High Concentration of Interface Traps in MOS Transistor Modeling

Zuhui Chen\textsuperscript{a),} Bin B. Jie\textsuperscript{b)} and Chih-Tang Sah\textsuperscript{c)}

\textsuperscript{a)}zh_chen@msn.com, Xiamen University, China; \\
\textsuperscript{b)}bb_jie@msn.com, Peking University, China; \\
\textsuperscript{c)}tom_sah@msn.com, University of Florida, USA.

Abstract

Steady-state recombination-generation-trapping of holes and electrons at one-electron neutral traps (charge states 0 and $-1$) located at the SiO$_2$/Si interface is employed to investigate the effects of high concentration of interface traps on the recombination dc base-terminal current vs gate-voltage (R-DCIV or $I_{B,IV}$) and gate capacitance vs gate-voltage (CV or $C_{gb,IV}$) properties of inversion n-channel Silicon MOS field-effect transistors. The shapes (called lineshapes) of the CV and IV curves are increasingly distorted as the interface trap concentration increases due to the $V_{GB}$ dependence of the trap-charge $Q_T$ from electrons trapped at the 1-electron neutral interface traps. The distortions are illustrated by four families of curves; each family covers a range of a device or trap parameter which is: interface trap concentration, basewell-body dopant impurity concentration, oxide thickness, or interface-trap energy level position in the Si energy gap. At high concentrations of interface traps, the low-frequency and high-frequency CV characteristics, distinguished by the trapping rate, $C_{gb-lf,IV}$ and $C_{gb-hf,IV}$, show extreme distortions, with even two minima or one maximum in the CV valley.

1. Introduction

Interface trap density and its spatial variations along the silicon surface channel at the SiO$_2$/Si interface affect the electrical characteristics and operation endurance or lifetime of Metal-Gate and Silicon-Gate field-effect MOS (Metal-Oxide-Semiconductor) transistors in memory, signal processing and switching applications. Extensive research has been undertaken recently to study the interfacial electronic traps at the SiO$_2$/Si interface in order to delineate their microscopic origins and to monitor their presence during manufacturing and operation of silicon MOS integrated-circuit chips. See [1-9] and references cited therein. It was demonstrated recently by experiments [10] that the recombination DC current-voltage (R-DCIV) characteristics can serve as a simple and accurate monitor for the manufacturing processes, such as diagnosing the operation reliability [10] and extracting the submicron MOS transistor parameters with nanometer spatial resolutions [11]. This method was reactivated in 1995 [10-12] after it was used 33 years earlier in 1962 [13]. Recombination DCIV uses the voltage applied to the gate (relative to the silicon basewell-body) $V_{GB}$ to change the concentrations of the Si electrons and holes at the SiO$_2$/Si interface. This then changes the electron and hole captures rates, $c_{ns}N_S$ and $c_{ps}P_S$, at the interface traps, which produces a gate-voltage-dependent base-terminal current, $I_B(V_{GB})$. The $I_{B,IV}$ and $V_{GB}$ curves have a bell shape (lineshape) with a maximum at a gate voltage that gives equal electron and hole capture rates, $c_{ns}N_S=c_{ps}P_S$. Device and materials parameters of the transistor can be extracted by fitting the experimental lineshapes to the theory computed from the analytical formula derived from the Shockley-Read-Hall electron-hole recombination-generation-trapping theory.

Both the IV and CV characteristics are distorted by the gate-voltage dependence of the density of the charge trapped at the interface traps. The CV characteristic is further distorted by the capacitance from charging and discharging the interface traps. The CV curves cover larger gate-voltage ranges than that of the IV curves, and the influence of the parasitic-capacitances on the CV curves is increasingly masked by that from the interface trapping capacitance as the interface trap concentration increases. Thus, the CV curve has traditionally been used to extract device and material properties [14-16], and interface trap parameters at high trap concentrations, such as in radiation damage of MOS transistors and now in floating-gate and MS/ONO/S memory transistors [1].

In this presentation we will report the theoretically anticipated effects of high concentration of interface traps on the R-DCIV, HFCV and LFCV characteristics. Section 2 gives a summary of the general theoretical analytical equations of recombination current-voltage and capacitance-voltage curves. Section 3 gives the four families of curves to illustrate the effects of interface trap on the R-DCIV and CV lineshapes.

