Lifetime Studies of 130nm nMOS Transistors Intended for Long-Duration, Cryogenic High-Energy Physics Experiments

J. R. Hoff[1], R. Arora[2], J.D. Cressler[2] *Fellow, IEEE*, G. W. Deptuch[1] *Senior Member, IEEE*, P. Gui[3], N.E. Lourenco[2], G. Wu[3], R. J. Yarema[1] *Life Member, IEEE*

***Abstract*– Future neutrino physics experiments intend to use unprecedented volumes of liquid argon to fill a time projection chamber in an underground facility. To increase performance, integrated readout electronics should work inside the cryostat. Due to the scale and cost associated with evacuating and filling the cryostat, the electronics will be unserviceable for the duration of the experiment. Therefore, the lifetimes of these circuits must be well in excess of 20 years. The principle mechanism for lifetime degradation of MOSFET devices and circuits operating at cryogenic temperatures is via hot carrier degradation. Choosing a process technology that is, as much as possible, immune to such degradation and developing design techniques to avoid exposure to such damage are the goals. This requires careful investigation and a basic understanding of the mechanisms that underlie hot carrier degradation and the secondary effects they cause in circuits. In this work, commercially available 130nm nMOS transistors operating at cryogenic temperatures are investigated. The results show that the difference in lifetime for room temperature operation and cryogenic operation for this process are not great and the lifetimes at both 300K and at 77K can be projected to more than 20 years at the nominal voltage (1.5V) for this technology.**

I. Introduction

Cryogenic operation improves all aspects of device performance such as speed, noise, and current drive [1]. However, these performance improvements may come at a cost of reduced device and circuit reliability. Hot carriers are well-known to degrade the performance of transistors. Hot carriers can arise with greater frequency and with higher energies at low temperatures. Thus the damage can occur much faster. The reason for the increase is straightforward: at low temperatures, phonon scattering is reduced and therefore the carrier mean-free path length is increased. Thus, the damage can occur much faster at cryogenic temperatures than at room temperature for a fixed stress condition [1][2]. With the increased mean-free path length, the probability that a carrier will obtain enough energy to cause impact ionization also increases. Scaling of CMOS processes inevitably produces larger electric fields within the device, contributing to generation of hot carriers. The combination of longer mean free path length and higher electric fields at the drain generally lessen reliability with scaling, and this is generally believed to be worse at cryogenic temperatures.

Future experiments at the Intensity Frontier in High Energy Physics will require cold electronics because the circuitry will be placed inside the cryogenic chambers. These experiments are planned for many years of operation and data collection. The electronics, installed as close as possible to the wires that are the sources of relevant signals, will be sealed inside the chamber and therefore will be completely inaccessible for the duration of the experiment. Therefore, the integrated readout circuits must be designed using processes and techniques that would yield the greatest likelihood for successful operation for 20 or more years. Clearly, a fundamental understanding of the lifetimes of cold electronics in commercially-available processes is essential and design rules must be adjusted as required in order to ensure that these lifetimes can be reached. In order to predict a degree of degradation after 5, 10, or 20 years of operation in a cold environment, accelerated degradation tests under overstress conditions must be carried out.

II. Hot Carrier and Cryogenic Effects

A. Hot Carrier Degradation and Its Effects on Device Performance

Hot carriers are energetic electrons and holes with higher effective temperatures than the carriers in the surrounding lattice [3]. Their higher energy is generated in and by the electric field *near the drain* in field-effect transistors. Therefore, damage is not uniformly distributed across the channel. It is located *near the drain*. Hot carriers can become energetic enough to cause impact ionizations. The carriers that are products of these ionizations, if they are the so-called "lucky electrons" [4], may overcome the Si/SiO_2 energy barrier, may damage bonds and may change the properties of the Si/SiO_2 interface. Hot holes generated in impact ionization in NFETs, for the most part, are swept away as substrate current. Hot electrons with enough energy to overcome the Si/SiO_2 barrier can damage the Si/SiO_2 interface itself causing interface traps that can exchange charge with the channel [3][5]. Alternatively, they can travel deeper into the oxide causing electron traps that can influence the channel but

Manuscript received November 4, 2011. This work was supported by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the United States Department of Energy.

[1] J. R. Hoff (telephone: 630-840-2398, e-mail: jimhoff@fnal.gov), G. W. Deptuch (e-mail: deptuch@fnal.gov), and R. J. Yarema (e-mail: yarema@fnal.gov), are with the Fermi National Accelerator Laboratory, Batavia, IL 60510 USA .

[2] R. Arora, N.E. Lourenco and J.D. Cressler (e-mail: cressler@ece.gatech.edu) are with the Georgia Institute of Technology, Atlanta, GA 30332 USA.

[3] G. Wu and P. Gui (e-mail: pgui@smu.edu) are with Southern Methodist University, Dallas, TX 75205 USA.

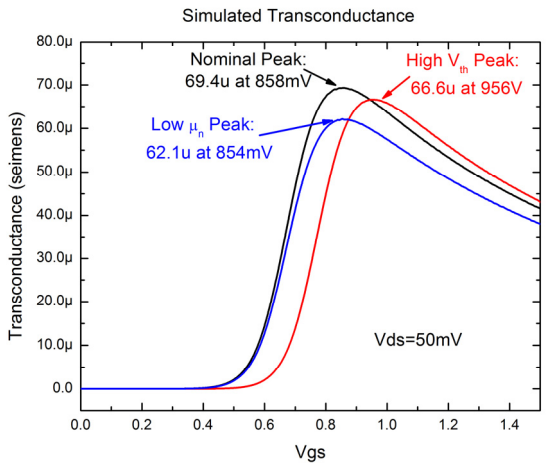

Fig. 1 Simulated transconductance (g_m) versus Gate-to-Source Voltage (V_{GS}) for different conditions in a 0.64µm/0.13µm NMOS. The black curve is for nominal room temperature process values as provided by the manufacturer. The blue curve shows the transconductance when only mobility is decreased by 10%. The red curve shows the transconductance when only the threshold voltage is increased by 100mV.

not necessarily exchange charge with the channel. Finally, they can travel through the oxide completely as gate current. Which alternative is followed depends on a number of factors, including the Si/SiO_2 barrier activation energy and oxide thickness. The Lucky Electron model [4] does not distinguish between the gate oxide over the channel, the spacer oxides over the Lightly-Doped Drains or the field-oxides in, for example, Shallow Trench Isolation (STI) regions. It is very possible that in modern, deep-submicron processes, the effects of Hot-Carrier Injection are more complex than in older processes with larger feature sizes.

