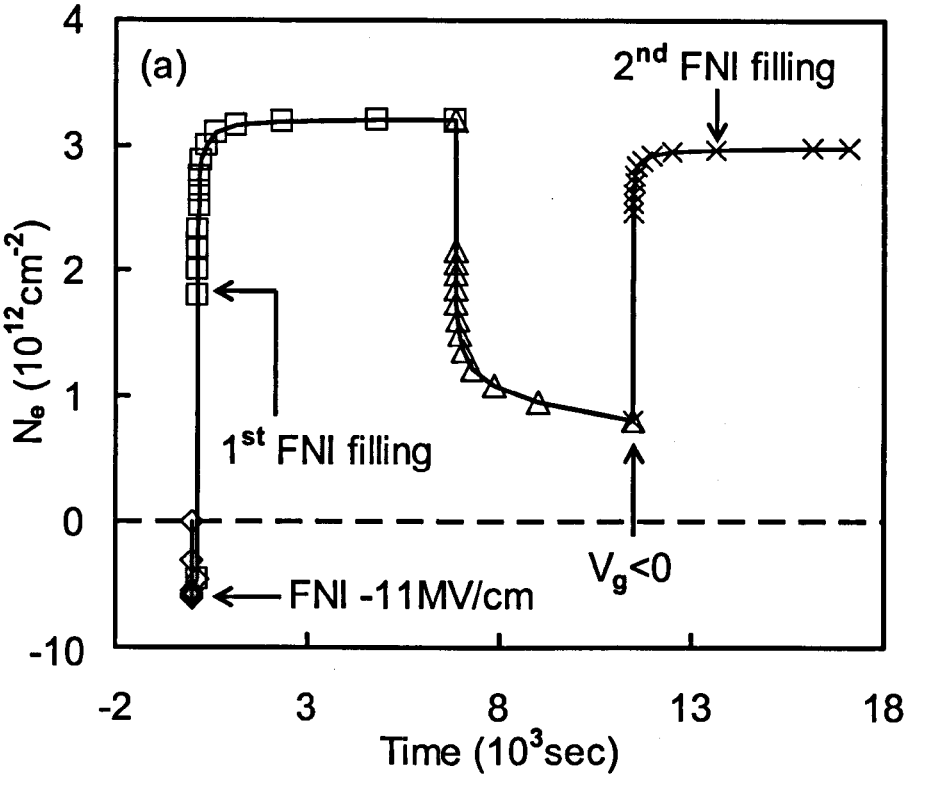
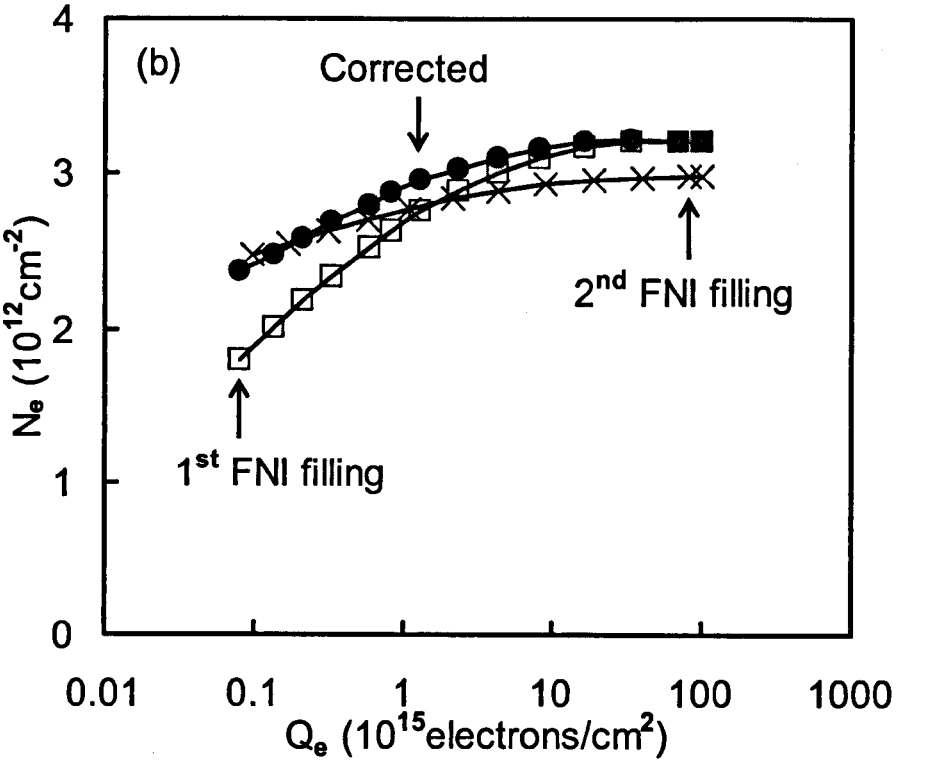


Figure 16. Generated electron traps after stressing a 13.8 nm oxide under -11 MV/cm. The solid arrow represents the increase of trapping at high Q_e, while the dashed arrow is a shift of the solid arrow to lower Q_e. It is apparent that the increase of trapping at high Q_e is larger than the increase at relatively low Q_e, indicating the creation of ANPC.


(a)
2nd FNI filling
1st FNI filling
FNI -11MV/cm
Vg<0
Ne (10^12cm^-2)
Time (10^3sec)
4
3
2
1
0
-10
-2
3
8
13
18


(b)
Corrected
2nd FNI filling
1st FNI filling
Ne (10^12cm^-2)
Qe (10^15electrons/cm^2)
4
3
2
1
0
0.01
0.1
1
10
100
1000

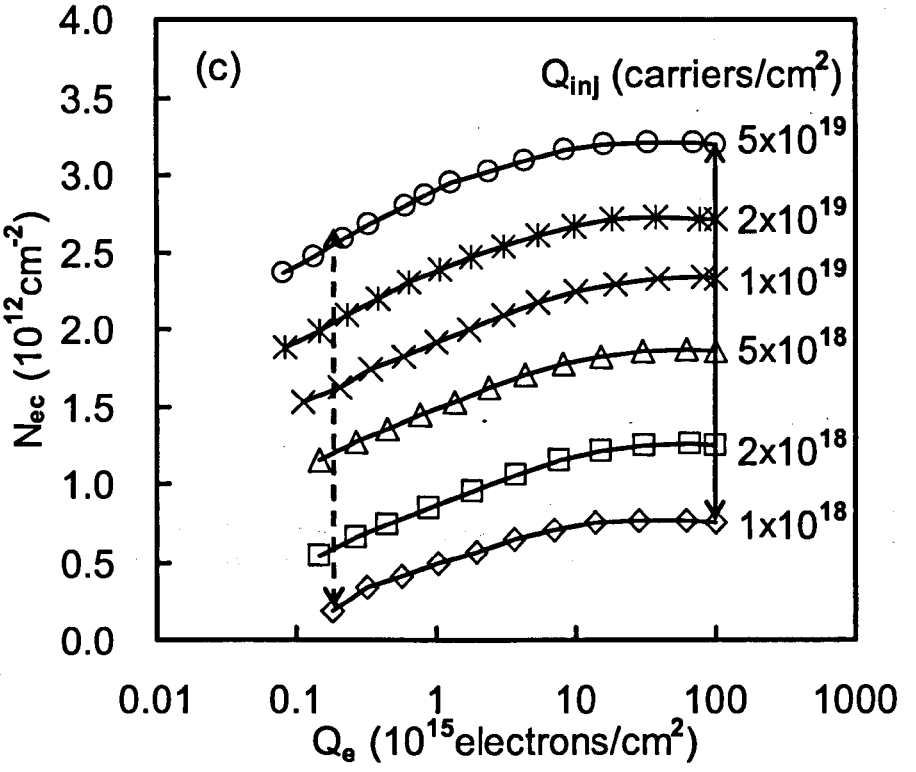
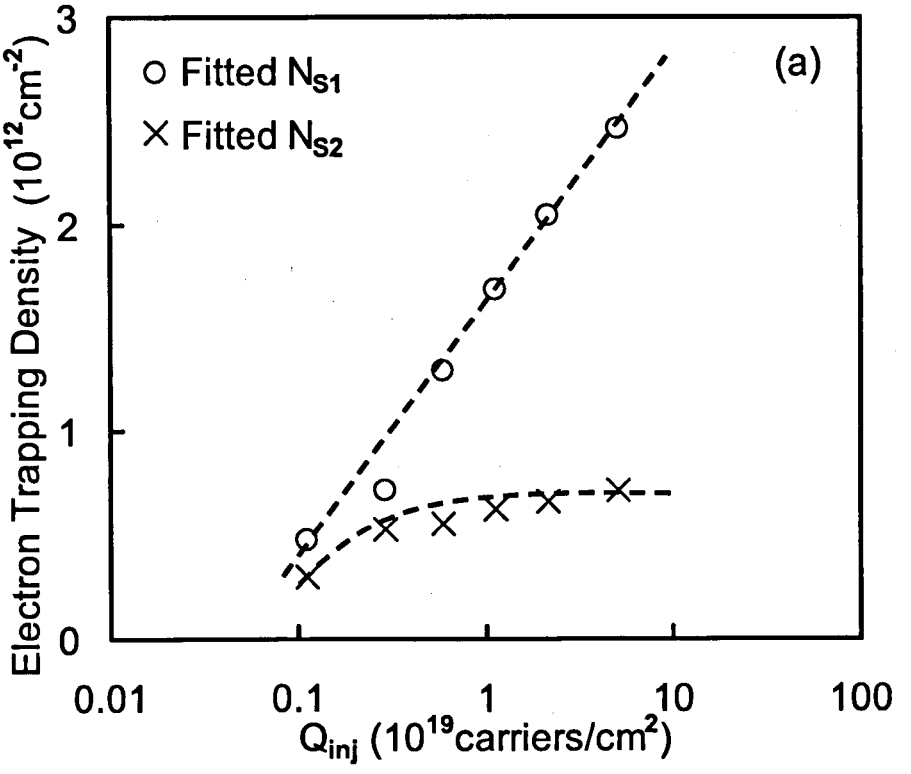


Figure 17. (a) The experiment sequence to determine ANPC effects on electron trapping kinetics. A negative bias was applied (symbol '△') after the 1st FNI filling. At the end of $V_g < 0$, electron injection was switched on again to re-neutralize generated ANPC (symbol 'x'). (b) Symbol '●' is after correcting ANPC offsetting effects on electron trapping kinetics. (c) After correction, the increase of trapping is similar over of the whole range of Q_e (the dashed arrow is a shift of the solid arrow to lower Q_e).


Fitted N_{S1}
Fitted N_{S2}
(a)
Electron Trapping Density (10^{12}cm^{-2})
0
1
2
3
0.01
0.1
1
10
100
Q_{inj} (10^{19}carriers/cm^2)

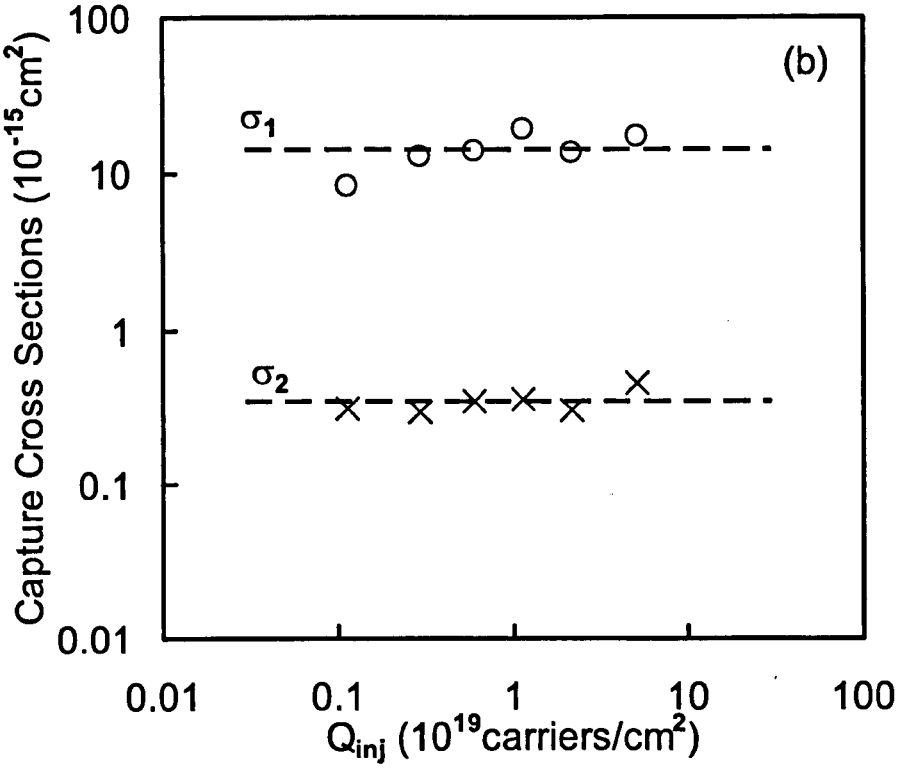

(b)
σ_1
σ_2
Capture Cross Sections (10^{-15}cm^2)
0.01
0.1
1
10
100
0.01
0.1
1
10
100
Q_{inj} (10^{19}carriers/cm^2)

Figure 18. (a). As stresses increase, the effective density of the larger trap (σ_1: $10^{-13} \sim 10^{-14}$ cm^2), N_{S1}, increases continuously, while the effective density of the smaller trap (σ_2: $10^{-15} \sim 10^{-16}$ cm^2), N_{S2}, tends to saturate. Different generation tendencies are suggesting that two different types of electron traps were created. (b) The two extracted capture cross sections are insensitive to stress levels, which again support the existence of two different electron traps, similarly as on 7.1 nm oxides.

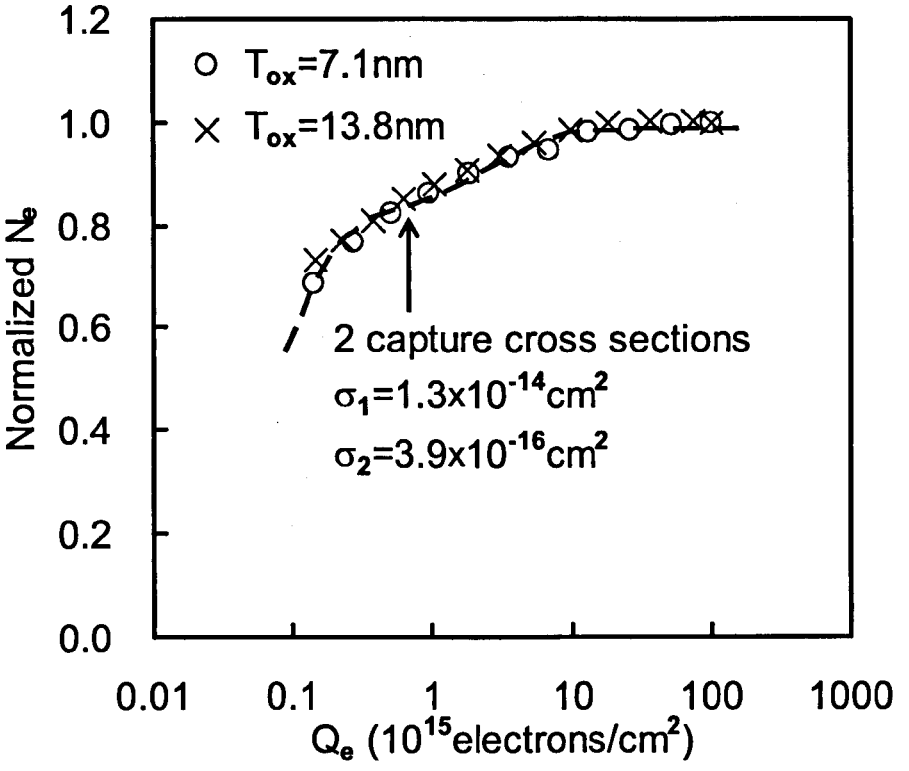


Figure 19. A comparison of electron traps generated in oxides of different thickness. N_e was normalized against its saturation value. The two sets of data can be fitted with the same capture cross sections (dashed line). This strongly supports that the same traps were generated in SiO_2 of different thickness.

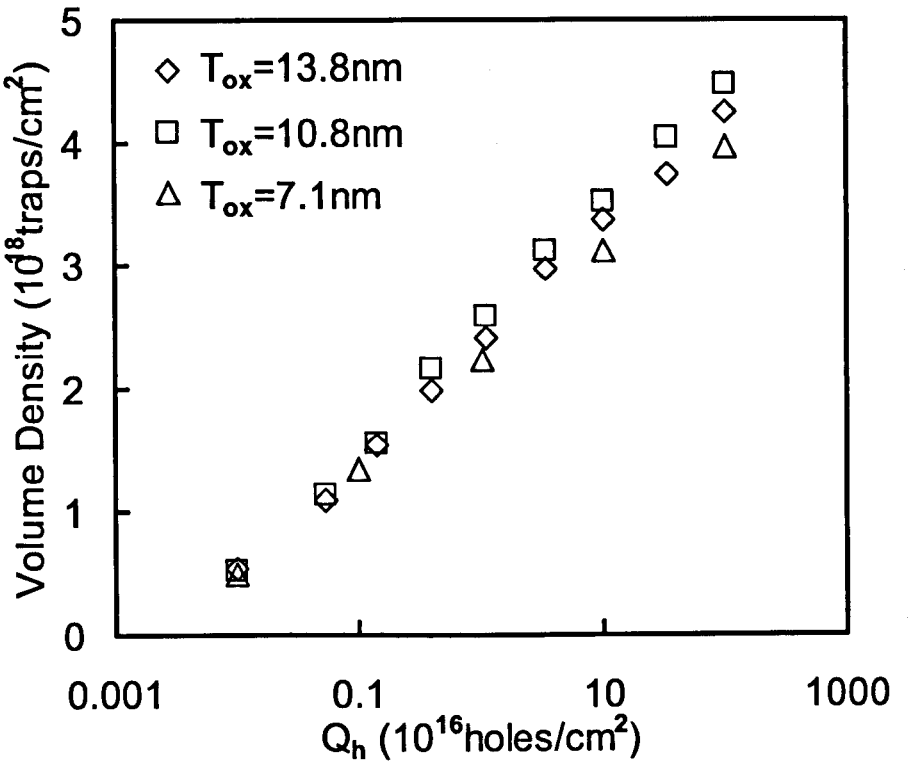


Figure 20. Electron trap generation per volume density as a function of Q_h. The volume density is clearly insensitive to oxide thickness ranging from 7.1 nm to 13.8 nm. This indicates a 'uniform' distribution of generated defects, and supports the percolation model for the oxide breakdown.

5 Negative Bias Temperature Instability

5.1. Introduction

Positive charge formation in gate oxides is one of the earliest instabilities identified for MOS devices [1-8]. For the sub 3 nm oxide used in modern MOSFETs, it together with the generation of interface states, is responsible for the negative bias temperature instability (NBTI) [9-12]. As the oxide becomes thinner, NBTI and positive charges become increasingly important, mainly because of two reasons. First, for each new generation of CMOS process, both operation temperature and electrical field increases. Second, to suppress boron penetration and increase the dielectric constant, the nitrogen density in gate oxide is increasing rapidly. One adverse effect of this increase in nitrogen density is a substantial enhancement of NBTI and positive charge formation [11-17].

The NBTI is so severe now that it limits the lifetime of pMOSFETs and becomes a pressing issue for the current CMOS industry [11,12,15-18]. This has motivated intensive research on NBTI recently [11,12,15-18]. Despite these efforts, there are a number of issues remaining to be solved.

One of these issues is the effect of measurement temperature on NBTI. Traditionally, after stressed at elevated temperature, the device was cooled down and defects were assessed at room temperature (RT) [3,19]. Recently, NBTI has often been measured at stress temperature [16,20-22]. One advantage for measuring at stress temperature is that it removes the cooling period and reduces the delay between the stress and measurement. This delay can decrease NBTI considerably. For example, it was reported that, when compared with the traditional transfer characteristics (TC) method [23,24], threshold voltage shift, ΔV_t, could be one order of magnitude higher

by using the 'On-The-Fly' measurement, where there was no delay between stress and measurement [20]. The ΔV_t measured at different temperatures is often used to extract the activation energy of defect generation, but how the measurement temperature affecting ΔV_t was not addressed. New experimental findings will be presented in this chapter to clarify the effect of measurement temperature on NBTI.

Additionally, for the positive charges formed during negative bias temperature stress (NBTS), it is not known whether all positive charges are the same or more than one type of positive charges are created. If more than one type of positive charges are created, will they have the same dependence on NBTS conditions, such as stress time and temperature? Moreover, we do not know if the positive charge formed during NBTS is the same as the positive charge created during other types of electrical stresses, such as the substrate hole injection [25-28]. Such issues will be addressed.

5.2. Investigation of NBTI using traditional measurement method

In this section, the traditional transfer characteristics (TC) method is used to monitor device degradation. The stress temperature is in the range of room temperature to 200°C, while the measurement temperature can be either the stress temperature or at some value between the room and stress temperature. The typical gate bias is around -3.17 V. The test follows the well-know 'stress-then-sense' procedure [29]. The NBTS was interrupted at pre-specified time, to monitor the positive charges and interface states. The interface states density is measured by using the subthreshold swing technique [30-32], as shown in section 2.4.1.2. Positive charges were measured from the gate voltage shift at the midgap point, where the contribution from the interface states is negligible [7,33,34], with more details in section 2.4.2.2. In this work, we follow the well-accepted practice by assuming that the charge centroid is at the dielectric/substrate interface [25-28].

5.2.1. Typical NBTI generation

Early works [1-3,12,19] reported at least three common features for NBTI. First, NBTI originates from two types of defects: generated interface states and positive charges [1-3,12,19]. Second, the build-up of NBTI follows a power law against stress time at a given temperature [3,9,12,35]. Most of reported power factors are within a narrow range between 0.2 and 0.3. Third, the number of positive charges was found to be approximately the same as the number of created interface states [3,10,19,36,37].

Before reporting new results and findings, the features mentioned are checked first for our samples. Figure 1 shows that both positive charges (symbol '○') and interface states (symbol 'x') were generated under a gate bias of -3.17 V and a temperature of 100°C. Their build-up follows the expected power law and a power factor of 0.26 can be extracted by fitting the generated positive charge. The density of effective positive charges, ΔN_{ot}, obviously agrees well with that of generated interface states, ΔD_{it}. As a result, the NBTI observed here has all the features reported earlier.

5.2.2. Types of positive charges

Among the generated positive charges, are all positive charges the same or more than one type of positive charges created? To answer this question, a device was first subjected to NBTS at 100°C and the symbol '◇' in Figure 2(a) shows the build-up of positive charges. After a pre-specified time, the NBTS was stopped and 1.7×10^{17} cm^{-2} of electrons were injected into the oxide under a positive gate bias (+6.5 MV/cm, symbol '○') in an attempt to neutralize the positive charges. It is clear that substantial amount of positive charges survived the neutralization. When a negative and positive gate bias ($|E_{ox}|$ = 5 MV/cm) was alternately applied with all other terminals grounded, part of positive charges could be repeatedly charged under $V_g < 0$ and discharged under $V_g > 0$. They will be referred to as 'cyclic positive charges (CPC)', as indicated in Figure 2(a). CPC has similar charging and discharging rates.

Figure 2(a) also shows that some positive charges are more difficult to neutralize and they will be referred to as 'anti-neutralization positive charges (ANPC)'. As a result, we can conclude that different types of positive charges are formed under NBTS.

To facilitate the comparison of positive charges formed under NBTS with that under other electrical stresses, Figure 2(b) shows the positive charge formed by the substrate hole injection (SHI) [27]. Here, the test sequence is similar to that in Figure 2(a), but the stress was at room temperature. Holes were supplied by forward biasing the pn junction underneath a pMOSFET, accelerated in the space charge region, and injected into the oxide [25-28]. The trapped positive charges were neutralized by an electron injection under E_{ox} = +8 MV/cm. When a negative and positive gate bias (|Eox| = 5 MV/cm) was alternately applied with all other terminals grounded, the behavior of the SHI induced positive charge is similar to that formed during NBTS. Both ANPC and CPC can be clearly seen here as well.

Figure 2(a) & (b) show that the level of positive charging achieved under $V_g < 0$ is well short of that reached during the NBTS or SHI. This means that, after neutralization, some defects cannot be positively recharged under $V_g < 0$ without hole injection. In reference 27, it has been shown that they were as-grown hole traps, while both ANPC and CPC was the generated hole trap.

In summary, the behavior of positive charges formed under NBTS is similar to that induced by SHI. This similarity suggests that positive charges formed under different stresses have common origins, which are determined by the nature of materials used in the testing sample. Apart from as-grown hole traps, two different types of positive charges, ANPC and CPC, can be generated. After neutralization, as-grown hole traps cannot be positively charged under $V_g < 0$ without hole injection, while both ANPC and CPC can [27,28]. Under the same magnitude of oxide field, the charging rate of CPC under $V_g < 0$ is similar to the discharging rate under $V_g > 0$. In contrast, the neutralization of ANPC is more difficult. More results will be given in the following sections to support the separation of ANPC from CPC.

5.2.3. Effects of experimental parameters on generated positive charges

5.2.3.1. Stress time

Figure 3 shows two devices subjected to NBTS at 200°C for different stress time, with symbol '●' stressed for 760 sec, and symbol 'o' stressed for only 10 sec. After stresses, both devices were cooled down to room temperature. Some charges could be lost during this cooling period of 4 min. A short NBTS was resumed at room temperature in an attempt to regain some of the lost charges. Afterwards, a negative and positive gate bias ($|E_{ox}|$ = 5 MV/cm) was alternately applied with all other terminals grounded. Firstly, the amount of CPC generated on both devices is similar, as pointed out by arrows labeled A. This suggests that CPC is pre-determined during the device fabrication process. Once all the defects are consummated, CPC would no longer respond to stress time. Secondly, ANPC generation is clearly stress time dependent, as pointed out by arrows labeled B. Longer stress time leads to greater generation of ANPC, in contrast with CPC.

Figure 4(a) shows the dependence of ANPC and CPC on the NBTS time at 100°C. Although CPC can be higher than ANPC initially, ANPC overtakes CPC at longer stress time. It is obvious that ANPC increases continuously, but CPC saturates as the stress time becomes longer. This confirms that there is only limited number of defects convertible into CPC, while such limitation was not observed for ANPC. ANPC and CPC must originate from different types of defects, therefore. Although Figure 1 shows that the density of total positive charges, ΔN_{ot}, increases continuously with NBTS time, one components of ΔN_{ot}, CPC, saturates.

As a comparison, Figure 4(b) shows the dependence of ANPC and CPC on the hole fluency, Q_h, during SHI. Here, an increase of stress time leads to higher hole fluency. The CPC saturates again, but ANPC does not. The similar dependence on stress levels supports our early suggestion that positive charges formed under different stresses have common origins.

In Figure 1, the build-up of positive charges clearly does not saturate, and follows a power factor of 0.26 against stress time. Among the generated positive charges, ANPC is the one with continuous generation characteristics, as shown in Figure 4(a). If ANPC is the main component responsible for positive charges generated during NBTS, one would expect its generation also follows a power factor within a range between 0.2 and 0.3. In Figure 5, ANPC generation is plotted in a log-log scale. After fitting the data, a power factor of 0.25 can be extracted. The similar power factor suggests that the build-up of positive charges during a long NBTS (> 100 sec) is dominated by ANPC generation.

5.2.3.2. Thermal stability

Next, we compared the annealing behavior of ANPC and CPC at different temperatures. After NBTS at 100°C, the device was cooled down, and ANPC and CPC were measured at room temperature. Afterwards, the device was exposed to elevated temperatures from 100°C to 200°C with all terminals floating. After the exposure, the device was again cooled down, so the annealing behavior of ANPC and CPC can be monitored at room temperature. Figure 6(a) shows that CPC remain stable throughout the annealing, while ANPC reduced considerably after temperature increased. In Figure 6(b), the annealing behavior of ANPC generated under NBTS is compared with ANPC generated under SHI. In both case, ANPC was normalized against initial value before annealing. The similar dependence of ANPC on annealing temperature further supports our early suggestion that positive charges formed under different stresses have common origins.

5.2.3.3. Stress temperature

Figure 7(a), (b), and (c) show the dependence of ΔN_{ot}, ANPC and CPC on stress temperature, respectively. The measurement temperature was the same as the stress temperature. Once again, CPC will saturate and the saturation level is insensitive to stress temperature. This reinforces our statement that the level of CPC is limited by

the number of defects available. However, both ANPC and the total charge density, ΔN_{ot}, behave differently.

When the temperature increases from 100°C to 150°C, ΔN_{ot} and ANPC rise substantially, because the NBTI is a thermally accelerated process. Based on this, one would expect that there should be a clear increase of ΔN_{ot} and ANPC, when the temperature rose from 25°C to 100°C. However, Figure 7(a) & (b) show that they are actually insensitive to temperature in this range. Unlike CPC, this insensitivity to temperature cannot be explained by the limitation of defects, since both ΔN_{ot} and ANPC can rise further for higher temperature (e.g. 150°C). This issue will be examined in details in the next section.

5.2.4. Effects of measurement temperature

5.2.4.1. On ANPC and CPC

Previous work at this university [27,28] showed that ANPC generated by SHI had an energy level above the bottom edge of silicon conduction band, when positively charged. This makes its neutralization a thermally activated process, as illustrated in Figure 8. A higher temperature will increase the number of electrons in the higher level of silicon conduction band, allowing them to reach and recombine with the positive ANPC. In contrast, the positive CPC has an energy level close to the bottom edge of silicon conduction band, so that its neutralization is not sensitive to temperature, as shown in Figure 9.

In previous sections, we have proposed that the ANPC created by NBTS is the same as that generated by SHI. Consequently, the neutralization of NBTS induced ANPC should be a thermally activated process as well. This is confirmed by Figure 10. After NBTS at 100°C, ANPC was first measured at the stress temperature. We then increased the temperature to 150°C. As expected, the measured ANPC is clearly

reduced because of the increased neutralization at 150°C. Similar to the SHI induced CPC, the NBTS created CPC is also insensitive to the measurement temperature.

To further confirm that the ANPC reduction at higher measurement temperature is indeed caused by measurement temperature effect, but not from ANPC annealing effect, the following experiment was carried out. A device was heavily stressed under NBTS at 200°C for 500 sec. A large amount of ANPC was created. After the stress, the device was cooled down to room temperature, ANPC and CPC were measured, symbol '○' in Figure 11(a). The measurement temperature was then raised to 150°C, ANPC and CPC were measured again (symbol '□'). As expected, CPC is insensitive to measurement temperature, but ANPC reduces substantially as measurement temperature increases. Figure 11(b) shows that when measurement temperature was lowered to 100°C, larger ANPC was measured compared to 150°C. More and more ANPC could be measured as measurement temperature drops further. This confirms that ANPC is highly dependent on the measurement temperature. Lower measurement temperature will reveal larger amount of ANPC.

The reduction of ANPC at higher measurement temperature could explain the 'insensitivity' of ANPC to the temperature in the range of 25°C to 100°C observed in Figure 7(b). As commonly used in NBTS tests [21,38,39], measurements in Figure 7(b) were carried out at the stress temperature. An increase of temperature from 25°C to 100°C has two opposite effects on ANPC. On one hand, it enhances the generation of ANPC. On the other hand, it accelerates its neutralization. The 'insensitivity' of ANPC to temperature between 25°C and 100°C result from the balance of these two competing processes.

5.2.4.2. On effective charge density and interface states

If the above explanation is correct, for a given number of defects, one can predict that more positive charges will be measured at lower temperature. To testify this prediction, one device was stressed at 200°C. The effective charge density, ΔN_{ot}, was measured first at 200°C and then at 25°C. The results are presented by the 1st two '○'

symbols from the left of Figure 12. Although a small increase can be seen when measurement temperature drops from 200°C to 25°C, the change is hardly significant, when compared with that suggested by Figure 11(b), where measurement temperature only varied between 25°C and 150°C. Indeed, one may say that the ΔN_{ot} is again insensitive to measurement temperature here.

To explain this newly found insensitivity, we measured the ΔN_{ot} when the measurement temperature was increased from 25°C to 200°C. Figure 12 clearly shows that there is a monotonic reduction of ΔN_{ot} as measurement temperature increases. When the measurement temperature returned to 200°C, the ΔN_{ot} became obviously smaller than that recorded at the same temperature earlier (the 1st 'o' from the left), as pointed out by the arrow. There is little doubt that some positive charges were annealed out, as shown in Figure 6(a). The question is how much the decrease of ΔN_{ot} in Figure 12 is caused by annealing and how much of it originates from the raise of measurement temperature.

To answer the above question, an annealing experiment was carried out and the result is given in Figure 13(a) and (b). Here, a device was first stressed at 200°C under a gate bias of V_g = -3.17 V. The V_g was then set to zero and the annealing was monitored against time at 200°C. As expected, the ΔN_{ot} was indeed unstable and dropped rapidly at the beginning of V_g = 0 V period. As the time increases beyond 20 sec, however, Figure 13(a) shows that ΔN_{ot} does not appear dropping further. To confirm that the ΔN_{ot} is indeed steady now, ΔN_{ot} is also plotted against logarithmic time in Figure 13(b). As one can see, ΔN_{ot} is stabilized at longer time.

The annealing behavior shown in Figure 13(a) and (b) can be used to explain the anomalous temperature dependence in Figure 12. After stressed and measured at 200°C, the sample was at 200°C without bias for at least 30 sec, for lifting probes and taking the sample off the 200°C probe station. This time is longer than that needed for the annealing process, as shown in Figure 13(a) and (b). As a result, at 25°C, we only measured the thermally stable part of ΔN_{ot}. It happens that an increase of the stable part of ΔN_{ot} at 25°C largely compensates the annealing effects. As a

result, the 'insensitivity' to temperature when it dropped from 200°C to 25°C in Figure 12 is an artifact.

According to the above explanation, the number of defects responsible for positive charges has been stabilized from the 2nd measurement point onward in Figure 12. As temperature rises from 25°C back to 200°C, the reduction of ΔN_{ot} must originate from the effects of measurement temperature. To ensure that further annealing is negligible, we measured ΔN_{ot} at 25°C again after the temperature returned to 200°C. Figure 12 confirms that the ΔN_{ot} is essentially the same as that measured for the 1st time at 25°C.

The annealing behavior of interface states, ΔD_{it}, is shown in Figure 13(c). It is apparent that little annealing of ΔD_{it} occurred during the $V_g = 0$ V period, similar to results reported by Huard et al [17]. This explains why similar ΔD_{it} was observed in Figure 12. In Figure 14, generated defect were normalized against their maximum value, with sympol 'o' represents the generation of positive charges, and symbol '□' represents the generation of interface states. It can be seen that when V_g was set to zero, the annealing of interface states (symbol '+') is significantly less than that of positive charges (symbol 'x'). The annealing is mainly dominated by positive charges. Once stabilized, Figure 12 shows that ΔN_{ot} is highly sensitive to measurement temperature, but ΔD_{it} is not.

5.2.5. Dynamic NBTI

The dynamic behavior of defects created at different temperatures, but measured at the same temperature is compared next.

5.2.5.1. The same NBTS time

Two devices were subjected to NBTS for 13 hours: one at 25°C and the other one at 100°C, as shown in Figure 15(a). Similar to Figure 7(a), ΔN_{ot} appears insensitive to

temperature within this range. After NBTS, the device stressed at elevated temperature was allowed to cool down to room temperature. Then, both devices went through a cycle of positive and negative gate biases ($|E_{ox}|$ = 5 MV/cm) with all other terminals grounded. Unlike Figure 7(b), where ANPC was measured at stress temperature, the arrow 'A' in Figure 15(b) points out significant increase of ANPC generation after NBTS at 100°C. This result suggests that ANPC is a thermally accelerated component generated under NBTS, even within the stress temperature range of 25°C to 100°C.

It has been shown earlier that ΔN_{ot} during NBTS has three components: as-grown hole traps, CPC and ANPC. ANPC is the only thermally accelerated component, and it is expected that the increase of ANPC is also reflected on ΔN_{ot}. However, Figure 15(b) shows that due to the measurement temperature effect, the difference in ΔN_{ot} (arrow 'B') is obviously smaller than the increase of ANPC (arrow 'A'). For a given number of defects, Figure 12 showed that less ΔN_{ot} could be measured at higher temperature. If one assumes that the ratio between ΔN_{ot} measured at 25°C and 100°C is a constant, regardless of the generation level, one could use this ratio to correct ΔN_{ot} measured at 100°C. In Figure 15(c), every data measured at 100°C were corrected to 25°C using a fixed ratio of 1.3 obtained from Figure 12. It is apparent that after correction, the difference in ΔN_{ot} (arrow 'B') is similar to the increase of ANPC (arrow 'A'). This further supports that ANPC is the only thermally enhanced component.

5.2.5.2. The same positive charge generation

One may not be convinced that the increase of ANPC (arrow 'A') observed in Figure 15(b) is entirely thermally accelerated, as larger ΔN_{ot} was measured at the end of NBTS at 100°C (arrow 'B'). Another experiment was carried out, where NBTS at 100°C was interrupted as soon as ΔN_{ot} reached the level of NBTS at 25°C. Again, when measured at stress temperature, ΔN_{ot} is 'insensitive' within this temperature range, as shown in Figure 16(a). After NBTS, the device stressed at elevated temperature was allowed to cool down to room temperature. Then, both devices went

through a cycle of positive and negative gate biases ($|E_{ox}|$ = 5 MV/cm) with all other terminals grounded. Despite having shorter NBTS time at 100°C, the arrow 'A' in Figure 16(b) clearly indicates an obvious increase of ANPC post 100°C NBTS, while ΔN_{ot} were similar at the end of NBTS. Similarly to Figure 15(c), every data measured at 100°C in Figure 16(b) were corrected to 25°C using a fixed ratio of 1.3. The result is given in Figure 16(c). Again, the increase of ANPC (arrow 'A') is approximately the same as the difference in ΔN_{ot} after correction (arrow 'B'). This strongly supports previous claim that only ANPC generation is thermally enhanced.

5.2.6. Implications to NBTI tests

We now explore the implication of our results to NBTI tests. Five points are worth of mentioning. Firstly, although positive charges recorded at stress and room temperature may appear similar, Figure 12 shows that it is misleading for suggesting that positive charges are insensitive to measurement temperature.

Secondly, the real density of defects responsible for positive charges can be underestimated at either stress [21,38,39] or room temperature [3,19,35,37,40]. At stress temperature, a higher temperature always enhances the neutralization of ANPC.

Thirdly, based on the 2nd point, any thermal activation energy extracted from NBTS at different temperatures will not result from a single physical process. Care should be exercised on its interpretation.

Fourthly, it was reported that approximately the same number of positive charges and interface states were generated by NBTS. This is also observed in Figure 1. However, ΔN_{ot} is clearly sensitive to measurement temperature, while the symbol 'x' in Figure 12 shows that ΔD_{it} is not. Although it is difficult for believing the good agreement between ΔN_{ot} and ΔD_{it} shown in Figure 1 is a coincidence, present results at least suggest that it is questionable.

Finally, for a given number of defects, since positive charges change with temperature, it introduces a temperature-dependent instability of threshold voltage for MOSFETs.