2. Theory of the R-DCIV and CV Methods

Figure 1 shows the cross-sectional view of the MOS transistors used in our previous reports as approximation to state-of-the-art bulk or thick-base (in contrast to thin base) and hence single-gate MOS transistors. Thin base will be reported in the future. The present calculation is for the basewell-channel region (BCR). The results are also applicable to the other four regions labeled in Fig. 1 as the drain and source junction and extension regions (SRJ, DRJ, SER and DER). The geometrical lengths of the five regions are labeled as $Y_{SR}$, $Y_{SI}$, $Y_{BC}$, $Y_{DJ}$ and $Y_{DE}$. We follow the IEEE-Standard notations which we have used in our previous reports [10].
The steady-state electron-hole recombination rate per unit interfacial area, $R_{SS}$, at interface traps of energy-level $E_{TI} = E_T - E_i$, distributed over the boundary of the SiO$_2$/Si interface, $x=0$ plane, with an areal density $N_I(y,z)$ (trap/cm$^2$), is given by the Shockley-Read-Hall (SRH) formula [2,3,8,10,11,13,17,18,22] (in the unit of electron-hole recombination events minus generation events per interface area cm$^2$ per second):

$$N_{IT} = \frac{c_{en} c_{ep} N_s P_s - c_{en} c_{ep} N_i}{c_{en} N_s + c_{en} c_{ep} + c_{ep} P_s + c_{ep}}$$  (1)

The basewell-terminal current $I_B$ is given by integrating (1) over the channel area $dydz$ [10,11,22]

$$I_B(V_{DG}) = \int R_{SS}(V_{DG},y) \, dydz = -q \left( \frac{c_{en} c_{ep}}{2} \right)^{1/2} n W \frac{\exp[U_n(y)] - 1}{\exp[U_n(y)]^2 \cosh[U_i(y)] + \cosh[U_i(y)]}$$  (2)

The effective surface potential and interface-trap energy level are defined by

$$U_s^* = U_s + \log\left( \frac{c_{en}}{c_{ep}} \right)^{1/2} - U_s + U_p$$  (2A)

$$U_i^* = (U_i - E_i)/kT + \log\left( \frac{c_{en}}{c_{ep}} \right)^{1/2}$$  (2B)

Here $c_{en}$, $c_{ep}$, $e_n$ and $e_p$ are the thermal rates of electron and hole capture and emission at the interface traps.

The surface potential, $U_s=qV_s/kT$, is related to the applied gate voltage by integrating the $x$-component of the Poisson equation twice, assuming no $y$-dependence [22], such as in a MOS capacitor or in a MOS transistor with zero applied voltage between the source and drain terminals, but with a finite voltage applied between the tied drain-source and the basewell-body terminals. This is known as the top-emitter bias configuration which we analyzed in our previous papers on the theoretical R-DCIV and G-DCIV characteristics. A future report will extend the results of Jie and Sah [2,3] for $V_{DS}=0$ in the MOS transistor operation range. For $V_{DS}=0$, we have

$$V_{DG} = V_s + V_mB - Q_{IT}/C_{OX} + e_i E_s/C_{OX}$$  (3)

$C_{OX}$ is the oxide capacitance per unit area. $E_s$ is the x-direct electric field normal to the SiO$_2$/Si interface on the semiconductor-Si side of the interface. $Q_{IT}$ is the areal density of total charge residing in the interface traps. $Q_{IT}$ is a function of surface potential and hence gate voltage since it depends on the electron and hole concentrations at Si side of the SiO$_2$/Si interface. For our demonstration model of the interface traps, we take the 1-electron neutral electron trap, which is defined as an electrically neutral trapping or binding potential that can bind only one electron, becoming negative after capturing an electron. Then $Q_{IT}$ is given by [19]

$$Q_{IT} = -qN_{IT} \times F_{IT} = -qN_{IT} \times \frac{c_{en} N_s + c_{en} + c_{ep} P_s + c_{ep}}{c_{en} N_s + c_{en} + c_{ep} P_s + c_{ep}}$$  (4)

The asymptotic capacitances at frequencies much higher and much lower than the trapping frequency ($2\pi$ times the reciprocal trapping time-constant) are of practical importance since they can be uniquely defined and quickly measured. The reciprocal trapping time constants or the capture and emission rate coefficients, $c_{en}$, $c_{ps}$, $e_{ns}$ and $e_{ps}$, at the interface traps, which impact the noise of the MOS transistor, can be extracted by measuring the frequency dependence of the capacitance, such as those reported in the historical first use of the CV to obtain the interface traps, known as the Terman Method [20]. However, the extraction requires a full-scale small-signal admittance analysis using the CTSA transmission-line equivalent-circuit formulation [19]. Thus, in this report we shall only give the theoretical capacitance-voltage characteristics at the asymptotic high-frequency and low-frequency. Their simplicity in underlying trapping physics enhances the illustration and brackets the dynamic or trapping properties of the interface traps. Thus, using the trapping frequency as the demarcation, the high-frequency and low-frequency capacitances are

$$C_{p-S} = C_s \times C_{ox} / (C_s + C_{ox})$$  (5)

$$C_{p-S} = (C_s + C_{IT}) \times C_{ox} / [(C_s + C_{IT}) + C_{ox}]$$  (6)

