A Distributed Bulk-Oxide Trap Model for Al_2O_3 InGaAs MOS Devices

Yu Yuan, Bo Yu, Jaesoo Ahn, Paul C. McIntyre, Peter M. Asbeck, *Fellow, IEEE*, Mark J. W. Rodwell, *Fellow, IEEE*, and Yuan Taur, *Fellow, IEEE*

***Abstract*—This paper presents a distributed circuit model for bulk-oxide traps based on tunneling between the semiconductor surface and trap states in the gate dielectric film. The model is analytically solved at dc. It is shown that the distributed bulk-oxide trap model correctly depicts the frequency dispersion in the capacitance– and conductance–voltage data of Al_2O_3–InGaAs MOS devices that do not fit the conventional interface state model. The slope degradation or stretch-out of the measured capacitance–voltage curve near flatband can be also explained by the distributed bulk-oxide trap model.**

***Index Terms*—Bulk-oxide trap, III–V, MOS, tunneling.**

I. Introduction

RECENTLY, III–V compound semiconductor MOSFETs have been intensely investigated to replace silicon CMOS for high-performance digital applications. In many reports in the literature [1]–[6], frequency dispersion is commonly observed in the capacitance– and conductance–voltage data of high-κ/III–V MOS devices. Examples are shown in Fig. 1 for atomic layer deposition (ALD) Al_2O_3 on an n-type $In_{0.53}Ga_{0.47}As$ MOS capacitor. Dispersion in the strong accumulation region cannot be explained by the conventional interface states whose time constant in such bias regions is far shorter than the period of typical measurement frequencies, i.e., 1 kHz–1 MHz [7], [8]. On the other hand, trap states inside the gate insulator, called bulk-oxide traps or border traps, do have long time constants as they interact with the conduction band electrons via tunneling [9]. In addition, when the conventional conductance method [7] for interface states is applied to the high-to-low transition region (maximum slope) of the $C-V$ data, the frequency dispersion of conductance does not follow the well-known peak behavior. In addition, the $C-V$ stretch-out with respect to the ideal curve indicates an interface state density far exceeding that extracted from the dispersion in that region. Such a discrepancy can be resolved by a bulk-oxide trap model in which the low-frequency component causing $C-V$ stretch-out is larger than the high-frequency component responsible for dispersion.

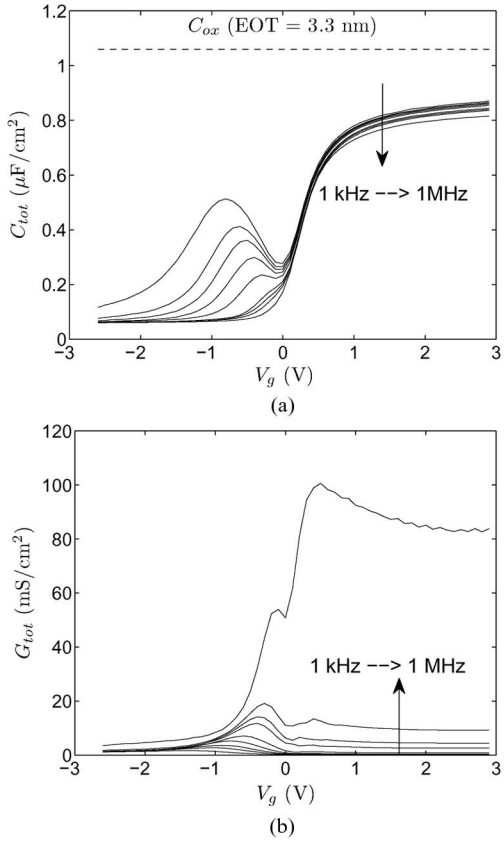


Fig. 1. Experimental (a) $C-V$ and (b) $G-V$ data of Al_2O_3/n-InGaAs MOS at 1 kHz, 2 kHz, 3 kHz, 5 kHz, 10 kHz, 30 kHz, 50 kHz, 100 kHz, and 1 MHz.