5.2.7. Conclusions

In this section, positive charges formed under negative bias temperature stress (NBTS) are investigated. Efforts were made to identify the types of positive charges and their dependence on stress conditions. It is found that NBTS could generate both anti-neutralization positive charges (ANPC) and cyclic positive charges (CPC). While CPC has similar charging and discharging rates, ANPC is more difficult to neutralize. As stress time or temperature increases, CPC saturates, but ANPC does not. This suggests that they have different origins. Both NBTS induced CPC and ANPC behave similarly to those created by substrate hole injection.

Effects of measurement temperature on positive charges were studied in details. It is found that, for a given number of defects, an increase of temperature leads to lower positive charges. Although CPC is insensitive to the measurement temperature, higher temperature enhances the neutralization of ANPC. Parts of NBTS induced positive charges are thermally unstable and can be annealed at stress temperature. When cooled for measurement at room temperature, some positive charges were lost. As a result, the density of defects responsible for positive charges is generally underestimated when measured at either stress or room temperature. The implication to NBTI tests is explored. For a given number of defects, the variation of positive charges with temperature results in a temperature-dependent instability of threshold voltage.

5.3. Investigation of NBTI using 'On-The-Fly' measurement method

Despite the fact that, recently, NBTI has often been measured at stress temperature [21,38,39] to remove the cooling period and reduces the delay between the stress and measurement, it has been shown that the standard ΔV_t measurement leads to a large recovery, proportional to the time spent between the stress and measurement [21]. This effect reduces the relevance of ΔV_t as a parameter to quantify the degradation. In this context, a new 'On-The-Fly' measurement technique has been proposed [20] to avoid the recovery between the stress and measurement.

In this section, the 'On-The-Fly' measurement method is used to monitor device degradation. The stress temperature is in the range of room temperature to 200°C, while the measurement temperature can be either the stress temperature or room temperature. The typical gate bias is -3.17 V. ΔV_t is obtained by monitoring the drain current degradation and transconductance variation [20]. Further information about 'On-The-Fly' method can be found in section 2.4.3.

5.3.1. Main differences between 'On-The-Fly' and traditional measurements

Three important differences have been observed:

(i) It has been reported that with the 'On-The-Fly' measurement method, ΔV_t can be over one order of magnitude higher than the ΔV_t by the traditional transfer characteristics (TC) method [20]. This phenomenon is also observed on the current testing sample. Figure 17 shows that the ΔV_t measured by the 'On-The-Fly' method is indeed over one order of magnitude higher than the ΔV_t measured by the traditional TC method, resulting from the elimination of NBTI recovery.

(ii) When temperature changes from room temperature to 200°C. The ΔV_t measured by the 'On-The-Fly' method only changes modestly with temperature, as shown in Figure 18. This is in agreement with the small activation energy (0.03 eV) obtained

in early work [20]. Consequently, this can easily lead to the assumption that the defect generation during NBTS is insensitive to temperature, which is against the results obtained by using the traditional method. In the following sections, it will be shown that this assumption is not correct.

(iii) The time dependence of degradation is given in Figure 19. After eliminating the recovery between stress and measurement, ΔV_t no longer follows the often quoted power factor of 0.25. Now the power factor varies from 0.48 to 0.17 between the measurement range of 2 ms to 1000 sec. This power factor variation has also been reported by other researchers [41,42]. It is suggested that the power factor of 0.25 was obtained with measurement interruptions. Consequently, the degradation is the net result of defect generation and recovery, and do not represent the whole degradation process.

5.3.2. Estimation of NBTI induced effective mobility variation

When compared with the ΔV_t measured by the traditional method, the one order of magnitude increase of ΔV_t measured by the 'On-The-Fly' method is surprisingly large. Care must be exercised to ensure that this large increase is not an artifact.

Apart from threshold voltage shift, stresses will also induce a degradation of low field carrier mobility [43]. For the traditional technique, ΔV_t was monitored at the threshold condition and the number of holes in the p-channel is negligible. The impact of mobility variation on ΔV_t is insignificant. For the 'On-The-Fly' technique, the ΔV_t was monitored at the stress voltage. The pMOSFET was in strong inversion when stressed at a voltage such as V_g = -3.17 V and there were a large amount of holes in the p-channel. The impact of mobility variation on ΔV_t can be substantial. From the equations given in the section 2.4.3, it is not explicit whether the effect of mobility variation on ΔV_t has been taken into account for the 'On-The-Fly' technique. This uncertainty is addressed in the following.

For the 'On-The-Fly' measurement, V_d was set at -0.025 V and $|V_d| << |V_g - V_t|$. This allows the use of the following equation:

$$I_d = A \cdot \mu(V_g, V_t, t) \cdot (V_g - V_t(t)) \tag{1}$$

Where $A = C_{ox} \cdot \frac{W}{L} \cdot V_d$

From equation (1), one can obtain,

$$\frac{dI_d}{dt} = A \cdot (V_g - V_t) \cdot \frac{d\mu}{dt} + A \cdot \mu \cdot \frac{d(V_g - V_t)}{dt} \tag{2}$$

For the 'On-The-Fly' technique, the time is divided into small steps. In this sense, it mimics the typical numerical simulation.

With numerical approximation, one can have,

$$\frac{\Delta I_d}{\Delta t} \approx A \cdot (V_g - V_t) \cdot \frac{\Delta\mu}{\Delta t} + A \cdot \mu \cdot \frac{\Delta(V_g - V_t)}{\Delta t} \tag{3}$$

or

$$\Delta I_d \approx A \cdot (V_g - V_t) \cdot \Delta\mu + A \cdot \mu \cdot \Delta(V_g - V_t) \tag{4}$$

It is clear that both mobility variation and a shift of $(V_g - V_t)$ can change I_d.

If one assumes,

$$f_1 = A \cdot (V_g - V_t) \cdot \Delta\mu \tag{5}$$

$$f_2 = A \cdot \mu \cdot \Delta(V_g - V_t) \tag{6}$$

and,

$$\alpha = \frac{f_1}{f_2} \tag{7}$$

where α can be a function of μ, $(V_g - V_t)$, and time.

Equation (4) becomes,

$$\begin{aligned} \Delta I_d &\approx f_1 + f_2 \\ &= \alpha \cdot f_2 + f_2 \\ &= (\alpha + 1) \cdot f_2 \\ &= (\alpha + 1) \cdot A \cdot \mu \cdot \Delta(V_g - V_t) \end{aligned} \tag{8}$$

Within the accuracy of numerical evaluation, one can make the following approximations: within one time step, Δt, between two measurement points, the relation between ΔI_d and $\Delta(V_g - V_t)$ is approximately linear. This is to assume $(\alpha + 1) \cdot A \cdot \mu \approx$ constant within a time step, which is a standard practice for numerical simulation.

The question is how to evaluate $(\alpha + 1) \cdot A \cdot \mu$. The 'On-The-Fly' technique imposes a change of V_g by $\pm$ DV at a given time, and measure the corresponding change in I_d. This leads to,

$$\Delta(V_g - V_t) \approx \Delta V_g = 2 \cdot DV \tag{9}$$

with a corresponding change in I_d by

$$\Delta I_d = I_{d(V_g + DV)} - I_{d(V_g - DV)} \tag{10}$$

The trans-conductance at a given time is,

$$g_m \approx \frac{\Delta I_d}{\Delta V_g} = \frac{I_{d(V_g+DV)} - I_{d(V_g-DV)}}{2 \cdot DV} \tag{11}$$

From equations (8) and (9), one has,

$$\begin{aligned}(\alpha+1)\cdot A \cdot \mu &= {}^{\Delta I_d}\!/\!_{\Delta(V_g - V_t)} \\ &= g_m\end{aligned} \tag{12}$$

It is worth of pointing out that g_m is evaluated 'locally' for each time point and it has not been assumed that g_m is a constant during NBTI.

During the NBTI, V_g is a constant. Between two neighboring measurement points, (n - 1) and (n), we have,

$$\Delta(V_g - V_t(n)) = -\Delta V_t(n) \tag{13}$$

where 'n' indicates the change is for one time step between the point (n - 1) and (n).

From equations (12) and (13), we have,

$$\begin{aligned}\Delta V_t(n) &= -\Delta[V_g - V_t(n)] = -\frac{\Delta I_d(n)}{g_m} \\ &= -\frac{I_d(n) - I_d(n-1)}{[g_m(n) + g_m(n-1)]/2}\end{aligned} \tag{14}$$

g_m is assumed to be $[g_m(n) + g_m(n-1)]/2$ for each time step.

The accumulative shift of threshold voltage is,

$$\Delta V_t(accumulative) \approx -\sum_{n=1}^{N} \frac{I_d(n) - I_d(n-1)}{[g_m(n) + g_m(n-1)]/2} \tag{15}$$

This is the equation (15) in section 2.4.3.

If mobility variation were neglected, $\alpha = 0$ would be required according to equations (4) ~ (7). It should be emphasized that the 'On-The-Fly' technique did not assume $\alpha = 0$ and the mobility variation with time has been taken into account when evaluating ΔV_t through measuring g_m at each point in time. If $\alpha = 0$ were assumed, equation (12) leads to $g_m = A \cdot \mu$, which is not valid under the present test conditions.

To evaluate the variation of mobility during NBTI, from equations (1) and (4), one has,

$$\frac{\Delta\mu(n)}{\mu(n)} = \frac{\Delta I_d(n)}{I_d(n)} + \frac{\Delta V_t(n)}{V_g - V_t(n)} \tag{16}$$

and

$$\frac{\Delta\mu(n)}{\mu_0} = \frac{\mu(n)}{\mu_0} \times \left[\frac{\Delta I_d(n)}{I_d(n)} + \frac{\Delta V_t(n)}{V_g - V_t(n)} \right] \tag{17}$$

The accumulative mobility variation is,

$$\frac{\Delta\mu(accumulative)}{\mu_0} \approx -\sum_{n=1}^{N} \frac{\Delta\mu(n)}{\mu_0} \tag{18}$$

Under the typical test conditions used in this project, the relative change of drain current, threshold voltage, and the effective mobility are given in Figures 20(a) ~ (c), respectively. The $\Delta V_t/(V_g - V_t)$ is over 20%, corresponding to a shift of V_t larger

than 0.5V. This only leads to a few percent degradation of drain current, because the effective mobility at the stress voltage actually increased.

During stresses, it is generally accepted that carrier mobility will degrade. This appears contradicting the mobility increase observed in Figure 20(c) and an explanation must be given. It should be point out that the 'effective mobility', μ, in Figure 20(c) is different from the low field mobility, $\mu_{(LF)}$. The relation between these two can be expressed as [44],

$$\mu = \frac{\mu_{(LF)}}{1+\theta\cdot\left(V_g - V_t - \frac{1}{2}\cdot V_d\right)} \tag{19}$$

The mobility degradation reported by early works [9,43] is for the $\mu_{(LF)}$. Under our test condition, there is little doubt that $\mu_{(LF)}$ is also degraded. To show this degradation, the $I_d \sim V_g$ before and after a typical NBTS is plotted in Figure 21(a) and the corresponding trans-conductance is given in Figure 21(b). At low $|V_g|$, $gm \approx A \cdot \mu \approx A \cdot \mu_{(LF)}$. It is clear that stress reduces the maximum g_m, indicates a degradation of $\mu_{(LF)}$.

As $|V_g|$ increases, the effective mobility reduces, leading to a reduction of g_m. After the stress, the $I_d \sim V_g$ is shifted toward higher $|V_g|$. Figure 21(b) shows that the g_m at a given $|V_g| = 3.17$ V actually increases. As a result, the reduction of effective $|V_g|$ over-compensates the degradation of $\mu_{(LF)}$, leading to an increase in the effective mobility at a given stress voltage.

5.3.3. Effects of measurement temperature

There are two possible scenarios for the weak dependence of NBTI induced ΔV_t on temperature shown in Figure 18. First, NBTI induced defect generation is indeed insensitive to temperature. Second, the generated defects have a weaker impact at elevated temperature. That means when temperature increases, the generated defect

density increases as well. However, due to higher measurement temperature, their effects on ΔV_t become smaller. As a result, these effects will balance each other out, and give us the impression of insensitivity. In the following, it will be shown that the latter scenario is correct.

In Figure 22, a device was first stressed at 200°C (symbol '◇'). The device was then allowed to cool down to room temperature. Some charges were inevitably lost during this cooling period of 4 min. When ΔV_t was measured after cooling, one would expect a smaller ΔV_t. The symbol 'o' in Figure 22, however, shows that ΔV_t measured at room temperature is actually higher. There can be two possible explanations for the higher ΔV_t. First, the number of defects is increased at room temperature. Second, for a given number of defects, their effects on ΔV_t are temperature dependent. This is based on the observation in section 5.2 that the charging of some defects can increase for lower temperature. To test these explanations, temperature was raised back to 200°C and ΔV_t was measured again (symbol '△'). ΔV_t was much lower initially, because the gate was floating and some charges were lost during the warming up period of 4 min. As the stress time increases, ΔV_t rose, but the point 'B' clearly failed to reach the level achieved at room temperature (point 'A'). Since ΔV_t recorded at 200°C on both sides of the room temperature measurement is smaller than ΔV_t at room temperature, it is strongly against the explanation that the higher ΔV_t at room temperature was caused by a larger number of defects. In conclusion, like the ΔV_t measured by the traditional method, for a given number of defects, the lower the temperature, the higher the 'On-The-Fly' ΔV_t becomes [27,28,45].

By using the test procedure shown in Figure 22, the points 'A' and 'B' in Figure 22 after different NBTS time were obtained and plotted against NBTS time in Figure 23. The difference between the curves 'A' and 'B' in Figure 23 is caused by different measurement temperatures. Figure 23 also compares 'A' with the ΔV_t stressed and measured at room temperature (symbol '◇'). The difference between the 'A' and symbol '◇' shows the 'real' effect of stress temperature on defect creation. It is

obvious that these differences are negligible at short stress time, but becomes substantial as stress time increases, which will be addressed in section 5.3.4.

To examine the defects responsible for the ΔV_t measured 'On-The-Fly', $V_g > 0$ (E_{ox} = +5 MV/cm) and $V_g < 0$ (E_{ox} = -5 MV/cm) was alternately applied after a typical NBTS in Figure 24. Similar to the ΔV_t measured by the traditional method and shown in Figure 2, ΔV_t(On-The-Fly) also consists of three components: as-grown hole trapping, cyclic positive charges (CPC), and anti-neutralization positive charges (ANPC).

To further compare the defects measured by the two different techniques, two tests were carried out. In the first test, the impact of measurement temperature on different defects measured by the 'On-The-Fly' technique is studied. In Figure 25, two devices were stressed under the same V_g for the same time, one at 200°C and the other at room temperature. The defects were then measured at room temperature for both devices. It is obvious that the CPC in these two devices are similar, but ANPC is significantly higher for the device stressed at 200°C. Moreover, the difference in ANPC (the arrow marked as 'B') is approximately the same as the difference in the total ΔV_t (the arrow marked as 'A'). This means that only ANPC creation is thermally accelerated. Both CPC and as-grown hole trapping is insensitive to temperature, in agreement with the observation when the traditional technique is used.

In the second test, the thermal stability of created defects is explored. A device was first stressed at room temperature (V_g = -3.17 V) for overnight. It was then exposed to elevated temperature for a fixed time of 800 sec. The impact of this exposure on ΔV_t, CPC and ANPC is shown in Figure 26(a) & (b). It is clear that CPC is stable, while ANPC can be annealed. This is again in agreement with the observation based on the traditional measurement. In conclusion, although the ΔV_t measured 'On-The-Fly' is over one order of magnitude higher, the electrical signatures of defects involved are the same. This implies that the same types of defects are responsible for the ΔV_t measured by both the traditional and the 'On-The-Fly' methods.

5.3.4. Effects of stress time

Figure 23 shows that the impact of measurement temperature on ΔV_t is negligible at short stress time (< 10 sec), but becomes increasingly important at longer stress time. It is possible that this is because different types of defects dominate ΔV_t at different stress time. Supporting evidence for this explanation will be presented below.

Since ANPC is the only defect where charging is sensitive to temperature, it is reasonable to assume that at short stress time, the as-grown hole traps and CPC are the dominant defects, leading to the insensitiveness to the measurement temperature in Figure 23. In Figure 27, the 'recoverable' part of ΔV_t is compared for two devices. One was stressed for only 10 sec, where the impact of measurement temperature is negligible. The other was stressed for 570 sec, where the impact of measurement temperature is substantial. Figure 27 shows that the recoverable part dominates ΔV_t after 10 sec stress. The recoverable part is the sum of CPC and as-grown hole trapping (see Figure 24). Since both of CPC and as-grown hole trapping is not sensitive to the measurement temperature, it explains the insensitiveness of ΔV_t to temperature at short stress time.

After a 570 sec stress, Figure 27(a) & (b) show that the recoverable part remains essentially the same as that after a 10 sec stress, but the ANPC is substantially increased. This confirms that ANPC is responsible for the observed effects of measurement temperature on ΔV_t.

The measurement in Figure 27(a) & (b) was carried out at the stress temperature. Figure 28(a) ~ (c) show that if the comparison after a short and long stress was based on the measurement at room temperature. The difference in ANPC in these two cases is even larger.

5.3.5. Effects of stress voltage

Up to now, the stress voltage used is V_g = -3.17 V, corresponding to an oxide field of -11 MV/cm. This is much higher than normal operation voltage. The question is whether the same types of defects will be created at lower voltage.

Figure 29 compares the effects of measurement temperature on ΔV_t after stressed at -11 MV/cm and -8 MV/cm. Unlike the -11 MV/cm case, the impact of measurement temperature is negligible when stressed at -8 MV/cm. This is because ANPC generation is sensitive to the voltage level during NBTS. Figure 30(a) & (b) show that ANPC reduces significantly at -8 MV/cm, but CPC is hardly changed.

One question is whether the impact of measurement temperature shown in Figure 22 on ΔV_t can ever be observed when stressed at -8 MV/cm. In another word, under -8 MV/cm, can sufficient ANPC be created to make the effects of measurement temperature clearly observable? When the stress time at -8 MV/cm is increased to 75000 sec, Figure 31(a) & (b) show that ANPC is increased to a level high enough that the impact of measurement temperature on ΔV_t becomes obvious. We conclude that the same types of defects are created at different gate biases.

5.3.6. As-grown hole traps in thin dielectric

For the ΔV_t measured by the 'On-The-Fly' method, the results presented in sections 5.3.1 ~ 5.3.4 indicate that 'as-grown hole trapping' plays a dominating role during the initial stage of NBTS. Under V_g = -3.17 V, this initial period is approximately 10 sec. As expected, Figure 32 shows that the as-grown hole trapping saturates at longer time. The properties of the as-grown hole traps will be explored.

Although as-grown hole traps have been investigated by many early works [25,26], relatively thick oxides were used. This is because trapping is highly unstable when the oxide is less than 3 nm, and one example is shown in Figure 24. If the traditional method is used, the detrapping will be significant, making it difficult for

characterizing the as-grown hole traps. The 'On-The-Fly' method suppresses the detrapping and gives an opportunity for characterizing as-grown hole traps in thin dielectric. In Chapter 4, it is shown that the electrical signature of electron traps is the same for 'thick' and 'thin' dielectric layer. It is interesting to study how hole traps depend on oxide thickness.

5.3.6.1. Capture cross sections evaluated from the measured hole current

An important signature of traps is their capture cross section. To estimate the capture cross section, one must know the carrier fluency through the dielectric.

During NBTS, both holes and electrons can flow through the ultra-thin gate dielectric layer, as illustrated in Figure 33. As a result, carrier separation measurement is needed to separate the hole and electron contribution to the total gate current [46,47]. Figure 34(a) illustrates the gate, n-well, and channel current of a pMOSFET under inversion condition, and Figure 34(b) shows the data. The currents are plotted as a function of gate voltage, V_g. The channel current is due to the tunneling of holes from the Si substrate valence band, and the n-well current is due to the tunneling of electrons from the poly-Si gate valence band, as illustrated in Figure 33. For $|V_g| < 1.5$ V, the hole tunneling current is larger than electron current. While the electron current is larger than hole current at higher gate biases. This observation is in agreement with that reported in earlier work [11]. Figure 34(c) shows the relative contribution of hole current to the gate current. When NBTS was carried out under a gate bias of $V_g = -3.17$ V, roughly 10% of the gate current is caused by holes.

The as-grown hole trapping is plotted against hole fluency, Q_h, in Figure 35. Early works at this university [26,48] show that there exist two well separated capture cross sections for as-grown hole traps: one in the order of $10^{-13} \sim 10^{-14}$ cm^2 and the other in the order of 10^{-15} cm^2. When the data in Figure 35 are fitted with two capture cross sections, the extracted values are 6.9×10^{-17} cm^2 and 1.8×10^{-18} cm^2. They are three orders of magnitude less than that reported early. This discrepancy will be addressed in the next section.

5.3.6.2. An explanation for the difference in capture cross section

The difference in the capture cross section could be caused by the different hole injection processes. In the early works for thick oxides [26,48], hot holes were injected into the oxide by using techniques such as substrate hole injection. Figure 36 shows that holes were injected directly into the valence band of SiO_2 to fill hole traps. For the present NBTS on thin dielectric, Figure 37 shows that holes were injected by direct tunneling. Whether these holes can fill traps depends on the trap energy level.

The energy level of defects in silicon dioxides has been reported [49-52]. Based on the theoretical calculation, the neutral oxygen vacancy is about 3 eV [49] or 2.4 eV [50] above the SiO_2 valence band edge. It is widely accepted that oxygen vacancies are hole traps. As illustrated in Figure 37, the energy of holes tunneling from the top edge of silicon valence band is too low to fill these traps. If the hole fluency in Figure 35 is dominated by these low energy holes, the filling efficiency will be low, leading to the reduction in capture cross sections.

As illustrated in Figure 38, holes can be injected from higher subbands. These holes could have enough energy for filling traps. An increase of energy will reduce hole density, and the 'effective' hole fluency for filling traps could be much lower than the total fluency in Figure 35. A reduction of hole fluency will lead to an increase of capture cross sections. As a result, the capture cross section in thin and thick dielectric could actually be the same.

5.3.7. Hole current calculation

In order to estimate the number of holes that can fill traps, it is necessary to separate the more energetic holes from the total hole injection. This requires calculating the direct tunneling hole current through ultrathin gate oxides from different subbands in the inversion layer. The hole fluency will then be calculated by only including holes of sufficient energy to fill traps. It is interesting to find out if the capture cross section extracted by using this new hole fluency agrees with the values reported earlier.

5.3.7.1. Inversion layer charge carrier calculation

With the increase of substrate doping, electric fields in silicon became large enough to cause significant quantization of the carrier energy at the Si/SiO_2 interface. The energy band diagram for a pMOSFET in the inversion region is illustrated in Figure 38. E_V is the Si valence band edge. E_F is the Fermi level in Si substrate. ϕ_B is the valence band offsets between the Si and Si/SiO_2. ϕ_S is the total surface potential energy. The zero energy reference point is the Si valence band edge at the Si/SiO_2 interface.

To simplify the estimation of the quantum mechanical effects in the inversion layer of a n-type Si substrate, a triangular-well approximation [53] is used, which assumes a linear variation of the potential in the semiconductor surface layer. The energy level of the j^{th} energy subband in the i^{th} valley, E_{ij}, is given by [54,55]

$$E_{ij} = \left(\frac{\hbar^2}{2 \cdot m_{zi}}\right)^{\frac{1}{3}} \cdot \left[\frac{3}{2} \cdot \pi \cdot q \cdot F_{eff} \cdot \left(j - \frac{1}{4}\right)\right]^{\frac{2}{3}} \tag{20}$$

where m_{zi} is the quantization effective mass of the i^{th} valley, F_{eff} is the effective electric field in the Si substrate. F_{eff} can be estimated by,

$$F_{eff} = \frac{q \cdot (N_{depl} + \eta \cdot N_{inv})}{\varepsilon_0 \cdot \varepsilon_{si}}, \tag{21}$$

where N_{depl} and N_{inv} are depletion and inversion charge densities, respectively. The value of η for hole inversion is 0.5 [56]. If we use the density of states for a 2-D electron gas (2-DEG) and Fermi-Dirac statistics, the inversion layer carrier density from each subband can be expressed as,

$$N_{ij} = \left(\frac{k \cdot T}{\pi \cdot \hbar^2}\right) \cdot m_{di} \cdot \ln\left[1 + \exp\left(\frac{E_F - E_{ij}}{k \cdot T}\right)\right], \tag{22}$$

where m_{di} is the density of states effective mass of the i^{th} valley.

Although these effective masses are weakly electric field dependent, they can be treated as constants in the first order approximation. It has been reported that : $m_z = 0.28 \times m_0$ and $m_d = 0.45 \times m_0$ for the heavy-hole band, and $m_z = 0.20 \times m_0$ and $m_d = 0.16 \times m_0$ for the light-hole band [57,58], where m_0 is the free electron mass. These values will be used here.

The total inversion layer carrier density is

$$N_{inv} = \sum_{ij} N_{ij} \tag{23}$$

The average inversion layer charge centroid thickness, z_{av}, is defined as

$$z_{av} = \sum_{ij} \frac{N_{ij} \cdot z_{ij}}{N_{inv}} \tag{24}$$

where z_{ij} is the average charge centroid for the j^{th} energy subband in the i^{th} valley, which is expressed as

$$z_{ij} = \frac{2 \cdot E_{ij}}{3 \cdot F_{eff}} \tag{25}$$

The depletion layer doping density, N_{depl}, can be evaluated from,

$$N_{depl} = \sqrt{2 \cdot \varepsilon_{si} \cdot \phi_d \cdot \frac{N_A - N_D}{q}}, \tag{26}$$

where the substrate band bending in the depletion layer, ϕ_d, is,

$$\phi_d = \phi_s - \frac{k \cdot T}{q} - \frac{q \cdot N_{inv} \cdot z_{av}}{\varepsilon_{si}}. \tag{27}$$

The equations (20) to (27) can be numerically solved. The flow chart of the program used is illustrated in Figure 39. The surface potential and substrate doping density

are two input parameters. The iteration sequence begins with an assumed total inversion layer carrier density and the depletion layer doping charge density. This permits a calculation of the inversion layer carrier density for each valley and subband as a function of the effective electric field, as well as the depletion layer doping charge density. A new total carrier density can then be obtained by using equation (23). The newly calculated values are compared with the assumed ones. If their differences are not acceptable, the assumed values are updated and the calculation starts again. The iteration continues, until the differences between the assumed and calculated values become acceptable.

The calculated E_{ij} and N_{ij} is plotted in Figure 40(a) & (b), respectively. Three subbands were evaluated here for two valleys. For a given subband, the energy level of the heavy hole valley (i = 1) is less than that of the light hole valley (i = 2). Figure 40(b) shows when F_{eff} = 1.65 MV/cm (V_g = -3.17 V), that the carrier density in the 2nd subband is approximately four orders of magnitude less than that in the 1st subband. And the N_{ij} in the 3rd subband is again roughly four orders of magnitude less than that in the 2nd subband. The relationship between the effective electrical field and surface potential is given in Figure 40(c).

5.3.7.2. Direct tunneling current calculation

The total direct tunneling current consists of the hole current from each subband [59],

$$J = \sum_{ij} J_{ij} = q \cdot \sum_{ij} N_{ij} \cdot f_{ij} \cdot T_{ij} \tag{28}$$

where f_{ij} is the hole impact frequency on the interface and T_{ij} is the tunneling probability for carriers of the j^{th} subband in the i^{th} valley. The carrier density, N_{ij}, is calculated in the section 5.3.7.1.

The impact frequency, f_{ij}, for a particle of energy, E_{ij}, can be evaluated by [60],

$$f_{ij} = \frac{E_{ij}}{3 \cdot \hbar \cdot \pi \cdot \left(j - \frac{1}{4} \right)}. \tag{29}$$

To calculate the tunneling probability, the WKB approximation is used,

$$T_{ij} = T_{R\_ij} \cdot T_{WKB\_ij} \tag{30}$$

where T_{WKB_ij} is the WKB approximation of the transmission probability and T_{R_ij} the correction factor accounting for the reflections from the oxide interfaces.

The details for evaluating T_{R_ij} can be found in [59]. Under the present test condition, the typical value of T_{R_ij} is close to one. T_{WKB_ij} can be evaluated by [57],

$$T_{WKB\_ij} = \exp\left[-\frac{2}{\hbar} \cdot \left| \int_0^{x_{ox}} \sqrt{2 \cdot m_{ox} \cdot (E_{ij} - \phi(x))} dx \right| \right] \tag{31}$$

where m_{ox} is the effective mass of holes in the oxide, x is the distance into the oxide from the interface and $\phi(x)$ the energy barrier height, as illustrated in Figure 41.

The relation between the gate bias and the surface potential is,

$$V_g = V_{FB} + V_{ox} + V_{poly} + \phi_s, \tag{32}$$

where V_{FB} is the flatband voltage. V_{ox} is the oxide voltage drop. V_{poly} is the voltage drop in polysilicon gate due to polysilicon depletion. The V_{ox} is determined by,

$$V_{ox} = \frac{q \cdot (N_{depl} + N_{inv})}{\varepsilon_{sio2} \cdot \varepsilon_0} \cdot t_{ox}. \tag{33}$$

The V_{poly} is determined by

$$V_{poly} = \frac{\varepsilon_{sio2}^{\ 2} \cdot \varepsilon_0^{\ 2} \cdot E_{ox}^{\ 2}}{2 \cdot q \cdot \varepsilon_{si} \cdot \varepsilon_0 \cdot N_{poly}} \tag{34}$$

The relation between V_g and F_{eff} is shown in Figure 42.

To verify the program, the calculation result is compared with those reported by Hou et al [54] in Figure 43. Hou et al [54] used a substrate doping concentration $N_A = 5\times10^{17}$ cm^{-3}, a valence band offset between Si and SiO_2 $\phi_B = 4.5$ eV, and an effective mass of hole in SiO_2 $m_{ox} = 0.40\times m_0$ (solid line). When the same parameter values were used, the present program produces very similar results (symbol 'x'). The symbol '◇' is the measured hole current by Hou et al [54]. The agreement at lower $|V_g|$ is good, but the calculated value is higher as $|V_g|$ increases. The reason is not clear at present.

After verifying the program, the calculation is compared with the hole current measured in this project. Figure 44 shows the direct tunneling hole current obtained from carrier separation measurement (symbol '◇'). To best fit the experimental result under the typical stress voltage (i.e. $V_g = -3.17$ V), the effective mass of holes in SiO_2 was found to be $m_{ox} = 0.35\times m_0$. This value is within previous results reported for valence band hole tunneling ($0.35 \sim 0.50\times m_0$) [54,57,61]. The calculated hole current is lower than the measured value at lower $|V_g|$. The current value below 10^{-12} A is not reliable in the present experimental setup, which may partially explain the higher measured value. A full understanding of the difference between test and calculation has not been achieved at present.

The hole current calculated for each subband is presented in Figure 45(a). As expected, the higher the energy level, the lower the hole current becomes, because of the reduction in carrier density. The largest current is from the 1st subband of the heavy hole valley. Figure 45(b) shows the ratio against this largest current component. Under the typical testing condition, it is clear that the current from the 2nd subband in the heavy hole valley is three orders of magnitude less. Figure 45(c)

shows the ratio against the current from the 2nd subband in the heavy hole valley. Again, under the typical testing condition, the current from the 2nd subband in the light hole valley is 34 times lower. The current from the 3rd subbands are three to five orders of magnitude less.

5.3.8. Estimation of capture cross sections

5.3.8.1. Room temperature

The calculated hole current is now used to estimate the capture cross sections of hole traps. When the tunneling of holes from three subbands in two valleys are used to calculate the hole fluency, Figure 46(a) shows that the resultant capture cross sections are $\sigma_1 = 3.8\times10^{-17}$ cm^2 and $\sigma_2 = 9.8\times10^{-19}$ cm^2. These values are three orders of magnitude less than those reported in literature [48]. Figure 45 shows that the total hole current is dominated by tunneling from the 1st subband. Figure 41 shows that the energy of holes from the 1st subband is too low to fill the as-grown hole traps, explaining the small capture cross section.

Figure 41 shows that holes from the 2nd subbands have an energy high enough to fill the as-grown hole traps. When the hole fluency is evaluated from the tunneling of holes from the 2nd subband of the heavy hole valley alone, Figure 46(b) shows that the extracted capture cross sections agree well with the reported values. In fact, if one uses the capture cross sections extracted from earlier work on relatively thick SiO_2 ($\sigma_1 = 7.3\times10^{-14}$ cm^2 and $\sigma_2 = 3.6\times10^{-15}$ cm^2, [48]), the present data can be fitted reasonably well.

The contribution of holes tunneling from subbands of higher energy should be estimated. Figure 46(c) shows the results, when holes tunneling from the 2nd subband of both valleys are taken into account. Since the hole current from the 2nd subband of the light hole valley is 34 times less than that from the heavy hole valley, the inclusion of the 2nd subband of the light hole valley has a weak impact on Q_h and

consequently, a weak effect on σ_1 & σ_2. Similarly, Figure 46(d) shows that the contribution from the 3rd subband to Q_h is negligible. These results indicate that the main source of holes for filling the as-grown hole traps is the 2nd subband of the heavy hole valley. This suggests that higher subbands than the 2nd can be neglected.

5.3.8.2. Effects of temperature

Since temperature is a key parameter for NBTI tests, its effects on the as-grown hole trapping should be evaluated. Figure 47(a) ~ (c) show the impact of temperature on E_{ij}, N_{ij} and hole currents. Figure 48 shows the impact of temperature on the capture cross sections. An increase of temperature from room temperature to 200°C reduces the capture cross section by a factor of fifty. Attempts will be made to justify this reduction.

Early work [62] also showed that the capture cross section reduces for higher temperature. An increase of temperature from 77K to 295K, however, only reduces the capture cross section by a factor of 2.4. The range of temperature used for this work is between room temperature and 200°C. To test the effects of temperature in this range on the capture cross section, the substrate hole injection was carried out on a relatively thick oxide (5.5 nm). Figure 49 shows that the capture cross section only reduces by a factor of four in this temperature range. This is much smaller than that in Figure 48.

The large reduction of the capture cross section in Figure 48 is not fully understood at present. To speculate on its origin, one may assume that the calculated energy level of subband may not be accurate. Figure 50 shows that, if the energy of the 2nd subband of the heavy hole valley is reduced from 0.64 eV to 0.57 eV, the calculated temperature effect will be in good agreement with the experimental observation. Given the various assumptions and simplification made in the calculation, one may argue that it is possible to have an uncertainty in the energy level less than 0.1 eV. At present, conclusions cannot be drawn and further work is needed.

5.3.9. Conclusions

The effects of measurement temperature on NBTI are also investigated for the newly proposed 'On-The-Fly' technique. For a given number of defects, it is found that ΔV_t increases for lower measurement temperature. The impact of measurement temperature on ΔV_t is stronger for longer stress time. This time-dependence results from that different defect dominate ΔV_t at different stress time. Although ΔV_t measured at stress temperature can appear insensitive to temperature, it does not imply that defect creation is not sensitive to temperature. The ΔV_t measured at stress temperature should not be used to assess the thermal acceleration of defect generation. The temperature-dependent instability of threshold voltage should be taken into account when estimating the NBTI-limited device lifetime.

Like the traditional measurement, it is found that three different types of defect contribute to ΔV_t measured 'On-The-Fly': one is the as-grown hole traps, the other two are created under NBTS. The as-gown hole traps can dominate the initial stage of NBTS, then saturate at longer time. The theoretical calculation indicates that holes tunneling from the 2^{nd} subband of the heavy hole valley dominate the filling of as-grown hole traps for thin SiON during typical NBTS. The capture cross sections are similar to those reported for relatively thick SiO_2: one in the order of $10^{-13} \sim 10^{-14}$ cm^2 and the other in the order of 10^{-15} cm^2.

Finally, the carrier mobility variation during NBTS was estimated. It is found that the effective mobility of holes at V_g = -3.17 V actually increases with NBTS time. This results from a reduction of $|V_g - V_t|$ during NBTI and the effective mobility increases for lower $|V_g - V_t|$.

References

1. B. E. Deal, M. Skalar, A. S. Grove and E. H. Snow, "Characteristics of the surface-state charge (Qss) of thermally oxidized silicon," J. Electrochem. Soc., **114**, 266 (1967).
2. A. Goetzberger, A. D. Lopez and R. J. Strain, "On the formation of surface states during stress aging of thermal Si-SiO interfaces," J. Electrochem. Soc., **120**, 90 (1973).
3. K. O. Jeppson and C. M. Svensson, "Negative bias stress of MOS devices at high electric fields and degradation of MNOS devices," J. Appl. Phys., **48**, 2004 (1977).
4. A. R. Stivers and C. T. Sah, "A study of oxide traps and interface states of the silicon-silicon dioxide interface," J. Appl. Phys., **51**, 6292 (1980).
5. W. Chen, A. Balasinski, and T. P. Ma, "Lateral profiling of oxide charge and interface traps near MOSFET," IEEE Trans. Electron Dev., **40**, 187 (1993).
6. D. M. Fleetwood, "Effects of hydrogen transport and reactions on microelectronics radiation response and reliability," Microelectronics Reliability, **42**, 523 (2002).
7. J. F. Zhang, S. Taylor and W. Eccleston, "Electron trap generation in thermally grown SiO_2 under Fowler-Nordheim stress," J. Appl. Phys., **71**, 725 (1992).
8. J. F. Zhang and W. Eccleston, "Effects of high-field injection on the hot-carrier-induced degradation of submicrometer pMOSFETs," IEEE Trans. Electron Dev., **42**, 1269 (1995).
9. D. K. Schroder and J. A. Babcock, "Negative bias temperature instability: road to cross in deep submicron silicon semiconductor manufacturing," J. Appl. Phys., **94**, 1 (2003).
10. G. Chen, M. F. Li, C. H. Ang, J. Z. Zheng and D. L. Kwong, "Dynamic NBTI of p-MOS transistors and its impact on MSFET scaling," IEEE Electron Dev. Lett., **23**, 734 (2002).
11. M. Houssa, M. Aoulaiche, J. L. Autran, C. Parthasarathy, N. Revil and E. Vincent, "Modeling negative bias temperature instabilities in hole channel metal-oxide-semiconductor field effect transistors with ultrathin gate oxide layers," J. Appl. Phys., **95**, 2786 (2004).