These are simple capacitance circuits, with $C_{ox}$ in series with the semiconductor space-charge layer capacitance $C_s$ for the high frequency given by (5), and the semiconductor space-charge capacitance $C_s$ in parallel with the interface-trap charging capacitance $C_{IT}$, i.e. $(C_s+C_{IT})$ for the low frequency given by (6).

The charge control analysis, $C=\partial Q/\partial V_s$, is used to calculate the capacitance. It neglects diffusion delay, not important for trapping at the infinitesimal-thin interfacial boundary layer, which is taken into account by the complicated exact theory using the CTSA method to give the admittance function $Y(\omega)=G+j\omega C$ from which the capacitance is obtained from $C(\omega)=\text{Im}Y(\omega)$ [19,23,24]. The charge-control capacitances of the semiconductor electron-hole and trapped charges are given by [15,19]

$$C_s = \frac{2}{E_s} \left[ -N_e \exp(-U_s - U_d + U_d) + N_s \exp(U_s + U_e - U_d + U_d) \right] + \frac{P_d}{(1 + g_s \exp(U_i - U_d - U_d))}$$  (7)

$$C_s = \frac{q}{kT} \left[ (c_{en} N_s + e_{ns}) (c_{en} P_s + e_{ps}) \right]$$  (8)

![Fig. 1 Cross sectional view of an inversion n-channel Si MOS transistor. Length of the five regions between the source and drain p/n junctions (source and drain extension and junction, and base-channel regions) labeled as (YSE, YSJ, YBC, YDJ, YDE).](image-url)
ranges are as follows. The interface-trap energy level is from gate covering the surface channel. The four parameters which are assumed spatially constant over the area of the varying one of the four device-material parameters, all of are computed to illustrate their lineshape distortions by trap and \( T = 300\text{K} \). Four families of IV and CV curves studied recently for \( V_{DS} \)

...the SiO\(_2\)/Si interface, \( x = 0 \), and their spatial variations the oxide thickness and dopant impurity concentrations at for p-Si body, \( U_{PN} \) is added to the and also Fermi distribution at high electron or/and hole electrons on [15] to include impurity deionization and degeneracy, and Fermi distribution at high electron or/and hole concentrations. For p-Si body, \( U_{PN} \) is added to the electron or p-Body minority-carrier term in \( E_S \) to give

\[
E_S = \frac{2T} {E_S} \left[ N_e \left( \exp \left( -\frac{U_S}{U_e} - U_{J} + U_{J} \right) - \exp \left( -U_{J} + U_{J} \right) \right) \right. \\
\left. + N_d \left( \exp \left( U_J + U_C - U_{J} + U_{V_H} \right) - \exp \left( U_J - U_{J} + U_{V_H} \right) \right) \right] \\
+ P_{AA} \left( U_J + \log \left[ \frac{1 + q_{eff} \exp \left( U_J - U_{J} - U_{V_H} \right) }{1 + q_{eff} \exp \left( U_J - U_{J} \right) } \right] \right) 
\]

(9)