The increase in negative charge at both the Si/SiO_2 interface and deeper in the oxide can alter the device threshold voltage V_{th} causing it to increase after stress. Consequently, there is a decrease in device current for any given bias condition. The increase in traps also decreases carrier mobility $\mu_{n,p}$. Together, the change in threshold voltage and the change in mobility have a measurable effect on channel transconductance. In strong inversion (for an nFET), this is:

$$g_m = \mu_n C_{ox} \frac{W}{L}(V_{GS} - V_{TH}) \qquad (1)$$

where μ_n is the electron mobility, C_{ox} is the oxide capacitance, W is the transistor width, L is the transistor length, V_{GS} is the gate-to-source voltage, and V_{TH} is the transistor's threshold voltage. Simulating transconductance versus V_{GS} for low drain-to-source voltage yields the familiar plots shown in Fig.1. Note that by the manufacturer's room temperature models, transconductance reaches a peak of 69.4 µS at a gate-to-source voltage of 858 mV (black curve). If the threshold voltage is increased by 100 mV, leaving the mobility unchanged, then the peak transconductance value decreases by approximately 4% and the peak position shifts by almost exactly 100 mV. If, on the other hand, mobility is decreased by 10%, leaving threshold voltage unchanged, then the peak transconductance value decreases by 10.5%, but the voltage of the peak position remains virtually unchanged.

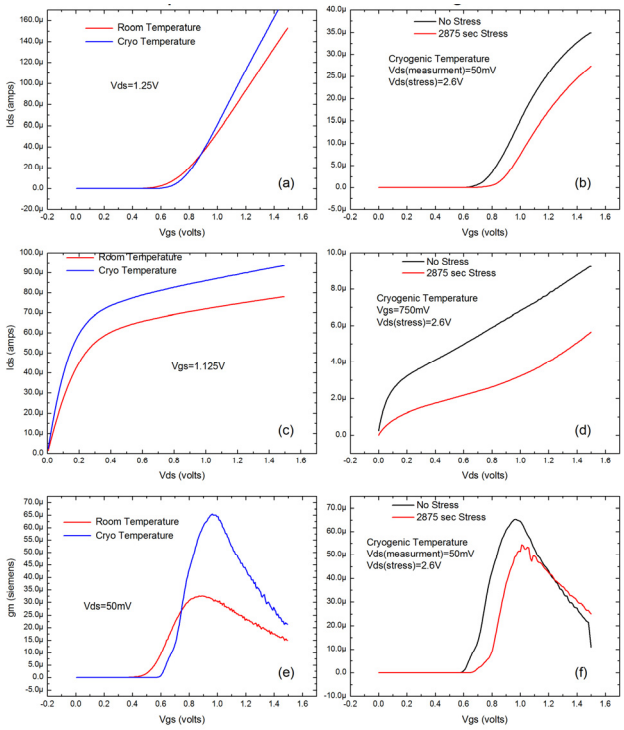

Fig. 2 The effect on device performance of cryogenic temperatures (a), (c), and (e) and the effect on device performance of hot carrier degradation (b), (d), and (f). (a) Drain current vs. V_{GS} for Room Temperature (red) and 77K (blue); (b) Drain current vs. V_{GS} for no stress (black) and 2875 seconds (as an example) of V_{DS}(stress)=2.6V at 77K (red); (c) Drain current vs. V_{DS} for Room Temperature (red) and 77K (blue); (d) Drain current vs. V_{DS} for no stress (black) and 2875 seconds (as an example) of V_{DS}(stress)=2.6V at 77K (red); (e) Small-signal transconductance vs. V_{GS} for Room Temperature (red) and 77K (blue); (b) Small-signal transconductance vs. V_{GS} for no stress (black) and 2875 seconds of V_{DS} (stress)=2.6V at 77K (red)

Therefore, it is possible to observe the change in both transconductance peak position and peak magnitude as a function of stress time. The decrease in peak magnitude is the result of both mobility decrease and threshold voltage increase. Although, in observing the change in peak location, to first order, the shift in threshold voltage can be extracted, throughout the rest of this paper, graphs showing the threshold voltage shifts use the more rigorous Constant Current Method [6] for V_{TH} extraction.

B. Cryogenic Temperature and Its Effects on Device Performance

With reduction in temperature comes a reduction in lattice vibration and, consequently phonon scattering [3]. This results in a substantial improvement in carrier mobility and a reduction in thermal noise. The improved mobility results in an increase in current drive capability for a given bias and improved transconductance. Another consequence of reduced temperature is an increase in the subthreshold slope and a reduction in drain-source leakage current when the device is cut off, resulting in improved switching times. Finally, threshold voltage increases, largely due to the temperature dependence of intrinsic carrier concentration.

Fig. 2 shows examples of experimental results of 0.64µm/0.13µ NMOS transistors displaying the effects of both cryogenic temperature (2(a), 2(c) and 2(e)) and hot carrier degradation (2(b), 2(d) and 2(f)).

III. Experimental Procedure

The Accelerated Stress experiments were carried out on two separate experimental setups, one for small devices and low currents and one for large devices and high currents. The low current setup used a HP4145B Semiconductor Parameter Analyser and the high current setup used a set of Keithley 237 and Keithley 2400 Source Measurement Units. Both setups were controlled by custom LabView programs that were created for these experiments and modified as needed. Ultimately, both followed the flow shown in Fig. 3.

The program applied definable stress voltages to the drains of the devices. The gate voltages were set to $\frac{1}{2}V_{DS}$ during stressing. The program stopped stressing at definable intervals to allow the devices to be measured. All measurements of g_m were conducted with transistors in the linear region with V_{DS}=50mV or V_{DS}=100mV, so, consequently, all measurements were stress-free. Moreover, measuring the device's parameters in the linear region ensures that charges crossing the channel must travel through the damaged region

