Direct-Current Measurements of Oxide and Interface Traps on Oxidized Silicon

Arnost Neugroschel, Fellow, IEEE, Chih-Tang Sah, Fellow, IEEE, K. Michael Han, Student Member, IEEE, Michael S. Carroll, Toshikazu Nishida, Member, IEEE, Jack Theodore Kavalieros, and Yi Lu, Member, IEEE

Abstract—A direct-current current-voltage (DCIV) measurement technique of interface and oxide traps on oxidized silicon is demonstrated. It uses the gate-controlled parasitic bipolar junction transistor of a metal-oxide-silicon field-effect transistor in a p/n junction isolation well to monitor the change of the oxide and interface trap density. The dc base and collector currents are the monitors, hence, this technique is more sensitive and reliable than the traditional ac methods for determination of fundamental kinetic rates and transistor degradation mechanisms, such as charge pumping.

I. INTRODUCTION

IT IS WELL recognized that the electrical characteristics of metal-oxide-semiconductor transistors (MOST’s) and bipolar junction transistors (BJT’s) degrade during circuit operation due to channel-hot-electron (CHE) and substrate-hot-electron (SHE) stresses which increase oxide (QOT) and interface trap (QIT) densities [1], [2]. In MOST’s, the trapped charges reduce the mobility (Δμ) and shift the threshold gate voltage (ΔVGT), both of which reduce drain saturation current (ΔID) which slows down the switching speed due to longer charging time of interconnect or load capacitances at lower currents. The trapped charges also shift the subthreshold gate voltage (ΔVGTSUB), and decrease subthreshold slope of the drain-current versus gate-voltage curve, which reduces the current cut-off sharpness, thereby increasing leakage current or standby power and decreasing the noise margin. In BJT’s, QIT and QOT will increase the minority carrier recombination rate in the base, thereby reducing its current gain, such as the common-emitter current gain, βP [3]. Thus, a quantitative separation of the effects of QOT and QIT is necessary to delineate the location and physical origin of the degradation in order to design and manufacture highly reliable integrated circuits with ten-year or longer operating life.

This research was performed by Neugroschel, C.-T. Sah, K. M. Han, M. S. Carroll, T. Nishida, and J. T. Kavalieros with the Department of Electrical Engineering, University of Florida, Gainesville, FL 32611-6200 USA.

Y. Lu was with the Department of Electrical Engineering, University of Florida, Gainesville, FL 32611-6200 USA. He is now with the Intel Corporation.

IEEE Log Number 9413284.

1For a brief tutorial review, see pp. 323325 and 674678 of [1] and Appendix B of [2].

The separation of QOT and QIT is generally difficult. It has not been reliably separated using the traditional capacitance and conductance methods or the transient methods because the test structures are two-terminal capacitors, or very small test transistors which give extremely small capacitances due to the very small device area. Many traditional methods for separating QOT and QIT were reviewed [4], and a two-step method was demonstrated. However, it uses the subthreshold slope to monitor QIT which is reliable only when there is not an inhomogeneous or lateral distribution of QIT and QOT. Hence, it is not reliable for monitoring the highly nonuniform QIT and QOT generated by CHE stress.

A novel method is demonstrated in this paper which measures the dc base and collector currents versus the gate voltage, to be known as DCIV method (in analogy to the traditional usage such as HFCV for high-frequency capacitance-voltage or QSCV for quasi-static CV), to monitor the QIT and QOT. The novel DCIV method contains two features: 1) The base current (IB) of the vertical BJT is used to measure the recombination current at the interface traps generated during fabrication or operation which avoids the error from lateral distribution or areal nonuniformity of QIT and QOT because IB is directly proportional to NIT or QIT/q. 2) The collector current (IC) of the vertical BJT is used to measure the QOT because IC increases sharply when the gate voltage passes the flat-band value toward depletion and inversion. The method will be described in this article using the nMOST and npnBJT of the BiMOS structure shown in Fig. 1. This BiMOS structure has been used previously to fabricate large test transistors with nearly 40000 μm² gate oxide area by Thompson [8] and [9] cited in [4]), but it is also present in the submicrometer nMOST’s in a p-well on n-substrate of production CMOS (Complementary MOS) inverter circuits. Thus, the novel DCIV method to be described can be easily applied to production test transistors and some examples to be given were data measured on micrometer and submicrometer MOST-BJT production structures.