3. R-DCIV and CV Lineshape Analysis

Wang and Sah [11] showed in 2001 that lineshape and width of R-DCIV \( I_B-V_{GB} \) curves are determined by the oxide thickness and dopant impurity concentration at the SiO\(_2\)/Si interface, \( x = 0 \), and their spatial variations from the source to the drain \( 0 \leq y \leq L \), denoted by \( X_{ox}(y,z) \) and \( P_{IM}(x=0,y,z) \) cm\(^{-3}\). It was shown rigorously that the spatial variation of the interface trap density, \( N_{IT}(y) \) cm\(^{-2}\), can affect the R-DCIV lineshape only when \( P_{IM}(x=0,y,z) = P_{IM}(y) \neq \) constant. However, it was not stated by them that this very general result is valid only when interface-trap charge density along channel region is zero, i.e. \( Q_{IT} = 0 \) or \( |Q_{IT}| < |Q_{SL}| \), and there is no voltage applied between the drain and source terminal, i.e. \( V_{DS} = 0 \) of the Top-Emitter bias configuration, which was extensively studied recently for \( V_{DS} \neq 0 \) by Jie and Sah [2, 3]. Our following analysis for large \( Q_{IT} \) or \( N_{IT} \) will use the Top-Emitter bias configuration, \( V_{DS} = 0 \). We will present results only for forward bias of \( V_{PIN} = V_{BD} = V_{BS} = 0 \text{mV} \) for CV curves and +300mV for IV curves. Reverse bias \( V_{PIN} < 0 \) and \( V_{DS} \neq 0 \) for MOS transistor operation will be reported in the future. We also assume a single species of neutral (charge states 0 and \(-1\)) one-electron interface trap and \( T = 300\text{K} \). Four families of IV and CV curves are computed to illustrate their lineshape distortions by varying one of the four device-material parameters, all of which are assumed spatially constant over the area of the gate covering the surface channel. The four parameters’ ranges are as follows. Interface-trap energy level is from \( E_T = E_I = 0 \text{eV} \) to \( \pm 0.5 \text{eV} \) to cover the \(-1.15\text{eV} Si energy gap with \( c_m = c_p = 10^{-8}\text{cm}^3/\text{s} \). The gate oxide thickness is from \( X_{OX}(y,z) = 1.2\text{nm} \) to 20nm. The basewell-body acceptor dopant impurity concentration is from \( P_{IM}(x=0,y,z) = 10^{10} \) to \( 10^{16} \text{cm}^{-3} \) to cover both the basewell and the drain and source extension regions. The gate-width channel length aspect ratio is 10 from \( W/L = 10\mu\text{m}/1\mu\text{m} \) which can be changed to reduce measurement noise since \( I_D \propto (W/L) \) while \( C_{gb} \propto W \times L \).

![Fig. 2](image-url)

Fig. 2 Effects of midgap (\( E_I = 0 \text{eV} \)) interface trap density on R-DCIV and CV lineshapes with \( N_{IT} = 1.0 \times 10^9, 1.0 \times 10^{10}, 1.0 \times 10^{11}, 1.0 \times 10^{12}, 5.0 \times 10^{12}, 1.0 \times 10^{13}, 3.0 \times 10^{13} \) and \( 5.0 \times 10^{13} \text{cm}^{-2} \): (a) Normalized \( I_B \) vs \( V_{GB} \); (b) Percentage deviation of \( I_B \); (c) high-frequency CV curves without \( C_{gb} \) and (d) low-frequency CV curve with \( \delta C_{gb} \).
The four families of curves for each parameter variation are given in Figures 2(a)-2(d) to 5(a)-5(b). The four Figure (a)’s give $I_B/I_{Bpk}$ vs $V_{GB}$. The four Figure (b)’s give the percentage deviation of $I_B$ from the reference indicated in each figure (b). The four Figure (c)’s and (d)’s give the CV curves at the high-frequency and low-frequency limits. The importance of the trapped charge density, $Q_{IT}$, at high concentrations of interface traps, $N_{IT} > \sim 10^{12}$cm$^{-2}$, is illustrated which was not investigated in all of our previous reports on low $Q_{IT}$.

Figures 2(a) to 2(d) shows the dependence of the IV and CV lineshapes on the interface trap density covering the range of $N_{IT} = 1 \times 10^9$ to $5 \times 10^{13}$ trap/cm$^3$.

Figures 3(a) to 3(d) shows the dependence of the IV and CV lineshapes on the spatially constant basewell-body impurity concentration, $P_{IM}=1 \times 10^{16}$ to $1 \times 10^{19}$ cm$^{-3}$, with $N_{IT}=1.0 \times 10^{12}$cm$^{-2}$. Such a high concentration is generated during repeated program-erase cycling of non-volatile floating-gate and SNONS memory transistors, recently reported by Victor Kuo of Taiwan Power Semiconductor Corporation. See reference [12] cited by us in [1]. Figure 3(b) shows huge deviations, $>5000\%$ for $P_{IM}=10^{19}$ cm$^{-3}$ when comparing the $I_B$-$V_{GB}$ curves using the range of $I_B/I_{Bpk}=0.1$ to 1.0. Figures 3(c) and 3(d) show that the HFCV and LFCV lineshapes broaden, and the linewidth increases with increasing basewell-body dopant impurity concentration at the SiO$_2$/Si interface.

Figures 4(a) to 4(b) shows the IV and CV distortions for oxide thickness from $X_{ox}=1.2$nm to 20nm at $N_{IT} = 1 \times 10^{12}$cm$^{-3}$.