In our previous publication [10], a distributed bulk-oxide trap model was proposed to explain the dispersion in strong accumulation of Al_2O_3/n-$In_{0.53}Ga_{0.47}As$ MOS data. In this paper, the model is completed by adding integration of the bulk-oxide trap density with respect to energy for calculating the equivalent admittance. A slightly different differential equation is derived and numerically solved to yield frequency-dependent capacitance and conductance of the MOS device. The model is validated and calibrated by the Al_2O_3/n-$In_{0.53}Ga_{0.47}As$ MOS data in both strong accumulation and flatband regions. The model is also applied to account for the $C-V$ stretch-out in the device.

II. $C-V$ and $G-V$ Data of Al_2O_3 on n-type InGaAs MOS

Fig. 1 shows the multiple-frequency $C-V$ and $G-V$ data measured from a Pt/Al_2O_3/n-$In_{0.53}Ga_{0.47}As$ MOS capacitor.

Manuscript received March 5, 2012; revised April 2, 2012; accepted April 20, 2012. Date of publication May 17, 2012; date of current version July 19, 2012. This work was supported in part by Semiconductor Research Corporation through the Nonclassical CMOS Research Center and in part by the Research Collaboration Project with Taiwan Semiconductor Manufacturing Company, Ltd. The review of this paper was arranged by Editor J. D. Cressler.

Y. Yuan, B. Yu, P. M. Asbeck, and Y. Taur are with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093 USA.

J. Ahn and P. C. McIntyre are with the Department of Materials Science and Engineering, Stanford University, Stanford, CA 94305 USA.

M. J. W. Rodwell is with the Department of Electrical and Computer Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106 USA.

Digital Object Identifier 10.1109/TED.2012.2197000

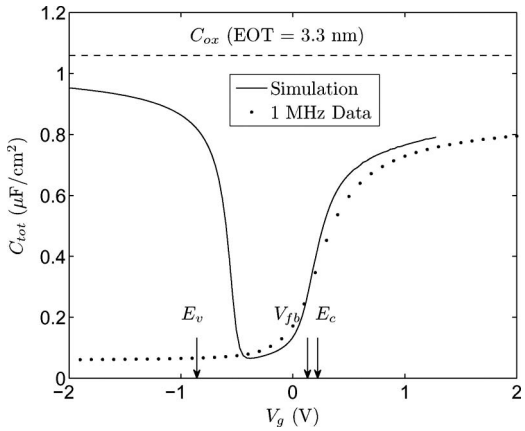


Fig. 2. $C-V$ of Pt/Al_2O_3/n-$In_{0.53}Ga_{0.47}As$ MOS measured at 1 MHz compared with quasi-static $C-V$ of ideal MOS without traps.

Since InGaAs is intended for making high-mobility nMOSFETs, bulk-oxide electron traps near the conduction band energy are investigated by biasing the n-type layer into accumulation. The capacitor is fabricated under similar processing procedures as in [13]. The semiconductor layer structure is 1-μm 2×10^{16} cm^{-3} doped n-$In_{0.53}Ga_{0.47}As$ on 100-nm 5×10^{18} cm^{-3} doped n-$In_{0.53}Ga_{0.47}As$ on n$^+$ InP substrate. The Al_2O_3 film is prepared by 50 cycles of ALD with trimethylaluminum precursor and water vapor oxidant. The sample is then annealed in forming gas for 30 min at 400 °C.

The Al_2O_3 data exhibit frequency dispersion characteristics in $C-V$ and $G-V$. To determine $C_{\rm ox}$, quasi-static $C-V$ for an ideal Al_2O_3/n-$In_{0.53}Ga_{0.47}As$ MOS is simulated and compared with the 1-MHz $C-V$ data, which are least affected by the traps in Fig. 2. Both the quantum confinement and nonparabolic band effects are incorporated in the simulation. By matching the simulated capacitance in accumulation to the 1-MHz data at somewhat higher gate voltages to account for the stretch-out effect due to traps, $C_{\rm ox}$ is extracted to be 1.06 μF/cm^2 (see the long dashed line near the top of the figure). Because of the low density of states of the $In_{0.53}Ga_{0.47}As$ conduction band, the accumulation capacitance is considerably lower than the oxide capacitance $C_{\rm ox}$. In addition, labeled in Fig. 2 are V_g values for E_c, E_v, and $V_{\rm fb}$, where the Fermi level crosses those energy values at the surface. Most of the steep transition region in the 1-MHz $C-V$ curve is actually near flatband or in moderate accumulation, not depletion. The degradation of the $C-V$ slope in the 1-MHz data in this region is due to stretch-out or slow charge/discharge of trap states during the gate bias sweep. The dispersion in strong accumulation cannot be explained by parasitic resistance in series with the MOS capacitor because the dispersion is observed at low frequencies of a few kilohertz, where the series resistance should have no effect. It cannot be explained by the conventional interface states either since their time constant in accumulation cannot be so long that the capacitance dispersion persists at a few kilohertz. Near the flatband or in moderate accumulation, the parallel conductance in the semiconductor G_p, as defined in Fig. 3, is calculated from the measured $C_{\rm tot}(\omega)$ and $G_{\rm tot}(\omega)$ for a given bias point. Instead of exhibiting a peak behavior as predicted in the standard interface state model, the G_p/ω versus $\log(\omega)$ plot in Fig. 3 is rather featureless, indicative of wide distribution of trap time constants more in line with the bulk-oxide trap model.