With reference to Fig. 1, the BJT can be measured before and after a stress in two configurations: The top-emitter (top-E) or bottom-emitter (bottom-E) measurement configurations, with the n+drain/p-base or n+substrate/n-epitaxy/p-base as the forward biased emitter/base junction. Our geometrical terminology deviates from the traditional, emitter-up and emitter-down, which confuses the geometrical location of the emitter.
with the emitted-charge-flux direction. In both configurations, the shape of the $I_B$-$V_{GB}$ curve and the magnitude of $I_B$ at a constant $V_{FE}$ will measure $Q_{IT}$ [5], [6]. However, we recently anticipated that the shape and magnitude of $I_C$ will also be a strong function of $V_{GB}$ in both configurations because $I_C$ increases sharply at the flat-band gate voltage, $V_{GB}$-flatband, from a low constant current to a high constant current at strong inversion voltage, $V_{GB}$-threshold. This sharp increase occurs when the electron-channel between the n+ drain and n-source appears at $V_{GB}$-flatband, which abruptly increases the emitter-base area in the top-E configuration and the collector-base area in the bottom-E configuration.

The stress-induced base current, $\Delta I_B$, is solely due to electron-hole recombination at the stress-generated interface traps [5], [6], hence, is a function of stress-induced interface charge and trap concentrations, $\Delta Q_{OT}$ and $\Delta N_{IT}$, or the stress-induced density-of-states of the interface traps and surface recombination velocity, $\Delta D_{IT}$ and $\Delta S_0$. However, the increase of the collector current with $V_{GB}$ is nearly all from geometrical increase in the emitter or collector area contributed from the nMOST’s electron channel. Therefore, the lateral shift in the $I_C$-$V_{GB}$ curve, $\Delta V - GB$, is mainly a function of the stress-induced change of flatband gate voltage, $\Delta V_{GB}$-flatband, and hence is a very sensitive monitor of $\Delta Q_{OT}$ and $\Delta Q_{IT}$. Thus, combining the $\Delta I_B - V_{GB}$ and $I_C - V_{GB}$ data will enable the separation of $\Delta Q_{OT}$ and $\Delta Q_{IT}$. Experimental data in the following section will demonstrate this capability of the novel DCIV method.

Minority carrier surface recombination rate or velocity $S_0$ at the Si/SiO$_2$ interface was studied extensively since the use of MOS-gate-controlled BJT was demonstrated by one of us in 1961–1962 [5], [6]. In the early and follow-up experiments, $I_B$ was measured in either the top-emitter configuration [5]–[9] or bottom-emitter configuration [10]–[13], to evaluate $S_0$. In [7] through [9], the BJT $\beta_p$ degradation during emitter-base reverse-bias stress at the junction breakdown voltage was also studied. In many of these earlier measurements, the $I_B - V_{GB}$ curve was also displaced along the gate-voltage axis due to stress, but the peak in $I_B - V_{GB}$ was not very sharp. In some cases no peak was observed. In addition, the magnitude of $I_B$ was greatly increased by the generated $N_{IT}$. Thus, an estimate of $\Delta Q_{OT}$ from the shift of $V_{GB}$ at the peak $\Delta I_B$ in the $\Delta I_B - V_{GB}$ curve cannot be very accurate and reliable.

II. DESCRIPTION AND DEMONSTRATION

Production n-channel MOST fabricated by state-of-the-art CMOS process is measured to demonstrate the proposed DCIV method. The starting n-Si wafer has a p-base well with surface concentration of $1 \times 10^{15}$ cm$^{-3}$, gate oxide thickness of $x_o \approx 150$ Å, channel length $L = 1.6$ μm, and the gate area of $A_G = 1.6 \times 10^2$ μm$^2$. The cross-sectional view was shown in Fig. 1.