12. S. S. Tan, T. P. Chen, J. M. Soon, K. P. Loh, C. H. Ang, and L. Chen, "Nitrogen-enhanced negative bias temperature instability: An insight by experiment and first-principle calculations," Appl. Phys. Lett., **82**, 1881 (2003).
13. N. Kimizuka, T. Yamamoto, T. Mogami, K. Yamaguchi, K. Imai and T. Horiuchi, "The impact of bias temperature instability for direct-tunneling ultra-thin gate oxide on MOSFET scaling," in Proc. IEEE VLSI Symp., Kyoto, p.73, Japan (1999).
14. N. Kimizuka, K. Yamaguchi, K. Imai, T. Iizuka, C. T. Liu, R. C. Keller and T. Horiuchi, "NBTI enhancement by nitrogen incorporation into ultrathin gate oxide for 0.10μm gate CMOS generation," in Proc. IEEE VLSI Symp., Honolulu, p.92, USA (2000).
15. N. K. Jha and V. R. Rao, "A new oxide trap-assisted NBTI degradation model," IEEE Electron Dev. Lett., **26**, 687 (2005).
16. B. Kaczer, V. Arkhipov, M. Jurczak and G. Groeseneken, "Negative bias temperature instability (NBTI) in SiO2 and SiON gate dielectrics understood through disorder-controlled kinetics," Microelectronic Eng., **80**, 122 (2005).
17. V. Huard, M. Denais, F. Perrier, N. Revil, C. Parthasarathy, A. Bravaix and E. Vincent, "A thorough investigation of MOSFETs NBTI degradation," Microelectronics Reliability, **45**, 83 (2005).
18. S. Mahapatra, P. B. Kumar and M. A. Alam, "Investigation and modeling of interface and bulk trap generation during negative bias temperature instability of p-MOSFETs," IEEE Trans. Electron Dev., **51**, 1371 (2004).
19. C. E. Blat, E. H. Nicollian and E. H. Poindexter, "Mechanism of negative-bias-temperature instability," J. Appl. Phys., **69**, 1712 (1991).
20. M. Denais, A. Bravaix, V. Huard, C. Parthasarathy, G. Ribes, F. Perrier, Y. Rey-Tauriac and N. Revil, "On-the-fly characterization of NBTI in ultra-thin gate oxide PMOSFET's," in IEDM Tech. Dig., San Francisco, p.109, USA (2004).
21. M. Ershov, S. Saxena, H. Karbasi, S. Winters, S. Minehane, J. Babcock, R. Lindley, P. Clifton, M. Redford and A. Shibkov, "Dynamic recovery of negative bias temperature instability in p-type metal-oxide-semiconductor field-effect transistors," Appl. Phys. Lett., **83**, 1647 (2003).

22. K. Onishi, R. Choi, C. S. Kang, H. J. Cho, Y. H. Kim, R. E. Nieh, J. Han, S. A. Krishnan, M. S. Akbar and J. C. Lee, "Bias-temperature instabilities of polysilicon gate HfO2 MOSFETs," IEEE Trans. Electron Dev., **50**, 1517 (2003).
23. J. F. Zhang and W. Eccleston, "Donor-like interface trap generation in pMOSFET's at room temperature," IEEE Trans. Electron Dev., **41**, 740 (1994).
24. J. F. Zhang and W. Eccleston, "Positive bias temperature instability in MOSFET's," IEEE Trans. Electron Dev., **45**, 116 (1998).
25. J. F. Zhang, H. K. Sii, G. Groeseneken and R. Degraeve, "Hole trapping and trap generation in the gate silicon dioxide," IEEE Trans. Electron Dev., **48**, 1127 (2001).
26. J. F. Zhang, H. K. Sii, A. H. Chen, C. Z. Zhao, M. J. Uren, G. Groeseneken and R. Degraeve, "Hole trap generation in gate dielectric during substrate hole injection," Semicond. Sci. and Technol., **19**, L1 (2004).
27. J. F. Zhang, C. Z. Zhao, A. H. Chen, G. Groeseneken and R. Degraeve, "Hole-traps in silicon dioxides - Part I: Properties," IEEE Trans. Electron Dev., **51**, 1267 (2004).
28. C. Z. Zhao, J. F. Zhang, G. Groeseneken and R. Degraeve, "Hole-traps in silicon dioxides - Part II: Generation mechanism," IEEE Trans. Electron Dev., **51**, 1274 (2004).
29. J. F. Zhang, H. K. Sii, R. Degraeve and G. Groeseneken, "Mechanism for the generation of interface state precursors," J. Appl. Phys., **87**, 2967 (2000).
30. J. R. Brews, "Subthreshold behavior of uniformly and nonuniformly doped long-channel MOSFET," IEEE Trans. Electron Dev., **26**, 1282 (1979).
31. S. Horiguchi, T. Kobayashi and K. Saito, "Interface-trap generation modeling of Fowler-Nordheim tunnel injection into ultra-thin gate oxide," J. Appl. Phys., **58**, 387 (1985).
32. C. Tan, M. Xu and Y. Wang, "Application of the difference subthreshold swing analysis to study generation interface-trap in MOS structure due to Fowler-Nordheim aging," IEEE Electron Dev. Lett., **15**, 257 (1994).
33. P. J. McWhorter and P. S. Winokur, "Simple technique for separating the effects of interface traps and trapped-oxide charge in metal-oxide-semiconductor transistors," Appl. Phys. Lett., **48**, 133 (1986).

34. Y. Park and D. K. Schroder, "Degradation of thin tunnel gate oxide under constant Fowler-Nordheim current stress for a flash EEPROM," IEEE Trans. Electron Dev., **45**, 1361 (1998).
35. S. Ogawa, M. Shimaya and N. Shiono, "Interface-trap generation at ultrathin SiO2 (4-6nm)-Si interfaces during negative-bias temperature aging," J. Appl. Phys., **77**, 1137 (1995).
36. S. S. Tan, T. P. Chen, J. M. Soon, K. P. Loh, C. H. Ang, W. Y. Teo and L. Chen, "Linear relationship between H+-trapping reaction energy and defect generation: Insight into nitrogen-enhanced negative bias temperature instability," Appl. Phys. Lett., **83**, 530 (2003).
37. X. J. Zhou, L. Tsetseris, S. N. Rashkeev, D. M. Fleetwood, R. D. Schrimpf, S. T. Pantelides, J. A. Felix, E. P. Gusev and C. D'Emic, "Negative bias-temperature instabilities in metal-oxide-silicon devices with SiO2 and SiOxNy/HfO2 gate dielectrics," Appl. Phys. Lett., **84**, 4394 (2004).
38. H. Katto, "Positive/negative BT instability in scaled N/P-MOSFETs and MOSCs," in IEEE International Integrated Reliability Workshop Final Report, Lake Tahoe, p.54, USA (2001).
39. K. Onishi, R. Choi, C. S. Kang, H. J. Cho, Y. H. Kim, R. E. Nieh, J. Han, S. A. Krishnan, M. S. Akbar and J. C. Lee, "Bias-temperature instabilities of polysilicon gate HfO2 MOSFETs," IEEE Trans. Electron Dev., **50**, 1517 (2003).
40. C. Schlunder, R. Brederlow, P. Wieczorek, C. Dahl, J. Holz, M. Rohner, S. Kessel, V. Herold, K. Goser, W. Weber and R. Thewes, "Trapping mechanisms in negative bias temperature stressed p-MOSFETs," Microelectronics Reliability, **39**, 821 (1999).
41. S. Rangan, N. Mielke and E. C. C. Yeh, "Universal recovery behavior of Negative Bias Temperature Instability," in IEDM Tech. Dig., Washington DC, p.341, USA (2003).
42. H. Aono, E. Murakami, K. Okuyama, A. Nishida, M. Minami, Y. Ooji and K. Kubota, "Modeling of NBTI saturation effect and its impact on electric field dependence of the lifetime," Microelectronics Reliability, **45**, 1109 (2005).
43. V. Huard, M. Denais and C. Parthasarathy, "NBTI degradation: From physical mechanisms to modeling," Microelectronics Reliability, **46**, 1 (2006).
44. Y. Tsividis, "Operation and modeling of the MOS transistor," WCB/McGraw-Hill, 1999.

45. C. Z. Zhao and J. F. Zhang, "Effects of hydrogen on positive charges in gate oxides," J. Appl. Phys., **97**, 073703-1 (2005).
46. C. H. Ang, C. H. Ling, Z. Y. Cheng and B. J. Cho, "Origin of temperature-sensitive hole current at low gate voltage regime in ultrathin gate oxide," J. Appl. Phys., **88**, 2872 (2000).
47. D. J. DiMaria and J. H. Stathis, "Anode hole injection, defect generation, and breakdown in ultrathin silicon dioxide films," J. Appl. Phys., **89**, 5015 (2001).
48. J. F. Zhang, C. Z. Zhao, H. K. Sii, G. Groeseneken, R. Degraeve, J. N. Ellis and C. D. Beech, "Relation between hole traps and hydrogenous species in silicon dioxides," Solid-State Elec., **46**, 1839 (2002).
49. E. P. O'Reilly and J. Robertson, "Theory of defects in vitreous silicon dioxide," Phys. Review B, **27**, 3780 (1983).
50. J. K. Rudra and W. B. Fowler, "Oxygen vacancy and the E'_1 center in crystalline SiO2," Phys. Review B, **35**, 8223 (1987).
51. Y. Lu and C. T. Sah, "Thermal emission of trapped holes in thin SiO2 films," J. Appl. Phys., **78**, 3156 (1995).
52. P. M. Lenahan and J. F. Conley Jr, "A comprehensive physically based predictive model for radiation damage in MOS systems," IEEE Trans. Nuclear Sci., **45**, 2413 (1998).
53. F. Stern, "Self-consistent results for n-type Si inversion layers," Phys. Review B, **5**, 4891 (1972).
54. Y. T. Hou, M. F. Li, Y. Jin and W. H. Lai, "Direct tunneling hole currents through ultrathin gate oxides in metal-oxide-semiconductor devices," J. Appl. Phys., **91**, 258 (2002).
55. H. H. Muller and M. J. Schulz, "Simplified method to calculate the band bending and the subband energies in MOS capacitors," IEEE Trans. Electron Dev., **44**, 1539 (1997).
56. S. Rodriguez, J. A. Lopez-Villaneuva, I. Melchor and J. E. Carceller, "Hole confinement and energy subbands in a silicon inversion layer using the effective mass theory," J. Appl. Phys., **86**, 438 (1999).
57. K. N. Yang, H. T. Huang, M. C. Chang, C. M. Chu, Y. S. Chen, M. J. Chen, Y. M. Lin, M. C. Yu, S. M. Jang, D. C. H. Yu and M. S. Liang, "A physical model for hole direct tunneling current in p+ poly-gate pMOSFETs with ultrathin gate oxides," IEEE Trans. Electron Dev., **47**, 2161 (2000).

58. J. Cai and C. T. Sah, "Gate tunneling currents in ultrathin oxide metal-oxide-silicon transistors," J. Appl. Phys., **89**, 2272 (2001).

59. L. F. Register, E. Rosenbaum and K. Yang, "Analytic model for direct tunneling current in polycrystalline silicon-gate metal-oxide-semiconductor devices," Appl. Phys. Lett., **74**, 457 (1999).

60. B. Govoreanu, P. Blomme, K. Henson, J. Van Houdt and K. De Meyer, "An effective model for analysing tunneling gate leakage currents through ultrathin oxides and high-k gate stacks from Si inversion layers," Solid-State Elec., **48**, 617 (2004).

61. R. K. Chananan, K. McDonald, M. Di Ventra, S. T. Pantelides, L. C. Feldman, G. Y. Chung, C. C. Tin, J. R. Williams and R. A. Weller, "Fowler-Nordheim hole tunneling in p-SiC/SiO2 structures," Appl. Phys. Lett., **77**, 2560 (2000).

62. G. Van Den Bosch, G. Groeseneken, H. E. Maes, R. B. Klein and N. S. Saks, "Oxide and interface degradation resulting from substrate hot-hole injection in metal-oxide-semiconductor field-effect transistors," J. Appl. Phys., **75**, 2073 (1994).

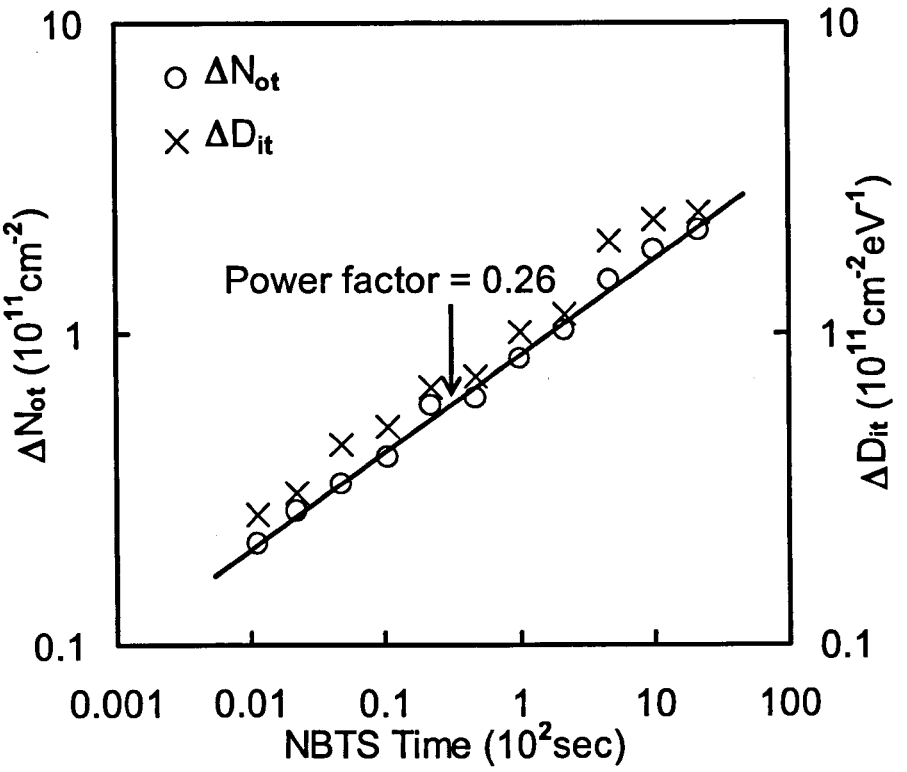
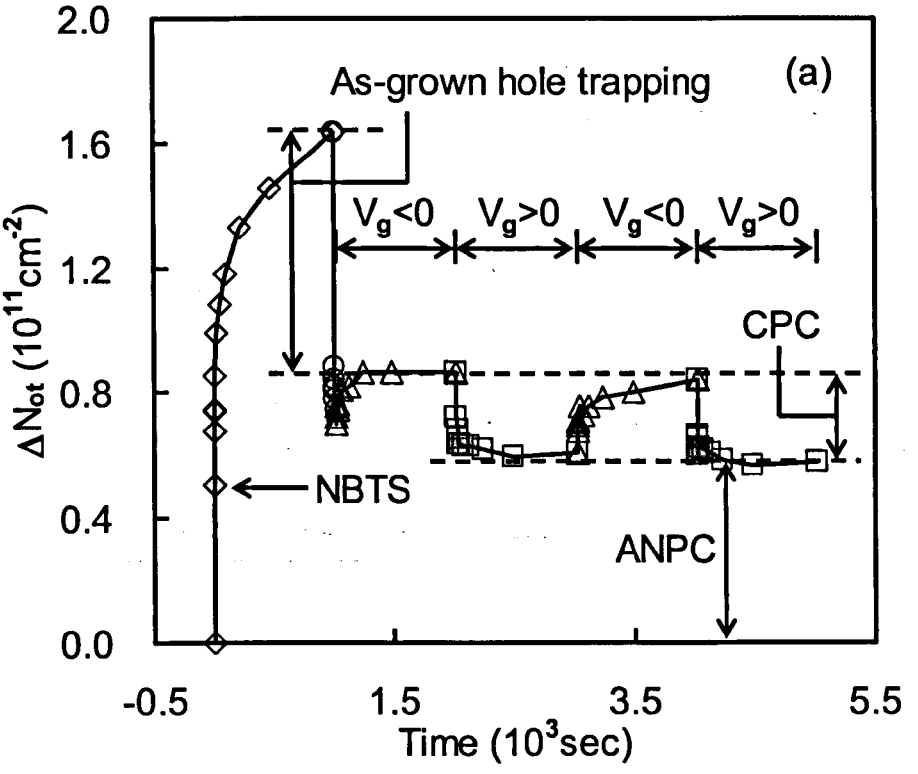


Figure 1. Build-up of the effective density of positive charges, ΔN_{ot}, and interface states, ΔD_{it}, during NBTS with a gate bias of -3.17 V and at 100°C. The solid line is obtained by fitting with ΔN_{ot}, and a power factor of 0.26 is found. It is apparent that ΔN_{ot} is in good agreement with ΔD_{it}.


2.0
1.6
1.2
0.8
0.4
0.0
ΔNot (10^11cm^-2)
As-grown hole trapping
(a)
Vg<0
Vg>0
Vg<0
Vg>0
CPC
NBTS
ANPC
-0.5
1.5
3.5
5.5
Time (10^3sec)

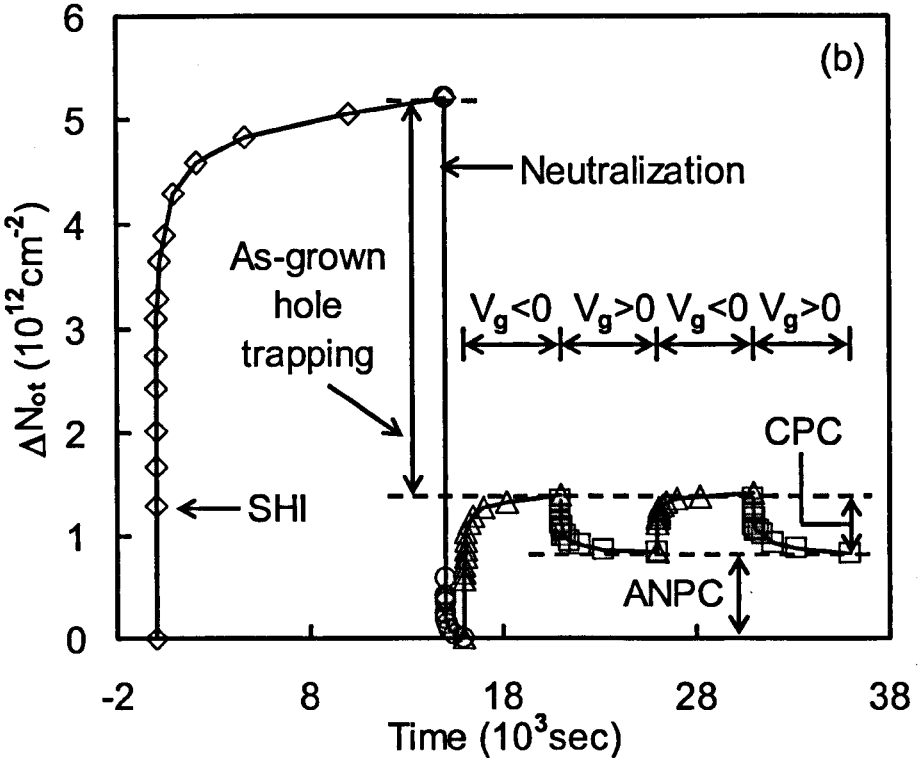

6
5
4
3
2
1
0
ΔNot (10^12cm^-2)
(b)
Neutralization
As-grown
hole
trapping
Vg<0
Vg>0
Vg<0
Vg>0
CPC
SHI
ANPC
-2
8
18
28
38
Time (10^3sec)

Figure 2. Testing procedure and types of positive charges formed during (a) NBTS and (b) substrate hole injection [27]. In (a), NBTS was carried out at V_g = -3.17 V and 100°C. It was followed by an electron injection at E_{ox} = +6.5 MV/cm and 100°C (symbol 'o'). A negative ($V_g < 0$) and positive ($V_g > 0$) gate bias ($|E_{ox}|$ = 5 MV/cm) was then applied alternately, with all other terminals grounded. In (b), SHI was carried out with E_{ox} = -5 MV/cm, a n-well bias of 6 V and a p-substrate bias of 7 V. The electron injection was at E_{ox} = +8 MV/cm. Different types of positive charges can be clearly identified. Apart from as-grown hole trapping, there are 'cyclic positive charges (CPC)' and 'anti-neutralization positive charges (ANPC)'. CPC can be cycled by changing gate bias polarity and it has similar charging ($V_g < 0$) and discharging ($V_g > 0$) rate. The discharging of ANPC is more difficult than its charging.

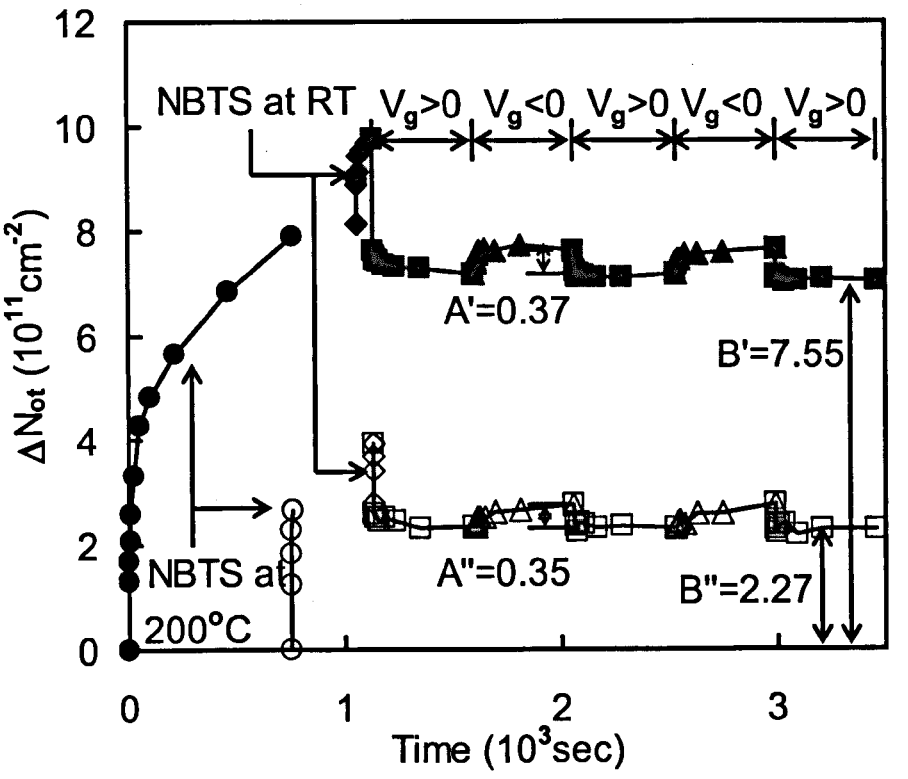
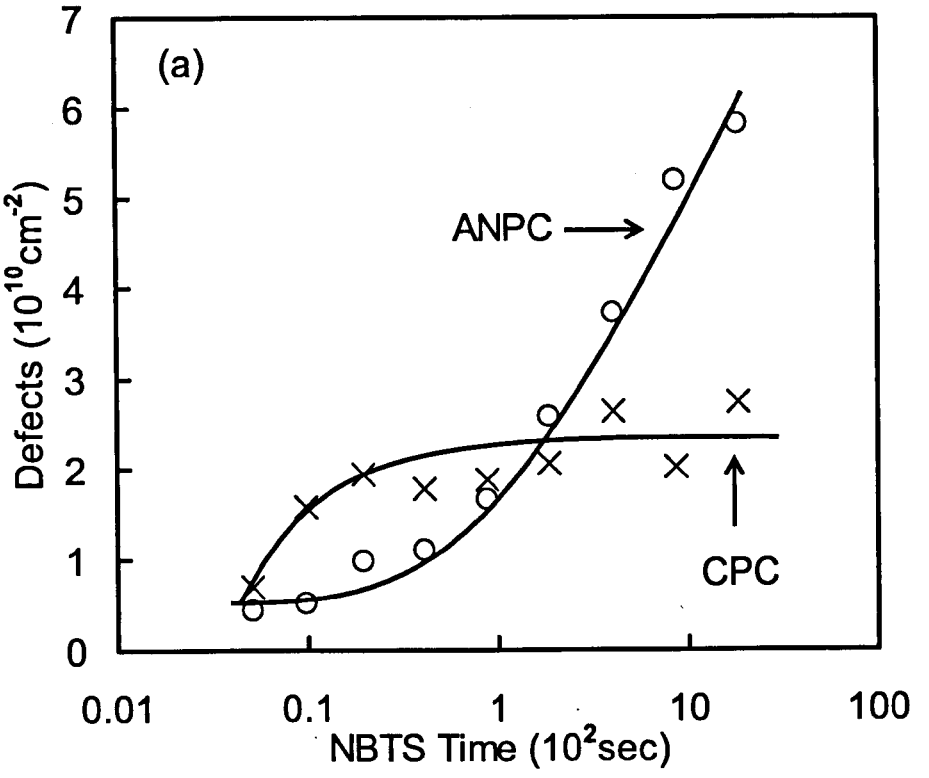


Figure 3. Two devices were subjected to NBTS at 200°C for 760 sec (symbol '●') and 10 sec (symbol '○'), respectively. After stresses, both devices were cooled down to room temperature. Some charges could be lost during this cooling period of 4 min. Therefore, a short NBTS was resumed at room temperature in an attempt to regain some of the lost charges. Afterwards, a negative and positive gate bias ($|E_{ox}|$ = 5 MV/cm) was alternately applied with all other terminals grounded. The arrows labeled A point out that CPC generation is similar on both devices. This suggests that CPC is pre-determined during the device fabrication process. Once all the defects are consummated, CPC would no longer respond to stress time. Moreover, the arrows labeled B point out that ANPC generation is clearly stress time dependent, longer stress time leads to greater generation of ANPC.


(a)
ANPC
CPC
Defects (10^{10}cm^{-2})
NBTS Time (10^2sec)
0
1
2
3
4
5
6
7
0.01
0.1
1
10
100

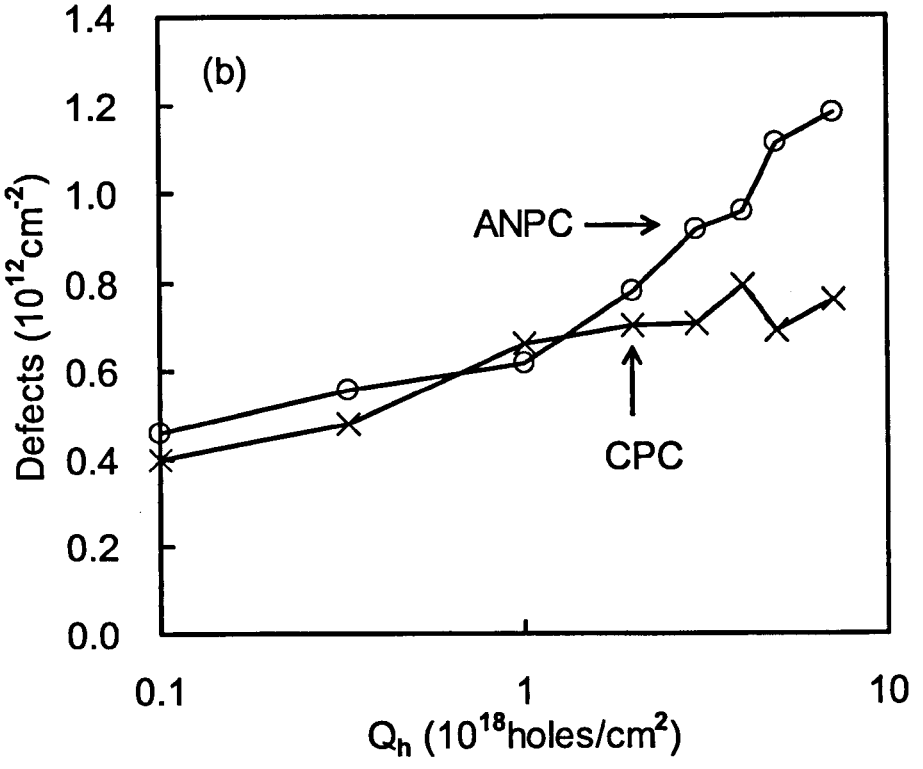

(b)
ANPC
CPC
Defects (10^{12}cm^{-2})
Q_h (10^{18}holes/cm^2)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
0.1
1
10

Figure 4. Dependence of different types of positive charges on stress levels. In (a), NBTS was carried out at Vg = -3.17 V and 100°C. In (b), SHI was carried out at E_{ox} = -5 MV/cm, a n-well bias of 6 V and a p-substrate bias of 7 V [27]. CPC saturates at high stress levels in both cases, while ANPC does not.

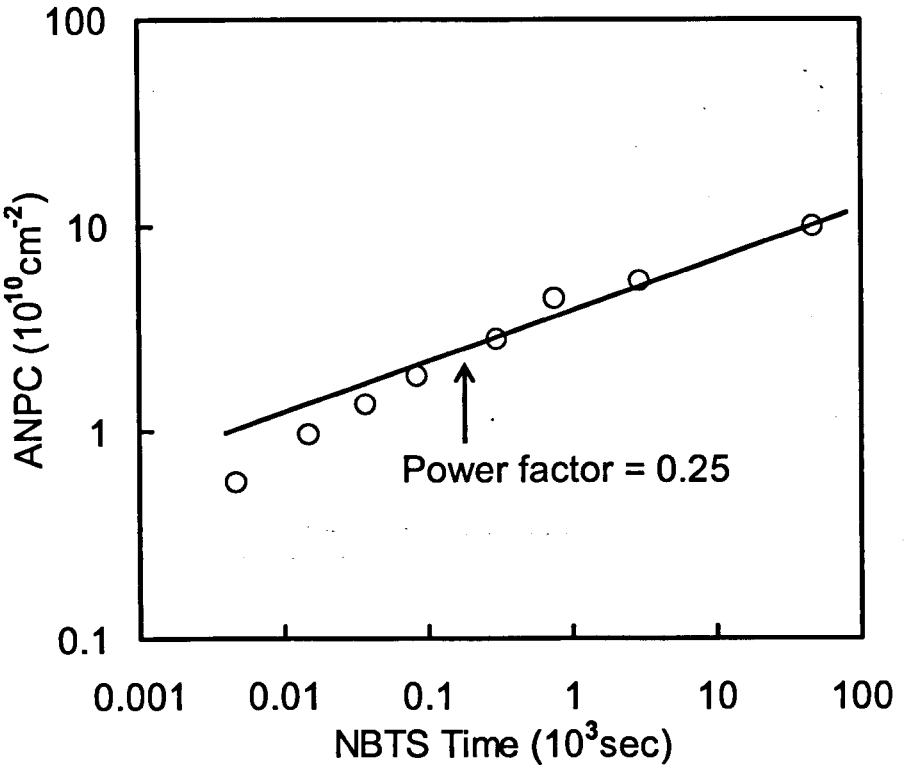
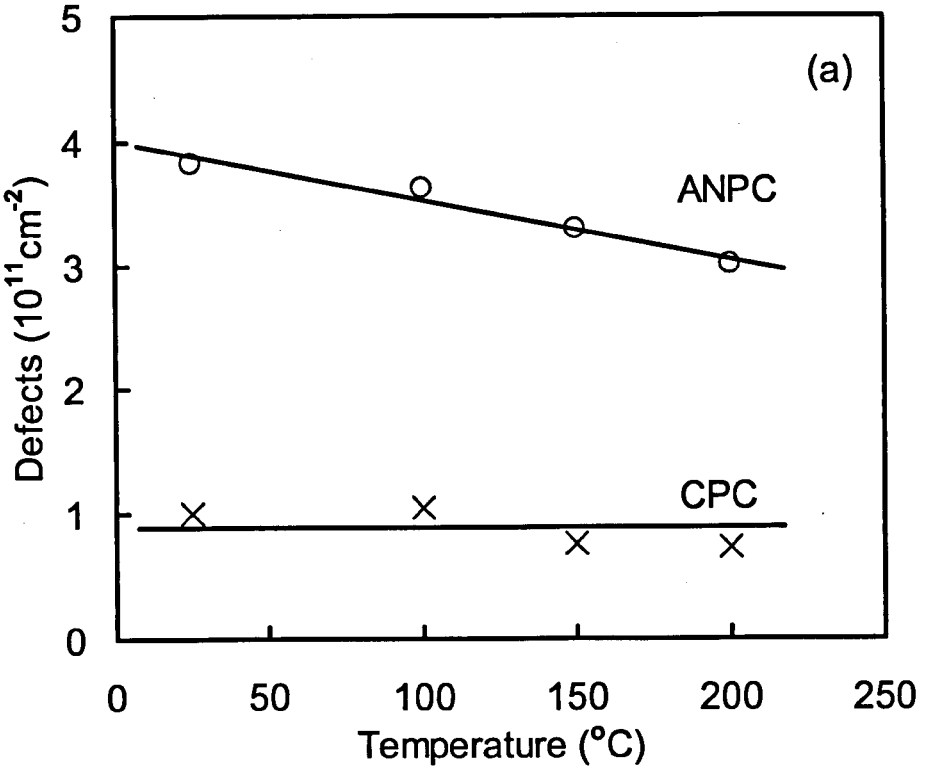


Figure 5. ANPC generation is plotted in log-log scale. After fitting the data, a power factor of 0.25 was extracted. When compared to the power factor extracted in Figure 1, it is apparent that the build-up of positive charges during NBTS is dominated by ANPC generation.


(a)
ANPC
CPC
Defects ($10^{11}cm^{-2}$)
Temperature (°C)
0
1
2
3
4
5
0
50
100
150
200
250

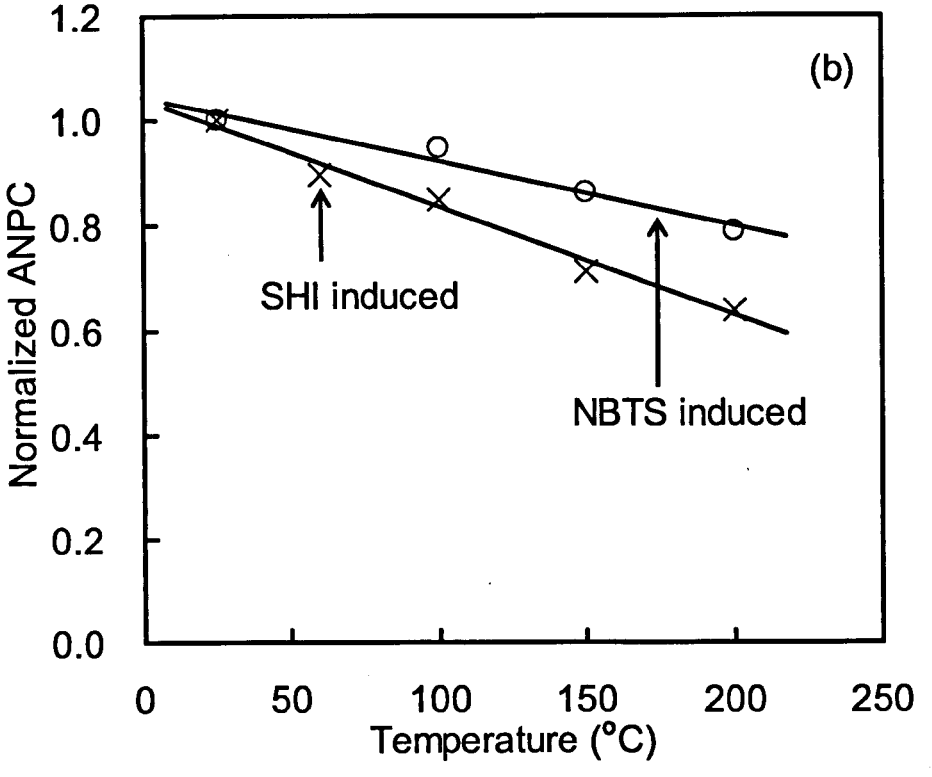
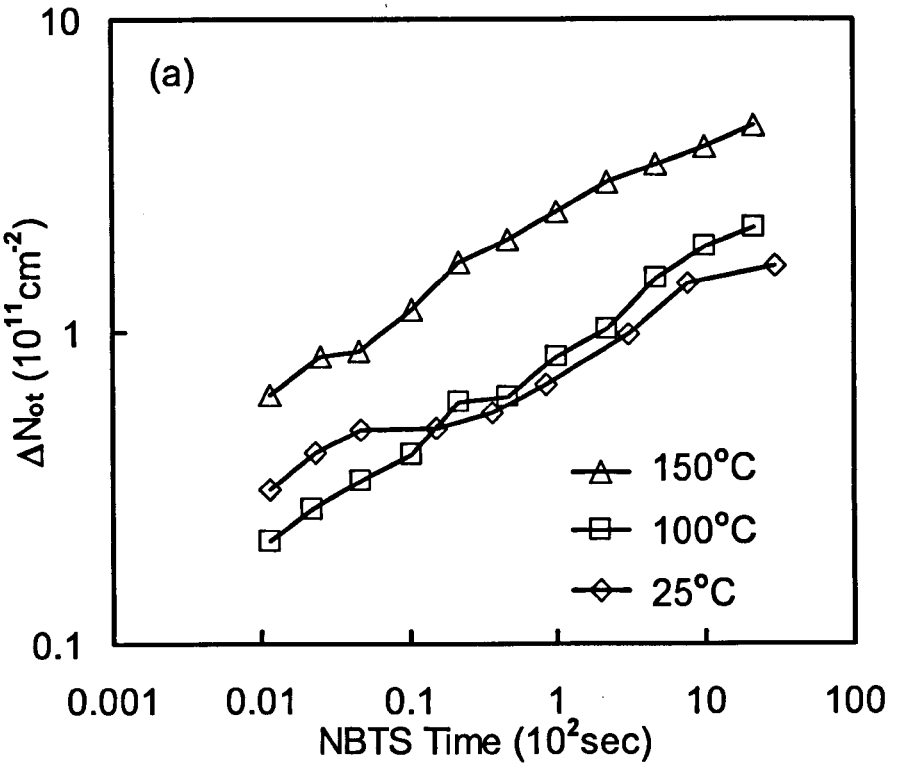
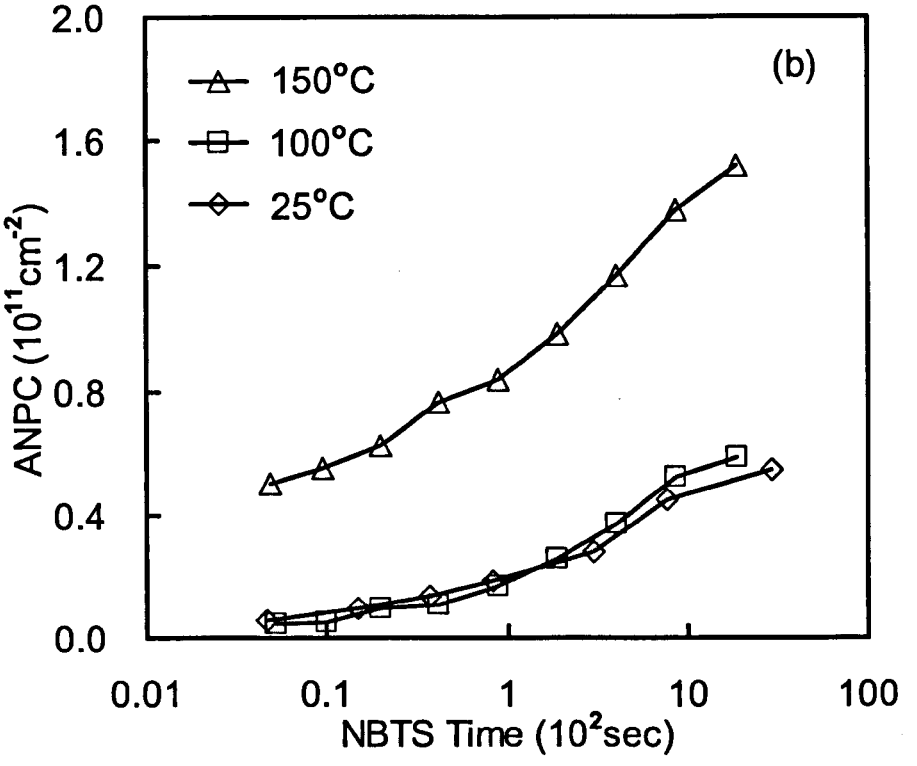

(b)
SHI induced
NBTS induced
Normalized ANPC
Temperature (°C)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
50
100
150
200
250

Figure 6. After NBTS, ANPC and CPC were measured at room temperature. Afterwards, the device was exposed to elevated temperature from 100°C to 200°C with all terminals floating. After the exposure, the device was cooled down, so the annealing behavior of ANPC and CPC can be monitored at room temperature. (a) Comparison of the annealing behavior of ANPC and CPC. CPC remain stable throughout the annealing, while ANPC reduced considerable after temperature increased. (b) Compared with SHI induced ANPC annealing behavior. In both case, ANPC was normalized against initial value before annealing. The similar dependence of ANPC on annealing temperature further supports our early suggestion that positive charges formed under different stresses have common origins.