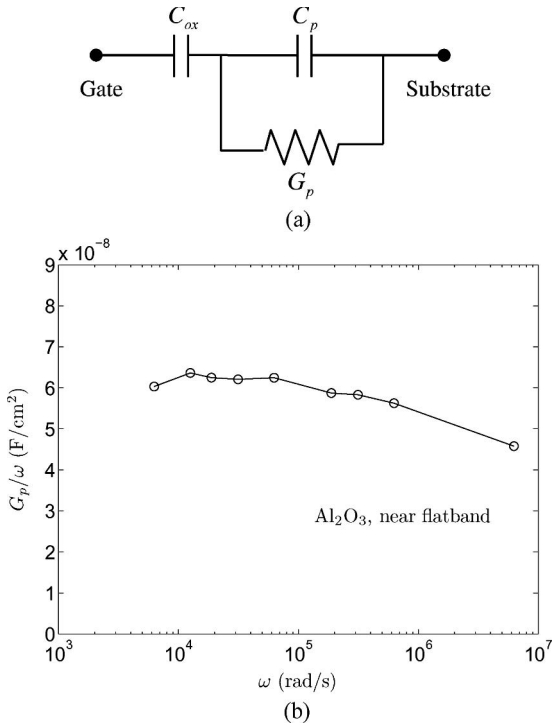


Fig. 3. (a) Equivalent circuit for C_p and G_p. (b) Experimental G_p/ω versus ω (log scale) at $V_g = 0.3$ V.

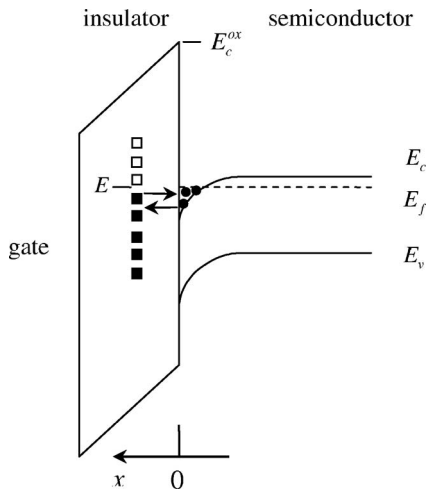


Fig. 4. Schematic of tunneling between bulk-oxide traps in the gate insulator and conduction band of the semiconductor.

The $C-V$ humps in the negative V_g bias region corresponding to depletion and weak inversion do exhibit a peak G_p/ω behavior and are attributed to a localized density of interface states. Analysis of these features is not covered in this paper.

III. Distributed Bulk-Oxide Trap Model

In MOS devices, traps in the bulk gate dielectric film can exchange charge with mobile carriers in the semiconductor bands through tunneling. Fig. 4 schematically shows the tunneling process between bulk-oxide traps and conduction band in an

n-type MOS device biased in accumulation. The time constant associated with charge exchange between bulk-oxide traps and semiconductor is governed by the tunneling mechanism that gives exponential dependence on the trap distance x from the interface [9], [11]

$$\tau(x) = f_0 \tau_0 e^{2\kappa x}. \tag{1}$$

Here, $\tau_0 = (n_s \sigma v_{\rm th})^{-1}$ is the time constant of the interface traps inversely proportional to the carrier density at the semiconductor surface n_s, σ is the trap cross-sectional area, and $v_{\rm th}$ is the electron thermal velocity. Other parameters in (1) are as follows: f_0 is the Fermi–Dirac function that a trap at energy E is occupied by an electron, and κ is the attenuation coefficient for an electron wave function of energy E decaying under an energy barrier $E_C^{\rm ox} > E$

$$\kappa = \sqrt{2m^* (E_C^{\rm ox} - E)}/\hbar. \tag{2}$$

m^* is the electron effective mass in the dielectric film and $E_C^{\rm ox}$ is the energy of the top of the dielectric barrier, as indicated in Fig. 4.