Fig. 2(a) and (b) shows the nnp-BJT’s $I_B - V_{GB}$ and $I_C - V_{GB}$ curves, measured in both the top-E and bottom-E configurations, before and after SHE stress as labeled. The oxide charges and interface traps were generated by areally uniform SHE stress with $V_{SB} = V_{DB} = 4$ V, and $V_{GB} = 7.5$ V. During the SHE stress, the bottom emitter junction (n+substrate/n-epitaxy/p-base shown in Fig. 1) was forward-biased to inject electrons into the p-base. Some of these electrons are accelerated (designated as hot electrons), by the reverse-biased surface space-charge layer ($V_{SB} = V_{DB} = 4$ V) of gate-induced collector/junction area, to >3.2 eV kinetic energy. These hot electrons are then injected into the gate oxide over the 3.12 eV SiO$_2$/Si electron potential barrier. Some of the injected electrons are captured by the neutral oxygen vacancy centers [14]–[15], giving $V_o + e^- \rightarrow V_o^-$ and the negative $Q_{OT}$ or positive $\Delta G_{ST}$. Because of their high kinetic energy (~4 eV from $V_{DB} = V_{SB} = 4$ V) which is greater than the bond energy (~3 eV) of the strained Si-Si and Si-O interfacial bonds and the interfacial Si-H and Si-O bonds, the hot electrons also created some new interface traps, $N_{IT}$ or $D_{IT}$, as indicated by the large increase of $I_B$ in Fig. 2(a) measured in both the top-emitter and bottom-emitter configurations. The build-up of $Q_{IT}$ also decreases the subthreshold slope of the nMOST’s $I_D - V_{GB}$ curve shown in Fig. 2(c), however, the $V_{GB}$ shift in $I_D - V_{GB}$ is due to the build-up of both $Q_{OT}$ and $Q_{IT}$.

$$\Delta V_{GB} = \Delta V_{GB-OT} + \Delta V_{GB-IT}$$

$$\equiv - (\Delta Q_{OT} + \Delta Q_{IT})/C_o$$

which cannot be separated by this MOST $I_D - V_{GB}$ measurement alone unless additional properties of the interface traps are known or assumed, a limitation also present in Terman’s method to obtain $D_{IT}$ from HFCV characteristics. The two BJT measurements just described in Fig. 2(a) and (b) can help to separate the $Q_{OT}$ and $Q_{IT}$, which are analyzed as follows.

The stress-generated increase of the $I_B$ shown in Fig. 2(a) gives a direct measure of the surface recombination velocity $S_0$ and the density-of-the-state of the interface traps, $D_{IT}$, because it is proportional to the maximum of the stress induced $I_B$, $\Delta I_B \equiv I_B$ (post-stress) $- I_B$ (pre-stress). Its peak and shape can be distorted by areal nonuniformity of $Q_{OT}, D_{IT}$, or other device parameters, such as base dopant concentration and oxide thickness. But areal nonuniformity alone cannot produce a base current which must come from electron-hole recombination, unlike the HFCV ($C_yb - V_{GB}$) used in Terman’s analysis and the $I_B - V_{GB}$ in the subthreshold slope analysis of the interface trap density, whose distortion could solely arise from areal inhomogeneity even when $D_{IT} = 0$. 
dependent density-of-state, for an energy distribution of interface traps with energy
traps are assumed to be independent of the binding energy
in the energy range

The density-of-state, cross sections, \( \sigma_n(E_{IT}) \) and \( \sigma_p(E_{IT}) \). However, \( S_0 \) calculated from measured \( \Delta I_B \) after stress using (2), can still be used to
monitor the build-up of the interface traps and the associated
\( \Delta V_{GB-ST} \). In the example shown in Fig. 2(a), the numerical
results are \( \Delta S_0 \approx 1600 \text{ cm}^2/\text{s} \) at the \( I_{B-peak} \) which occurs at
the gap energy position of \( V_S - V_F = -0.24 \text{ V} \) below the
midgap for the top-E curve stressed with a fluence of \( 5 \times 10^{18} \text{ electron/cm}^2 \), and \( \Delta S_0 \approx 40 \text{ cm}^2/\text{s} \) at \( V_S - V_F = -0.26 \text{ V} \) for the bottom-E curve stressed at a fluence of \( 1 \times 10^{17} \text{ electron/cm}^2 \). For many devices measured, \( I_{B-peak} \) of the
bottom-E was about five times smaller than that of top-E.