(a)
10
1
0.1
ΔN_{ot} ($10^{11}cm^{-2}$)
150°C
100°C
25°C
0.001
0.01
0.1
1
10
100
NBTS Time (10^2sec)


(b)
2.0
1.6
1.2
0.8
0.4
0.0
ANPC ($10^{11}cm^{-2}$)
150°C
100°C
25°C
0.01
0.1
1
10
100
NBTS Time (10^2sec)

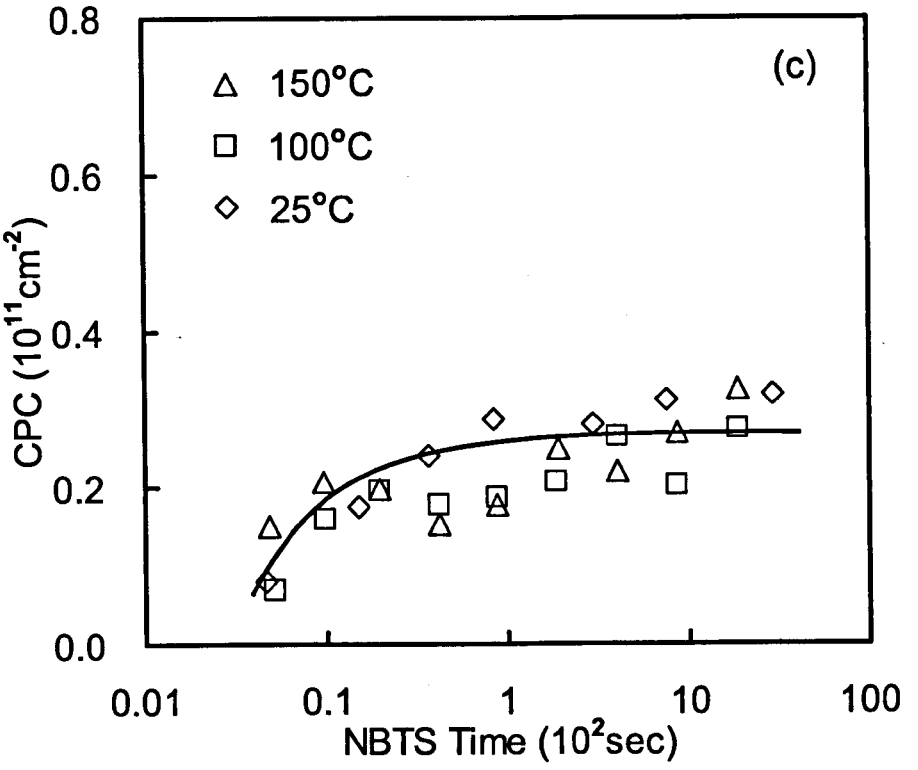


Figure 7. Generation of positive charges during NBTS at different temperatures. The measurement temperature was the same as the stres temperature. The gate bias is -3.17 V in all cases. The effective density of total positive charges, ΔN_{ot}, ANPC, and CPC are shown in (a), (b) and (c), respectively. CPC is insensitive to the stress temperature, while both ΔN_{ot} and ANPC increases substantially when temperature raised from 100°C to 150°C.

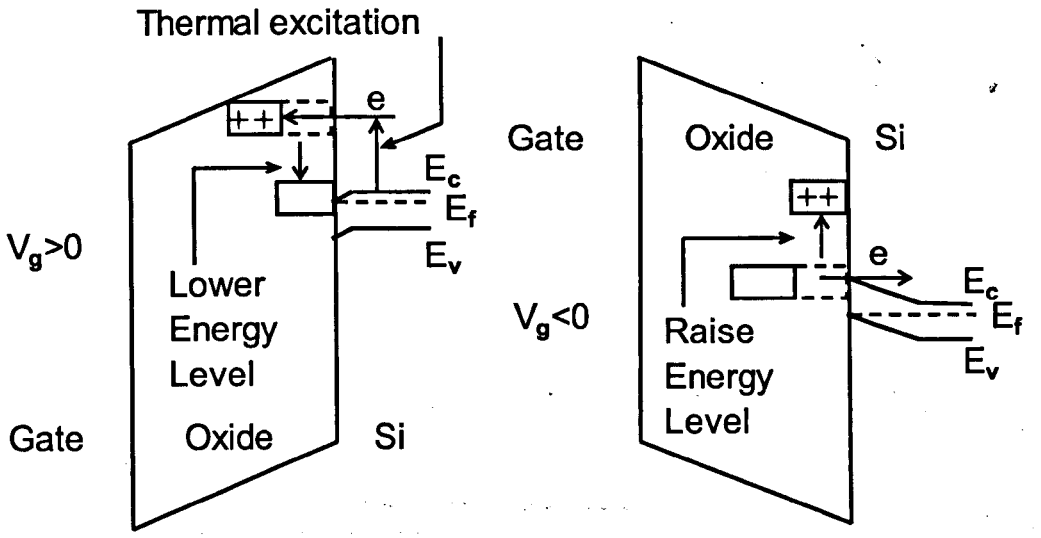


Figure 8. Schematic energy band diagrams for anti-neutralization positive charges (ANPC) during (a) discharging (b) charging. When positively charged, the energy level of ANPC is above the bottom edge of Si conduction band (E_c). After neutralization, however, it is lowered.

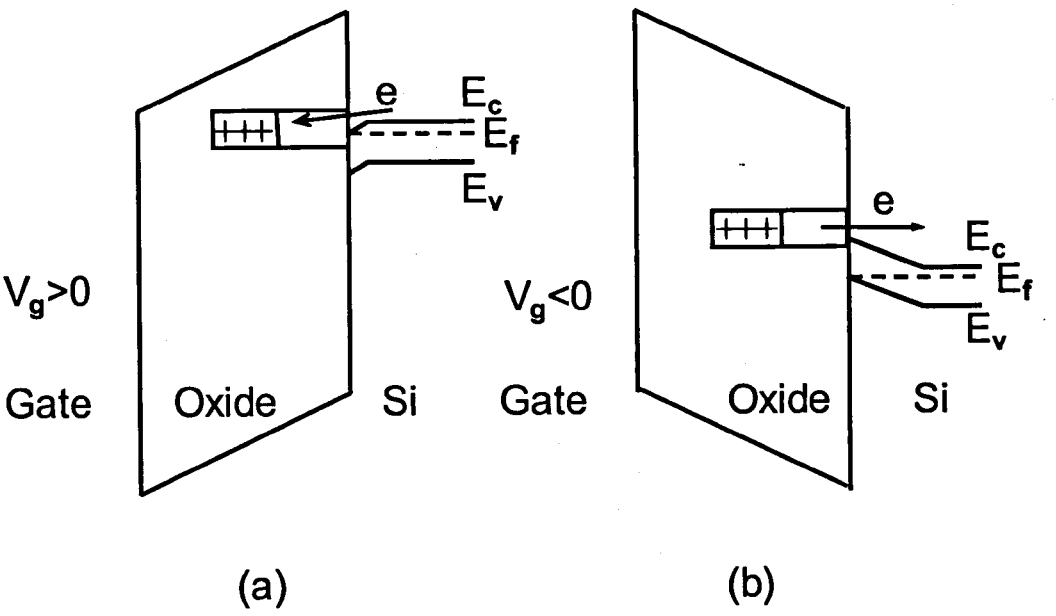


Figure 9. Schematic energy band diagrams for cyclic positive charges (CPC) during (a) discharging (b) charging. The energy level of CPC is close to the bottom edge of Si conduction band (E_c) and does not appear to change for different CPC charge status.

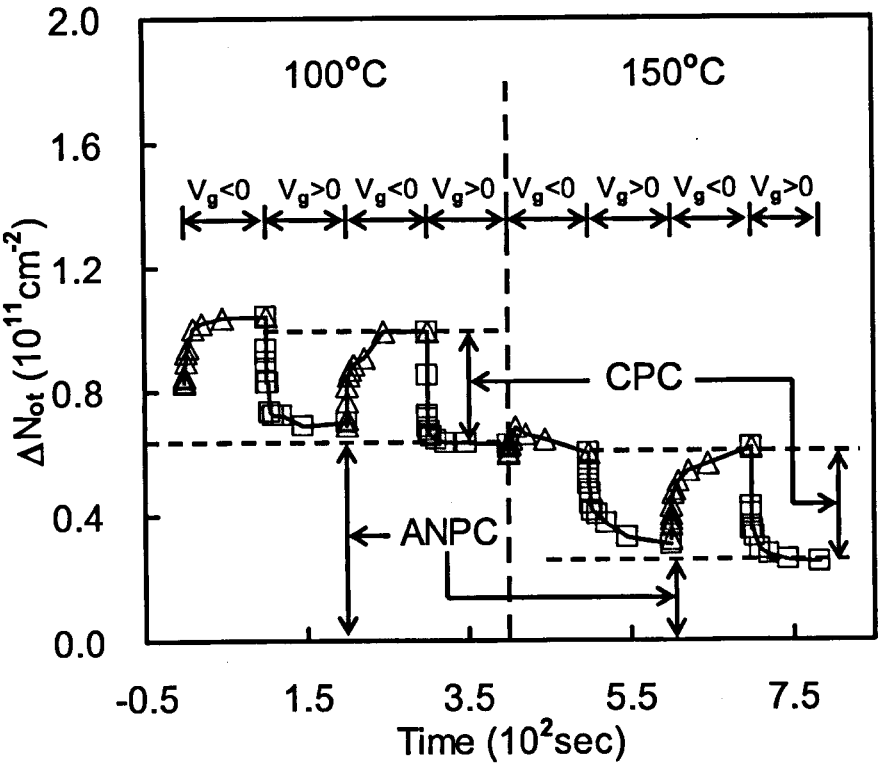
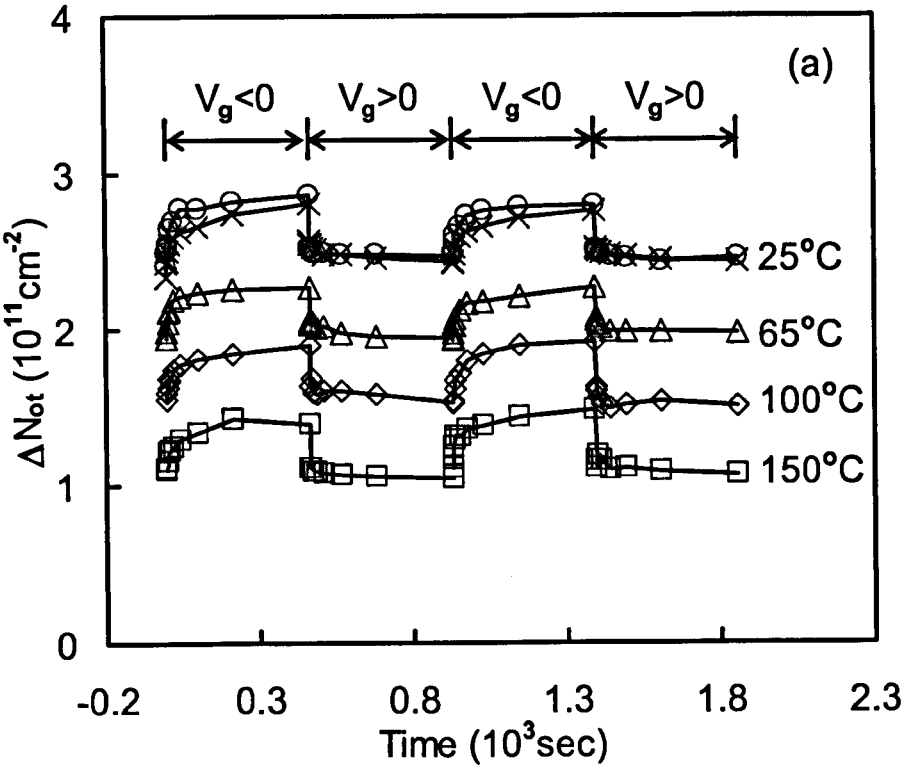


Figure 10. Effective density of positive charges measured at different temperatures. The device was stressed by NBTS at V_g = -3.17 V and 100°C for 2000 sec. CPC is clearly insensitive to measurement temperature, but ANPC reduces for higher temperature.


(a)
V_g<0
V_g>0
V_g<0
V_g>0
25°C
65°C
100°C
150°C
ΔN_{ot} ($10^{11}cm^{-2}$)
Time (10^3sec)
0
1
2
3
4
-0.2
0.3
0.8
1.3
1.8
2.3

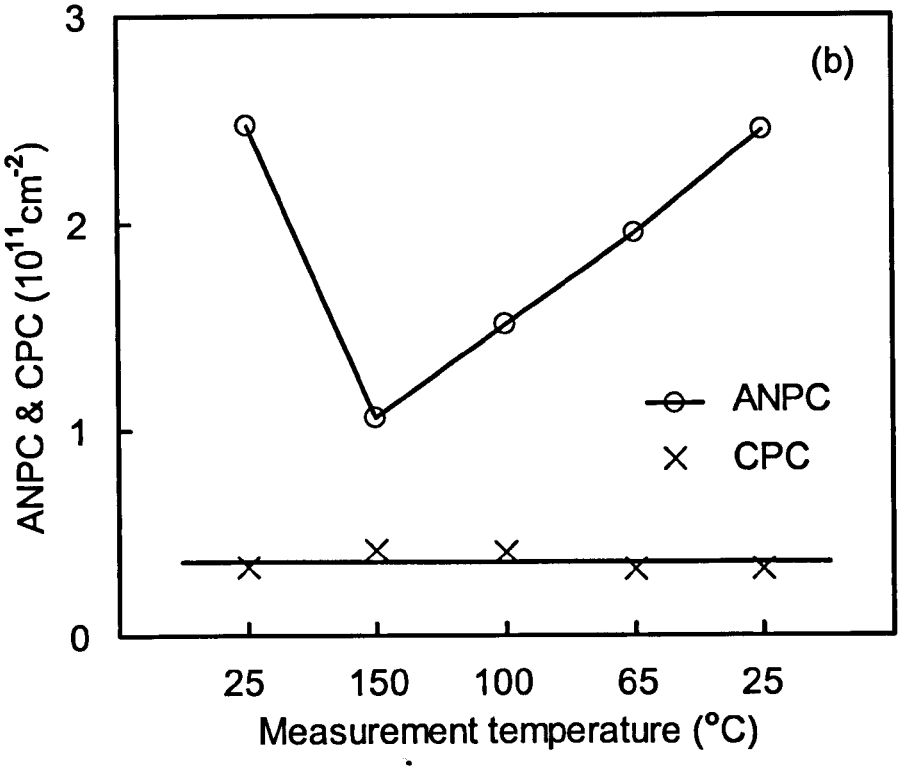

(b)
ANPC
CPC
ANPC & CPC ($10^{11}cm^{-2}$)
Measurement temperature (°C)
0
1
2
3
25
150
100
65
25

Figure 11. A device was heavily stressed under NBTS at 200°C for 500 sec, in order to create a large amount of ANPC. (a) ANPC and CPC were firstly measured at room temperature (symbol 'o'), then the measurement temperature was raised to 150°C (symbol '□'). As expected, CPC is insensitive to measurement temperature, but ANPC reduces substantially as measurement temperature increases. (b) When measurement temperature was lowered to 100°C, larger ANPC was actually measured compared to 150°C. More and more ANPC could be measured as measurement temperature drops further. This confirms that ANPC is highly dependent on the measurement temperature. Lower measurement temperature will reveal larger amount of ANPC.

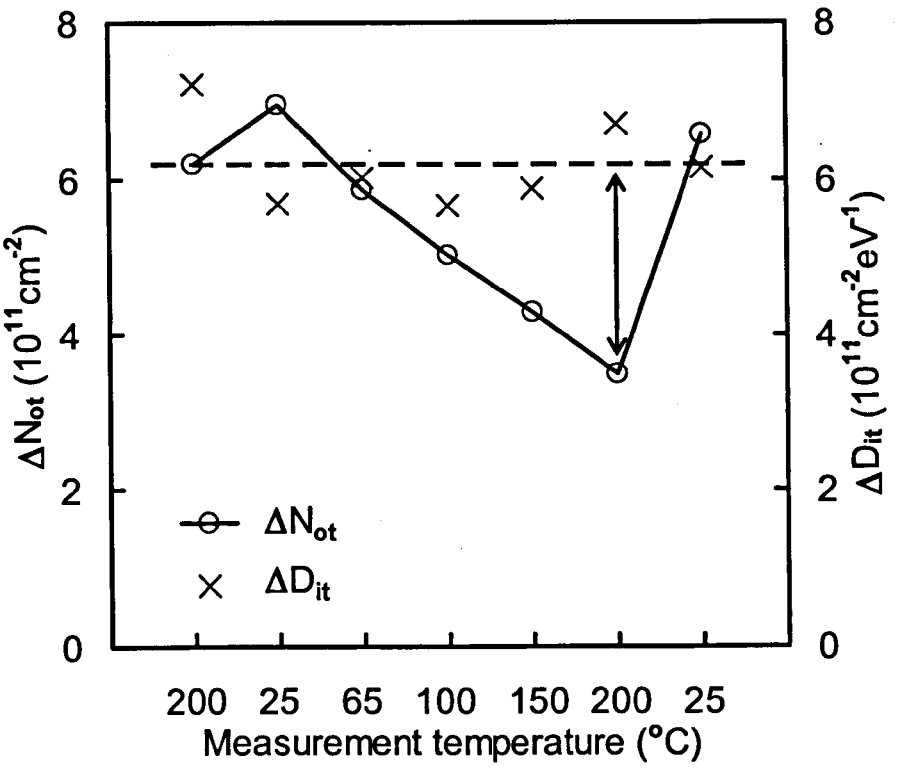
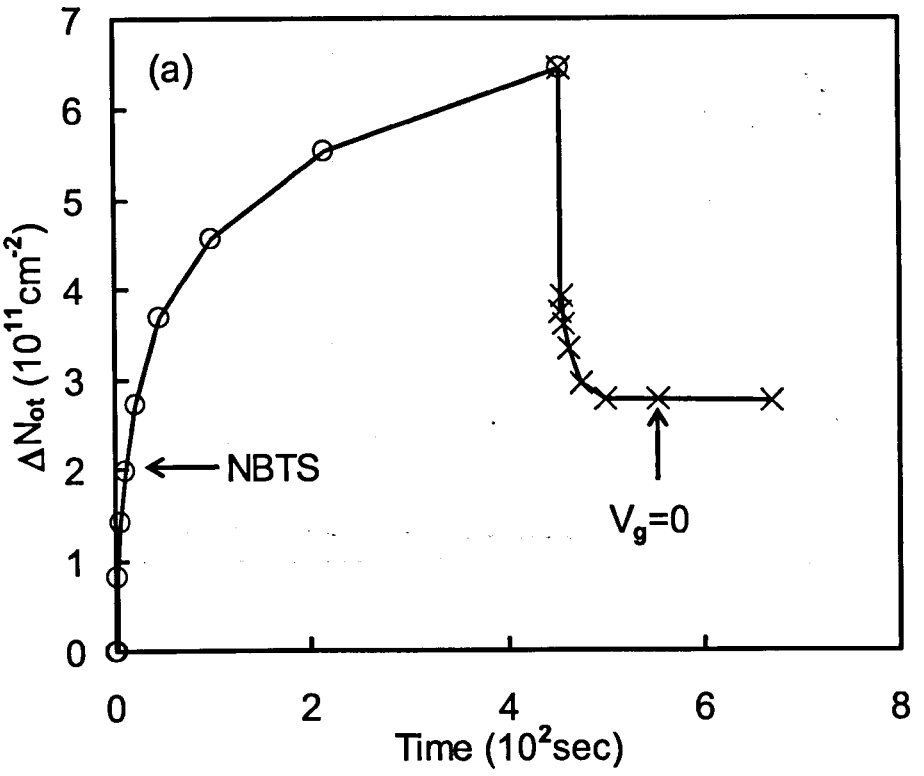
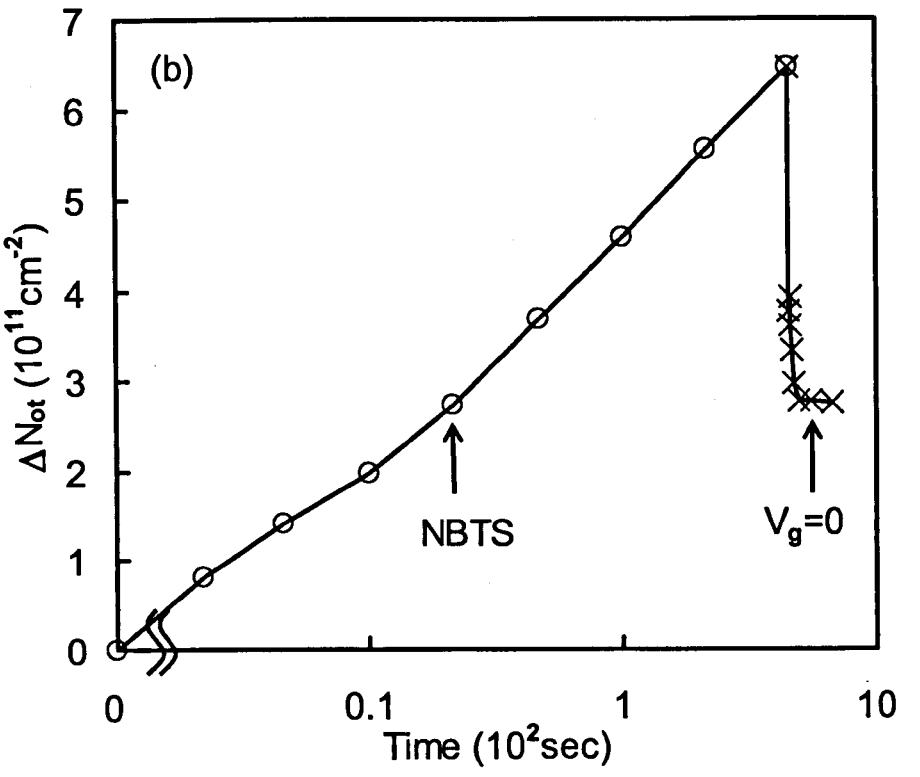


Figure 12. Effects of measurement temperature on the density of positive charges and interface states. The device was stressed at -3.17 V and 200°C for 200 sec (not shown). It was then measured at 200°C to give the first data point, 'o' and 'x', in the figure from the left. This is followed by measurement at different temperatures, in the sequence of 25°C, 65°C, 100°C, 150°C, and 200°C. Finally, the device was cooled down and measured at 25°C again. The two measurements for positive charges are considerably different at 200°C, but approximately the same at 25°C. The measurement of interface states is insensitive to temperature.


(a)
ΔN_{ot} ($10^{11}cm^{-2}$)
NBTS
V_g=0
Time (10^2sec)
0
1
2
3
4
5
6
7
0
2
4
6
8


(b)
ΔN_{ot} ($10^{11}cm^{-2}$)
NBTS
V_g=0
Time (10^2sec)
0
1
2
3
4
5
6
7
0
0.1
1
10

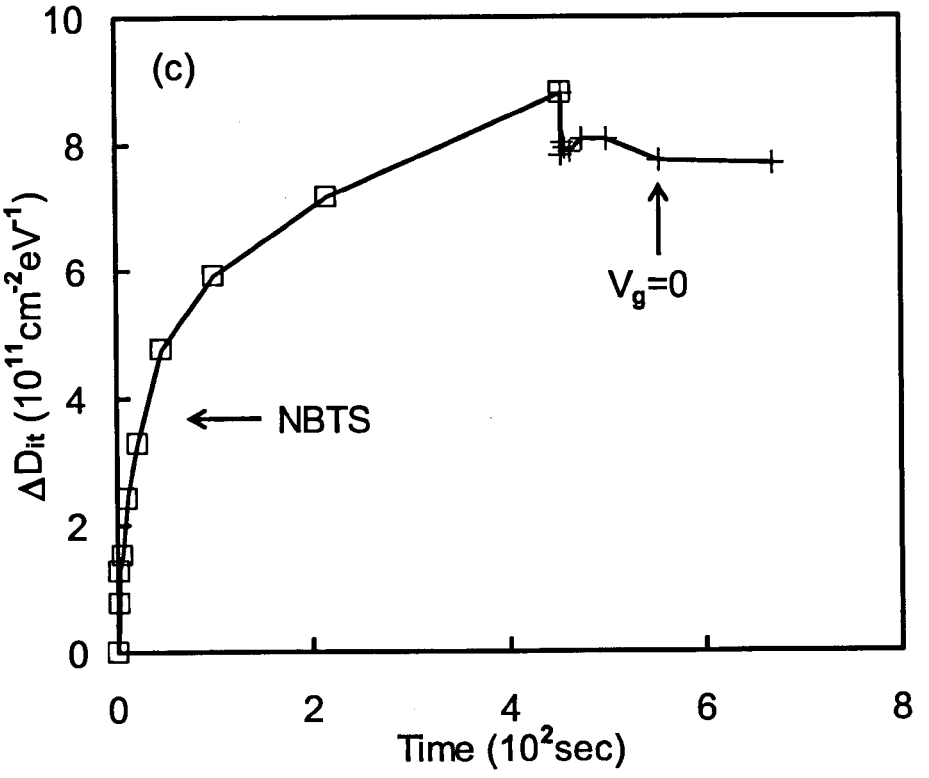


Figure 13. Generation and annealing of positive charges. The transient behavior is shown against time in (a) linear and (b) logarithmic scale. The NBTS was carried out at V_g = -3.17 V and 200°C for 450 sec. The device was then kept at 200°C with V_g = 0. It can be seen that positive charges were annealed rapidly when V_g was set to zero. They were stabilized at longer time. (c) Generation and annealing of interface states. It appears that little interface states were annealed when V_g was set to zero.

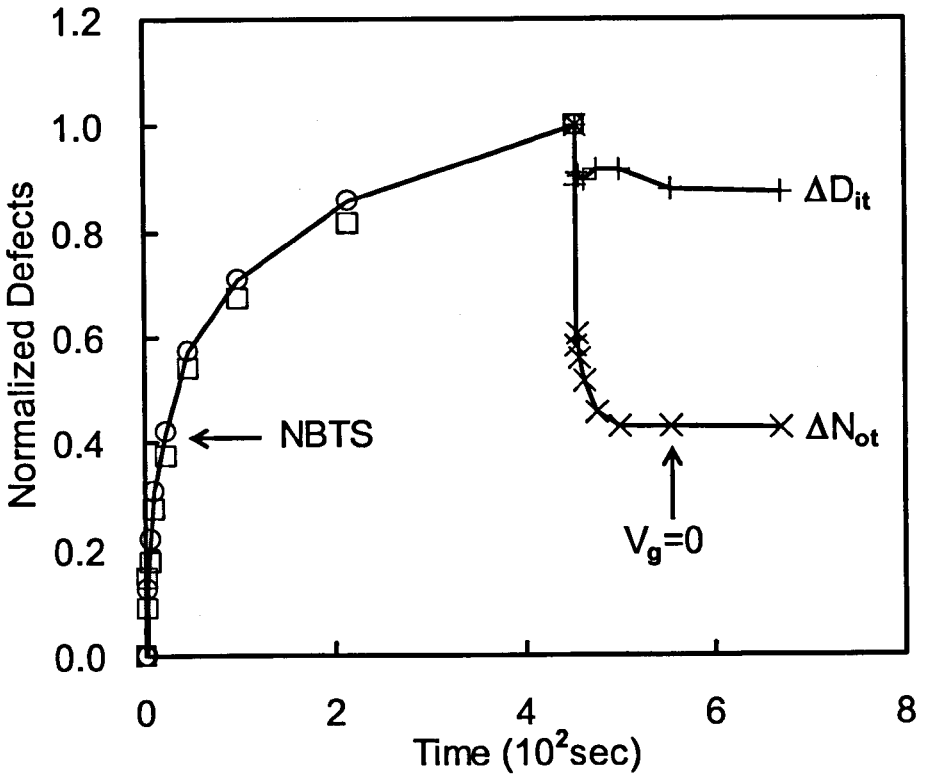
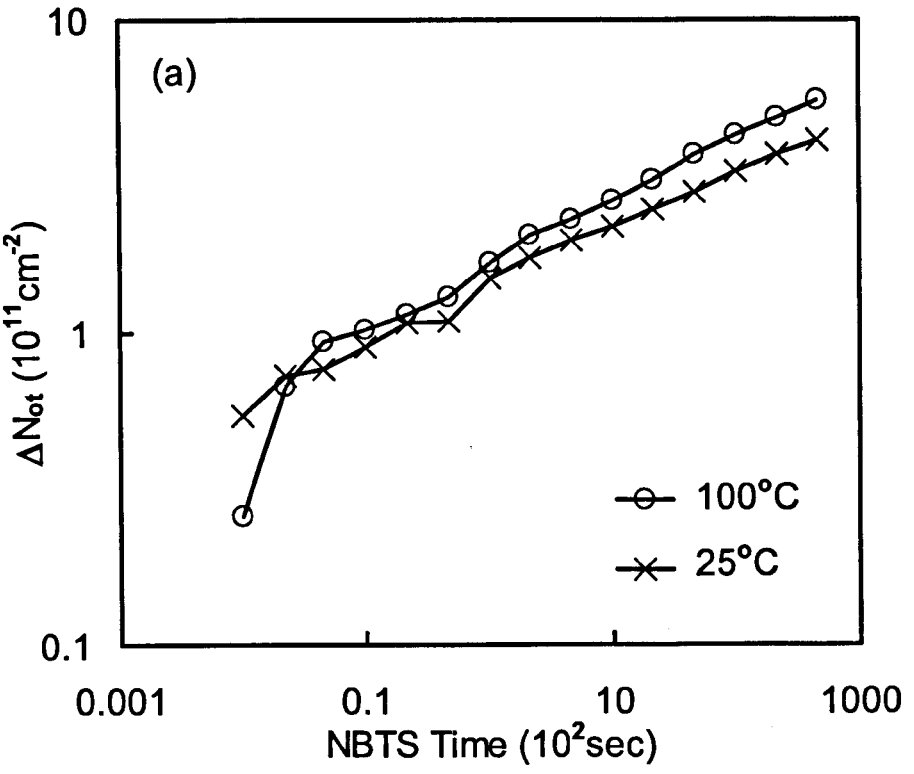
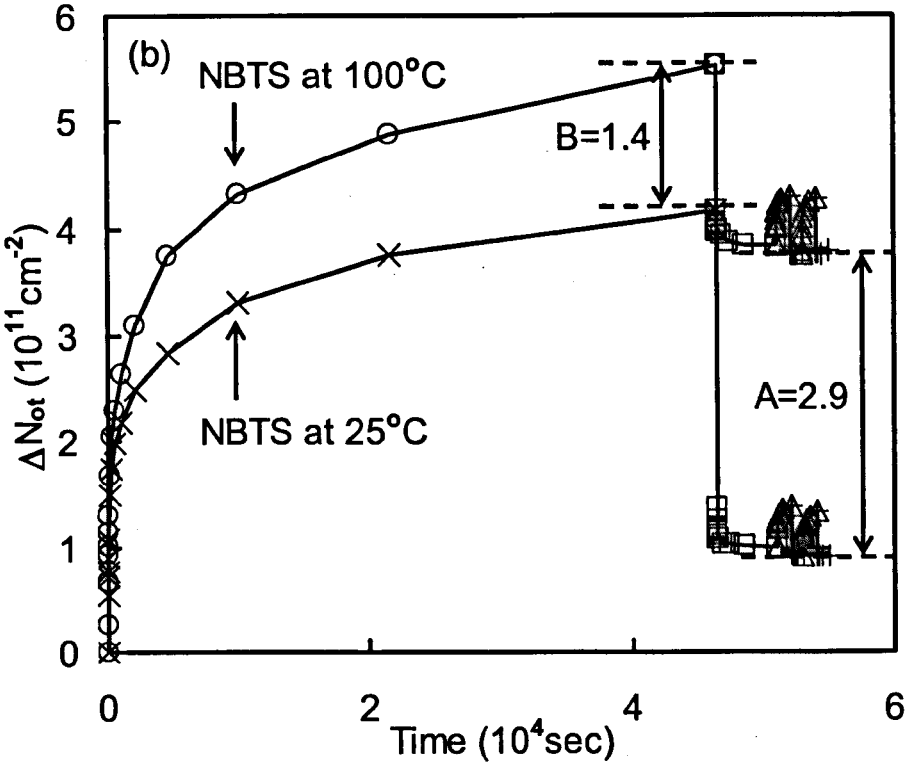


Figure 14. Generated defect were normalized against their maximum value. Sympol '○' represents the generation of positive charges, and symbol '□' represents the generation of interface states. It can be seen that when V_g was set to zero, the annealing of interface states (symbol '+') is significantly less than that of positive charges (symbol 'x'). The annealing is mainly dominated by positive charges, as they are unstable. But they will both stabilize at longer time.


(a)
10
1
0.1
ΔN_{ot} ($10^{11}cm^{-2}$)
0.001
0.1
10
1000
NBTS Time (10^2sec)
100°C
25°C


(b)
NBTS at 100°C
NBTS at 25°C
B=1.4
A=2.9
6
5
4
3
2
1
0
ΔN_{ot} ($10^{11}cm^{-2}$)
0
2
4
6
Time (10^4sec)

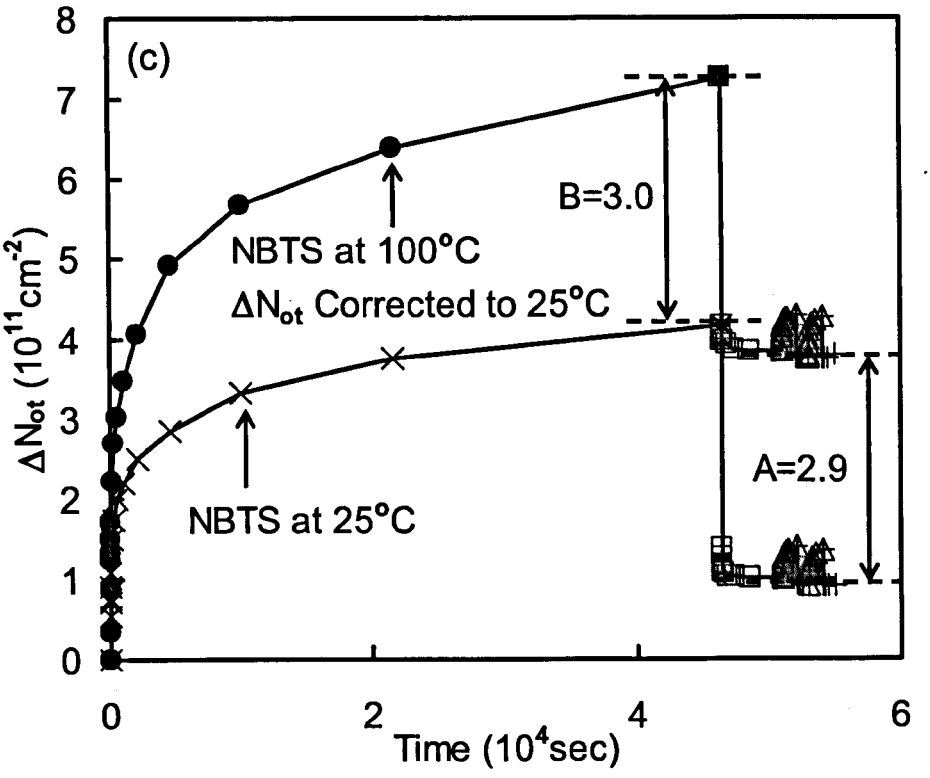
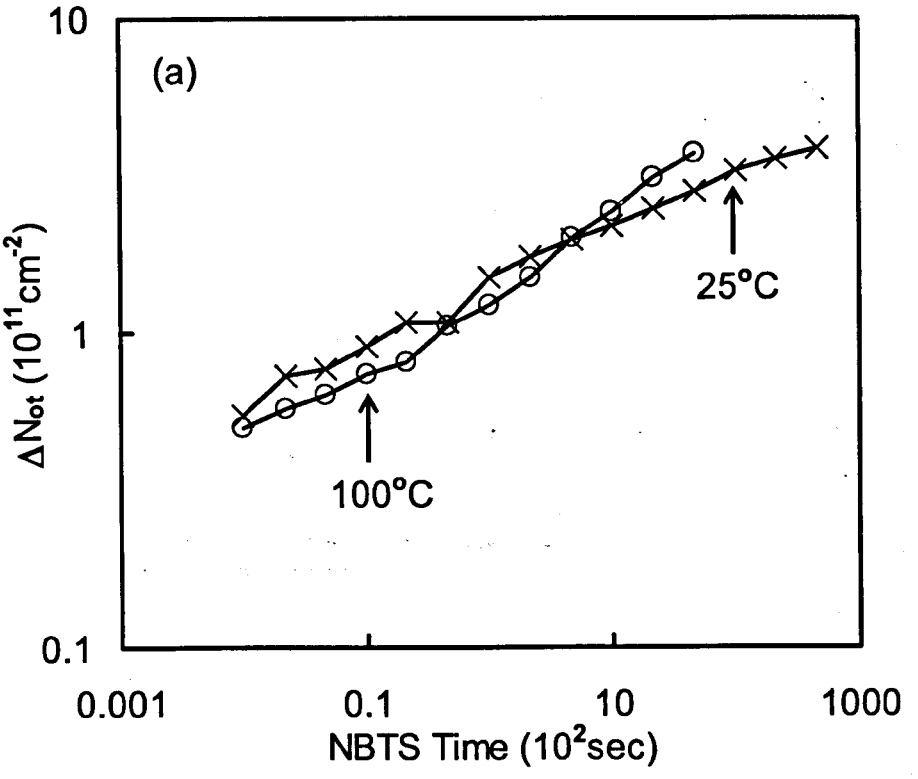
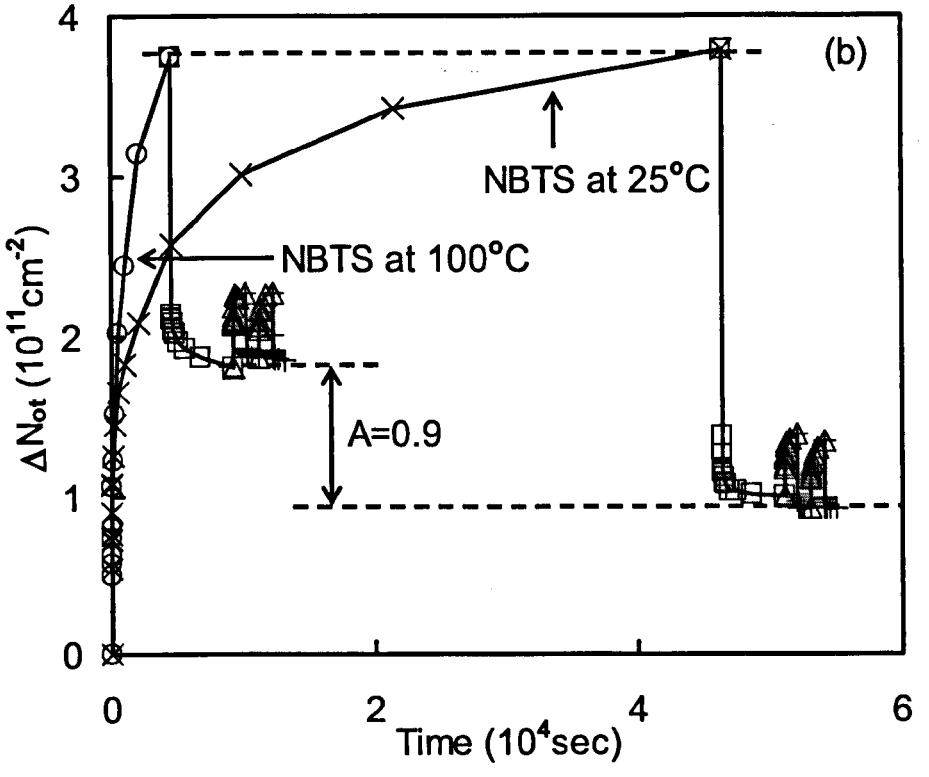


Figure 15. (a) Two devices were subjected to NBTS for 13 hours: one at 25°C and the other one at 100°C. (b) After NBTS, the device stressed at elevated temperature was allowed to cool down to room temperature. Then, both devices went through a cycle of positive and negative gate biases ($|E_{ox}|$ = 5 MV/cm) with all other terminals grounded. The arrow 'A' points out significant increase of ANPC generation after NBTS at 100°C. (c) Every data measured at 100°C were corrected to 25°C using a fixed ratio obtained from Figure 12. It is apparent that after correction, the difference in ΔN_{ot} (arrow 'B') is similar to the increase of ANPC (arrow 'A'). This further supports that ANPC is the only thermally enhanced component.