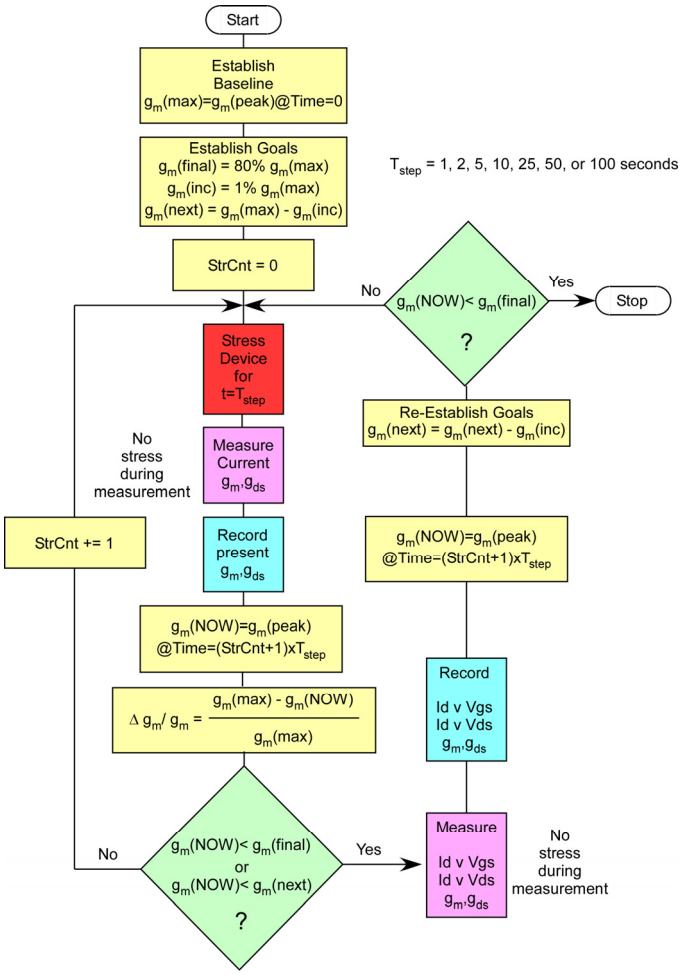

Fig. 3 The experimental procedure used in the Lifetime measurements.

[5]. As V_{DS} increases and the device moves into saturation, current is injected across the pinched-off region, and charges will not necessarily be exposed to the conditions of the damaged region.

The transconductance was determined by a numerical differentiation of the device's I_{DS} vs. V_{GS} curve at each V_{GS}. The peak of this curve was first determined by a simple numerical procedure and then, to eliminate noise, the transconductance curve was fitted by a 5th order polynomial in the vicinity of the coarsely defined peak. The transconductance peaks used in this paper were taken from the peaks of the fitted polynomials. The method demonstrated the desired stability and immunity to noise of individual measurements. Room temperature measurements were conducted in an air-conditioned room and cryogenic measurements were conducted in a liquid nitrogen bath.

The ultimate goal of these experiments is the prediction of lifetime and the development of guidelines for the design of circuits. End-of-life has been defined as a 10% reduction in transconductance and lifetime (τ) is the time required to take a device from virgin to a defined end-of-life. Lifetime is a function of stress voltage. Lifetime estimation is a long-standing procedure [7] and can be assumed to be correct if the degradation follows a power law. It is accomplished through stress experiments by plotting the log of lifetime versus $1/V_{DS}$, determining the rate-of-change of lifetime with inverse stress voltage and then projecting that slope to a lifetime of 20 years. The stress voltage that corresponds to such a lifetime represents an upper limit to possible power supplies for the circuitry. Classically, the stress step time was not critical, provided the step time was granular enough to observe the end-of-life condition. However, it was discovered over the course of this experiment that very fine time steps were necessary to avoid missing fast components of the Hot-Carrier induced damage processes.

IV. Experimental Results

The test transistors were all fabricated on the same wafer in a low power version of a 1.5V 130nm CMOS process with 2.3nm oxide thickness and Shallow Trench Isolation (STI). The available transistor sizes are 80µm/0.13µm, 80µm/0.5µm, and 0.64µm/0.13µm.

The substrate current, I_{SUB}, is an established indicator in the monitoring of Hot-Carrier Effects in MOSFETs. Fig. 4 shows the change in maximum I_{SUB} with V_{DS} as well as the change in V_{GS} corresponding to I_{SUB}(MAX) for the 80µm/0.5µm transistors. Note that $V_{GS}@I_{SUB}$(MAX) does move with increasing V_{DS}, but the I_{SUB} peak also broadens (not shown in the figure) with increasing V_{DS} making selection of $V_{GS}@I_{SUB}$(MAX)= $V_{DS}/2$ an adequate approximation of the most severe stressing conditions. Therefore, it was decided to use V_{GS}= $V_{DS}/2$ throughout the experiment.

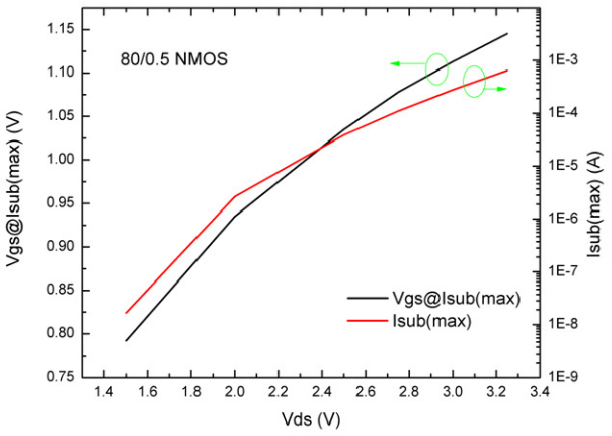

Fig. 4 A dual graph showing to the right the gate-to-source voltage V_{GS} that results in maximum substrate current and to the left the magnitude of maximum substrate current. Both graphs are with respect to drain-to-source voltage (V_{DS}).

The following sub-sections (A-C) provide a detailed examination of the results in each available transistor size. Section V discusses the gain factor for analog amplifiers. Section VI is a discussion of the results. Section VII gives lifetime predictions that can be made from the available data

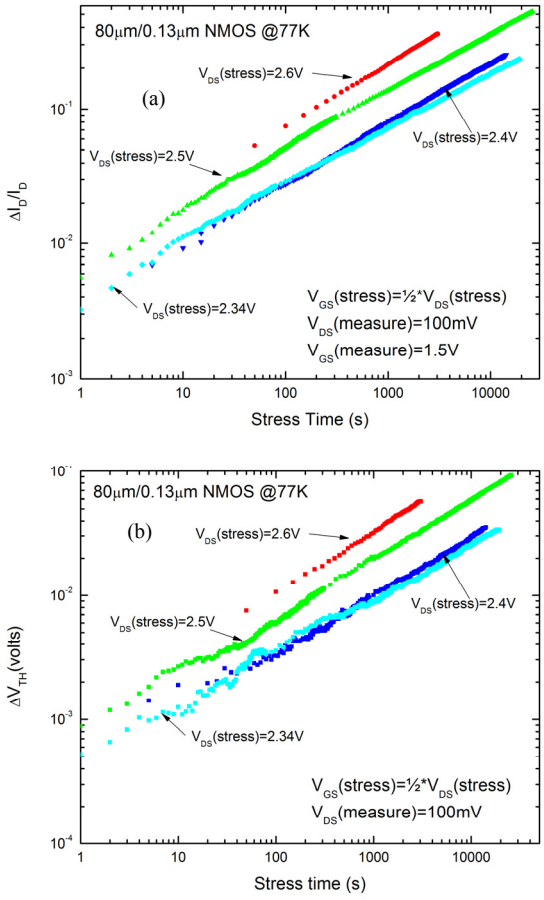

Fig. 5(a) The normalized change in current with stress time for different stress voltages. (b) The change in threshold voltage with stress time for different stress voltages.