The poststress-priorstress \( I_C - V_{GB} \) curves of both the top-E and
bottom-E configurations shown in Fig. 2(b) give a very
sensitive measure of the stress-generated \( V_{GB} \) shift. \( I_C \) is flat
in the accumulation range and is proportional to the area of
the n+drain/p-base well junction (or the sum of the area of
n+drain and n+source if drain and source are tied together
during the \( I_C \) measurement). When \( V_{GB} \geq V_F \approx -0.55 \text{ V} \)
(Greater sign is for nMOST.), an electron surface channel
begins to form which will collect the electrons injected by
the bottom-emitter and pass the collected electrons to the n+drain
or/n+source, causing an increase of \( I_C \) (or \( I_D + I_S \)). The
\( I_C \) quickly reaches a higher plateau as \( V_{GB} \) increases further
to about \(-0.15 \text{ V} \). This increase of \( I_C \) is proportional to the
added collector area from the gate-induced electron-channel.
The three characteristic Si surface potentials or Si energy band
bendings (FB = flatband at \( V_S = 0 \text{ V} \), INV = inversion
at equal electron-hole surface concentration \( N_S = P_S \) or
\( V_S = V_F - V_{BE}/2 \), and TH = threshold or strong inversion
at \( N_S = P_{base} \) or \( V_S = 2V_F - V_{BE} \) are marked by dots
on the pre-stress \( I_C - V_{GB} \) curve in Fig. 2(b). They show
that \( I_C \) starts to rise sharply at \( V_{FB} \approx -0.55 \text{ V} \) at flatband
in this example, reaching the higher plateau about halfway
between inversion \( V_{GB-SO} \approx -0.25 \text{ V} \) and the MOST
threshold voltage, \( V_{GB-th} \approx +0.05 \text{ V} \). Thus, the rise of \( I_C \)
is sharp and occurs in a short range of \( V_{GB} \), in this case,
\(-0.05 \text{ V} \leq V_{GB} \leq 0.05 \text{ V} \).

It was asserted in the preceding discussion that \( I_C \) is not caused by carrier recombination or generation of the
newly generated interface traps, but solely by the increase of
the emitter or collector area from the gate-induced electron channel described above. This is now experimentally proven
in Fig. 2(b) by the nearly parallel \( I_C \) shift of the post-stress
\( I_C \) from its pre-stress range with the nearly identical height
for both the top-E and bottom-E measurement configurations,
although this stress has generated a large \( N_{IT} \) to give the
large increase of \( I_B \) shown in Fig. 2(a). This model is further
supported by the observed and anticipated reduction of slope of
the post-stress \( I_C - V_{GB} \) at higher \( V_{BE} \) bias reflecting a larger
negative \( Q_{IT} \) (\( \approx \Delta D_{IT} \)). This is expected from the
higher surface electron concentration injected by the emitter
to charge the interface traps negatively due to i) the added stress-induced \( \Delta D_{IT} \), and ii) a larger energy range of \( N_{IT} \) towards
the Si conduction band edge, estimated by \( \Delta E_{IT} \sim V_{BE} \).

A quantitative analysis of the \( Q_{IT} \) contribution to \( I_C - V_{GB} \)
shift in Fig. 2(b) can be made from

\[
\Delta V_{GB-ST} = -\Delta Q_{IT}(V_F - V_{FN})/C_0
\]
bound electron states which are split-off states from the valence band edge, because they are localized or bound in the SiO₂/Si interface. This charge state assignment was implied by Bardeen when he introduced the concept of neutral Fermi level.