For a given gate dc bias, bulk-oxide traps at a certain depth x and energy E change occupancy in response to a small-signal ac modulation. Bulk-oxide traps at energy close to $E = E_f$ are most responsible for the small-signal capacitance. The effects of bulk-oxide traps at a specific depth and energy on the small-signal MOS admittance can be modeled by a serial combination of capacitance and conductance. The bulk-oxide traps within an incremental depth Δx at x and an incremental energy value ΔE at E are represented by incremental capacitance $\Delta C_{\rm bt}(E, x)$ and incremental conductance $\Delta G_{\rm bt}(E, x)$ connected in series. If the density per volume per energy of bulk-oxide traps is $N_{\rm bt}$ in units of cm^{-3}Joule^{-1}, then [7], [9]

$$\Delta C_{\rm bt}(E, x) = \frac{f_0(1 - f_0) q^2 N_{\rm bt}}{kT} \Delta E \Delta x. \tag{3}$$

$\Delta G_{\rm bt}(E, x)$ and $\Delta C_{\rm bt}(E, x)$ are related by time constant $\tau(x)$

$$\Delta C_{\rm bt}(E, x)/\Delta G_{\rm bt}(E, x) = \tau(x) = f_0 \tau_0 e^{2\kappa x}. \tag{4}$$

To integrate for a continuous energy distribution of bulk-oxide traps, the serial connection of $\Delta C_{\rm bt}(E, x)$ and $\Delta G_{\rm bt}(E, x)$ at a given x must be first converted to a parallel combination of incremental admittance. Because the factor $f_0(1 - f_0)$ is sharply peaked at $E = E_f$, κ in (2) is set to be a constant with $E = E_f$ in the integration. The total incremental admittance at x is then

$$\begin{aligned} \Delta Y_{\rm bt}(x) &= \int_E \frac{1}{\frac{1}{j\omega \Delta C_{\rm bt}(E,x)} + \frac{1}{\Delta G_{\rm bt}(E,x)}} \\ &= \frac{q^2 N_{\rm bt} \ln(1 + j\omega \tau_0 e^{2\kappa x})}{\tau_0 e^{2\kappa x}} \Delta x. \end{aligned} \tag{5}$$

For a continuous distribution of bulk traps throughout the oxide thickness, the equivalent circuit of the MOS device is of a distributed form shown in Fig. 5, where the oxide capacitance is broken into an infinite number of serial segments with branches of $\Delta Y_{\rm bt}(x)$ connected at different depths. Here, $\epsilon_{\rm ox}$ is the permittivity of the insulator and C_s is the semiconductor capacitance.

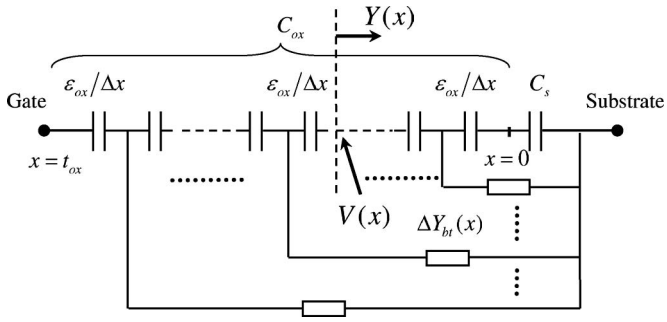


Fig. 5. Equivalent circuit for bulk-oxide traps distributed over the depth of the insulator. The semiconductor capacitance is represented by C_s.