(a)
10
1
0.1
ΔNot (10^11cm^-2)
25°C
100°C
0.001
0.1
10
1000
NBTS Time (10^2sec)


(b)
4
3
2
1
0
ΔNot (10^11cm^-2)
NBTS at 25°C
NBTS at 100°C
A=0.9
0
2
4
6
Time (10^4sec)

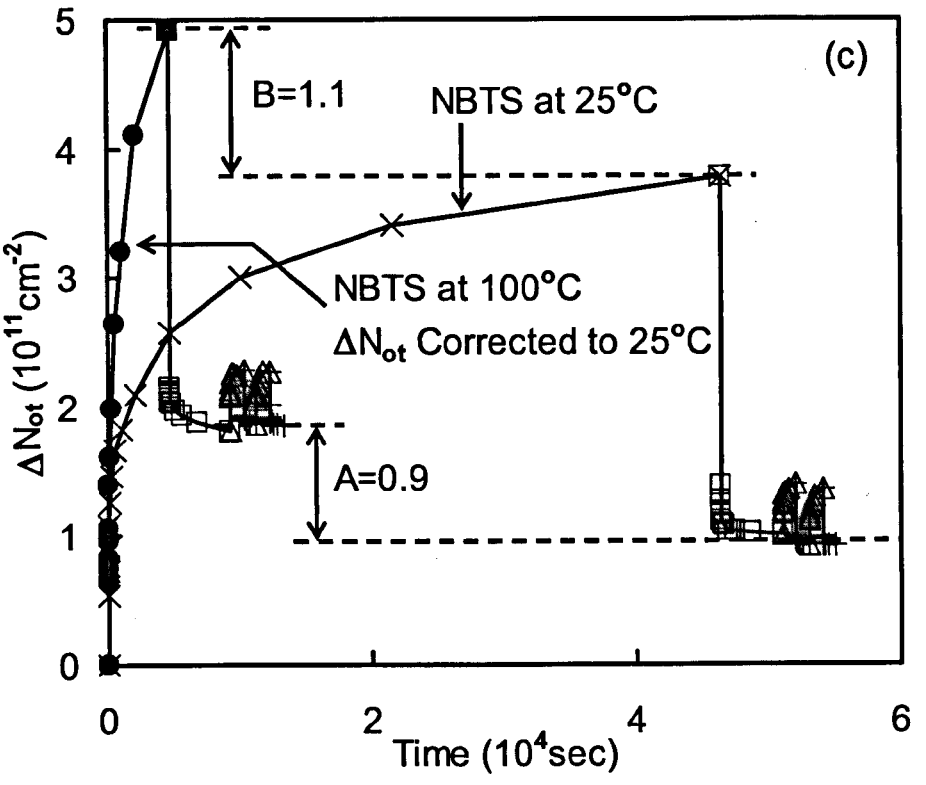


Figure 16. Similar to previous experiment, where two devices were subjected to NBTS: one at 25°C and the other one at 100°C. (a) The one stressed at elevated temperature was interrupted as soon as ΔN_{ot} reached the level of NBTS at 25°C. (b) After NBTS, the device stressed at elevated temperature was allowed to cool down to room temperature. Then, both devices went through a cycle of positive and negative gate biases ($|E_{ox}|$ = 5 MV/cm) with all other terminals grounded. Despite having shorter NBTS time at 100°C, the arrow 'A' clearly indicates an obvious increase of ANPC post 100°C NBTS. (c) Every data measured at 100°C were corrected to 25°C using a fixed ratio. Again, the increase of ANPC (arrow 'A') is approximately the same as the difference in ΔN_{ot} after correction (arrow 'B'). This strongly supports previous claim that only ANPC generation is thermally enhanced.

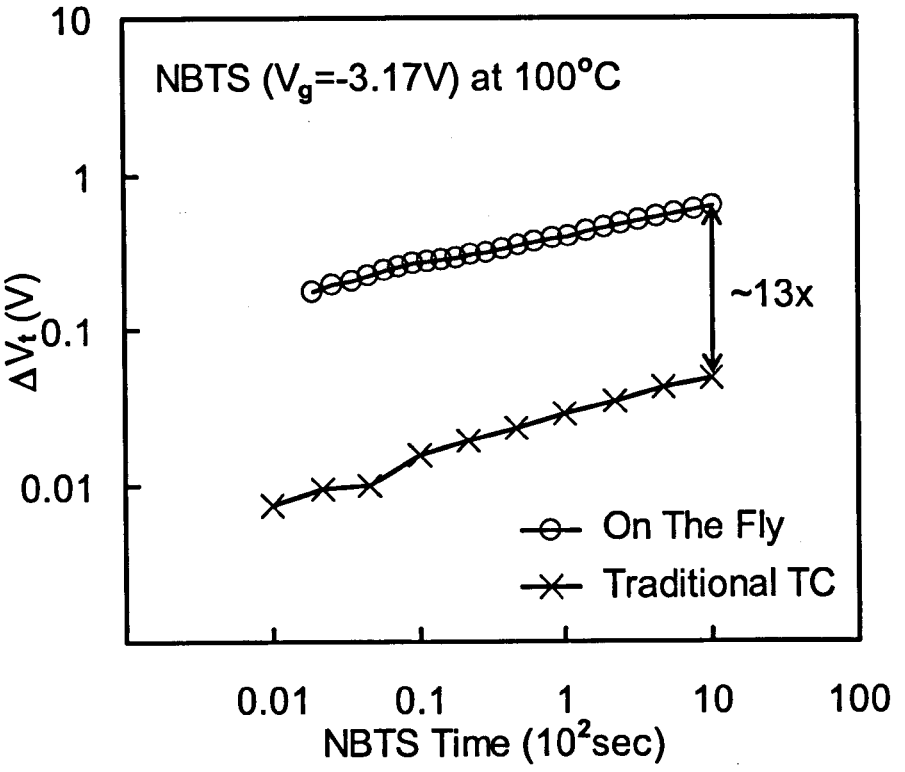


Figure 17. A comparison of ΔV_t measured by the 'On-The-Fly' method with that by the traditional transfer characteristics (TC) method. With the 'On-The-Fly' method, ΔV_t is one order of magnitude higher, resulting from the elimination of NBTI recovery.

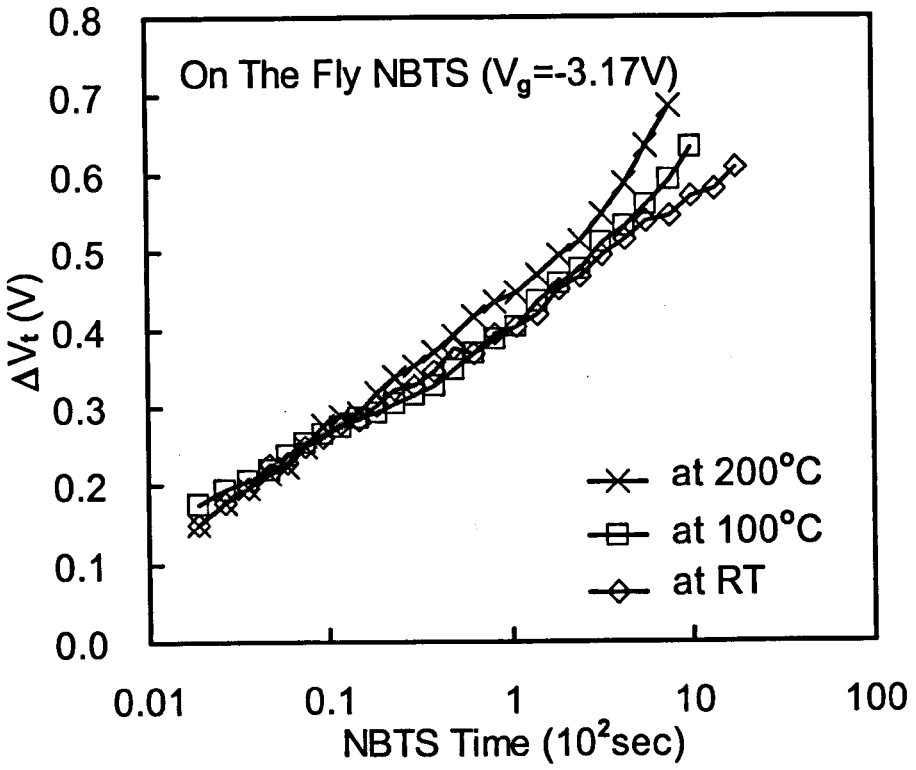


Figure 18. A comparison of ΔV_t measured by the 'On-The-Fly' method, when devices were stressed at different temperatures. The ΔV_t was measured at stress temperatures and appears insensitive to temperature.

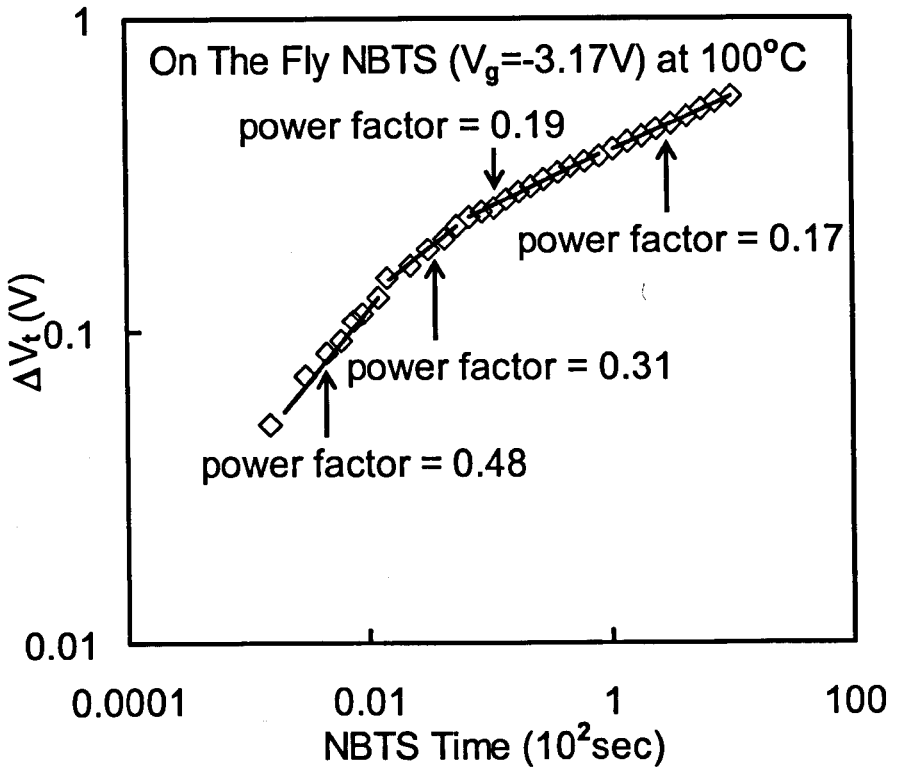
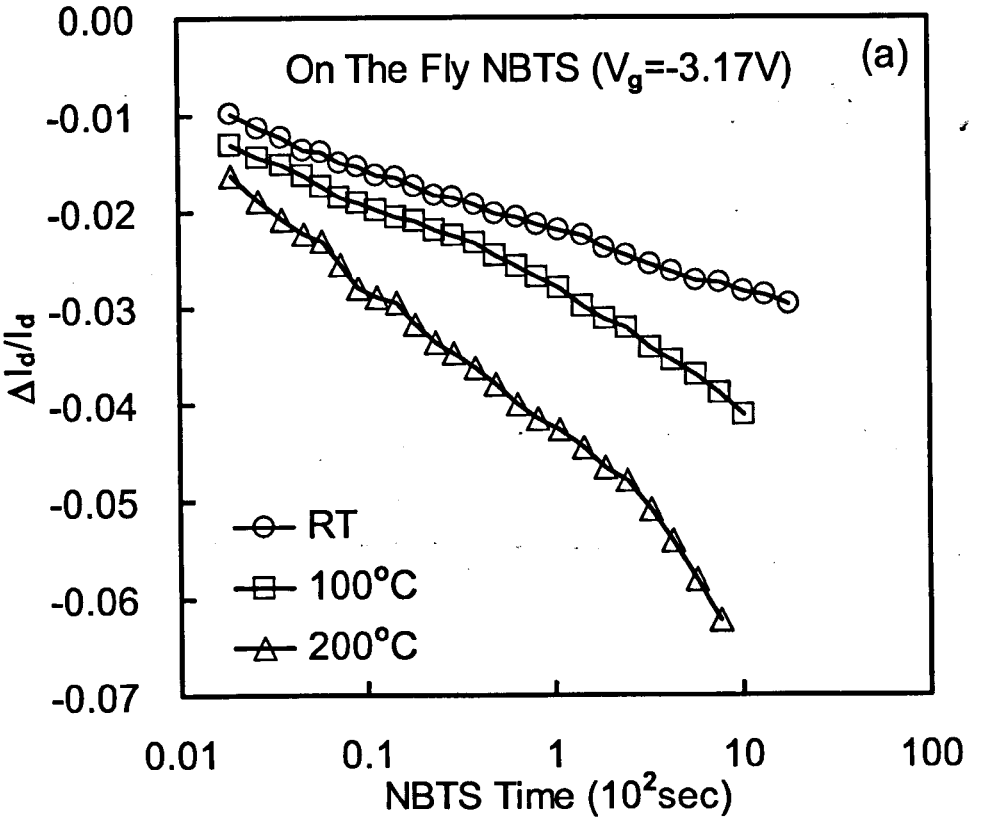
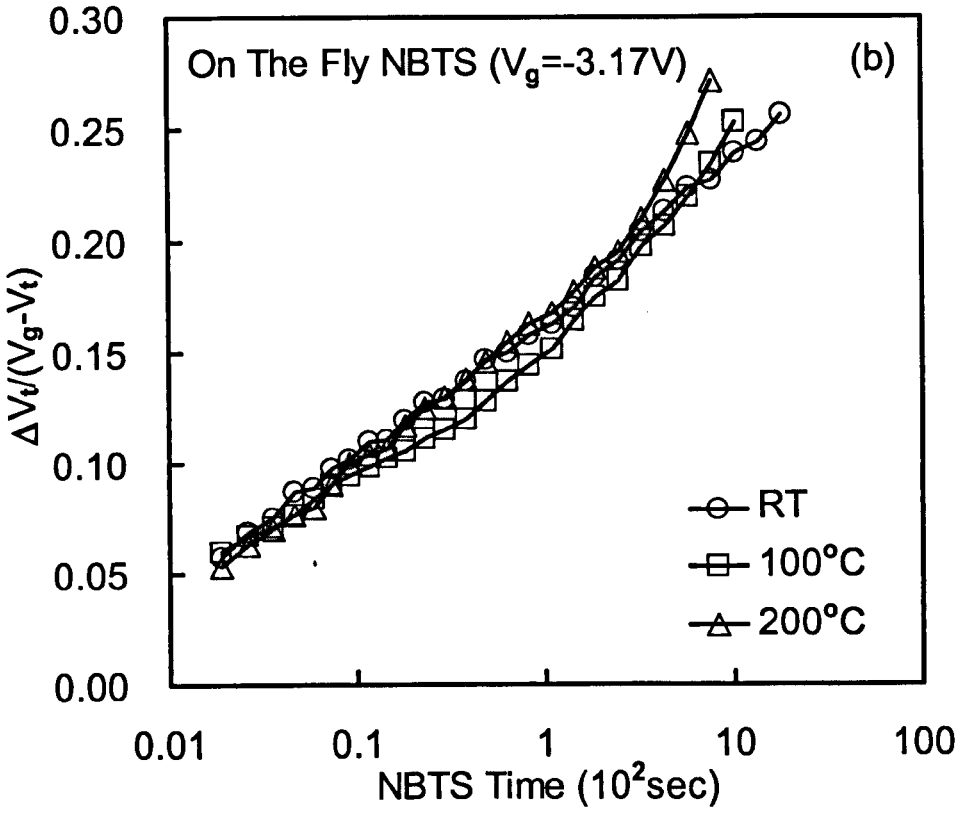


Figure 19. Time dependence of degradation. After eliminating the recovery between the stress and measurement, ΔV_t no longer follows the often quoted power factor of 0.25. Now the power factor varies from 0.48 to 0.17 between the measurement range of 2 ms to 1000 sec.


On The Fly NBTS (V_g=-3.17V)
(a)
$\Delta I_d/I_d$
0.00
-0.01
-0.02
-0.03
-0.04
-0.05
-0.06
-0.07
RT
100°C
200°C
0.01
0.1
1
10
100
NBTS Time (10^2sec)


On The Fly NBTS (V_g=-3.17V)
(b)
$\Delta V_t/(V_g-V_t)$
0.30
0.25
0.20
0.15
0.10
0.05
0.00
RT
100°C
200°C
0.01
0.1
1
10
100
NBTS Time (10^2sec)

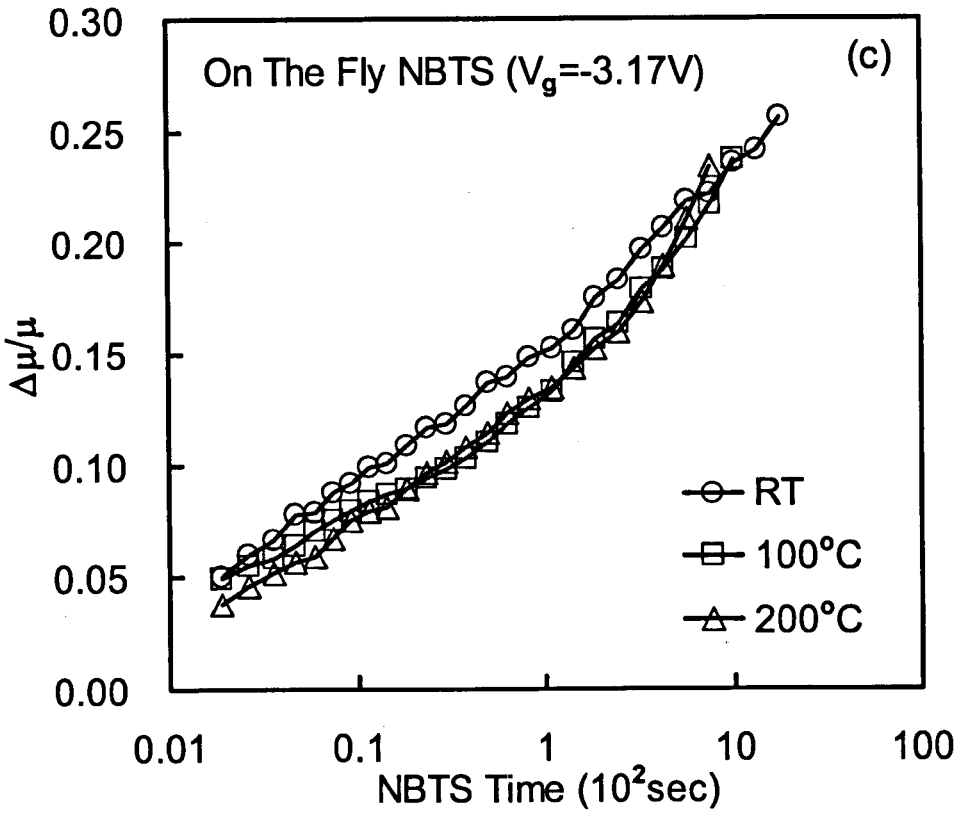
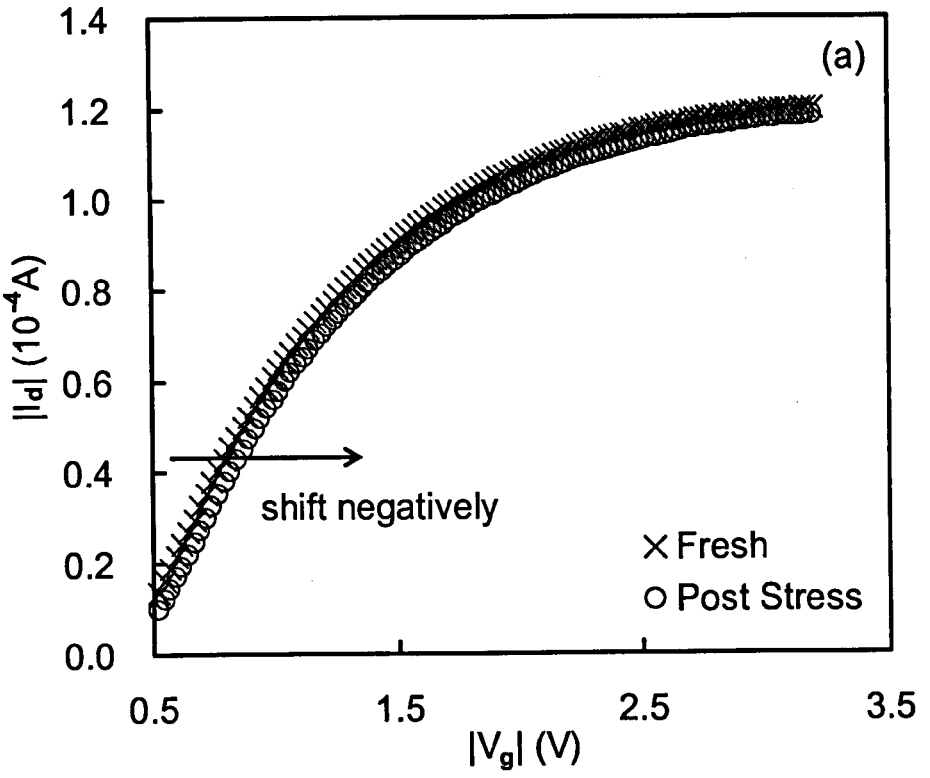


Figure 20. Three sets of data have been used here, from room temperature to 200°C. (a) I_d degradation is temperature dependent. At room temperature, I_d degradation is only ~ 3% at the end of stress, whereas at 200°C, it is almost doubled. (b) $\Delta V_t/(V_g - V_t)$ only changes modestly with temperature. (c) The effective mobility at the stress voltage is actually increased.


(a)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
$|I_d|$ (10^{-4}A)
shift negatively
× Fresh
○ Post Stress
0.5
1.5
2.5
3.5
$|V_g|$ (V)

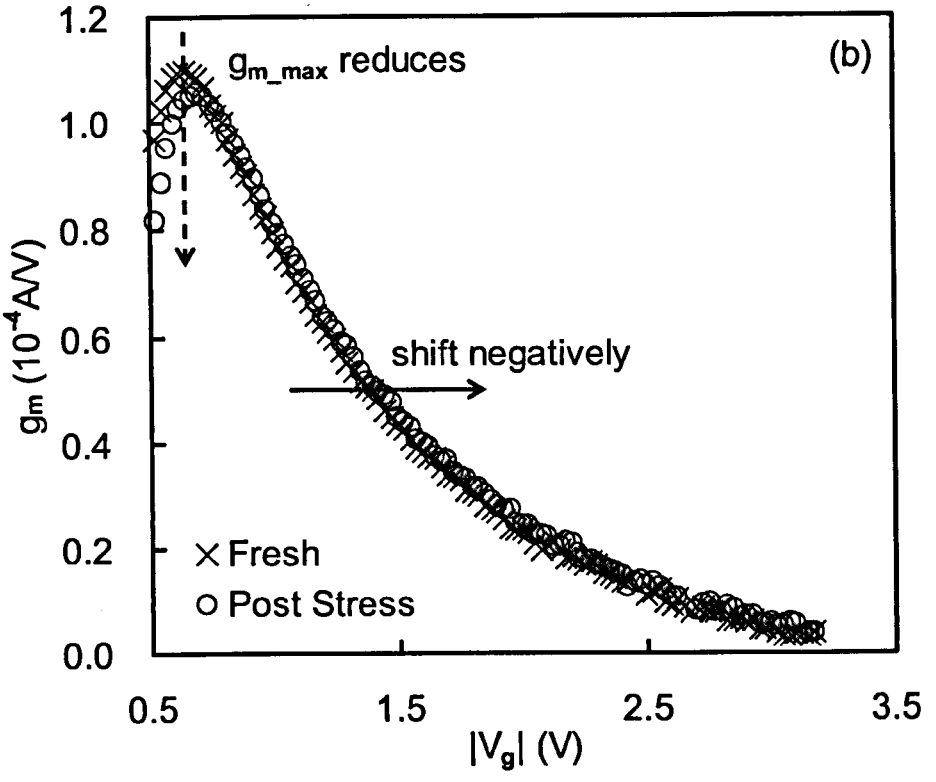

(b)
g_{m_max} reduces
1.2
1.0
0.8
0.6
0.4
0.2
0.0
g_m (10^{-4}A/V)
shift negatively
× Fresh
○ Post Stress
0.5
1.5
2.5
3.5
$|V_g|$ (V)

Figure 21. (a) $I_d \sim V_g$ scan was performed on a fresh device, from V_g = -0.4 V to -3.17 V (symbol 'x'). The device was then subjected to V_g = -3.17 V at 100°C for 1000 sec. Afterwards, another $I_d \sim V_g$ scan was carried out (symbol 'o'). The post stress $I_d \sim V_g$ curve shifts negatively, indicating the generation of positive charges. (b) From the $I_d \sim V_g$ curves, the g_m before (symbol 'x') and after stress (symbol 'o') are obtained. The maximum trans-conductance, g_{m_max}, is indeed reduced after NBTS, as pointed out by the dashed arrow. After the stress, the $I_d \sim V_g$ is shifted toward higher $|V_g|$. The g_m at a given $|V_g|$ = 3.17 V actually increased.

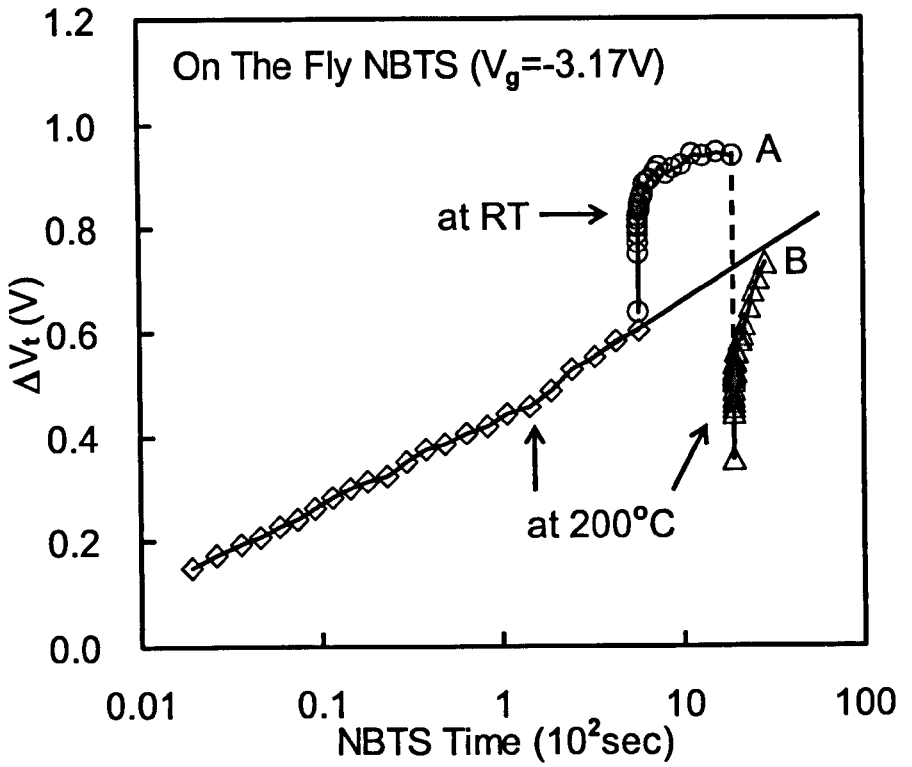


Figure 22. Effects of temperature on ΔV_t. The device was stressed and measured first at 200°C, then at room temperature, followed by 200°C again. It is clear that ΔV_t depends on measurement temperature and is higher at room temperature than at 200°C. The solid and dashed lines are guides for the eye.

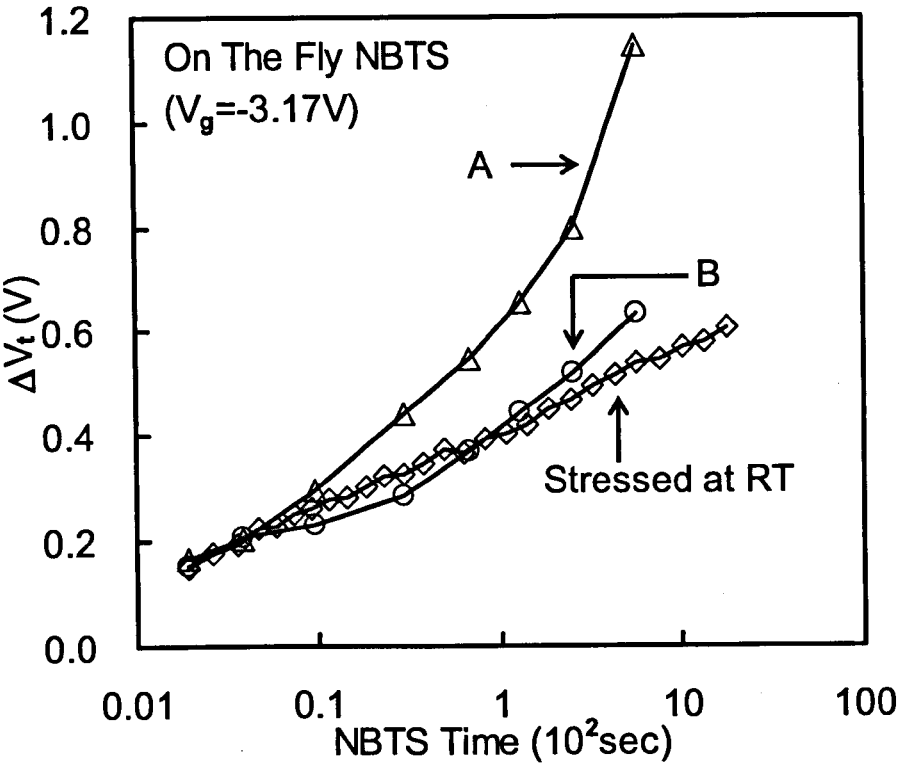


Figure 23. A comparison of ΔV_t measured at different temperatures. The test procedure used is shown in Figure 22. The curves 'A' and 'B' correspond to the points 'A' and 'B' in Figure 22 after different NBTS time. The symbol '◇' represents the ΔV_t for a device stressed and measured at room temperature. The difference between 'A' and 'B' is caused by the effect of different measurement temperatures. The difference between 'A' and symbol '◇' is the 'real' effect of stress temperature on defect generation. Although this difference is modest at short NBTS time, but becomes substantial for a long NBTS time.

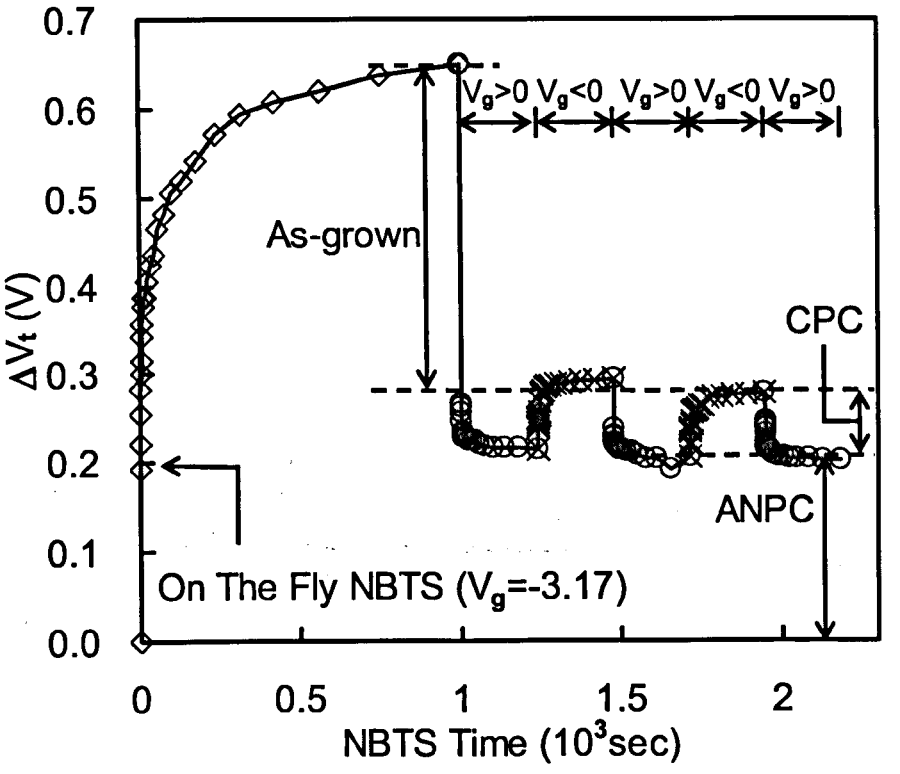


Figure 24. A device was first subjected to NBTS at 100°C (symbol '◇'). Followed by a cycle of positive and negative gate biases ($|E_{ox}|$ = 5 MV/cm) with all other terminals grounded. The stress and V_g cycle were at 100°C. ΔV_t was measured at V_g = -3.17 V during all sequences. It is clear that part of ΔV_t could be repeatedly charged under $V_g < 0$ and discharged under $V_g > 0$. This characteristic is similar to the 'cyclic positive charges' (CPC). In addition, some of ΔV_t are 'unrecoverable' under $V_g > 0$. The 'unrecoverable' defects behave similarly to the 'anti-neutralization positive charges' (ANPC).

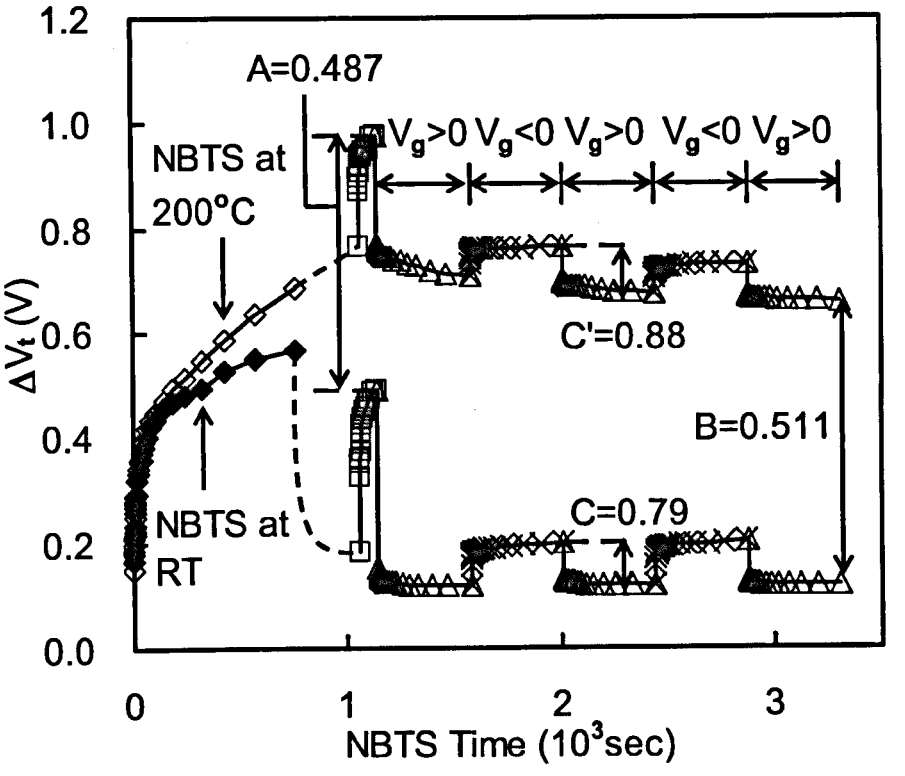
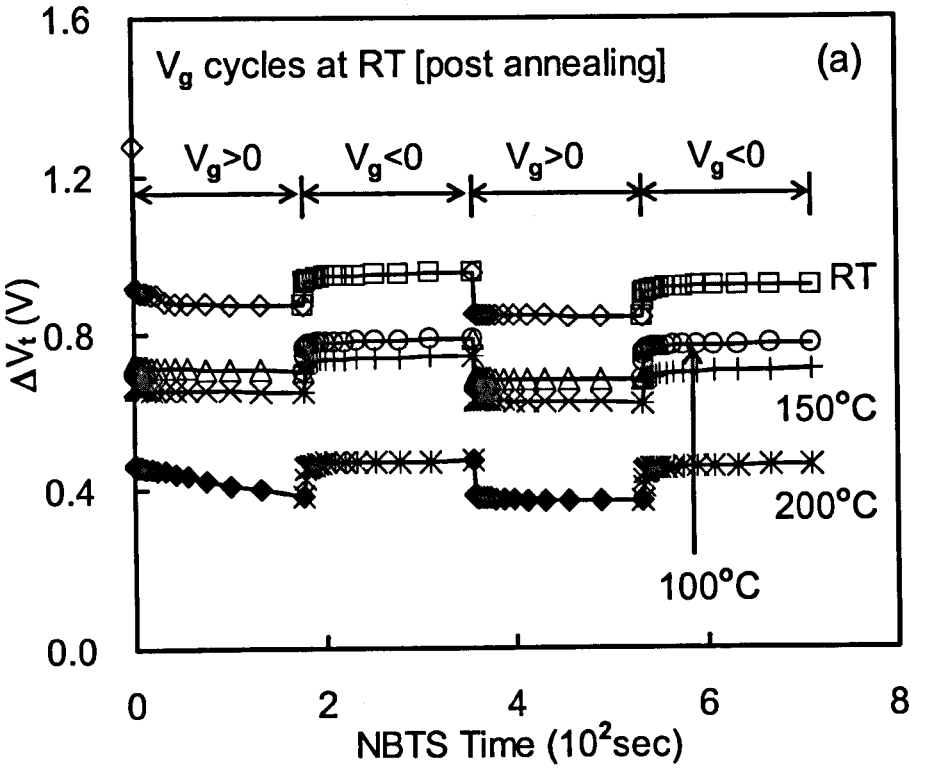


Figure 25. Dynamic behaviour of NBTI. Two devices were stressed by the same V_g (-3.17 V) for the same time (760 sec) at 200°C and room temperature, respectively. V_g was then removed for both devices for 4 min, during which the device stressed at 200°C was cooled down to room temperature. To regain the lost charges, both devices were stressed by V_g = -3.17 V for a short period at room temperature (symbol '□'). This is followed by applying $V_g > 0$ and $V_g < 0$, alternately. It is obvious that the CPC in these two devices are similar, but ANPC is significantly higher for the device stressed at 200°C. Moreover, the difference in ANPC is approximately the same as the difference in the total ΔV_t ($A \approx B$). This means that only ANPC creation is thermally accelerated.