In the present example, DIT was generated by the areally uniform CHE stress at 10⁻¹⁶ cm⁻² SHEi stress at VDB = 12 V, VCB = 10 V, and IC = 1 nA. CHEi stress at VGB = VFB = 16 V and floating VSS and I_D = 1 μA for 1 s (curve 1) and 500 s (curve 2).

Additional examples are given in Fig. 3(a)-(c) for the bottom-Emitter configuration which use IC to monitor negative, positive, and turn-around ΔQOT induced by stress. Fig. 3(a) is identical to Fig. 2(b) showing positive ΔVGB from negative ΔQOT. Fig. 3(b) shows negative ΔVGB after SHEi stress (curve 2) with VGB = 12 V and VDB = VSB = 10 V, due to positive ΔQOT, as anticipated [15] by the electron-impact emission of electrons trapped at the neutral oxygen vacancy, V_G = e⁺ + V_G = e⁺ + 2e⁻. Fig. 3(c) demonstrates the successive stresses that gave negative ΔQOT first and then positive ΔQOT, which is the so-called turn-around effect coined by Young [21]. Curve 1, showing positive ΔVGB, was measured after a short (~1 s) CHEi stress at VGB = VFB ≈ 15 V with the source floating, indicating negative ΔQOT due to capture of the electrons injected into the oxide.
along the entire length of the strongly inverted n-channel because $V_{GB} > V_{TB}$. Curve 2, showing negative $\Delta V_{GB}$, was measured after an additional 500 s stress, indicating that some originally trapped electrons (not the captured electrons during the short stress) are emitted via a second pathway, the interface traps in oxidized silicon is presented. It has several degradation mechanisms in BJT even without a separated inhomogeneity and hence usable for profiling dc method resulting in ease of instrumentation and extremely measurement configuration can be used to study the fundamental degradation mechanisms in BJT even without a separated gate over the emitter-base junction [3].

The sensitivity of this new DCIV is demonstrated experimentally in Fig. 4 which gives a sensitivity limit or minimum measurable $S_{0} < \sim 1$ cm/s and $N_{IT} \leq 10^9$ cm$^{-2}$.

III. SUMMARY

A new DCIV method for separating the oxide charge and interface traps in oxidized silicon is presented. It has several unique features difficult to attain previously. 1) It is a purely dc method resulting in ease of instrumentation and extremely high detection sensitivity. 2) It gives true $D_{IT}$, not affected by inhomogeneity and hence usable for profiling $D_{IT}$ and $Q_{IT}$. 3) It has very high $D_{IT}$ sensitivity in presence of large $Q_{OT}$. 4) It is applicable to submicrometer area devices in conventional production CMOS and nMOST junction well structure. 5) It can monitor the degradation kinetics of both MOST’s and BJT’s.

REFERENCES

Chih-Tang Sah (F’69) received two B.S. degrees, in electrical engineering and physics, in 1953 from the University of Illinois, Urbana-Champaign, and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, in 1954 and 1956, respectively.

His industrial career began with William Shockley in 1956, and continued at the Fairchild Semiconductor Corporation from 1959 to 1964, where he directed and managed a 65-member physics department which developed the first-generation silicon integrated circuit manufacturing technology. He was a Professor of Electrical Engineering and a Professor of Physics at the University of Illinois, from 1961 to 1988, where he guided 40 doctoral theses to completion in physics or electrical engineering. He has been the Robert C. Pittman Eminent Scholar and a Graduate Research Professor at the University of Florida, Gainesville, since 1988. He has published about 300 journal articles and given more than 100 invited lectures in China, Europe, Japan, Taiwan, and the United States on transistor physics, technology, and evolution history. At the University of Florida, he taught the undergraduate device core course and wrote the textbook *Fundamental of Solid-State Electronics* in 1991 and its *Study Guide* in 1993. His recent research interests have focused on the fundamental degradation mechanisms in deep submicrometer silicon MOS and bipolar transistors.