If we define $Y(x)$ to be the equivalent admittance at a point x looking into the semiconductor in Fig. 5, the recursive nature of the distributed circuit gives the admittance of the next point $x + \Delta x$ as

$$Y(x + \Delta x) = \Delta Y_{\rm bt}(x) + \frac{1}{\frac{\Delta x}{j\omega \epsilon_{\rm ox}} + \frac{1}{Y(x)}}. \tag{6}$$

Substituting (5) for $\Delta Y_{\rm bt}(x)$, the first-order terms in Δx then yield a differential equation for $Y(x)$

$$\frac{dY}{dx} = -\frac{Y^2}{j\omega \epsilon_{\rm ox}} + \frac{q^2 N_{\rm bt} \ln(1 + j\omega \tau_0 e^{2\kappa x})}{\tau_0 e^{2\kappa x}}. \tag{7}$$

The boundary condition is $Y(x = 0) = j\omega C_s$. This differential equation is the correct one to use over the one we derived earlier without energy integration [10], although either equation can fit to the same sets of data by readjusting parameters $N_{\rm bt}$, κ, and τ_0. In general, a slightly lower $N_{\rm bt}$ is needed in the new equation to produce the same amount of dispersion.

In general, (7) needs to be numerically solved to obtain the total admittance seen by the gate

$$Y(x = t_{\rm ox}) \equiv G_{\rm tot} + j\omega C_{\rm tot}. \tag{8}$$

A typical example of the solutions $C_{\rm tot}$ versus $\ln \omega$ and $G_{\rm tot}$ versus ω is given in Fig. 6. In the high frequency limit, $\omega \tau_0 \geq 1$, none of the bulk-oxide traps respond to the ac signal and $C_{\rm tot}$ is equal to $C_{\rm ox}$ in series with C_s as expected. For the measurement frequencies of 1 kHz–1 MHz, $1.4 \times 10^{-6} < \omega \tau_0 < 1.4 \times 10^{-3}$, $C_{\rm tot}$ linearly varies with $\ln(1/\omega)$, and $G_{\rm tot}$ linearly varies with ω, i.e., $G_{\rm tot}/\omega \approx$ constant. Both are consistent with the data trends in Fig. 1. Constant $G_{\rm tot}/\omega$ reflects the fact that, for a given gate bias, response of bulk-oxide traps spans a wide spectrum of frequencies due to their depth distribution, i.e., a clear distinction from conventional interface traps [7]. For a given frequency of $\omega < 1/\tau_0$, the depth of traps that respond to the small signal can be estimated by letting the factor $\omega \tau_0 e^{2\kappa x}$ in (7) equal unity, i.e., $x \sim (2\kappa)^{-1} \ln(1/\omega \tau_0)$. This is typically in the range of 0.1–1 nm.

For $\omega = 0$ or dc, Fig. 5 becomes a purely capacitive circuit and (7) is reduced to a real equation for capacitance $C(x)$

$$\frac{dC}{dx} = -\frac{C^2}{\epsilon_{\rm ox}} + q^2 N_{\rm bt}. \tag{9}$$

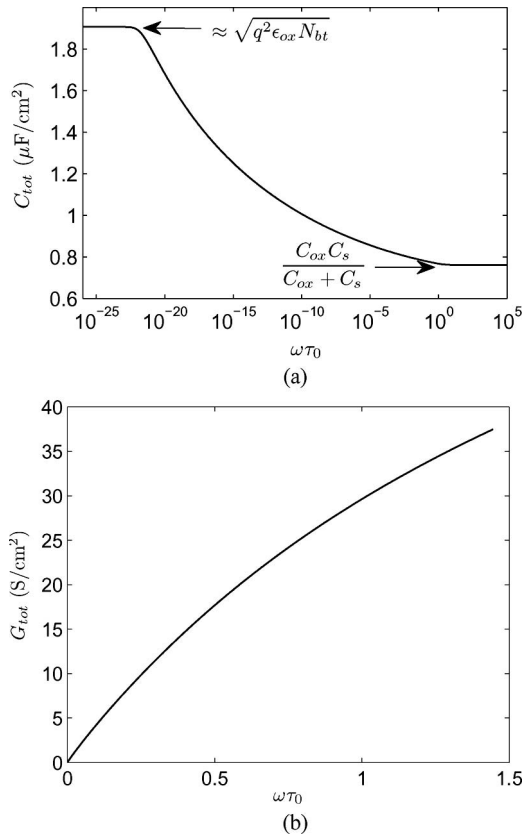


Fig. 6. Example of numerical solution to (7): (a) real and (b) imaginary parts of $Y(x = t_{\rm ox})$ versus $\omega\tau_0$ with $N_{\rm bt} = 4.2 \times 10^{19}$ cm^{-3} eV^{-1} and $\tau_0 = 2.3 \times 10^{-10}$ s.