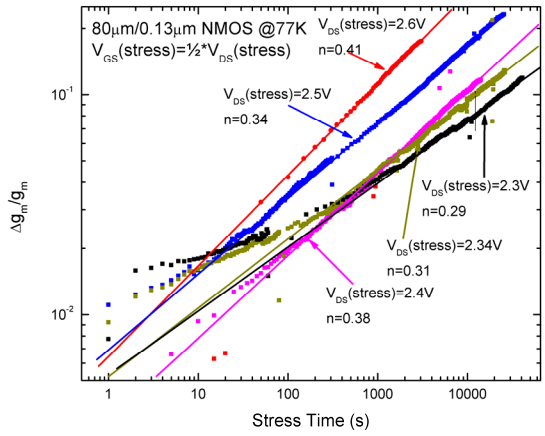

Fig. 6 The change in transconductance with stress time for different stress voltages

and finally Section XI will provide conclusions.

A. Results from 80µm/0.13µm

The tested devices are two-finger 80µm/0.13µm transistors with the drain in the middle. Fig. 5(a) shows the degradation of drain current with stress time for different stress magnitudes by plotting the change in drain current with time normalized to the initial drain current magnitude as shown in Eq. 2.

$$\frac{\Delta I_D}{I_D}(t) = \frac{(I_D(T=0) - I_D(T=t))}{I_D(T=0)} \quad (2)$$

For all measurements, the devices were characterized in the linear region with V_{DS} set to 100mV. For I_{DS} measurements, V_{GS} was set to 1.25V.

Fig. 5(b) shows the increase in threshold voltage with stress time for different stress magnitudes. All plots seem to follow the classical allometric (power law) format [4][5][8][9]

$$y = At^n \quad (3)$$

where A is a free fitting parameter that depends on V_{DS}, t is time, and n is a parameter that depends on stress, transistor size, technology and temperature.

The progression of the degradation of transconductance was traced in a manner equivalent to Eq. 2. Fig. 6 shows the degradation of transconductance with stress time for different stress magnitudes. At longer stress times and moderate stress voltages, these transconductance plots also appear to follow allometric formulae to the time limits of these stressing runs. However, their behavior at short stress times and particularly at lower stress voltages shows a departure from classical laws. Moreover, the initial measurements, even those after only 1 second of stress reveal some initial degradation up to 1%. At both 2.5 and 2.34V of stress voltage the normalized transconductance graphs seem to follow one another before splitting at approximately 20 seconds. The 2.4V curve was measured at 5 second initial intervals, while the other curves used 1 second initial curves. It is possible that a knee is not visible in the 2.4V curve due to under-sampling of the damage progression in this curve. At 2.3V, 2.34V and 2.5V of stress, the change in transconductance with time is small until a sudden change between 20 and 30 seconds. A fast initial drop

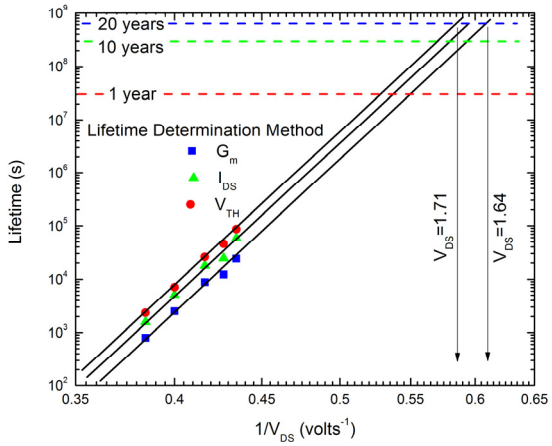

Fig. 7 Preliminary Lifetime predictions for the 80µm/0.13µm transistors at 77K by degradation of G_m and I_{DS} and by increase of V_{TH}

of g_m is observed and the knee observed in Fig. 6 moves to shorter stress times for higher stressing voltages.

Fig. 7 shows preliminary lifetime predictions based on three different (arbitrary) lifetime criteria, namely a 10% degradation of I_{DS}, a 10% degradation of g_m and a 50mV increase in V_{TH}. A definition of lifetime is arbitrary and a 10% degradation of I_{DS} is not quantitatively equivalent to, for example a 50mV increase in V_{TH}. Since all three lifetime predictions in Fig. 7 show similar change with stress voltage, it is concluded that the experimental method is sound.

B. Results from 80µm/0.5µm

Fig. 8(a) shows the degradation of drain current with stress time for different stress voltages for one-finger 80µm/0.5µm transistors. Note that here the degradation of drain current does not strictly maintain an allometric structure at high stress times.

Fig. 8(b) shows the increase in threshold voltage with stress time for different stress voltages for the 80µm/0.5µm transistors. Again, unlike the 80µm/0.13µm transistors, these transistors do not clearly follow an allometric trend. Moreover, unlike the current degradation, the threshold increase does not show a particular tendency towards super-allometric trends that change with increasing stress.

Fig. 8(c) shows the degradation of transconductance with stress time for different stress magnitudes. The rate of damage (i.e. the power law exponent, n) is changing (increasing) with stress voltage.

In comparing the 80µm/0.13µm results to the 80µm/0.5µm results, it is apparent that increased stress results in quicker degradation in both cases and that longer channel length results in longer lifetimes. For example, a common V_{DS}(stress) for 80µm/0.13µm and 80µm/0.5µm is 2.6V and the damage is one order of magnitude smaller on the device with the longer channel for the same stress times.