1.6
1.2
0.8
0.4
0.0
ΔVt (V)
Vg cycles at RT [post annealing]
(a)
Vg>0
Vg<0
Vg>0
Vg<0
RT
150°C
200°C
100°C
0
2
4
6
8
NBTS Time (10²sec)

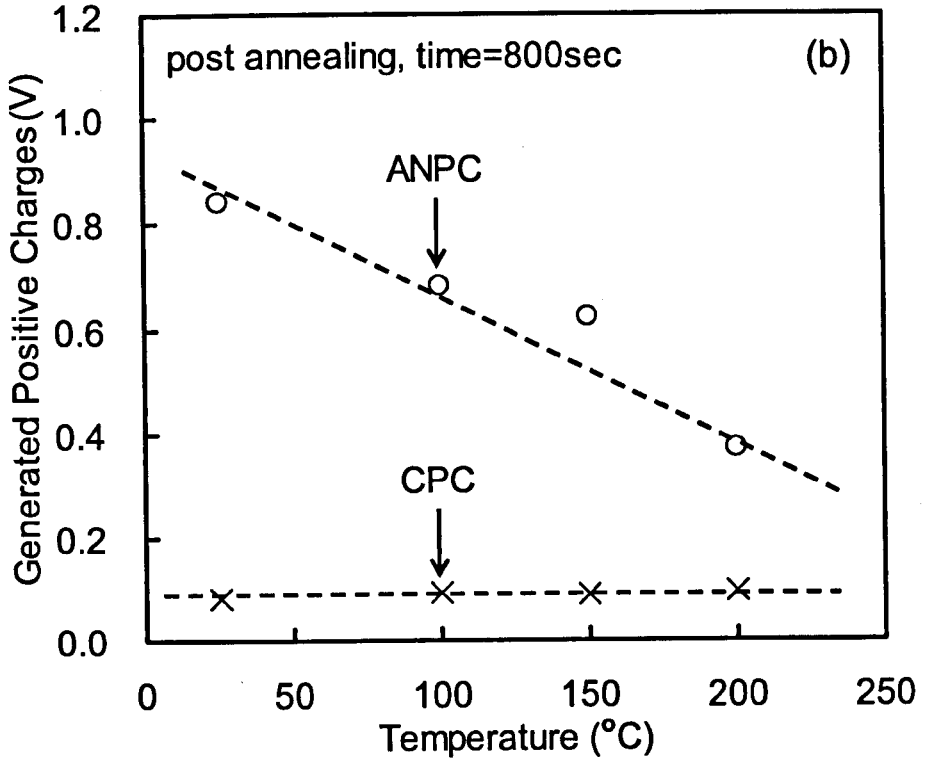
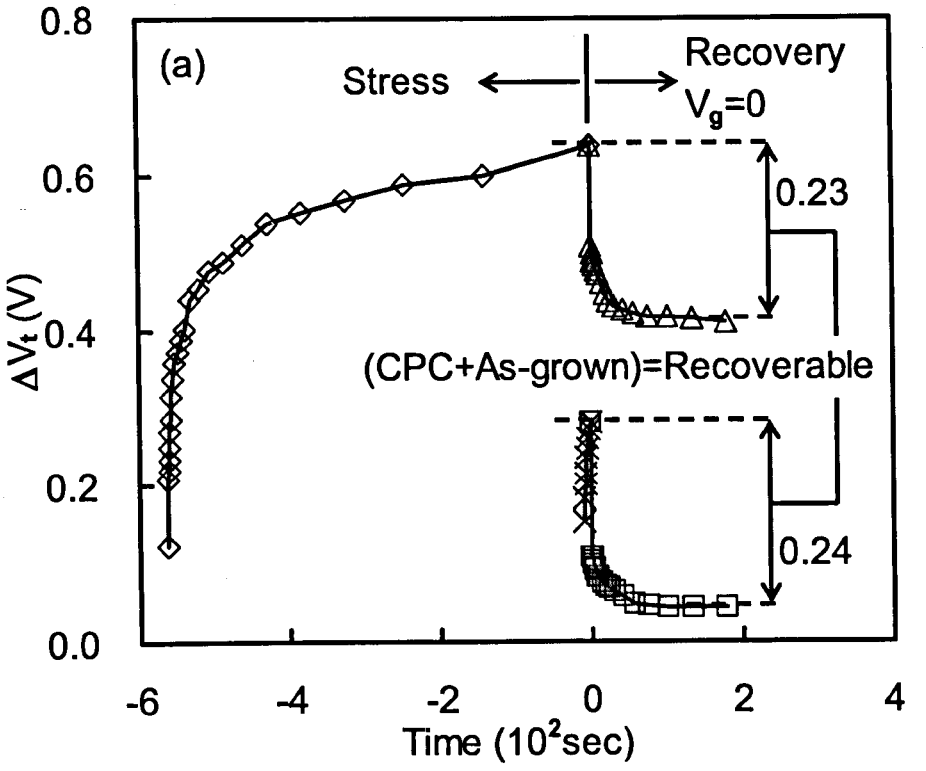

1.2
1.0
0.8
0.6
0.4
0.2
0.0
Generated Positive Charges (V)
post annealing, time=800sec
(b)
ANPC
CPC
0
50
100
150
200
250
Temperature (°C)

Figure 26. Thermal stability of generated positive charges. A device was stressed at room temperature, V_g = -3.17 V for overnight (75000 sec), followed by a sequence of $V_g > 0$ and $V_g < 0$. The device was then exposed to elevated temperature for 800 sec with all terminals floating. After the exposure, the device was cooled down to room temperature, so another sequence of V_g cycles could be performed. (a) Compares the V_g cycles post different annealing temperatures. The amplitude of ΔV_t goes down gradually as annealing temperature increases. (b) CPC remains stable after exposing to a temperature up to 200°C, while ANPC reduces considerably as temperature increases.


(a)
Stress
Recovery
V_g=0
0.23
(CPC+As-grown)=Recoverable
0.24
ΔV_t (V)
0.0
0.2
0.4
0.6
0.8
-6
-4
-2
0
2
4
Time (10^2sec)

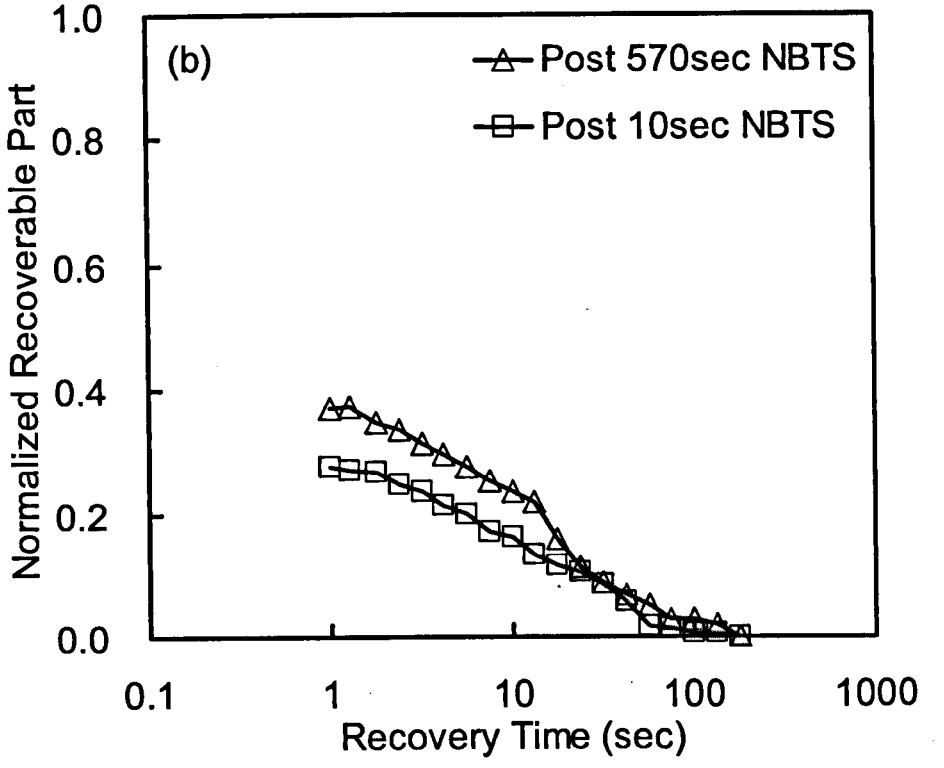
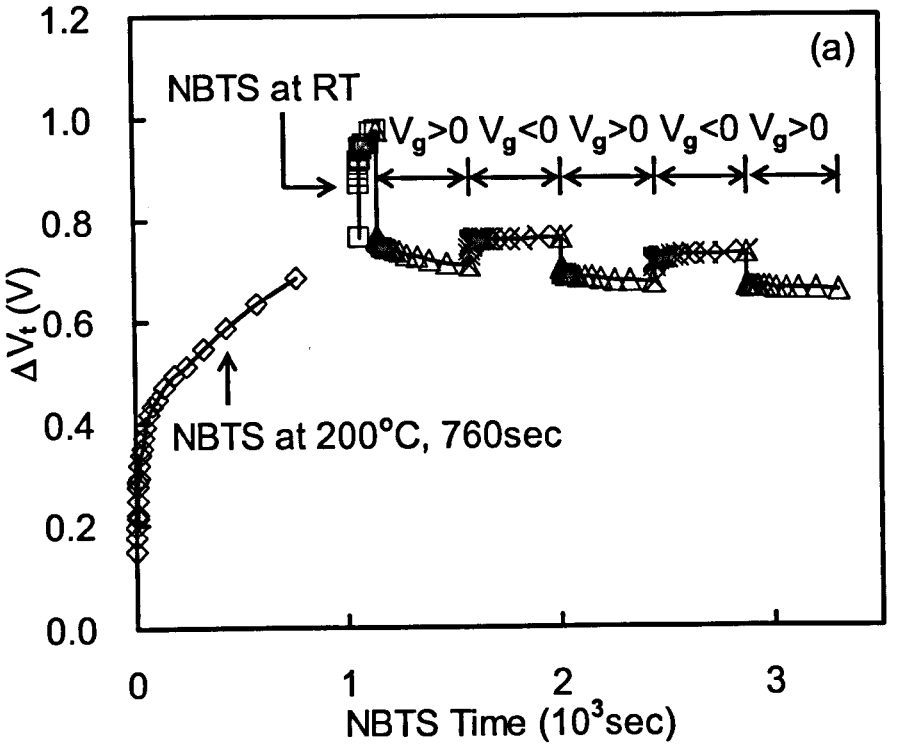
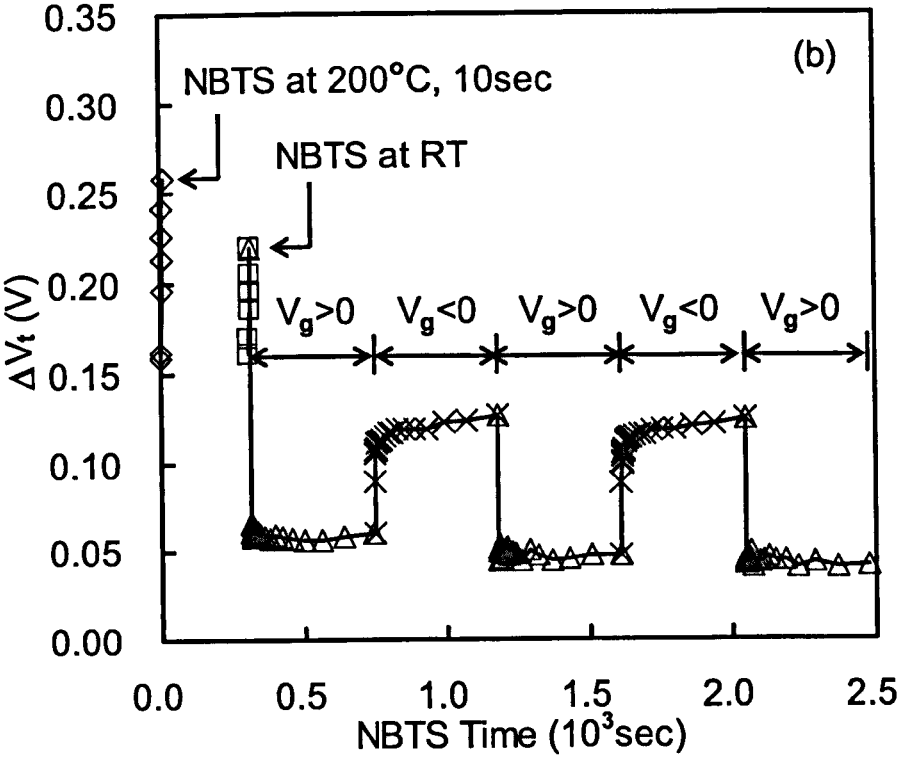

(b)
Post 570sec NBTS
Post 10sec NBTS
Normalized Recoverable Part
0.0
0.2
0.4
0.6
0.8
1.0
0.1
1
10
100
1000
Recovery Time (sec)

Figure 27. (a) A comparison of the recovery behavior of ΔVt after a short (10 sec) and long (570 sec) NBTS under V_g = -3.17 V. The time was set to zero at the end of NBTS. The stress and recovery were at 200°C. The V_g = 0 was applied for the recovey period. ΔV_t was measured at V_g = -3.17 V for both the stress and recovery periods. After 10 sec NBTS, majority of charges (> 80%) were recovered. After 570 sec NBTS, ΔV_t was more than doubled, but the 'recoverable' part did not increase. As NBTS time increases, the generation of 'unrecoverable' defects leads to the increased difference between the curve 'A' and 'B' in Figure 23. (b) Time dependence of the recoverable part. Each curve was normalized against its total recoverable part. It is apparent that the recovery speed is similar on both devices, and is not dependent on the stress level.


(a)
NBTS at RT
V_g>0 V_g<0 V_g>0 V_g<0 V_g>0
NBTS at 200°C, 760sec
ΔV_t (V)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
NBTS Time (10^3sec)


(b)
NBTS at 200°C, 10sec
NBTS at RT
V_g>0 V_g<0 V_g>0 V_g<0 V_g>0
ΔV_t (V)
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.0
0.5
1.0
1.5
2.0
2.5
NBTS Time (10^3sec)

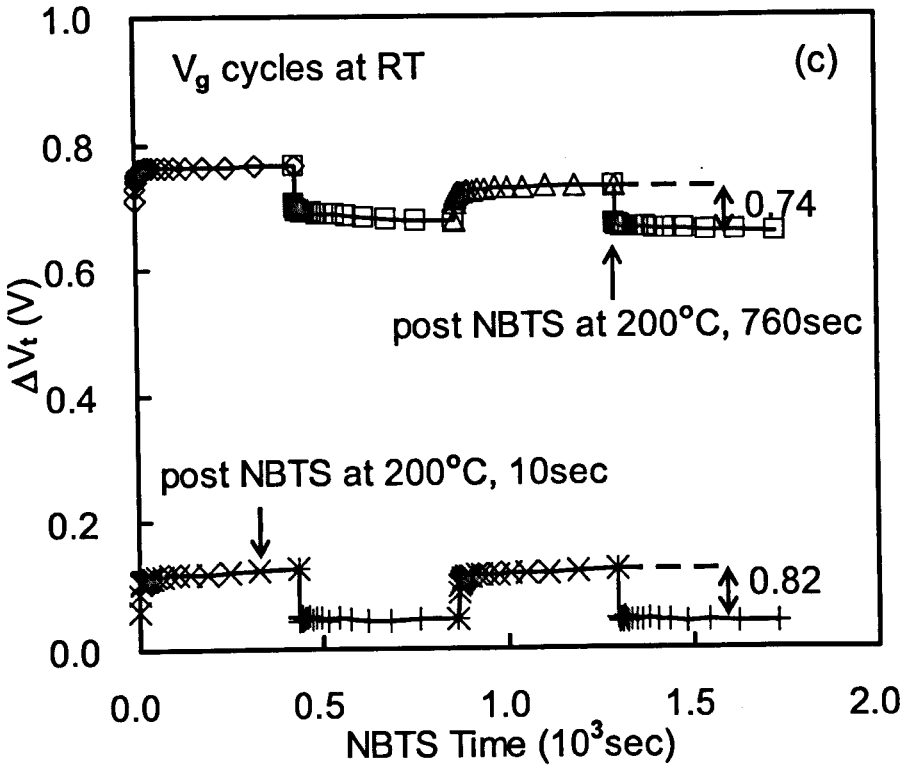


Figure 28. (a) & (b) Two devices were stressed at 200°C (V_g =-3.17 V) for 760 sec and 10 sec, respectively. V_g was then removed for both devices for 4 min, during which the devices were allowed to cool down to room temperature. To regain the lost charges, both devices were stressed by V_g = -3.17 V for a short period at room temperature (symbol '□'). This is followed by applying $V_g > 0$ and $V_g < 0$, alternately. (c) The V_g cycles post different NBTS time are compared. As expected, ANPC is NBTS time dependent. The longer the NBTS time, the larger ANPC is generated.

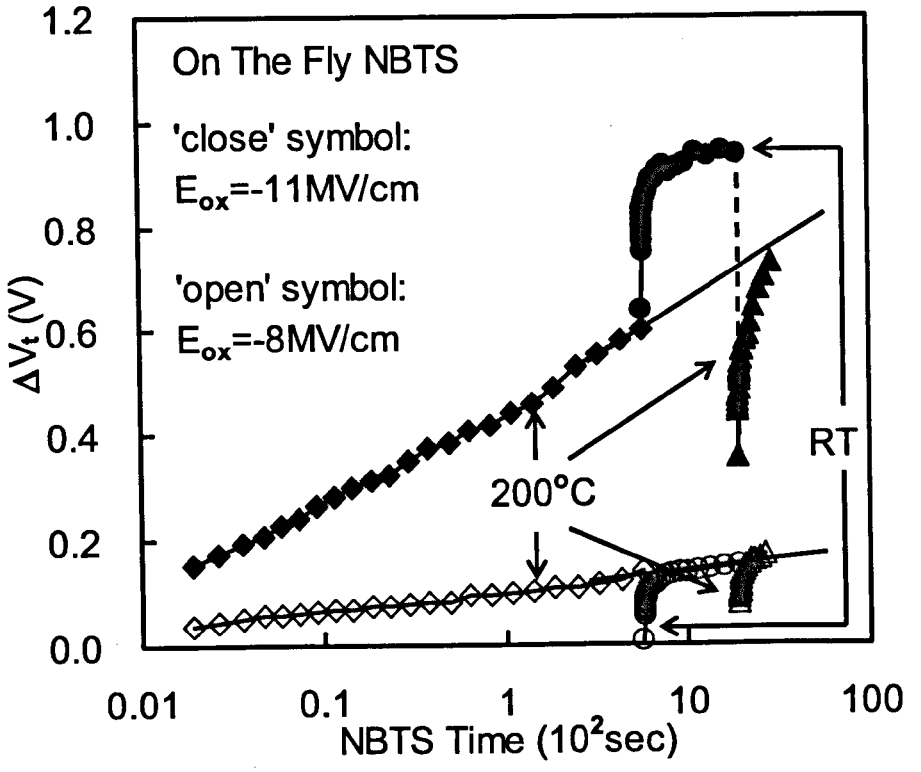
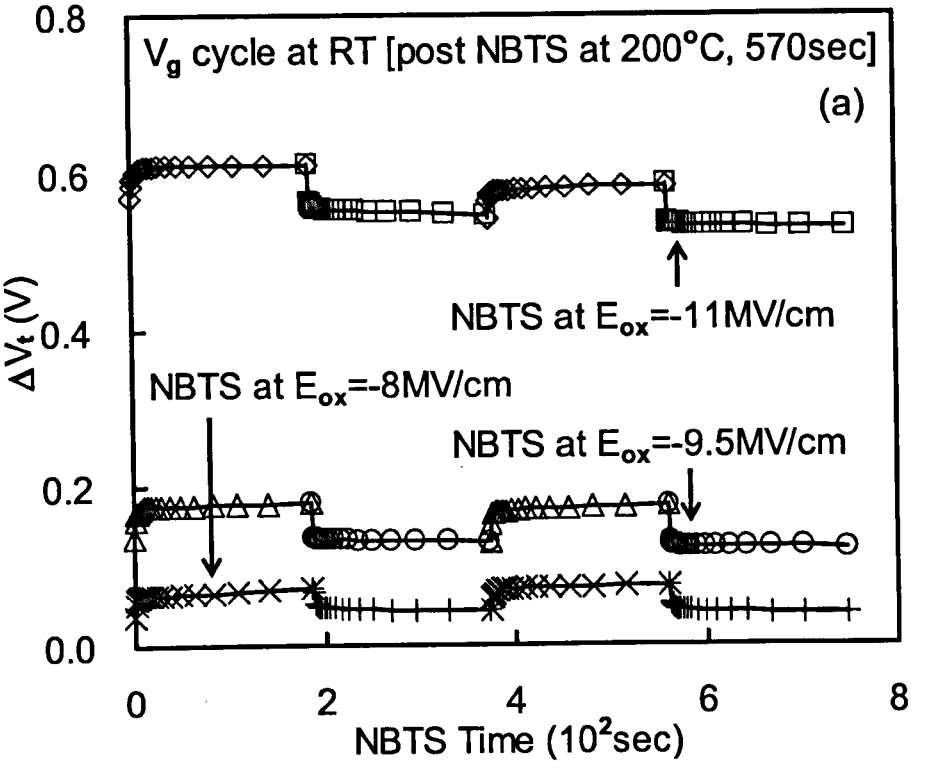


Figure 29. Effects of measurement temperature on ΔV_t after stressed at -11 MV/cm and -8 MV/cm. The devices were stressed at 200°C for 570 sec, and then switched to room temperature. Unlike the -11 MV/cm case, the impact of measurement temperature is negligible when stressed at -8 MV/cm.


V_g cycle at RT [post NBTS at 200°C, 570sec]
(a)
NBTS at E_{ox}=-11MV/cm
NBTS at E_{ox}=-8MV/cm
NBTS at E_{ox}=-9.5MV/cm
ΔV_t (V)
0.0
0.2
0.4
0.6
0.8
0
2
4
6
8
NBTS Time (10^2sec)

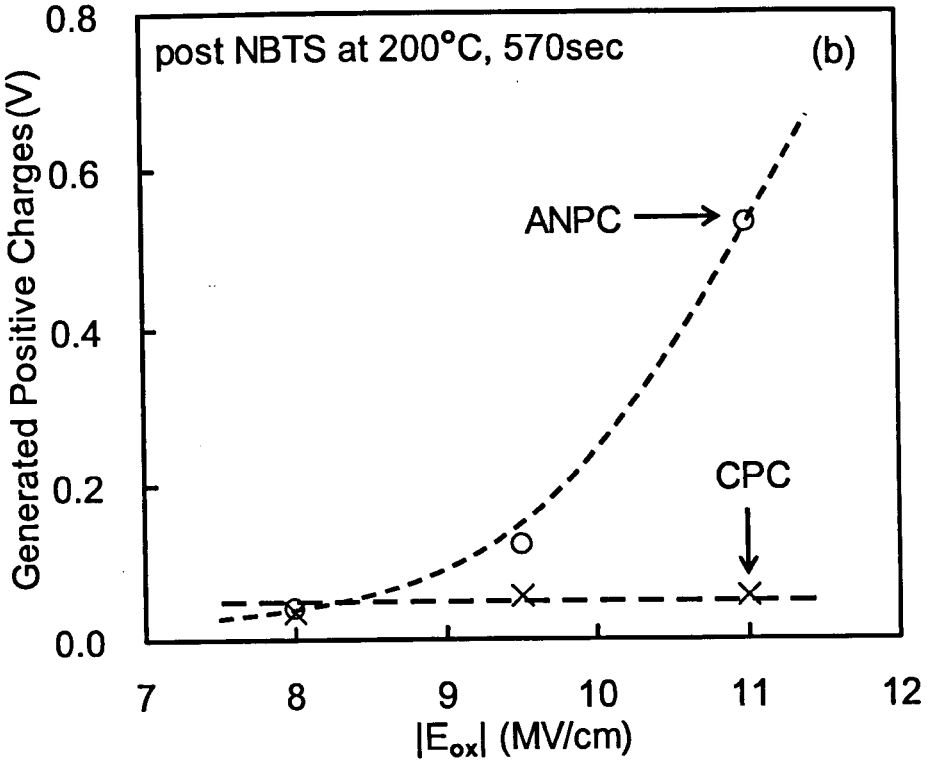
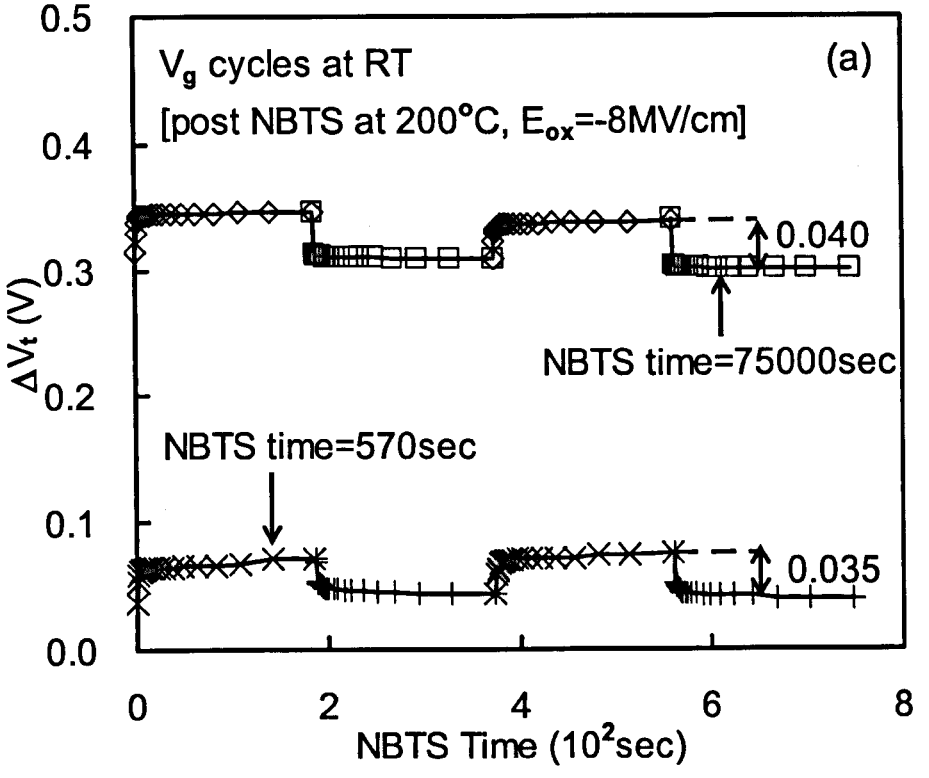

post NBTS at 200°C, 570sec
(b)
ANPC
CPC
Generated Positive Charges (V)
0.0
0.2
0.4
0.6
0.8
7
8
9
10
11
12
$|E_{ox}|$ (MV/cm)

Figure 30. Three devices were subjected to NBTS at 200°C for 570 sec. The oxide fields were set at -11 MV/cm, -9.5 MV/cm, and -8 MV/cm. The devices were then cooled down to room temperature, before a sequence of $V_g > 0$ and $V_g < 0$ was applied to determine ANPC and CPC generation. (a) Compares the V_g cycles post different stress oxide fields. (b) The generated positive charges are plotted against oxide field. It is clear that ANPC reduces significantly at -8 MV/cm, but CPC is hardly changed.


V_g cycles at RT
(a)
[post NBTS at 200°C, E_{ox}=-8MV/cm]
0.040
NBTS time=75000sec
NBTS time=570sec
0.035
ΔV_t (V)
NBTS Time (10^2sec)

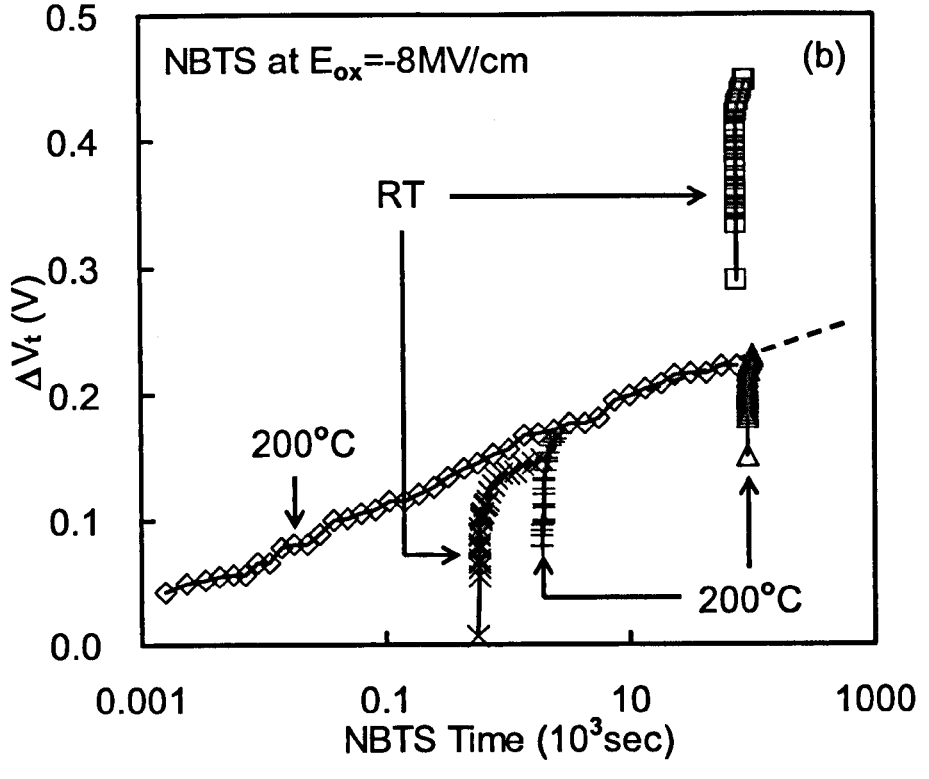

NBTS at E_{ox}=-8MV/cm
(b)
RT
200°C
200°C
ΔV_t (V)
NBTS Time (10^3sec)

Figure 31. A device was stressed under E_{ox} = -8 MV/cm at 200°C for overnight (75000 sec), in an attempt to create sufficient ANPC to make the effects of measurement temperature clearly observable. (a) After stress at 200°C, the devices were cooled down to room temperature, before a sequence of $V_g > 0$ and $V_g < 0$ was applied to determine ANPC and CPC generation. ANPC is significantly increased after 75000 sec stress, while CPC remains similar. (b) The effects of measurement temperature on ΔV_t now become obvious (symbol '□'), similar to Figure 22. This confirms ANPC generation is responsible for the effects of measurement temperature, and the same types of defects are created at different gate biases.

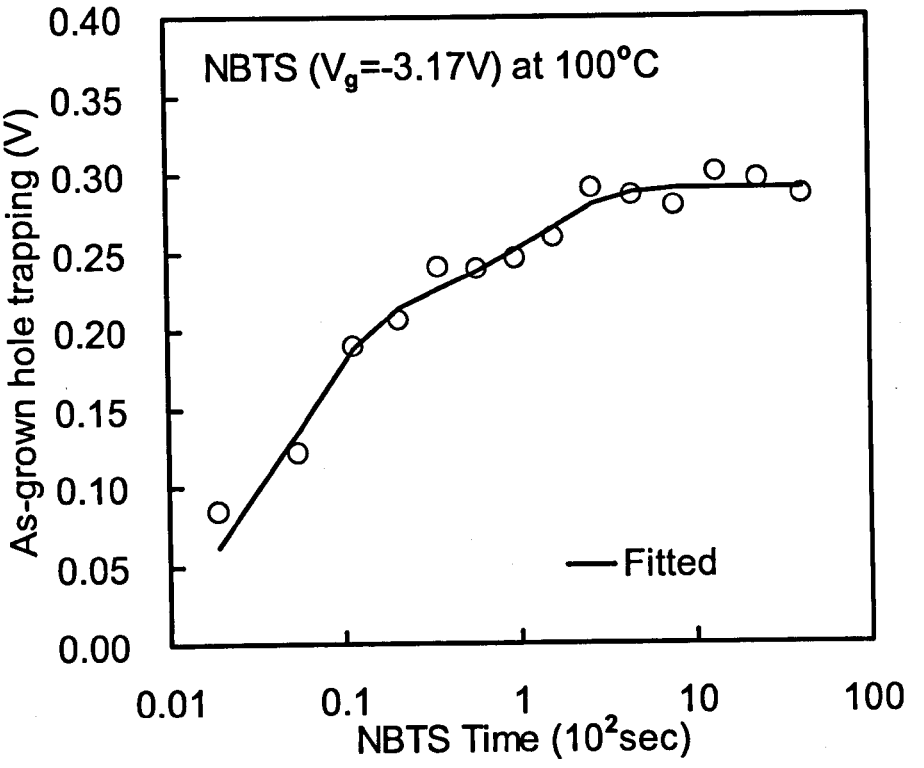


Figure 32. As-grown hole trapping plotted against NBTS time. The test procedure is similar to Figure 24, where a device was subjected to NBTS at 100°C, followed by a sequence of $V_g > 0$ and $V_g < 0$. The as-grown hole trapping can be obtained after each cycle. Under V_g = -3.17 V, the as-grown hole trapping increases with NBTS time, but starts to saturate around 10 sec of stress. The saturation level is 0.29 V. The solid line is obtained by fitting the data with first order model using two capture cross sections.

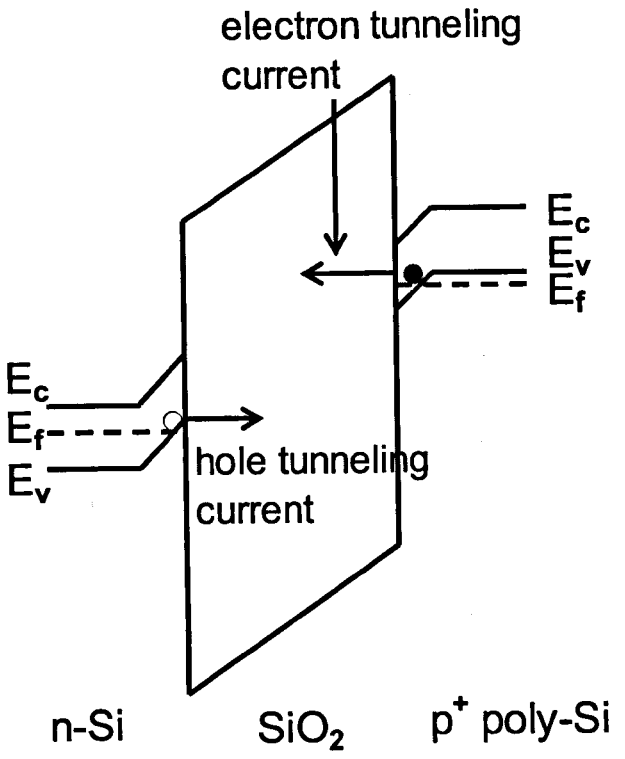
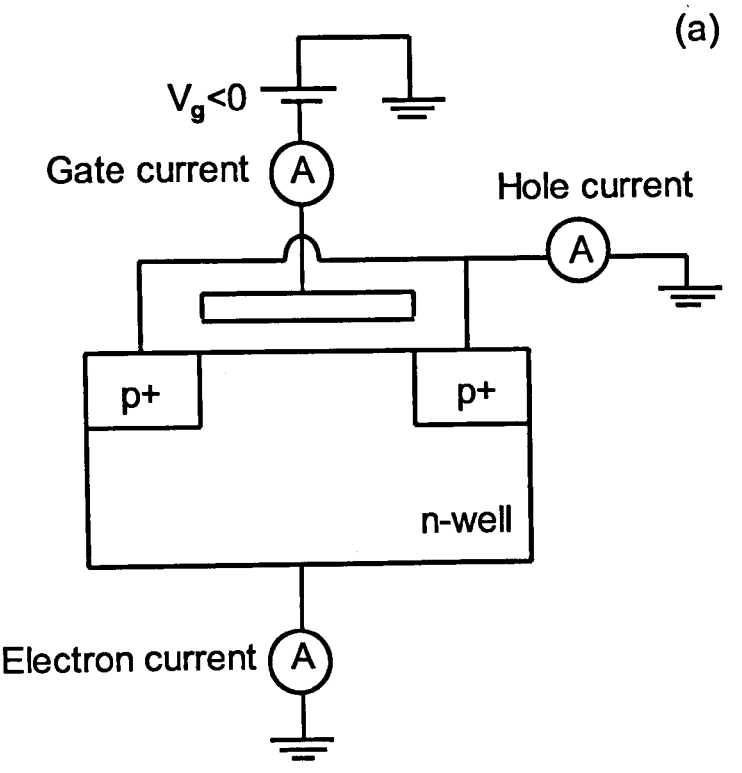
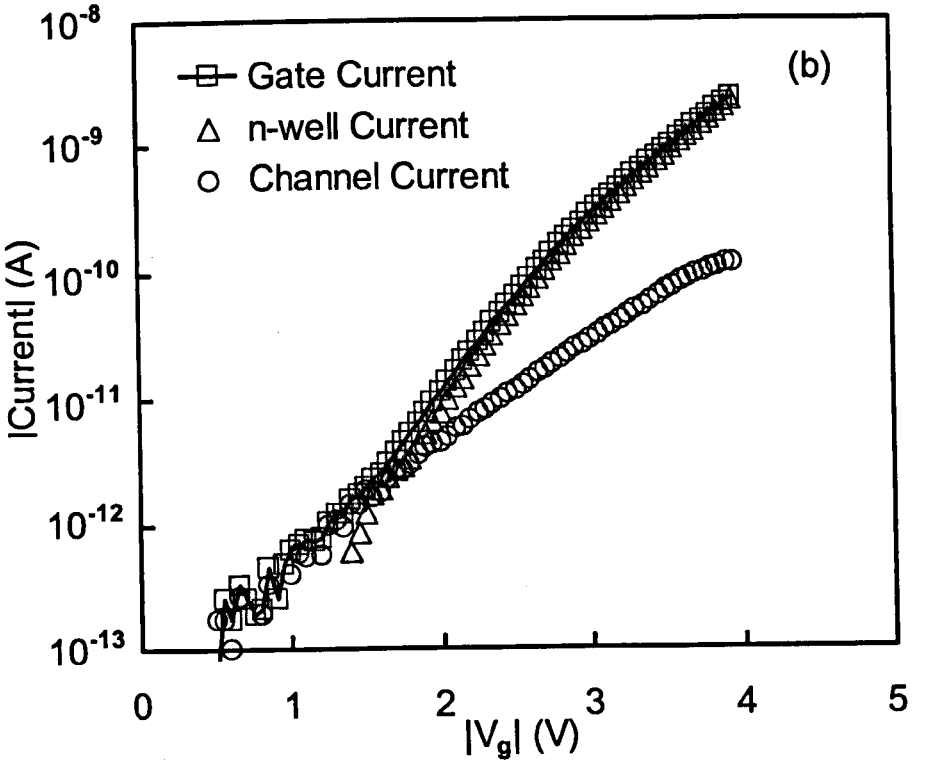


Figure 33. Schematic energy diagram of a pMOSFET under inversion condition, illustrating the tunneling of holes from the n-Si substrate valence band and the tunneling of electrons from the p^+ poly-Si gate valence band.