The boundary condition is $C(x = 0) = C_s$. For uniform $N_{\rm bt}$, (9) can be analytically solved to yield

$$C(x) = C_0 \frac{(C_s + C_0)\exp\left(2qx\sqrt{N_{\rm bt}/\epsilon_{\rm ox}}\right) + (C_s - C_0)}{(C_s + C_0)\exp\left(2qx\sqrt{N_{\rm bt}/\epsilon_{\rm ox}}\right) - (C_s - C_0)}. \quad (10)$$

Here, $C_0 = q\sqrt{\epsilon_{\rm ox}N_{\rm bt}}$. If $2qt_{\rm ox}\sqrt{N_{\rm bt}/\epsilon_{\rm ox}} \gg 1$, then $C(x = t_{\rm ox}) \approx \sqrt{q^2\epsilon_{\rm ox}N_{\rm bt}}$ [see the left plateau in Fig. 6(a)], insensitive to C_s. This is, of course, only a matter of theoretical interest, as in practice, it would take much longer than the age of the universe to charge up all the bulk traps in the oxide!

Of particular interest is the case in accumulation where C_s is very high. From (10), $C_{\rm tot}({\rm DC}) \approx C_0 \coth(C_0/C_{\rm ox})$, always larger than $C_{\rm ox}$. This is in contrast with the interface state or lumped-circuit border trap models, which do not produce dispersion when shorted out by large semiconductor capacitance. Dispersion in accumulation is therefore a good indicator of distributed bulk-oxide traps.

IV. Correlation of the Model With Multifrequency $C-V$ and $G-V$ Data in Strong Accumulation and Near Flatband

The experimental capacitance and conductance versus frequency data in strong accumulation in Fig. 1 (Al_2O_3/n-InGaAs at $V_g = 2.9$ V) are compared with model calculations in Fig. 7. For model parameters, semiconductor capacitance C_s

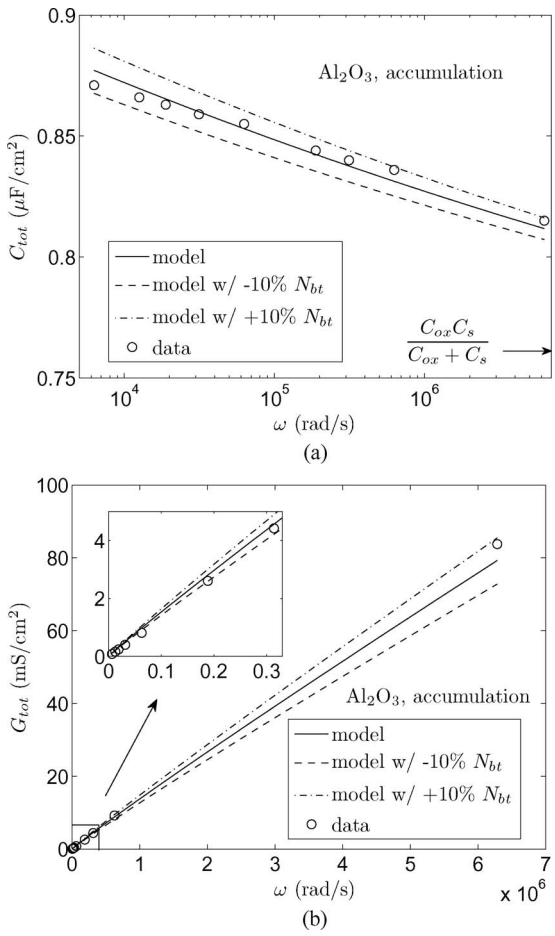


Fig. 7. Al_2O_3 MOS experimental (a) $C_{\rm tot}(\omega)$ and (b) $G_{\rm tot}(\omega)$ dispersion data (open circles) at $V_g = 2.9$ V in Fig. 1 compared with those calculated from the distributed bulk-oxide trap model (solid lines). A single bulk-oxide trap density $N_{\rm bt} = 4.2 \times 10^{19}$ cm^{-3} eV^{-1} is assumed in both $C_{\rm tot}(\omega)$ and $G_{\rm tot}(\omega)$ calculations. The rest of the model parameters are $C_{\rm ox} = 1.06$ μF/cm^2, $t_{\rm ox} = 5$ nm, $C_s = 2.7$ μF/cm^2, $\kappa = 5.1$ nm^{-1}, and $\tau_0 = 2.3 \times 10^{-10}$ s.