Figures 5(a) to 5(b) shows the IV and CV distortions from interface traps at different energy levels, covering the slowest and deepest traps at the Si-midgap to the fastest and shallowest traps near the Si-band-edges, $E_{IT}\equiv E_T-E_I=0$ to $\pm 0.5$eV, again at the very high interface trap concentration $N_{IT}=10^{12}$cm$^{-2}$. Figure 5(a) shows that if forward bias is less than interface trap energy, i.e. $V_{PN}<|E_T/q|$, the R-DCIV or $I_B$-$V_{GB}$ lineshape is almost symptomatically broadened by the shallower-energy-level traps. Note also that the broad top of the R-DCIV is a distinct signature of a shallow energy-level interface trap. For sufficiently high measurement frequency, the $C_{gb-lf}$-$V_{GB}$ curves are not changed much from charging and discharging $C_a$ because $C_a$ cannot keep up or respond to the small signal at such high signal frequency. Therefore, the stretch-out of the $C_{gb-lf}$ towards the inversion range (positive $V_{GB}$) is entirely from $Q_{IT}(V_{GB})$ which could be completely missed if $N_{IT}$ is small. While at sufficiently low signal frequency, $C_a$ would give a peak in the $C_{gb-lf}$-$V_{GB}$ valley, which is a distinct signature of the presence of an interface traps.
Fig. 4 Effect of oxide thickness on R-DCIV and CV lineshapes with ETI = 0.0eV and X\textsubscript{OX} = 12, 20, 35, 50, 100, 150 and 200A. (a) Normalized I\textsubscript{B} vs V\textsubscript{GB}; (b) Percentage deviation of I\textsubscript{B}; (c) High-frequency CV curves, no C\textsubscript{it}; and (d) Low-frequency CV curves with C\textsubscript{it}. P\textsubscript{IM} = 1E17cm\textsuperscript{-3}, N\textsubscript{IT} = 1E12cm\textsuperscript{-2}, and ETI = 0.0eV.

Fig. 5 Effect of discrete energy-level position of interface traps in silicon energy gap on R-DCIV and CV lineshapes with ETI = -0.5 to +0.5eV in 0.1eV steps: (a) Normalized I\textsubscript{B} vs V\textsubscript{GB}; (b) Percentage deviation of I\textsubscript{B}; (c) High-frequency CV curves with no C\textsubscript{it}; and (d) Low-frequency CV curves with C\textsubscript{it}. P\textsubscript{IM} = 1E17cm\textsuperscript{-3}, X\textsubscript{OX} = 35A, N\textsubscript{IT} = 1E12cm\textsuperscript{-2}.
4. Summary
This paper presents a study of the effects of high concentration of interface traps in MOS transistors on the lineshape of the R-DCIV (Recombination DC-Current-Voltage) curves and the asymptotic $C_{gb} - V_{GB}$ (gate-capacitance-Voltage) curves at high- and low-frequency limits. The electron-hole trapping frequency or rate at the SiO$_2$/Si interface traps is the demarcation frequency for the high-frequency and low-frequency capacitances. The Shockley-Read-Hall theory of recombination, generation and trapping of electrons and holes is applied to a model interface trap, the neutral (charge-states of 0 and −1) 1-electron trap. The trapped charges at the interface traps dominate the distortion of the R-DCIV and the CV curves. This was neglected in all of our previous studies which were all aimed at the low concentrations of interface traps in MOS transistors used in logic circuits during initial applications before substantial interface traps are generated. High interface (and oxide or insulator) trap concentrations generated during program-erase cycles is the main mechanism that limits the endurance of non-volatile MOS memory transistors such as the floating gate and recent SONOS.

HFCV curves excluding the trapping capacitance $C_{it}$ and LFCV curves including the $C_{it}$ show that these two asymptotic capacitances offer a simple visual means to track the generation of the interface traps and their properties, both static, such as trapping energy level and concentration, and dynamic, such as the emission and capture rates of electrons and holes. The experimental R-DCIV lineshape broadening and distortion, previously attributed by us solely to spatial variation of the dopant impurity concentration between the source and drain, can be account for, at least in part, by the gate-voltage dependent trapped-charge density at the interface traps when a high concentration of interface trap is generated during operation stress, such as those encountered in MOS memory transistors.

5. Acknowledgments
The authors thank Dr. Colin McAndrew of the Freescale Semiconductor Corporation, formerly Motorola, for his comments and suggestions. We also like to thank Professors Gennady Gildenblatt of Arizona State, Mitiko Miura-Mattausch of Hiroshima, and Xing Zhou of Nanyang Technology Universities, and also Dr. Colin McAndrew for encouraging us to investigate aspects of compact transistor modeling. This research was supported by CTSAh Associates (CTSA) which was founded by the late Linda Su-Nan Chang Sah.

6. References