C. Results from 0.64µm/0.13µm

Fig. 9 shows the degradation of drain current with stress time for different stress voltages for 0.64µm/0.13µm transistors. Note that here also the degradation of drain current does not maintain an allometric structure throughout the stress experiment. Higher stress results in quicker

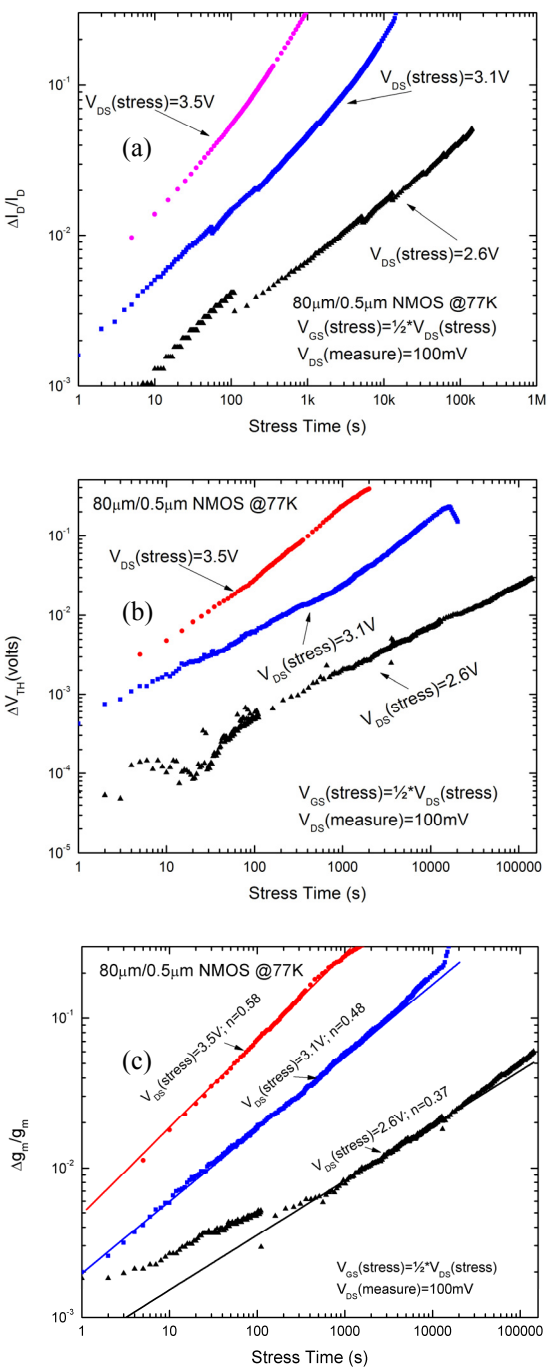

Fig. 8 (a) The normalized change in current with stress time for different stress voltages. (b) The change in threshold voltage with stress time for different stress voltages. (c) The normalized change in transconductance with stress time for different stress voltages

degradation, as expected, but like the 80µm/0.13µm there appears to be a knee in each curve that points to the activation of a degradation mechanism. This knee is visible between 10 and 100 seconds of stress for the 2.6V curve and between 50 and 100 seconds of stress for the 2.5V curve.

Fig. 10 shows the degradation of transconductance with stress time for different stress voltages for 0.64µm/0.13µm transistors. Here even more features are evident, i.e. activation knees and suggestion of saturation roll-over points

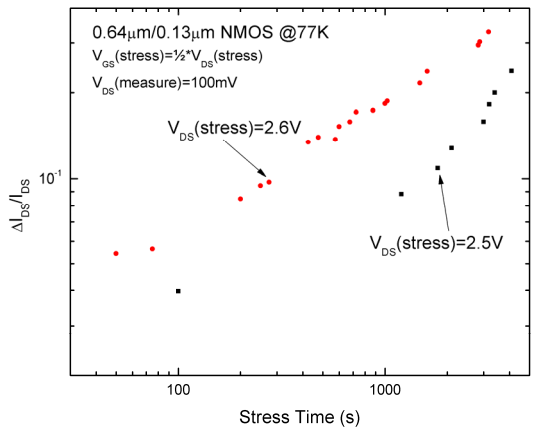

Fig. 9 The normalized change in current with stress time for different stress voltages for the 0.64μm/0.13μm transistors.

on some graphs. The curves point to initial degradations of up to 7% occurring at times below 10 seconds. A similar behavior was seen in Fig. 7 for the 80μm/0.13μm transistors but the initial degradation is more severe by a factor of several times for these narrower devices.

The narrow width of 0.64μm/0.13μm transistors requires the measurement of small currents. Additionally, smaller transistors exhibit poor matching. Currents of virgin transistors varied by up to 20% for the same bias conditions. These two facts resulted in a larger scattering of the results. Nevertheless, dependence of degradation on transistor width was observed. Comparing the 80μm/0.13μm transistors (Fig. 6) with the 0.64μm/0.13μm transistors (Fig. 10), there is a noticeable difference in device lifetime with device width. Room temperature measurements also showed a clear lifetime dependence on width. This observation is new and is even contrary to statements made in previous publications [1][10]. The smaller transistors were subject to a larger degree of degradation having verifiably shorter lifetimes.

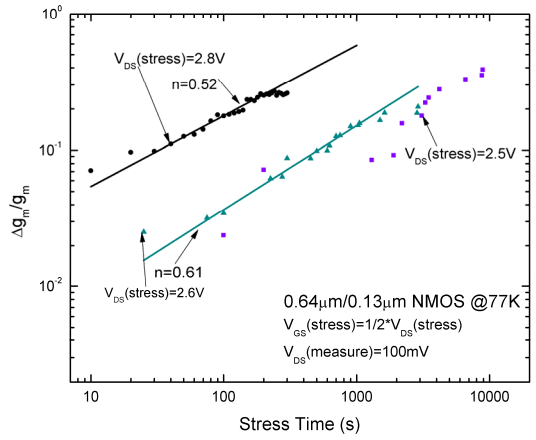

Fig. 10 The normalized change in transconductance with stress time for different stress voltages for the 0.64μm/0.13μm transistors. Note, there were too few points in the 2.5V stress experiment to provide an unambiguous power law fitting

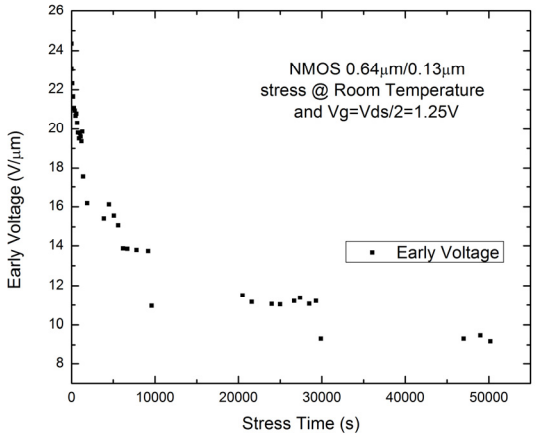

Fig. 11 The degradation of Early Voltage with stress time.