is chosen such that the serial combination of $C_{\rm ox}$ and C_s gives $C_{\rm tot}$ slightly below the measured 1-MHz capacitance at $V_g = 2.9$ V. κ is calculated from (2) with $m^* = 0.5m_0$ and $E_C^{\rm ox} - E = 1.99$ eV. Both the slopes of $C_{\rm tot}$ versus $\ln(1/\omega)$ and $G_{\rm tot}$ versus ω are sensitive to bulk-oxide trap density $N_{\rm bt}$. By choosing a single fitting parameter, uniform $N_{\rm bt} = 4.2 \times 10^{19}$ cm^{-3} eV^{-1}, good agreement is achieved between the model and the measured $C_{\rm tot}$ and $G_{\rm tot}$ data from 1 kHz to 1 MHz in Fig. 7(a) and (b). Parameter τ_0 is chosen so that $C_{\rm tot}$ is consistent with $C_{\rm ox}C_s/(C_{\rm ox} + C_s)$ at a frequency $\omega\tau_0 \to 1$, where the traps have no effect [see lower right of Figs. 6(a) and 7(a)]. For an assessment of sensitivity to the fitting parameters, Fig. 7 also shows two curves (dashed) calculated with 10% variation of $N_{\rm bt}$ for best fitting.

Fig. 8 shows the model correlation with data near the flatband voltage, i.e., in the region of steep $C-V$ transition in Fig. 1. A lower bulk-oxide trap density of $N_{\rm bt} = 2.2 \times 10^{19}$ cm^{-3} eV^{-1} is found to fit both the capacitance and conductance versus frequency data. Note that the zero-depth trap time constant τ_0 here is much longer than that in strong accumulation (see Fig. 7) such that nonlinearity starts to show up in $G_{\rm tot}$ at 1 MHz. This is consistent with the lower surface electron density near the flatband.