(a)
V_g<0
Gate current
A
Hole current
A
p+
p+
n-well
Electron current
A


(b)
Gate Current
n-well Current
Channel Current
|Current| (A)
10^{-8}
10^{-9}
10^{-10}
10^{-11}
10^{-12}
10^{-13}
0
1
2
3
4
5
$|V_g|$ (V)

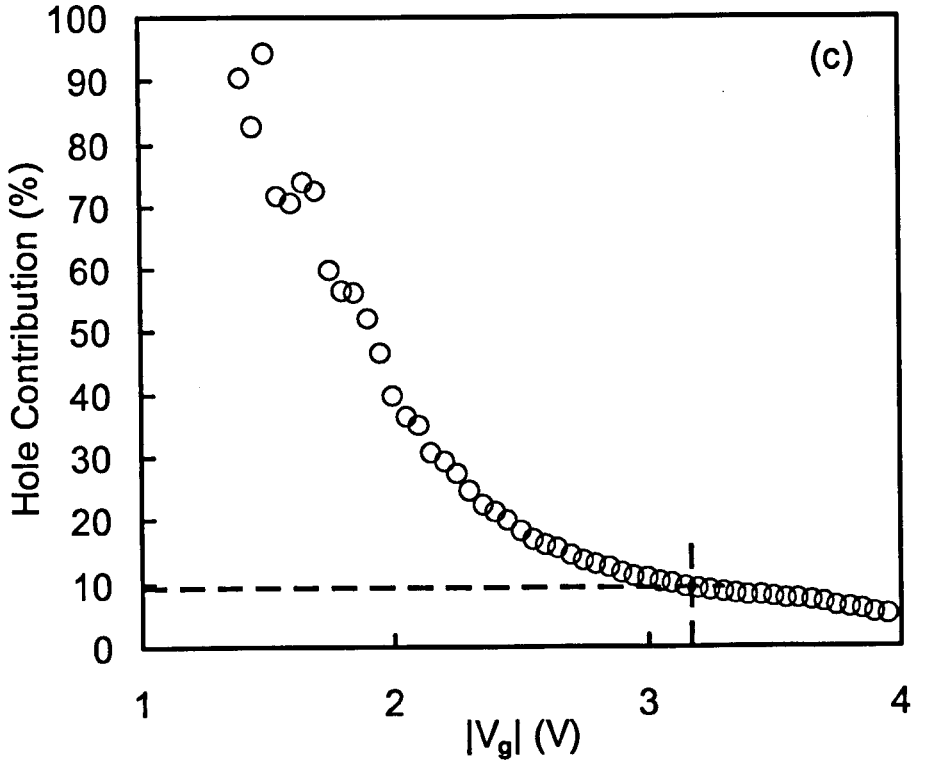


Figure 34. (a) Experimental set up of carrier separation measurement. A pMOSFET is under inversion condition. (b) The gate, n-well, and channel current flowing through the device are plotted as a function of gate voltage, V_g. The channel current is due to the tunneling of holes from the Si substrate valence band, and the n-well current is due to the tunneling of electrons from the poly-Si gate valence band. The hole tunneling current is larger than electron current when $|V_g| < 1.5$ V. While the electron tunneling current is larger than hole current at higher gate biases. (c) When NBTS was carried out at $V_g = -3.17$ V, roughly 10% of the gate current is caused by holes.

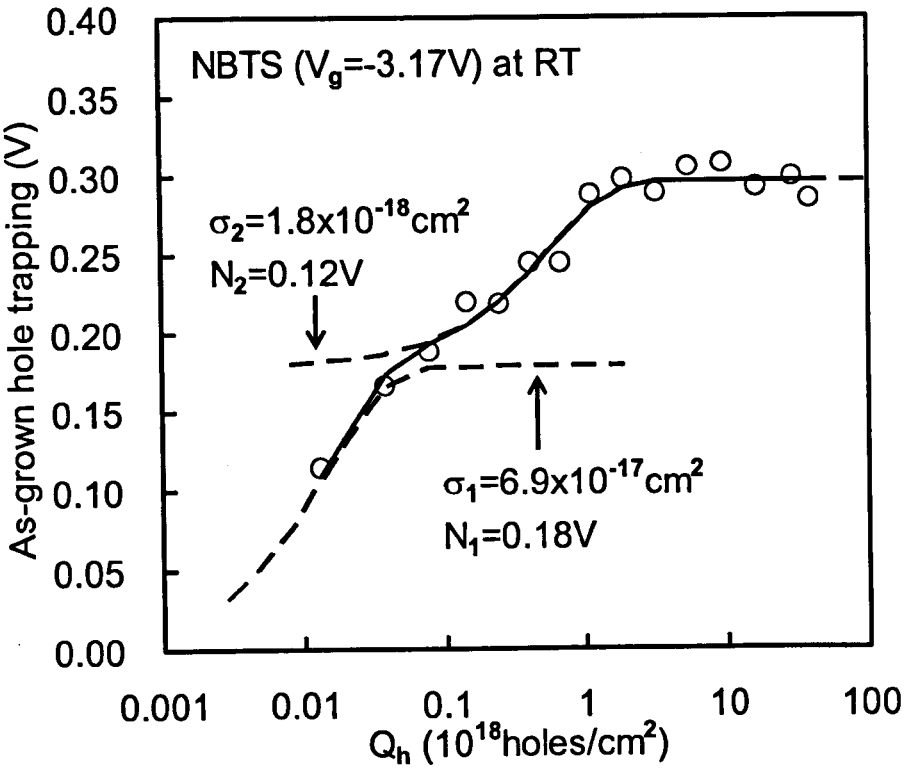


Figure 35. As-grown hole trapping is plotted against hole fluency, Q_h. The hole fluency is obtained through carrier separation measurement. The solid line is obtained by fitting the data with two capture cross sections. The extracted values are 6.9×10^{-17} cm^2 and 1.8×10^{-18} cm^2, which are three order of magnitudes less than that reported for as-grown hole traps.

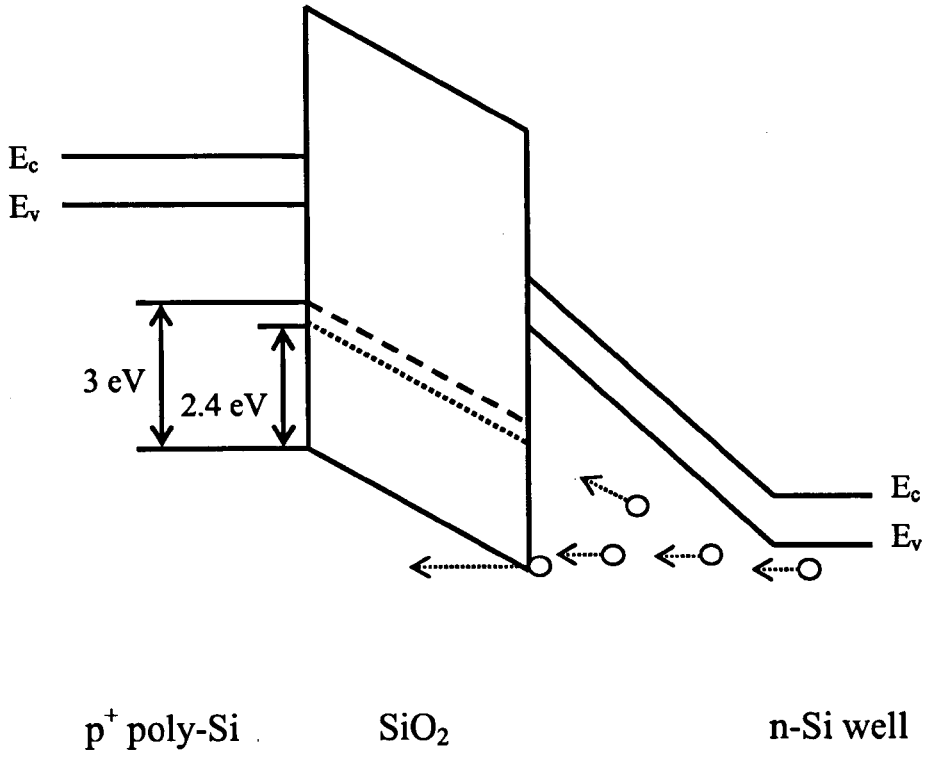


Figure 36. The energy band diagram showing during SHI, a small fraction of holes can gain enough energy, and then be injected into the gate oxide. The injected holes have higher energy than the neutral oxygen vacancy (dashed line). As a result, the as-grown hole traps can be easily filled during SHI.

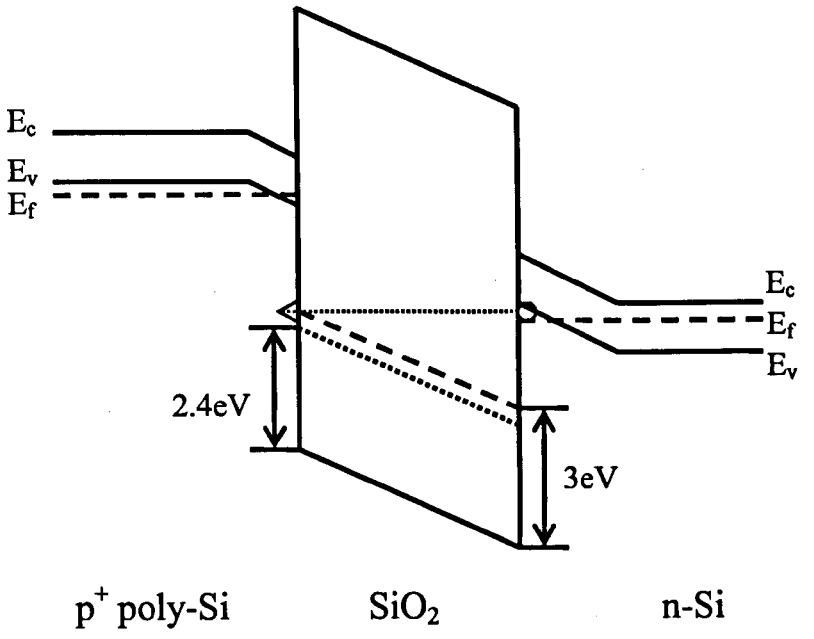


Figure 37. The energy band diagram showing during NBTS, hole transport was accomplished by direct tunneling (DT). The injected holes have lower energy than the neutral oxygen vacancy (dashed line). As a result, the majority of holes will flow through the gate oxide, and only a small fraction of holes with higher energy will be able to partially fill the as-grown hole traps.

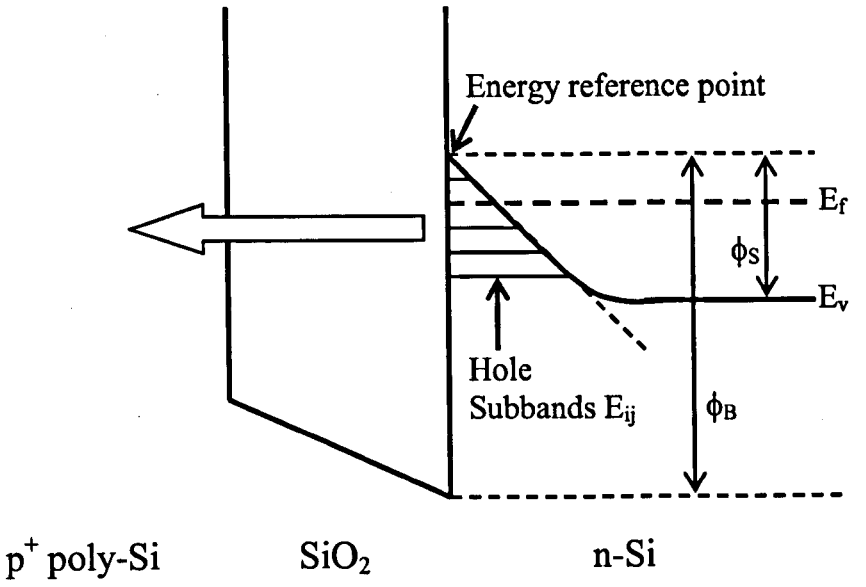


Figure 38. A schematic band diagram of a p^+ polysilicon/SiO_2/n-Si MOS structure showing the hole quantization effect in the substrate and direct tunneling (DT) of holes from the substrate inversion layer to the polysilicon gate.

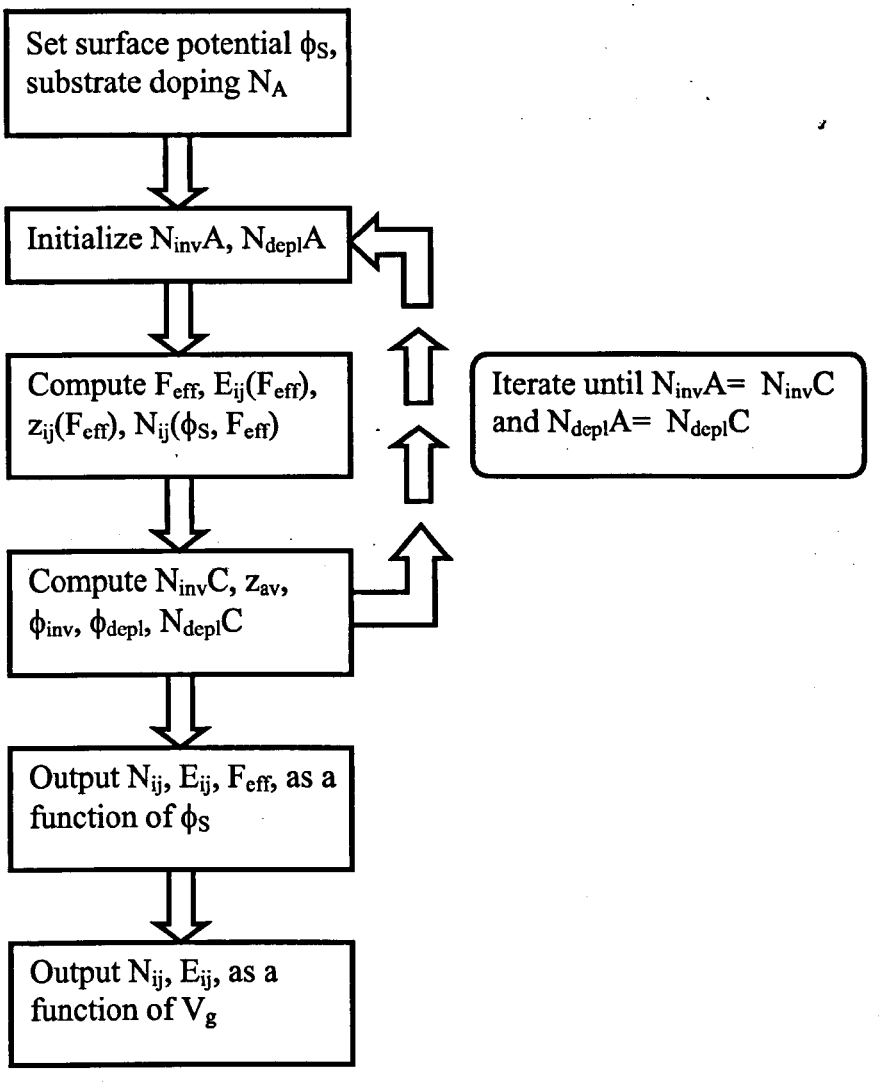
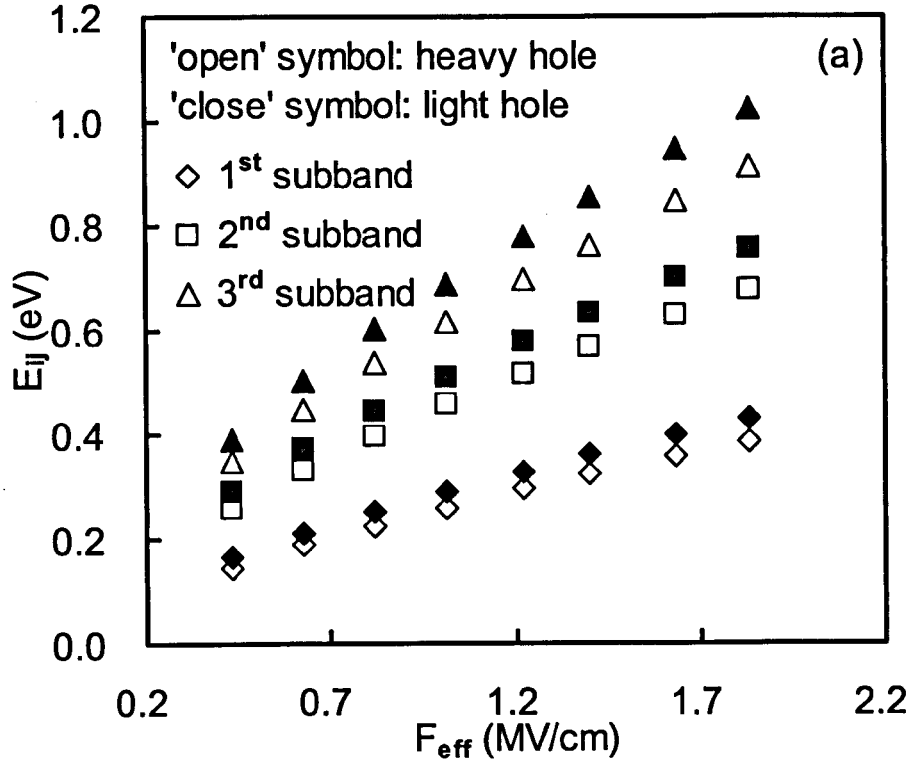
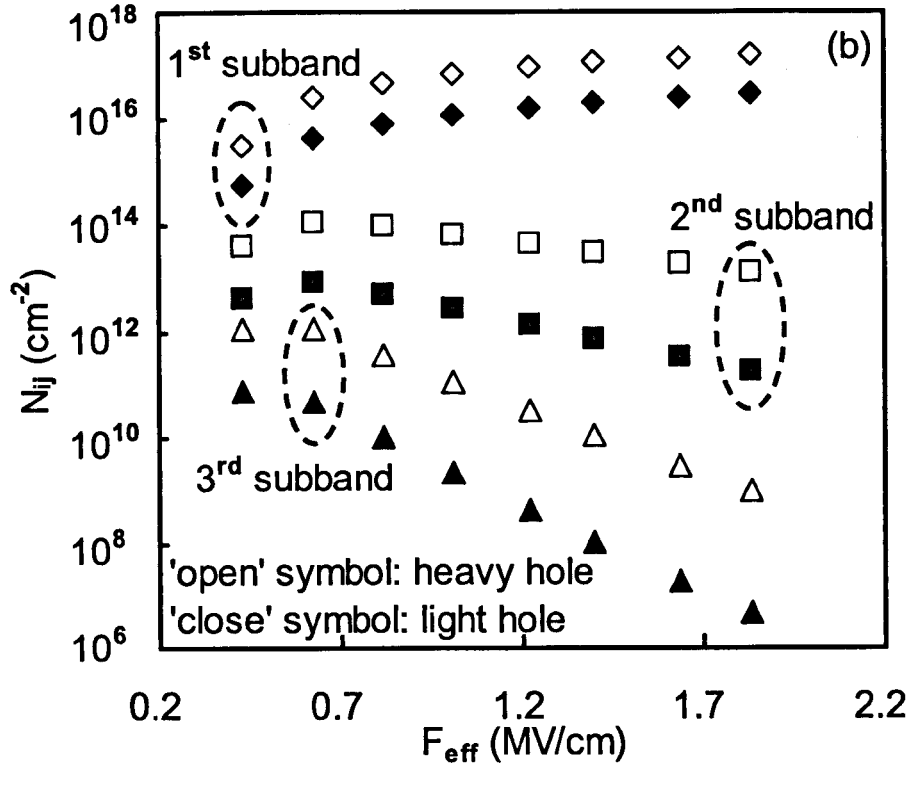


Figure 39. The flow chart of the program. The surface potential and substrate doping density are two input parameters. The iteration sequence begins with an assumed total inversion layer carrier density and the depletion layer doping charge density. This permits a calculation of the inversion layer carrier density for each valley and subband as a function of the effective electric field, as well as the depletion layer doping charge density. The newly calculated values of total carrier density are compared with the assumed ones. The iteration will resume until the differences between the assumed and calculated values become acceptable.


'open' symbol: heavy hole
'close' symbol: light hole
(a)
1st subband
2nd subband
3rd subband
Eij (eV)
Feff (MV/cm)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0.2
0.7
1.2
1.7
2.2


(b)
1st subband
2nd subband
3rd subband
'open' symbol: heavy hole
'close' symbol: light hole
Nij (cm-2)
Feff (MV/cm)
10^18
10^16
10^14
10^12
10^10
10^8
10^6
0.2
0.7
1.2
1.7
2.2

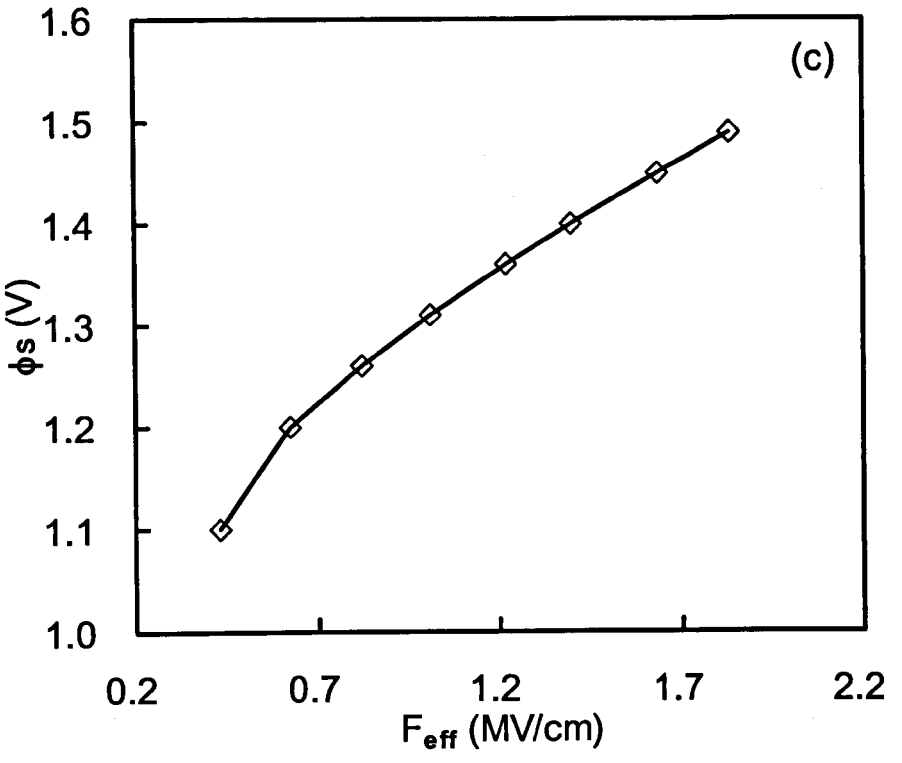


Figure 40. Three subbands were evaluated for two valleys. (a) For a given subband, the energy level of the heavy hole valley (i = 1) is less than that of the light hole valley (i = 2). (b) When F_{eff} = 1.65 MV/cm (V_g = -3.17 V), the carrier density in the 2nd subband is approximately 4 orders less than that in the 1st subband. And the N_{ij} in the 3rd subband is roughly 4 orders less than that in the 2nd subband. (c) The relationship between the effective electrical field and surface potential.

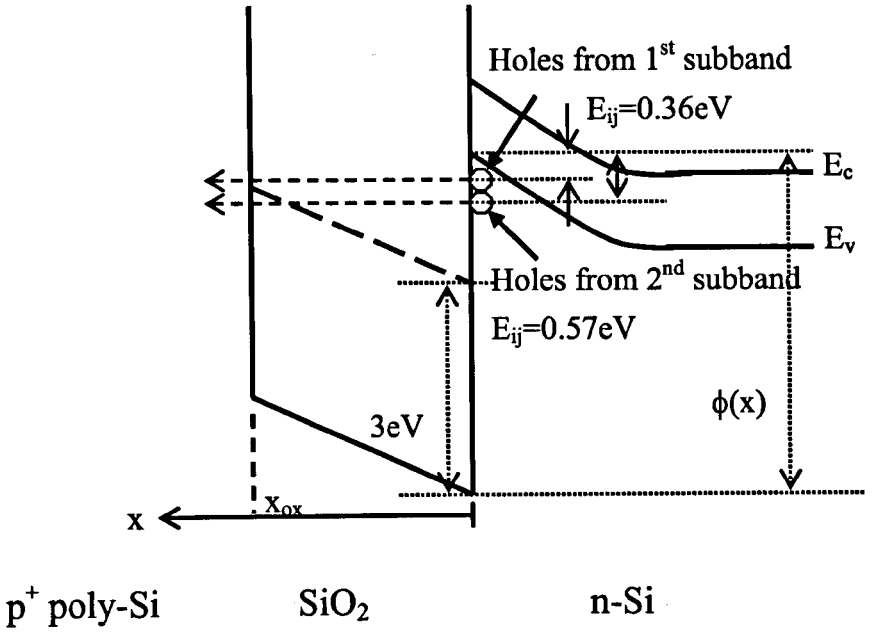


Figure 41. The energy band diagram under typical NBTS (V_g = -3.17 V). x is the distance into the oxide from the interface, $\phi(x)$ the energy barrier height. Only holes tunneled from the 2nd subband (E_{ij} = 0.57 eV) can partially fill the as-grown hole traps (dashed line).

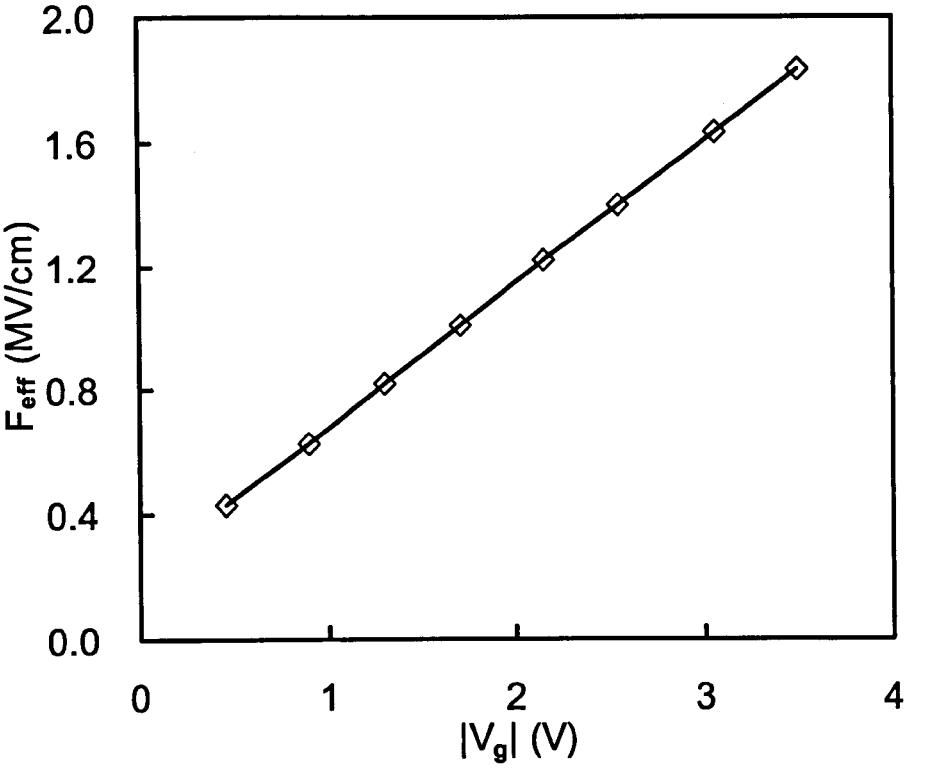


Figure 42. The relationship between the gate voltage, V_g, and the effective electric field in the Si substrate, F_{eff}. Under the typical testing condition (V_g = -3.17 V), F_{eff} is about 1.65 MV/cm.

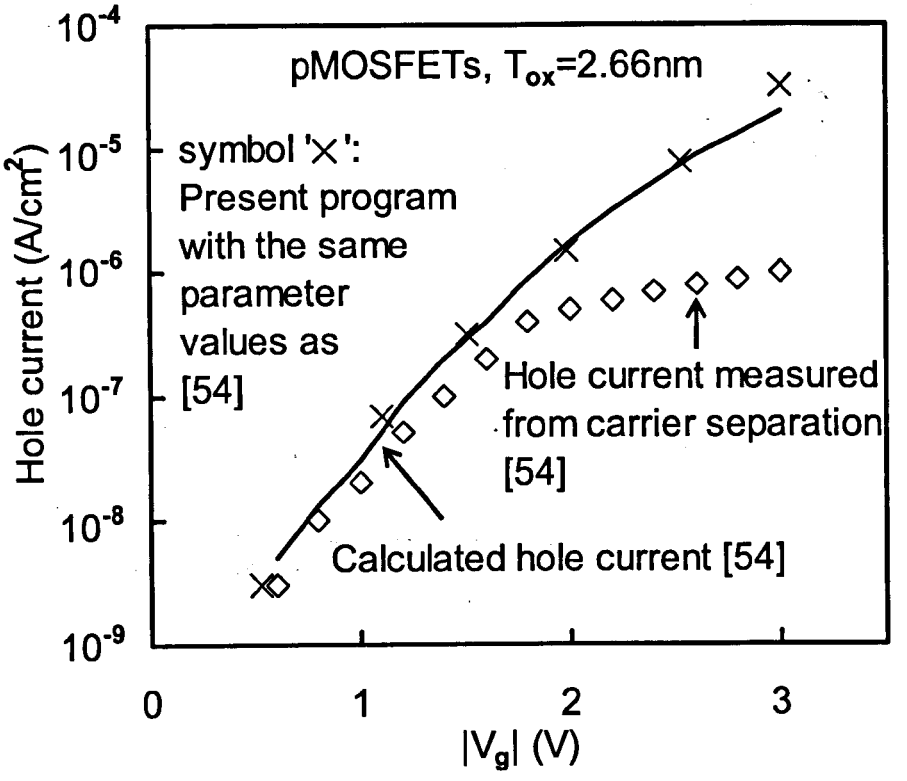


Figure 43. Symbol '◇' is the measured hole current by Hou et al [54]. The solid line represents the calculated hole current by Hou et al [54]. When the same parameter values were used, the present program produces very similar results (symbol 'x').

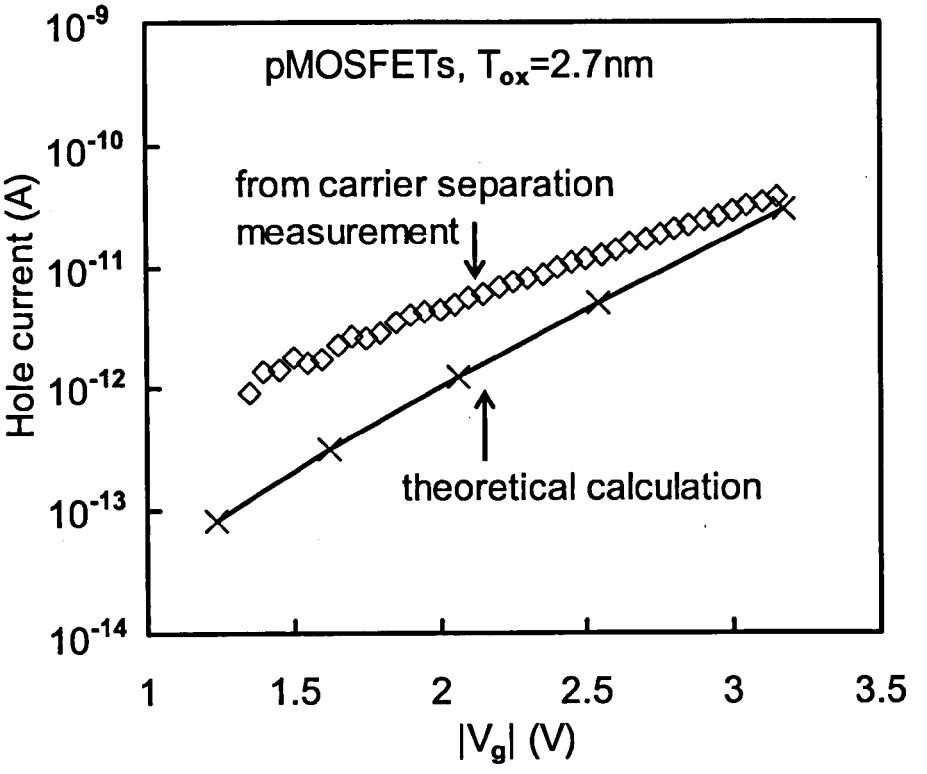
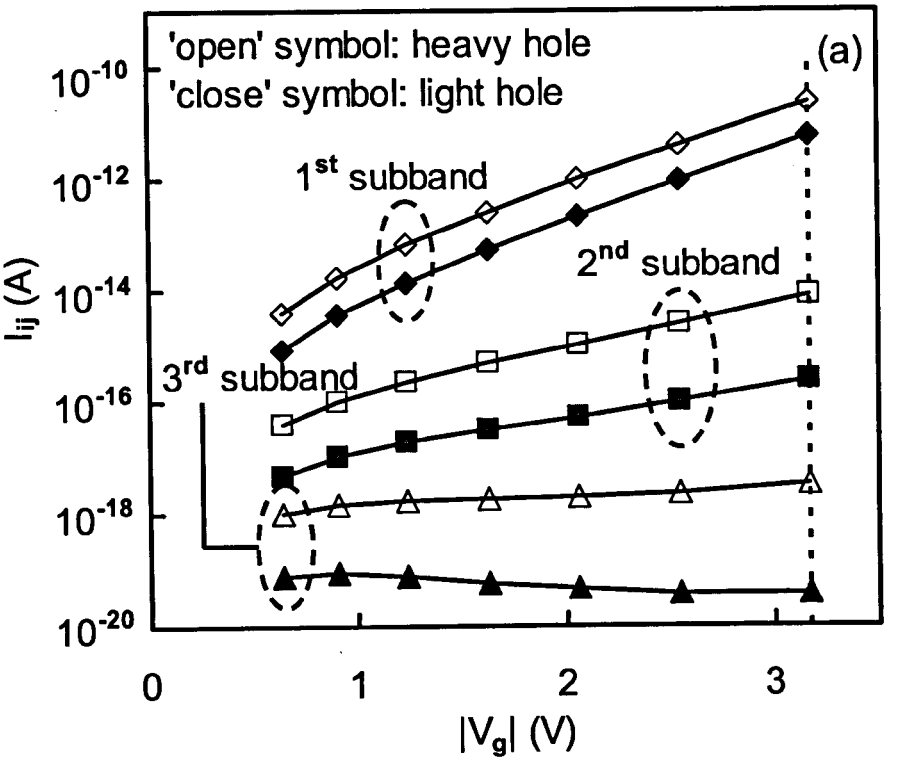
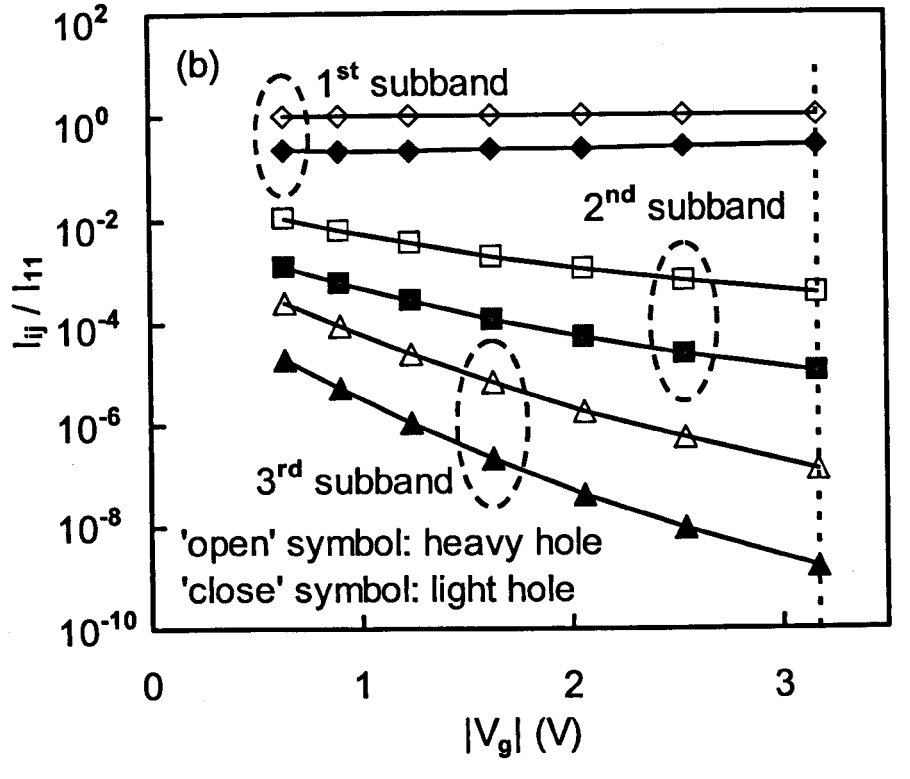


Figure 44. The calculated hole current (symbol 'x') is compared with the hole current measured from carrier separation measurement (symbol '◇'). To best fit the experimental result under the typical stress voltage (i.e. V_g = -3.17 V), the effective mass of holes in SiO_2 was found to be $m_{ox} = 0.35 \times m_0$.