Prof. Sah was the recipient of the IRE Browder J. Thompson Best Paper Prize for an author under 30, the J. J. Ebers and Jack Morton Awards from the IEEE Electron Device Society, the Franklin Institute Award for stable MOS Prize for an author under 30, the J. J. Ebers and Jack Morton Awards from the IEEE Electron Device Society and the United States on transistor physics, technology, and evolution history. At the University of Florida, he taught the undergraduate device core course and wrote the textbook *Fundamental of Solid-State Electronics* in 1991 and its *Study Guide* in 1993. His recent research interests have focused on the fundamental degradation mechanisms in deep submicrometer silicon MOS and bipolar transistors.

Prof. Sah was the recipient of the IRE Browder J. Thompson Best Paper Prize for an author under 30, the J. J. Ebers and Jack Morton Awards from the IEEE Electron Device Society, the Franklin Institute Award for stable MOS Prize for an author under 30, the J. J. Ebers and Jack Morton Awards from the IEEE Electron Device Society and the United States on transistor physics, technology, and evolution history. At the University of Florida, he taught the undergraduate device core course and wrote the textbook *Fundamental of Solid-State Electronics* in 1991 and its *Study Guide* in 1993. His recent research interests have focused on the fundamental degradation mechanisms in deep submicrometer silicon MOS and bipolar transistors.

K. Michael Han (S’95) was born in Singapore in 1966. He received the B.S.E.E. degree (high honors) and the M.S.E.E. degree from the University of Florida, Gainesville, in 1989 and 1992, respectively. He is currently pursuing the Ph.D. degree in electrical engineering at the University of Florida. His research interests include reliability studies of advanced MOST/BT devices, semiconductor device physics, advanced semiconductor processings, and analog circuit designs.

Toshihazu Nishida (S’84-M’88) was born in Uruna, IL, in 1963. He received the B.S. degree in engineering physics in 1984, the M.S. degree in 1985, and the Ph.D. degree in August 1988 in electrical engineering, all from the University of Illinois at Urbana-Champaign.

He is currently an Associate Professor of Electrical and Computer Engineering at the University of Florida, Gainesville. He has authored or co-authored more than 20 journal and conference articles, on topics such as the recombination properties of gold impurity centers in silicon, the effect of oxide trap generation on the reliability of a MOS inverter circuit, the threshold bond-breaking voltage and electric field of oxide traps, the electron and hole mobilities in the silicon surface inversion layer, profiling of thermal donors in Crochralski silicon, and hydrogenation of boron acceptor in silicon exposed to hydrogen-containing liquids and gases.

Dr. Nishida is a member of the American Physical Society, the Materials Research Society, and the American Association for the Advancement of Science.

Jack Theodore Kavalieros was born in Thessaloniki, Greece, on January 20, 1967. He received the B.S. and M.S. degrees in electrical engineering from the University of Florida, Gainesville, in 1989 and 1991, respectively. Since 1991, he has been working towards the Ph.D. degree at the University of Florida. His present interests include MOS device physics and in particular the reliability of the gate oxide layer.

He was the recipient of the Robert C. Pittman Fellowship in 1991. During the summers of 1989 and 1994, he interned at National Semiconductor Corporation and at the Intel Corporation, respectively.

Michael S. Carroll was born in Madison, WI, in 1966. He received the B.S. degree in electrical engineering (high honors) from the University of Florida, Gainesville, in 1990 and 1992, respectively. He is currently pursuing the Ph.D. degree in electrical engineering at the University of Florida, Gainesville. His doctoral research involves the modeling of reliability of advanced submicrometer silicon bipolar transistors. His interests are in submicrometer transistor technology development and reliability.

Yi Lu (M’95) was born in China in 1964. He received the B.S. degree from the University of Science and Technology of China in 1986, the M.A. degree in physical chemistry from Washington University, St. Louis, MO, in 1988, and the M.S. and Ph.D. degrees in electrical engineering from the University of Florida, Gainesville, with an Intel Foundation Graduate Fellowship under the direction of Prof. Chih-Tang Sah.

He joined the Intel Corporation in February 1995.