V. Early Voltage Degradation

The amplification factor μ_f is a fundamental parameter in analog circuit design. It is defined as

$$\mu_f = g_m r_{ds} \tag{4}$$

where r_{ds} is the drain to source dynamic resistance of the transistor in saturation, which is the inverse of the transistor's output conductance. The transistor's resistance is related to the drain current and the Early Voltage as follows:

$$V_A = \frac{I_{DS}\, r_{ds}}{L} \tag{5}$$

where L is the transistor's length. Therefore, given a known device length and a specific current, knowledge of the degradation of transconductance and the degradation of the Early Voltage is all that is needed to predict the degradation of the amplification factor.

Fig. 11 shows the degradation of the Early Voltage with stress time for a 0.64μm/0.13μm transistor at a stress voltage of 2.5 volts. Initially, there is a steep drop in Early Voltage that ultimately flattens out to a constant. The Early voltage in the figure decreases by a factor of two. Therefore, the dynamic resistance of the channel follows.

VI. Discussion

The plots presented in Section IV subsections A-C of this paper reveal a complex background of various mechanisms underlying the total degradation. The damage occurs at different rates and at different stages and depends on transistor dimensions and stressing voltages. A simple, classical, single-slope power law cannot explain the features seen in these transistors. The multiple slopes seen in the figures result from a set of damage processes that have dominant times and weaken after saturation. It should also be noted that certain types of damage are detectable depending on whether the characterization is performed in the linear or saturation region. The measurements shown thus far have all been performed in the linear region. However, Fig. 12 shows one stressing run that was executed on an 80μm/0.5μm device in CT in which the measurements were performed in both the linear (V_{DS}=50mV, black curve) and saturation (V_{DS}=750mV , red

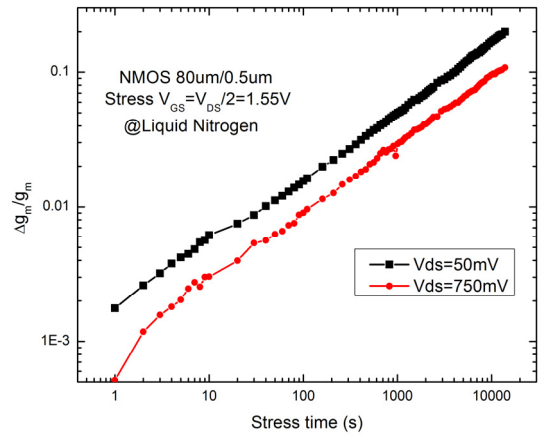

Fig. 12 The degradation of transconductance of an 80µm/0.5µm device estimated in the saturation and linear region.

curve) regions. It is evident that estimations of lifetime differ by a factor of 3 in this stressing experiment, resulting in a longer lifetime prediction for the measurements in saturation. The degradation damage occurs near the drain, and the pinch-off of the channel of a transistor in saturation permits charge transport to bypass the damage to a degree dependent on the saturation depth [5]. Thus, for devices in general, the estimation of lifetime predicted in the saturation region could be overly optimistic. On the other hand, for transistors that will remain in saturation during operation, the estimation of lifetime predicted in the linear region is pessimistic. The estimation of lifetime predicted by measurements in the linear region can certainly be viewed as a worst case and, therefore, it is used in Section VII for the conclusions.

To illuminate the different degradation stages, the various sections of the degradation curves were fitted to power law formulae. Some of these results were shown in the figures of the preceding sections. The first degradation stage happens early and is over quickly. Its presence is most often seen indirectly in the first measurements (typically performed after 1s or more of stress) where the initial damage can be as high as a few percent. The first damage is seen if the characterization is done in the linear region of transistor operation and is attributed to the degradation of the series resistance of the drain. This first degradation stage may transition directly into the second stage for higher stress voltages. Alternatively, if the stress voltage is low enough, some flat portion of the curve may remain visible for a short period of time until the second degradation stage takes over. This second degradation stage is attributed to mobility degradation. In some experiments, because of the granularity of the stressing time, only a few points in this stage are seen. During the flat transition between the first and second stage, if it is visible, n is relatively low. For example, see Fig. 6 for the 2.34V and 2.5V curves. The 2.4V curve in Fig. 6, unfortunately, uses a 5 second stress period, which is not fine enough to see the plot truly flatten out. Fig. 10 is also flat at low stress times, but here the stress time steps were 10 seconds for the 2.8V stress and 50 sec for the 2.6V curve, so only a few points are seen. After this first stage, an activation knee is seen after which the n parameters rise considerably.

Dual Mode or Early Mode Degradation has been reported in literature [5][8][10][11][12]. It has been used to explain observations of dual-slope degradation in other technology nodes. Dual-Mode degradation has been characterized as a phenomenon unique to Lightly-Doped Drain (LDD) NMOSFETs [5][8] and it appears to depend on the stress condition (stress magnitude, V_{GS}=V_{DS}, V_{GS}=1/2*V_{DS}, etc.) [5][8][11], fabrication technology [5], and device temperature [11]. One commonly accepted model for Dual-Mode Degradation is put forth in [8] and attributes the early mode to drain resistance modulation due to Hot-Carrier Injection into the LDD spacer oxides. The phenomena observed in these 130nm devices resemble these reports of Early Mode.

The extreme early stage of degradation has also been reported in literature [14] and has been attributed to damage to the spacer oxide which happens on a sub-second scale. The latter "conventional" and ultimately more dominant mode is attributed to Hot-Carrier Injection into the gate oxide over the channel. The ultimate lifetime has been suggested to be determined by the latter degradation mode [8] because the early mode saturates and ceases to affect degradation.