'open' symbol: heavy hole
'close' symbol: light hole
(a)
1st subband
2nd subband
3rd subband
10^{-10}
10^{-12}
10^{-14}
10^{-16}
10^{-18}
10^{-20}
I_{ij} (A)
0
1
2
3
$|V_g|$ (V)


(b)
1st subband
2nd subband
3rd subband
'open' symbol: heavy hole
'close' symbol: light hole
10^{2}
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
I_{ij} / I_{11}
0
1
2
3
$|V_g|$ (V)

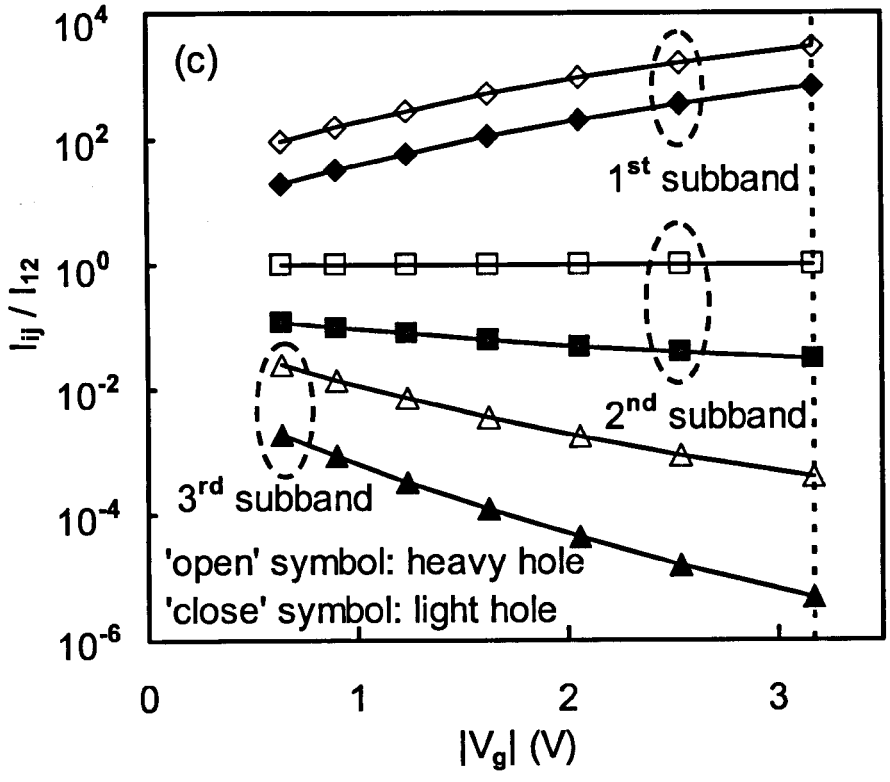
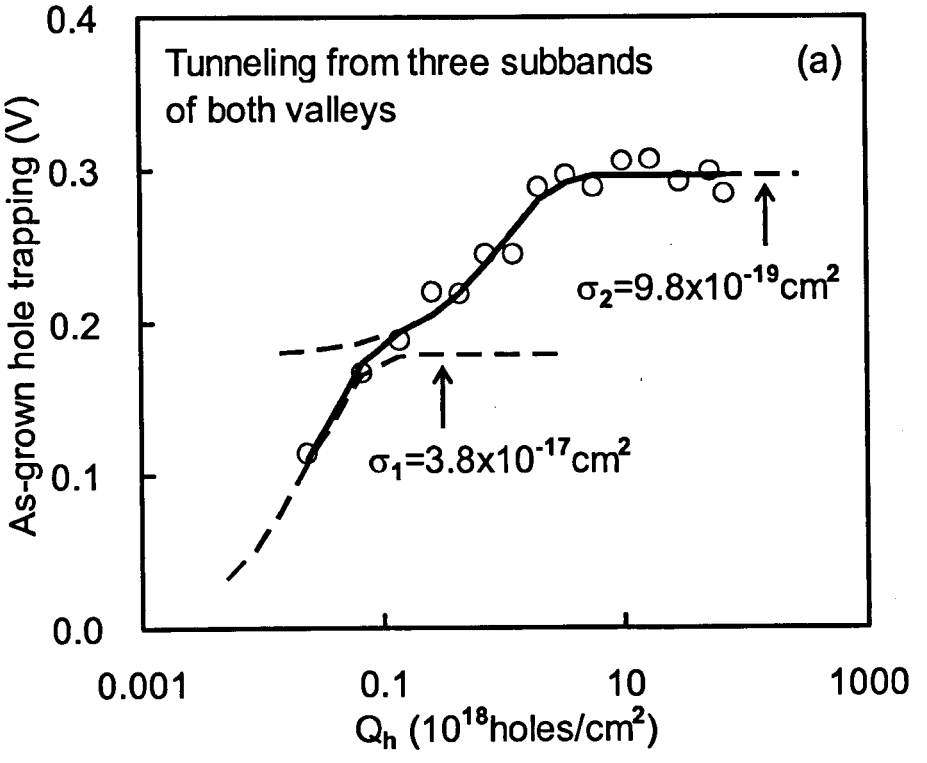
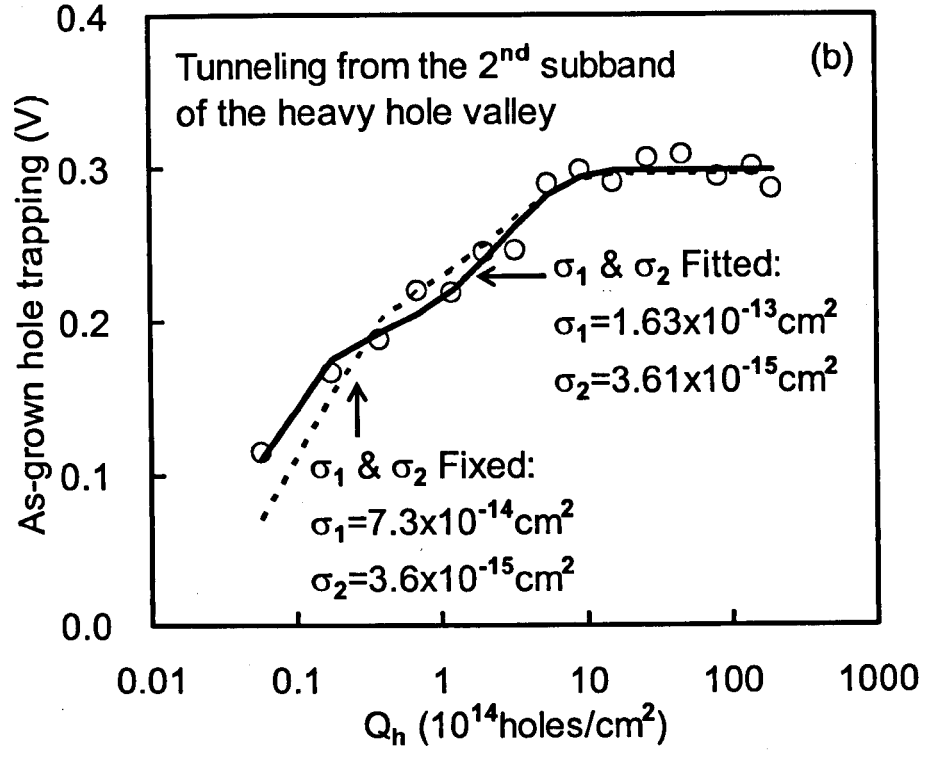
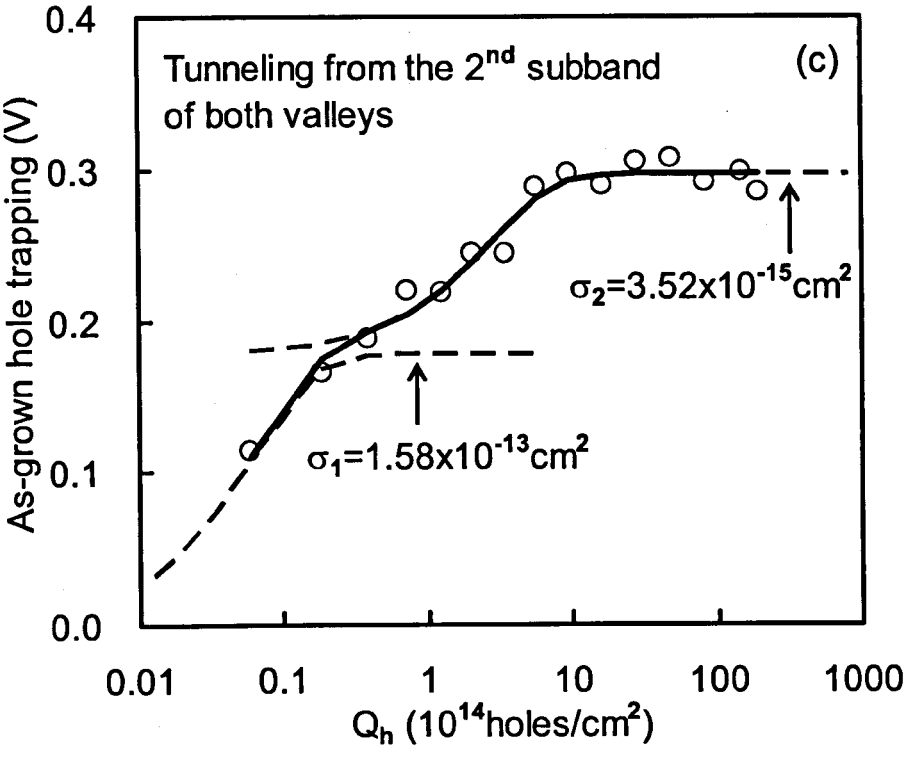


Figure 45. (a) The calculated hole current from each subband. The higher the energy level, the lower the hole current becomes, because of the reduction in carrier density. (b) The ratio against the current from the 1st subband of the heavy hole valley (I_{11}). (c) The ratio against the current from the 2nd subband of the heavy hole valley (I_{12}). The vertical dot lines represent the typical testing condition (V_g = -3.17 V).


Tunneling from three subbands
of both valleys
(a)
σ_2=9.8x10^{-19}cm^2
σ_1=3.8x10^{-17}cm^2
As-grown hole trapping (V)
0.0
0.1
0.2
0.3
0.4
0.001
0.1
10
1000
Q_h (10^{18}holes/cm^2)


Tunneling from the 2nd subband
of the heavy hole valley
(b)
σ_1 & σ_2 Fitted:
σ_1=1.63x10^{-13}cm^2
σ_2=3.61x10^{-15}cm^2
σ_1 & σ_2 Fixed:
σ_1=7.3x10^{-14}cm^2
σ_2=3.6x10^{-15}cm^2
As-grown hole trapping (V)
0.0
0.1
0.2
0.3
0.4
0.01
0.1
1
10
100
1000
Q_h (10^{14}holes/cm^2)


Tunneling from the 2nd subband
of both valleys
(c)
σ2=3.52x10-15cm2
σ1=1.58x10-13cm2
As-grown hole trapping (V)
0.0
0.1
0.2
0.3
0.4
0.01
0.1
1
10
100
1000
Qh (1014holes/cm2)

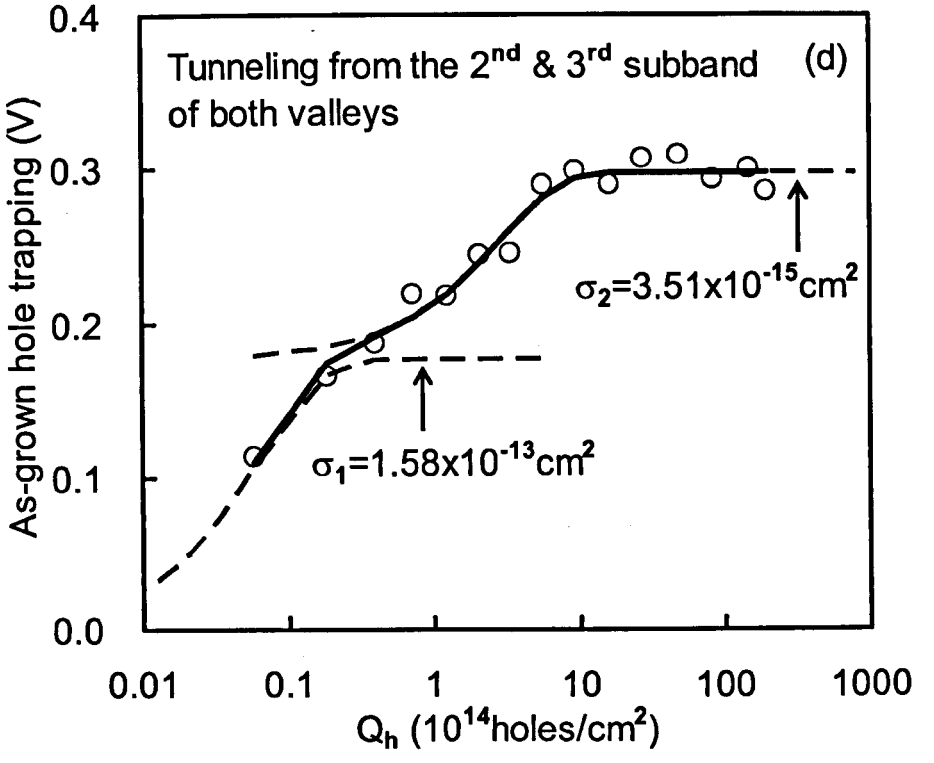
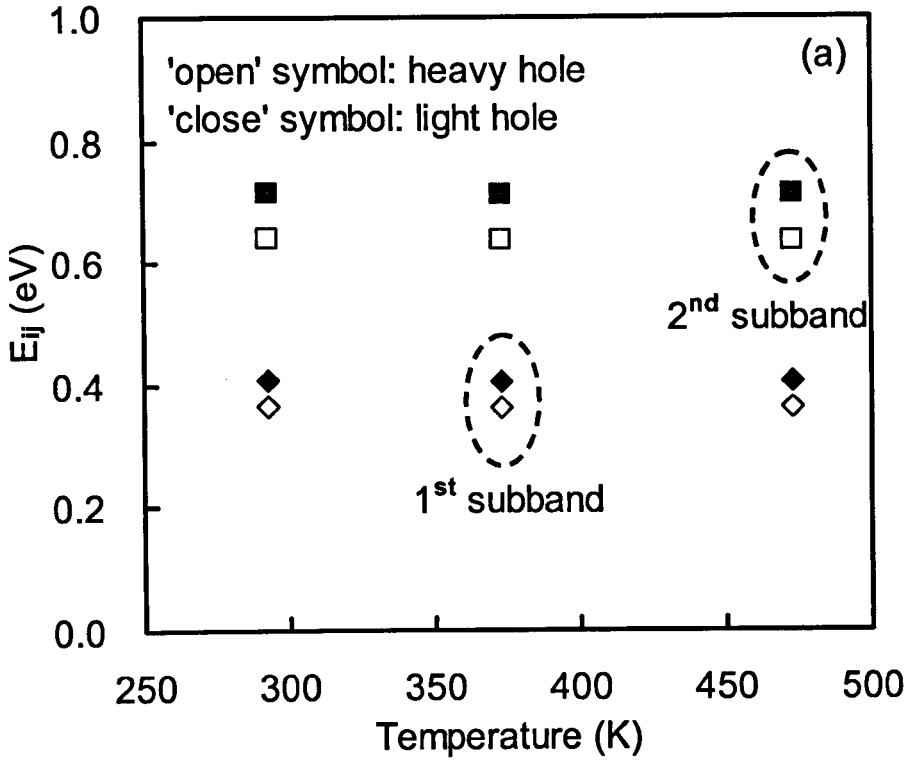
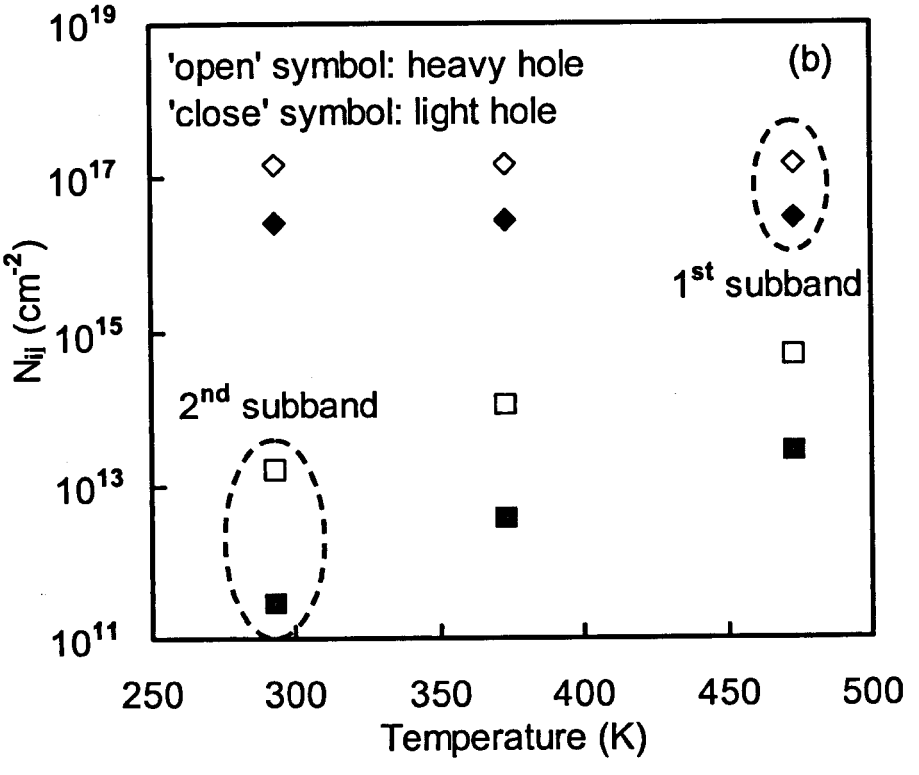

Tunneling from the 2nd & 3rd subband
of both valleys
(d)
σ2=3.51x10-15cm2
σ1=1.58x10-13cm2
As-grown hole trapping (V)
0.0
0.1
0.2
0.3
0.4
0.01
0.1
1
10
100
1000
Qh (1014holes/cm2)

Figure 46. The calculated hole current is used to estimate the capture cross sections of hole traps. (a) Holes tunneling from three subbands of both valleys. The capture cross sections found are: $\sigma_1 = 3.8\times10^{-17}$ cm^2 and $\sigma_2 = 9.8\times10^{-19}$ cm^2. These values are three orders of magnitude less than those reported in literature [48]. (b) Holes tunneling from the 2nd subband of the heavy hole valley alone. The extracted capture cross sections agree well with the reported values (solid line). The present data can also be reasonably fitted with the capture cross sections extracted from earlier work on relatively thick SiO_2 ($\sigma_1 = 7.3\times10^{-14}$ cm^2 and $\sigma_2 = 3.6\times10^{-15}$ cm^2, [48]) (dot line). (c) Holes tunneling from the 2nd subband of both valleys are taken into account. The 2nd subband of the light hole valley has a weak effect on σ_1 & σ_2. (d) Holes tunneling from the 2nd & 3rd subband of both valleys are taken into account. The contribution from the 3rd subband to Q_h is negligible. These results indicate that the main source of holes for filling the as-grown hole traps is the 2nd subband of the heavy hole valley.


'open' symbol: heavy hole
'close' symbol: light hole
(a)
2nd subband
1st subband
Eij (eV)
1.0
0.8
0.6
0.4
0.2
0.0
250
300
350
400
450
500
Temperature (K)


'open' symbol: heavy hole
'close' symbol: light hole
(b)
1st subband
2nd subband
Nij (cm-2)
10^19
10^17
10^15
10^13
10^11
250
300
350
400
450
500
Temperature (K)

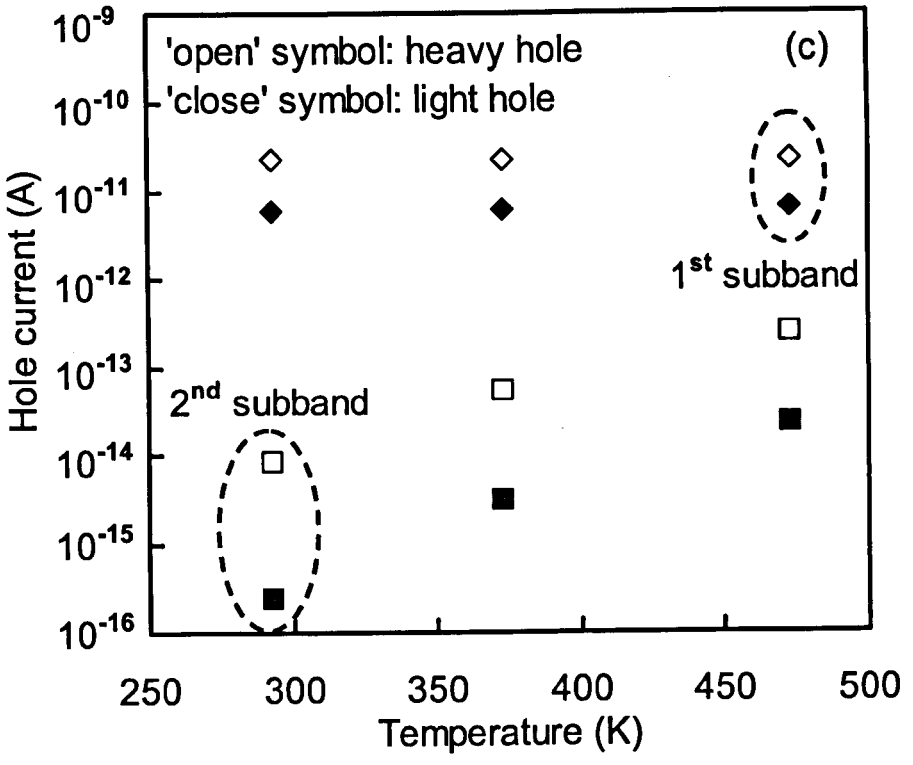


Figure 47. Calculation of (a) energy levels, (b) carrier concentration, and (c) hole current under different temperatures. The same parameter values were used for all temperature (V_g = -3.17 V, $N_A = 5\times10^{17}$ cm^{-3}, ϕ_B = 4.5 eV, and $m_{ox} = 0.35\times m_0$). Although elevated temperature will not increase the energy level of carriers, but it will increase the concentration of carrier at higher energy level. The transmission probability is independent of the temperature, resulting larger hole current for the 2nd subband at elevated temperature.

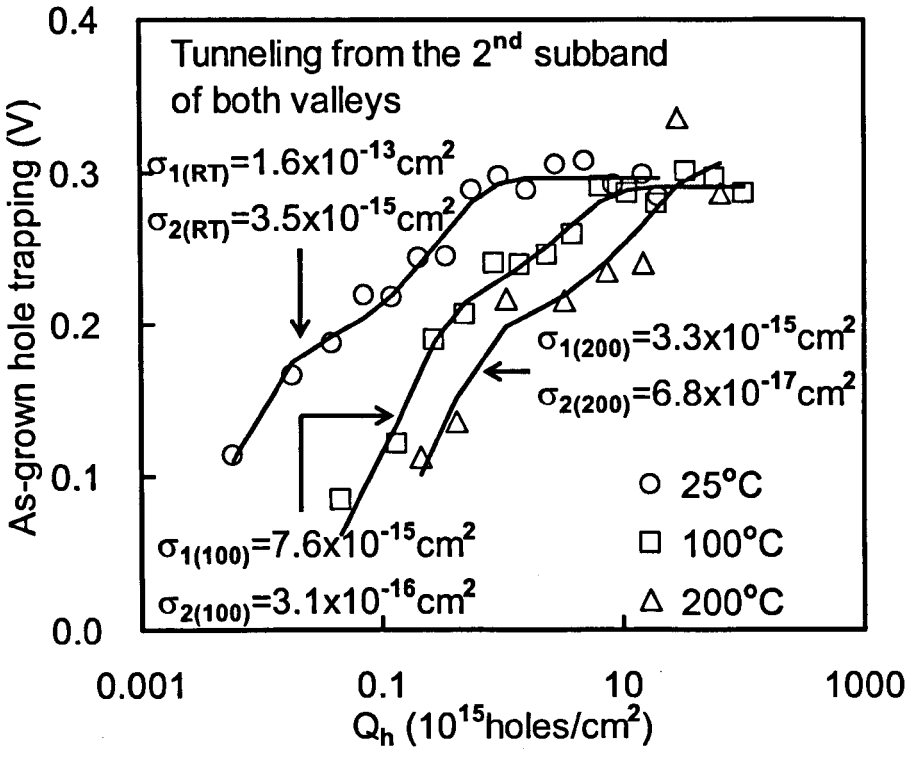


Figure 48. The effects of temperature on the capture cross sections. An increase of temperature from room temperature to 200°C reduces the capture cross sections by a factor of fifty.

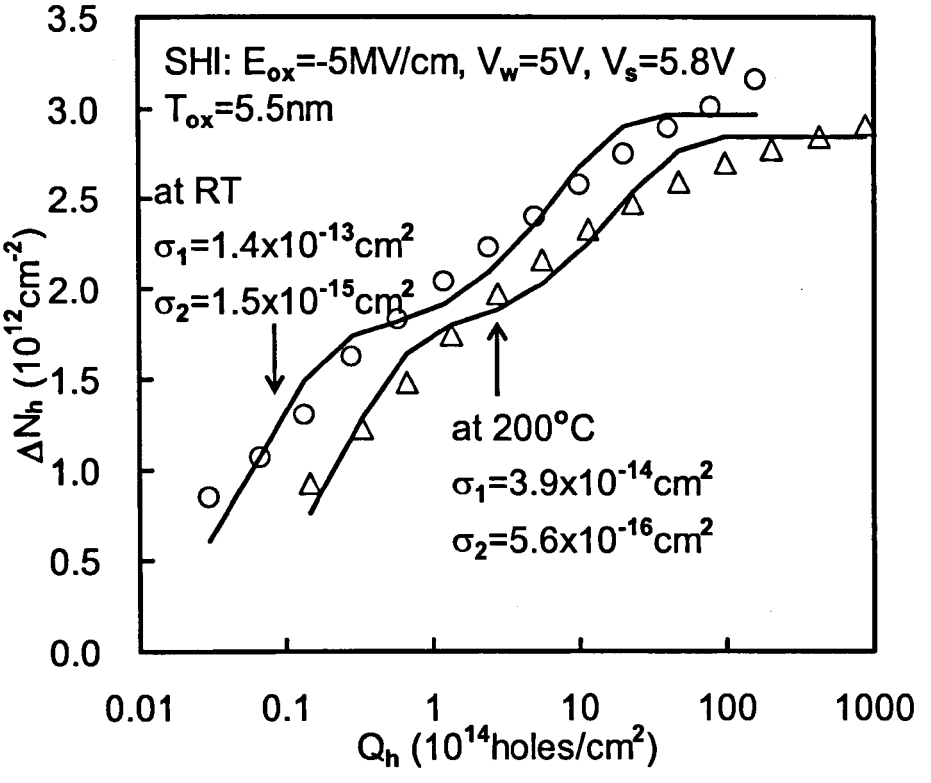


Figure 49. The effects of temperature on the capture cross sections for a relatively thick oxide (5.5 nm). The device was first heavily stressed ($Q_h = 10^{18}$ holes/cm^2), followed by a short electron injection to empty the trapped holes. Hole injection was switched on again to fill the as-grown hole traps at room temperature (E_{ox} = -5 MV/cm, V_w = 5 V, V_s = 5.8 V). Another short electron injection was applied before the as-grown hole trap filling at 200°C. The difference in capture cross sections is about a factor of four within this temperature range.

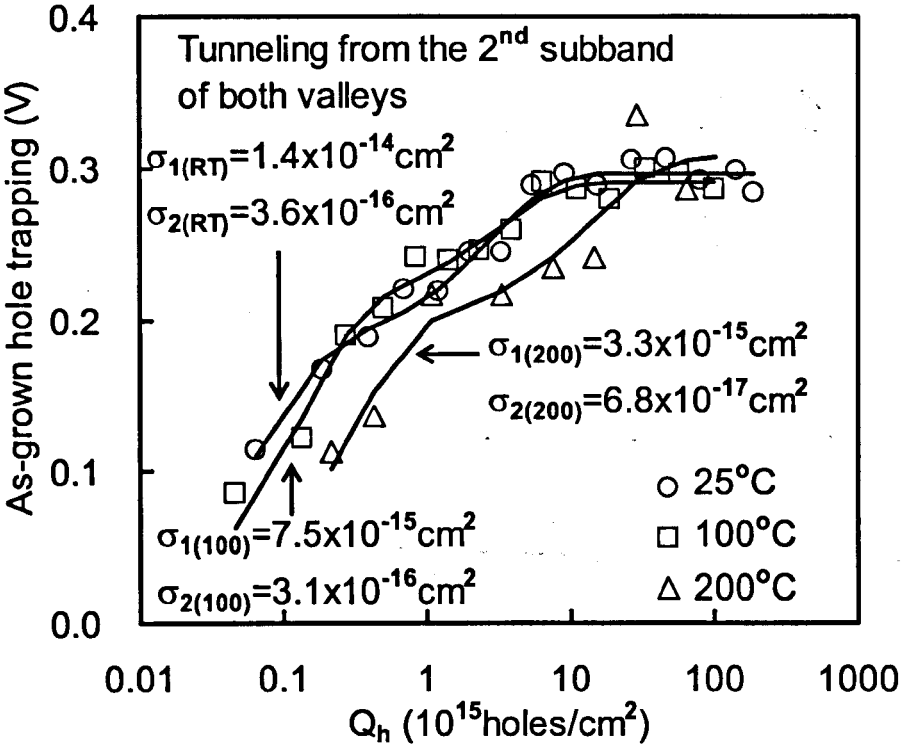


Figure 50. The energy level of the 2nd subband of the heavy hole valley is reduced from 0.64 eV to 0.57 eV. An increase of temperature from 25°C to 200°C only reduces the capture cross sections by a factor of four. A good agreement with the experimental observation on relatively thick is obtained.

6 Conclusions and Future Work

6.1. Conclusions

The electrical stress induced degradation of gate dielectrics in MOSFETs has been systematically investigated in this project. The work can be divided into three parts: the role of hydrogen in hole-induced electron trap generation, the trapping kinetics of electron traps generated in silicon dioxides, and the negative bias temperature instability. Conclusions for each part are given below:

6.1.1. Conclusions on the role of hydrogen in hole-induced electron trap generation

The focus was on whether holes can generate electron traps without going through hydrogen as intermediate species. Under substrate hole injection, hydrogenous species can potentially be released from three regions: Si/oxide interfacial region, oxide bulk and gate/oxide interfacial region. The impact of hydrogenous species released from these three regions on the generation was examined one-by-one.

Near the Si/oxide interface, it was shown that electron trap generation could be different for the same bombardment of interface and, consequently, the same hydrogen release. The threshold n-well voltage for electron trap generation is also different from that for interface states generation resultant from breaking the hydrogen bond. These results do not support that hydrogenous species released from the Si/oxide interface are important for hole-induced electron trap generation.

Under present experimental conditions, electron trap generation is strongly correlated with hole injection. These two also have the same threshold n-well voltage, indicating that hole injection controls the generation. After hole injection, its acceleration in the oxide is not responsible for creating electron traps. An increase of hydrogen density in a device has little impact on the hole-induced generation.

Finally, near the gate/oxide interface, any released H^+ would not travel through the oxide under negative bias. Since hydrogen-release and the subsequent transportation are thermally accelerated processes, the insensitivity of hole-induced electron trap generation to temperature in the range of 77K to 150°C suggests that these processes are not important for hole-induced electron trap creation. Present results support that holes can interact directly with the oxide to generate electron traps without going through hydrogen as intermediate species.

6.1.2. Conclusions on the trapping kinetics of electron traps generated in silicon dioxides

To improve the understanding of generated electron traps, trapping kinetics is investigated in this work. By stressing the oxide with substrate hole injection, the generation phase is separated from the trap filling phase. The use of uniform stress and filling removes the uncertainty in lateral distribution. The selection of relatively thick SiO_2 layer allows the direct measurement of trapped electron density and the uncertainty caused by the use of SILC is avoided. An achievement of this work is the development of a new method, which successfully corrected the effect of anti-neutralization positive charges (ANPC).

After the correction, electron capture cross section as large as $10^{-13} \sim 10^{-14}$ cm^2 is observed for generated acceptor-like electron traps. The filling kinetics follows the first order model and there are two genuine and well separated capture cross sections. The smaller capture cross section is in the region of $10^{-15} \sim 10^{-16}$ cm^2. For the first time, it is clearly shown that the density of the smaller trap saturates, while the density of the larger trap does not.

It is found that the traps created by FNI are similar to those by SHI in terms of both capture cross sections and the dependence of trap density on stress levels. This result supports that the same types of traps were created under different stresses, and the electrical signatures of traps were the same with or without hydrogen during the generation.

It is also shown that the electrical signature of generated electron traps does not depend on oxide thickness. The volume density is also insensitive to oxide thickness. This is in agreement with early assumption that the generation is a statistically random process. This work should provide useful information for modeling the SILC and oxide breakdown in the future. For example, the continuous generation of the larger traps implies that they will dominate the breakdown, while the saturation of the smaller traps makes them less important.

6.1.3. Conclusions on the negative bias temperature instability

Positive charges formed by negative bias temperature stress (NBTS) are investigated using the traditional transfer characteristics (TC) technique. Efforts were made to identify the types of positive charges and their dependence on stress conditions. We found that NBTS could generate both anti-neutralization positive charges (ANPC) and cyclic positive charges (CPC). While CPC has similar charging and discharging rates, ANPC is more difficult to neutralize. As stress time or temperature increases, CPC saturates, but ANPC does not. This suggests that they have different origins. Both NBTS induced CPC and ANPC behave similarly to those created by the substrate hole injection.

Effects of measurement temperature on positive charges were studied in details. It is found that, for a given number of defects, an increase of temperature leads to lower positive charges. Although CPC is insensitive to the measurement temperature, higher temperature enhances the neutralization of ANPC. Parts of NBTS induced positive charges are thermally unstable and can be annealed at stress temperature. When cooled for measurement at room temperature, some positive charges were lost. As a result, the density of defects responsible for positive charges is generally

underestimated when measured at either stress or room temperature. The implication to NBTI tests is explored. For a given number of defects, the variation of positive charges with temperature results in a temperature-dependent instability of threshold voltage.

The effects of measurement temperature on NBTI are also investigated for the newly proposed 'On-The-Fly' technique. For a given number of defects, it is found that ΔV_t increases for lower measurement temperature. The impact of measurement temperature on ΔV_t is stronger for longer stress time. This time-dependence results from that different defect dominate ΔV_t at different stress time. Although ΔV_t measured at stress temperature can appear insensitive to temperature, it does not imply that defect creation is not sensitive to temperature. The ΔV_t measured at stress temperature should not be used to assess the thermal acceleration of defect generation. The temperature-dependent instability of threshold voltage should be taken into account when estimating the NBTI-limited device lifetime.

Like the traditional measurement, it is found that three different types of defect contribute to ΔV_t measured 'On-The-Fly': one is the as-grown hole traps, the other two are created under NBTS. The as-gown hole traps can dominate the initial stage of NBTS, then saturate at longer time. The theoretical calculation indicates that holes tunneling from the 2^{nd} subband of the heavy hole valley dominate the filling of as-grown hole traps for thin SiON during typical NBTS. The capture cross sections are similar to those reported for relatively thick SiO_2: one in the order of $10^{-13} \sim 10^{-14}$ cm^2 and the other in the order of 10^{-15} cm^2.

Finally, the carrier mobility variation during NBTS was estimated. It is found that the effective mobility of holes at V_g = -3.17 V actually increases with NBTS time. This results from a reduction of $|V_g - V_t|$ during NBTI and the effective mobility increases for lower $|V_g - V_t|$.

6.2. Future work

Despite the effort made in this project, there are many problems remaining to be solved. These include, but are not limited to, the followings:

The mechanism of electron traps generation in gate oxides:

The result in this thesis show that the hole injection alone can generate acceptor-like electron traps, without going through hydrogen as intermediate species. However, under a typical Fowler-Nordheim injection, both electron and hole are injected. The number of injected electrons is generally higher than that of hole, and electrons can release hydrogen and introduce additional generation processes. The relative importance of hole and hydrogen induced generation will depend on their relative densities. Under the typical operation conditions, it is not clear if one of these two can dominate the generation for an industrial grade device. How hydrogen inducing electron traps is not known, either. These issues should be addressed in future works.

The relationship between generated electron traps and oxide degradation:

This work clearly shows that two types of electron traps are created under both substrate hole injection and Fowler-Nordheim injection. The generated defect densities change independently as stresses increase. Further works are needed for using this knowledge in modeling device instability. For example, it is interesting to study the relation between the generated two types of electron traps and the SILC. How the small traps contributing to breakdown is not clear. Furthermore, there is a lack of knowledge in the atomic structure of generated electron traps.

The negative bias temperature instability:

The present work shows that different types of defect are created during negative bias temperature stress. However, based on the commonly accepted hydrogen release model, it is unclear how hydrogen transportation could create two different types of positive charges, which requires further investigation.

Moreover, the hydrogen release model seems incapable in explaining the measurement temperature effects observed in this project. As a result, a new defect based model would be required to fully explain the degradation under negative bias temperature stress.

According to the hydrogen release model, both positive charges and interface states are created during negative bias temperature stress. However, some group of researchers believed that the generation of interface states is the main defect. In this thesis, when 'On-The-Fly' measurement technique is used to avoid the recovery effects, the shift in threshold voltage can be up to $5 \sim 6\times10^{12}$ cm^{-2}. The generation of interface states is relatively small when compared with the generation of positive charges. However, the interface states cannot be measured 'On-The-Fly' at present and new measurement technique should be developed.

The recovery phenomenon of negative bias temperature instability also needs further investigation. Once the stress voltage is removed, both the drain current and the threshold voltage shift would recover against time. This complicates the characterization of dynamic negative bias temperature instability, especially when attempting to extract the device lifetime. Neglecting the recovery effects will lead to a significant underestimation of negative bias temperature instability.

Finally, the calculation of hole currents through the dielectric is at a preliminary stage. More accurate models should be used in the future.

List of publications

Journal:

1. M. H. Chang and J. F. Zhang, "On the role of hydrogen in hole-induced electron trap creation," Semicond. Sci. Tech., **19**, 1333 (2004).
2. M. H. Chang, J. F. Zhang and W. D. Zhang, "An assessment of capture cross sections and effective density of electron traps generated in silicon dioxides," IEEE Trans. Elec. Dev., **53**, 1347 (2006).
3. W. D. Zhang, J. F. Zhang, C. Z. Zhao, M. H. Chang, G. Groeseneken and R. Degraeve, "Electrical signature of the defect associated with gate oxide breakdown," IEEE Elec. Dev. Lett., **27**, 393 (2006).
4. J. F. Zhang, M. H. Chang and G. Groeseneken, "Effects of measurement temperature on NBTI," submitted to IEEE Elec. Dev. Lett.
5. M. H. Chang and J. F. Zhang, "On positive charges formed under negative bias temperature stresses," submitted to J. Appl. Phys.

Conference:

1. M. H. Chang, J. F. Zhang, G. Groeseneken and R. Degraeve, "Damaging species for the hole injection induced electron trap generation," in 13th INFOS Conference, Barcelona, p.PS17, Spain (2003).
2. M. H. Chang and J. F. Zhang, "On the trapping kinetics of electron traps created in silicon dioxides," in 207th Meeting of the Electrochemical Society, Quebec City, p.199, Canada (2005).
3. M. H. Chang and J. F. Zhang, "An investigation of positive charges formed during negative bias temperature stress," in 207th Meeting of the Electrochemical Society, Quebec City, p.293, Canada (2005).

4. M. H. Chang, J. F. Zhang and G. Groeseneken, "Effects of 'On-The-Fly' measurement temperature on negative bias temperature instability," in 36th IEEE SISC, Arlington, USA (2005).
5. M. H. Chang, Y. Wang, J. F. Zhang, C. Z. Zhao, W. D. Zhang and M. Xu, "Contribution of as-grown hole traps to NBTI," abstract submitted to 211th Meeting of the Electrochemical Society, Chicago, USA (2007).