In extremely narrow devices, this early stage degradation

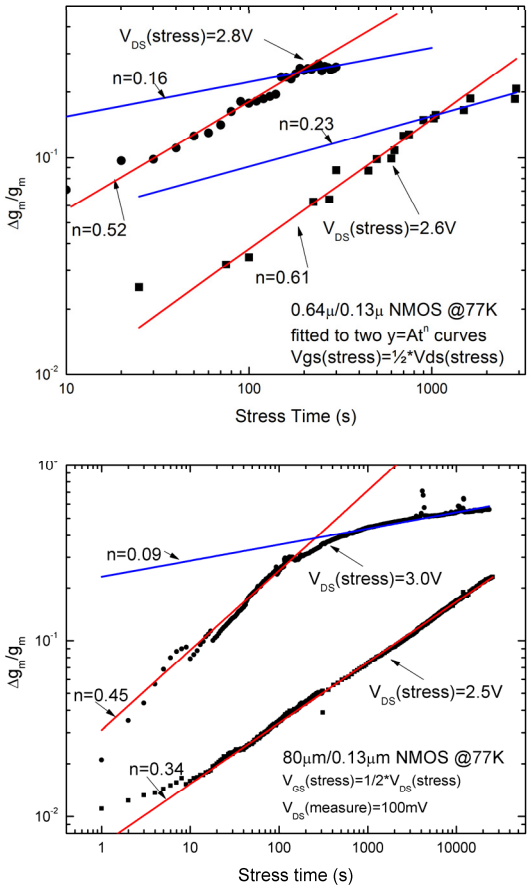

Fig. 13 (a) Evidence of change of slope in 0.64µm/0.13µm transistors (b) Evidence of change of slope in 80µm/0.13µm transistors

could be exacerbated by the presence of the oxides of the STI regions. In the wider devices typically studied, hot carrier injection into these field oxides would not influence the entire channel, only the edges. However, narrower devices would naturally experience this effect to a greater degree. This could account for the fast initial degradation seen in Fig. 10 where, after a single 10 second stress stage at V_{DS}(stress)=2.8V, 7% transconductance degradation was observed and after a single 25 second stress stage at V_{DS}(stress)=2.6V, almost 3% transconductance degradation was observed.

In the second degradation stage, the n coefficient changes to n≈0.4-0.6, which is close to the value of n=0.5 typically attributed to the formation of interface states [13]. According to [14], after the fast initial stage saturates, mobility degradation starts and damage spreads forward through the gate-drain overlap of the transistor and into the channel at a rate dependent on the magnitude of the electric field. This implies the formation of interface states, so power law formulae with n≈0.4-0.6 would be consistent with [14] as well. This second stage can be seen in the allometric slopes in Fig. 6 at times greater than 100 seconds, and in the allometric slopes seen in Fig. 10 at around 100 seconds for the 2.8V plot and 800 seconds for the 2.6V plot.

Finally, for long stress times, a final saturation occurs in which the n parameters change to values typically between 0.1 and 0.2 as shown in Fig. 13. This is consistent with the observations of [12] and is attributed to an increase Si/SiO_2 activation energy due to the filling of all available defect sites. The change of slope occurs earlier for larger stress voltages. It is also apparent that degradation would achieve saturation for all stressing runs if only tests could continue long enough.

The prediction of end-of-life is complicated by the fact that results may vary significantly due to the change of the n parameter, and also due to the arbitrariness of the end-of life goal. If 10% degradation of g_m is targeted, the goal may be achieved before or after a particular change of slope. Thus the lifetime can be underestimated or overestimated depending on the choice of end-of-life goal. To illustrate the range of possible underestimation, it is enough to analyze the device lifetime τ given by the power law that is expressed by the following proportionality [9][13]:

$$\tau \propto e^{\frac{1}{n}} e^{\frac{a}{V_{DS}}}, \qquad (6)$$

where a is a constant parameter. The value of the $\exp(1/n)$ component changes from about 150 to about 7 when n changes its value from 0.5 to 0.2. Thus the estimation of lifetime can vary by a factor greater than 20 if the end of life goal is changed from before to after a slope change - for example, if the end-of-life goal was changed from 10% to 20% degradation in transconductance.

The ultimate lifetime in devices exhibiting a single slope in degradation has always been determined by setting the target degradation level. In modern devices, when the slope of degradation changes in the course of the stressing measurements, the predictions may be imprecise but it is prudent to err on the side of caution.

Note that from Fig. 7, Fig. 9, and Fig. 11, the end-of-life goal, set at a 10% degradation of transconductance, is reached after the first slope change, but long before the second slope change. Therefore, in the interest making a conservative and yet realistic estimation, the classical lifetime graph was formed using lifetimes extracted from after the first slope change and before the second slope change – i.e. from when the n parameter was in the range of 0.5. This choice avoids the pessimistic predictions that would be based on the saturating early mode degradation present in the high-stress experiments. More importantly, it avoids the danger of making optimistic predictions based on the saturated, low-slope trend seen at the ends of stressing experiments.

VII. Prediction of Lifetime

Fig. 14 shows a series of very conservative lifetime predictions based on the 10% degradation of transconductance.

The blue squares in Fig. 14(a) are cryogenic 0.64µm/0.13µm transistors that have a lifetime prediction of 20 years of operation if a power supply rail of 1.54V or less is used. The red triangles in Fig. 14(a) are room temperature 0.64µm/0.13µm transistors that have a lifetime prediction 20 years of operation if a power supply rail of 1.59V or less is used.

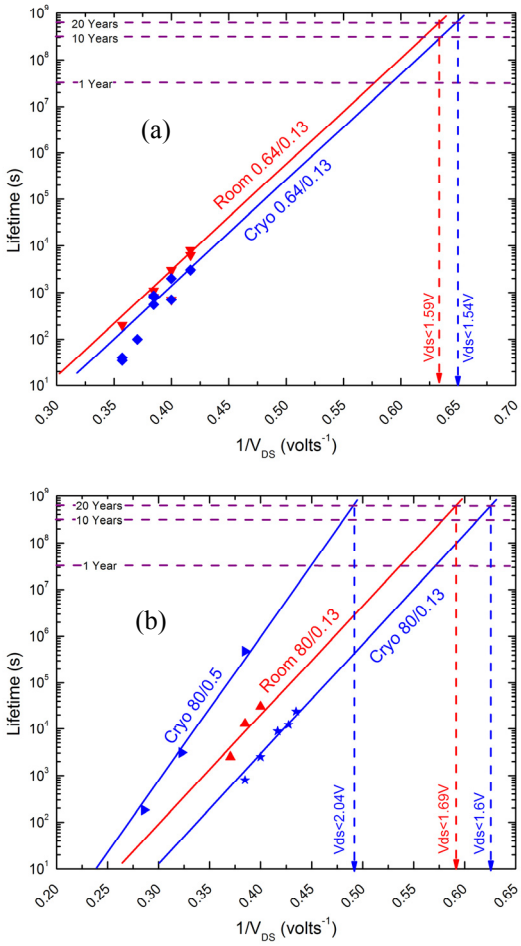

Fig. 14 Shows the projected lifetimes for various geometries and temperatures. Red corresponds to room temperature predictions. Blue corresponds to Cryogenic.

The blue right-pointing triangles in Fig. 14(b) are cryogenic 80µm/0.5µm transistors that have a lifetime prediction of 20 years of operation if a power supply rail of 2.04V or less is used. The red up-pointing triangles are room temperature 80µm/0.13µm transistors that have a lifetime prediction of 20 years of operation if a power supply rail of 1.69V or less is used. The blue stars are cryogenic 80µm/0.13µm transistors that have a lifetime prediction of 20 years of operation if a power supply rail of 1.6V or less is used.

VIII. Conclusions

Longer channel length results in longer lifetimes which is in keeping with previous research. This is very important for circuit designers because it provides one necessary tool by which longer device lifetimes can be guaranteed. A second lever would be to reduce the power supply rails. In this particular case, both the room temperature and cryogenic measurements predict that the devices can be run at their nominal voltage (1.5V) and still reach the desired 20 years of operation. Reduction of the power rails is, therefore, not strictly necessary, but it is still an effective means of extending lifetime. It should not be overlooked that transistors at cryogenic temperatures have improved transconductance, subthreshold slope, noise, and drive capability and that, for this process, these gains come at very little cost to device lifetime. It indicates that this deep-submicron process is remarkably resistant to hot-carrier degradation and this is in keeping with past predictions related to deep sub-micron processes.

The role of STI could be similar to the role of the spacer oxide in device degradation and this will be particularly significant to narrow devices and therefore particularly significant to digital designs. The degradation of devices in the linear region is relevant to digital switching times. Testing in the saturation region alone will not reveal this.

It is worth noting again that the chosen definition of lifetime is very significant. Had the definition of end-of-life been different, for example, a 20% degradation of transconductance, then predictions of lifetime would have included the saturated, low-n slopes of the degradation graphs, leading to significantly different conclusions. Therefore, a deeper understanding of exactly what circuit parameters are critical and exactly how much change can be tolerated in those parameters are essential.

IX. Acknowledgment

The authors would like to thank Dr. Veljko Radeka of Brookhaven National Laboratory for thorough discussions of the material, valuable comments and suggestions on interpreting the results and, ultimately, helping to direct the article to its final soundness and form.

Thanks go to Albert Dyer of the Fermi National Accelerator Laboratory for wire bonding and board assembly.

References

[1] Wang-Ratkovic, J., Lacoe, R., MacWilliams, K., Song, M., Brown, S., and Yabiku, G., "New Understanding of LDD CMOS Hot-Carrier Degradation and Device Lifetime at Cryogenic Temperatures", *35th Annual Proceedings of the Reliability Physics Symposium*, 1997 , pp. 312 – 319

[2] Song, M., MacWilliams, K., and Woo, J., "Comparison of NMOS and PMOS Hot Carrier Effects from 300 to 77K", *IEEE Transaction of Electronic Devices*, Vol. 44, No. 2, p. 268, 1997.

[3] Kim, S.A., "Hot Carrier Reliability of MOSFETs at Room and Cryogenic Temperature", Ph.D. Dissertation, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Massachusetts, USA, 1999

[4] Hu, C., Tam, S.C., Hsu F.-C., Ko, P. K., Chan, T. Y., and Terrill, K. W., "Hot-electron induced MOSFET degradation-model, monitor, and improvement," *IEEE Trans. Electron Devices,* vol. ED-32, pp. 375-385, 1985.

[5] Naseh, S. and Deen, M. J., "RF CMOS Reliability", *International Journal of High Speed Electronics and Systems*, Vol. 11, No. 4 (2001) 1249-1295

[6] Loke, A. L., Wu, Z-Y, Moallemi, R., Cabler, C. D., Lackey, C.O., Wee, T.T., Doyle, B. A., "Constant-Current Threshold Voltage Extraction in HSPICE for Nanoscale CMOS Analog Design", Synopsis Users Group (SNUG) 2010, San Jose, CA, USA (2010)

[7] Chen, T., Najafizadeh, L., Zhu, C., Ahmed, A., Diestelhorst, R., Espinel, G., Cressler, J., "CMOS Device Reliability for Emerging Cryogenic Space Electronics Applications", *2005 International Semiconductor Device Research Symposium*, pp 328-329, 2005

[8] Chan, V-H. and Chung, J. E., "Two-Stage Hot-Carrier Degradation and Its Impact on Submicrometer LDD NMOSFET Lifetime Prediction", *IEEE Transactions on Electron Devices*, Vol. 42, No. 5, (1995) 957-962

[9] Takeda, E. and Suzuki, N., "An Empirical Model for Device Degradation Due to Hot Carrier Injection", *IEEE Electron Device Letters*, VOL. EDL-4, No. 4, (1983) 111-113

[10] Raychaudhuri, A., Deen, M. J., King, M. I. H., and Kwan, W. S., "A Simple Method to Qualify the LDD Structure Against the Early Mode of Hot Carrier Degradation", *IEEE Transactions on Electron Devices*, Vol. 43, No. 1, (1996) 110-115

[11] King, E., Lacoe, R., and Wang-Ratkovic, J., "The Role of the Spacer Oxide in Determining Worst-case Hot-Carrier Stress Conditions for NMOS LDD Devices", *38th Annual International Reliability Physics Symposium*, San Jose, California, (2000) 83-92.

[12] Raychaudhuri, A., Deen, M. J., Kwan, W. S., and King, M. I. H., "Features and Mechanisms of the Saturating Hot-Carrier Degradation in LDD NMOSFET's", *IEEE Transactions on Electron Devices*, Vol. 43, No. 7, (1996) 1114-1122

[13] Renn, S-H, Pelloie, J-L, and Balestra, F., "Hot-Carrier Effects and Reliable Lifetime Prediction in Deep Submicron N- and P-Channel SOI MOSFET's", *IEEE Transactions on Electron Devices*, Vol. 45, No. 11, (1998) 2335-2342

[14] Manhas, S. K., De Souza, M. M., Oates, A.S., Chetlur, S.C., and Sankara Narayanan, E. M., "Early Stage Hot Carrier Degradation of state-of-the-art LDD N-MOSFETs", *38th Annual International Reliability Physics Symposium*, San Jose, California, (2000) 108-111