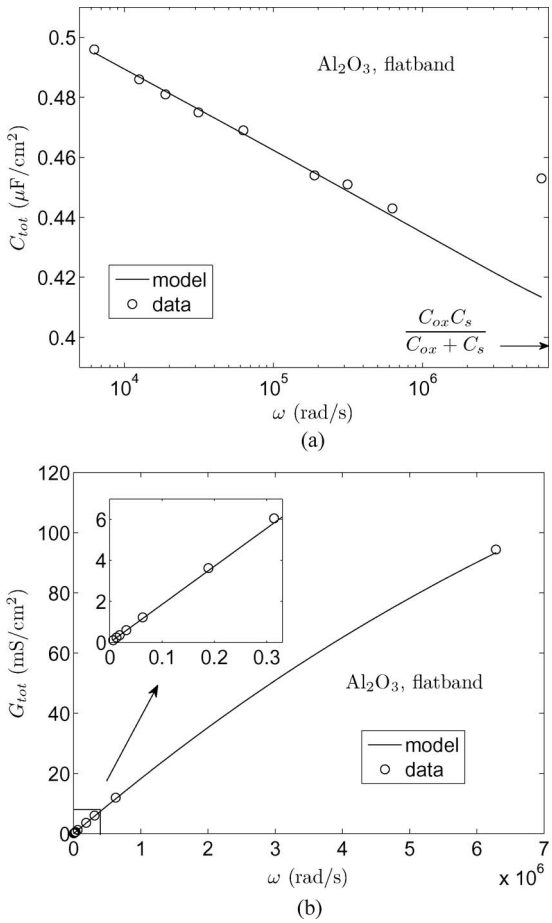


Fig. 8. Al_2O_3 MOS experimental (a) $C_{\rm tot}(\omega)$ and (b) $G_{\rm tot}(\omega)$ dispersion data (open circles) at $V_g = 0.3$ V in Fig. 1 compared with those calculated from the distributed bulk-oxide trap model (solid lines). A single bulk-oxide trap density $N_{\rm bt} = 2.2 \times 10^{19}$ cm^{-3} eV^{-1} is assumed in both $C_{\rm tot}(\omega)$ and $G_{\rm tot}(\omega)$ calculations. The rest of the model parameters are $C_{\rm ox} = 1.06$ μF/cm^2, $t_{\rm ox} = 5$ nm, $C_s = 0.635$ μF/cm^2, $\kappa = 5.47$ nm^{-1}, and $\tau_0 = 1.35 \times 10^{-7}$ s.

V. Correlation of the Model With C–V Stretch-Out Near Flatband

It is shown in Fig. 2 that, in addition to giving rise to C and G dispersions, oxide traps also cause a degradation of the $dC_{\rm tot}/dV_g$ slope near the flatband and moderate accumulation compared to the ideal simulated curve. Such a "stretch-out" of the 1-MHz C–V is commonly employed to extract the interface state density (Terman method) based on the relation $\Delta V_g/\Delta\Psi_s = 1 + (C_s + C_{\rm it})/C_{\rm ox}$, where Ψ_s is the surface potential and $C_{\rm it}$ is the capacitance due to interface states. Applying this method to the observed C–V slope degradation in Fig. 2 would yield an interface state density of 5.4×10^{12} cm^{-2} eV^{-1}. This level of interface state density would result in a much higher dispersion than the experimental data. For example, it leads to a G_p/ω peak value of 0.35 μF/cm^2, i.e., about 5.5 times the data in Fig. 3(b).

The discrepancy between low-frequency stretch-out and high-frequency dispersion is readily resolved with the bulk-oxide trap model that predicts much richer low-frequency component than high frequency [see Fig. 6(a)]. Physically, the slow sweep rate of V_g allows charging and discharging of more traps by tunneling deeper into the oxide. An expression for

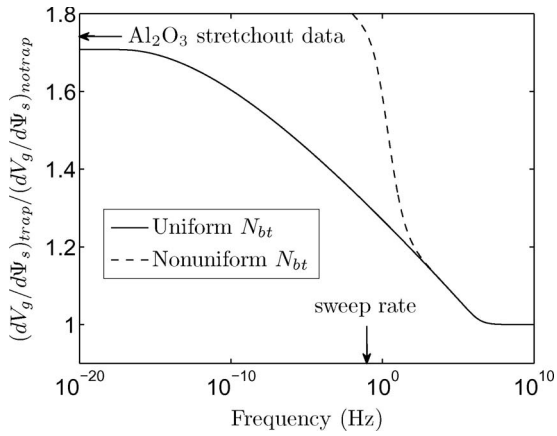


Fig. 9. Real part of $(\Delta V_g/\Delta\Psi_s)_{\rm trap}/(\Delta V_g/\Delta\Psi_s)_{\rm no\,trap}$ versus frequency at $V_g = 0.3$ V for Al_2O_3 MOS with uniform $N_{\rm bt}$ and other parameters extracted from high-frequency C and G dispersions in Fig. 8 (solid line). (Dashed line) Ratio calculated with a nonuniform $N_{\rm bt}$ that produces the observed stretch-out at a frequency in the range of the C–V sweep rate.

the frequency-dependent $\Delta V_g/\Delta\Psi_s$ can be derived from the distributed bulk-oxide trap model. We define the differential potential at a point x in the oxide in Fig. 5 as $V(x)$, then

$$V(x+\Delta x) - V(x) = \frac{Y(x)V(x)}{j\omega(\epsilon_{\rm ox}/\Delta x)} \tag{11}$$

where $Y(x)$ is the admittance at x solved by (7). Therefore

$$\frac{dV}{dx} = \frac{Y(x)V(x)}{j\omega\epsilon_{\rm ox}} \tag{12}$$

$$\frac{\Delta V_g}{\Delta\Psi_s} = \frac{V(x=t_{\rm ox})}{V(x=0)} = \exp\left[\frac{1}{j\omega\epsilon_{\rm ox}}\int_0^{t_{\rm ox}} Y(x)dx\right]. \tag{13}$$

For dc, $Y(x) = j\omega C(x)$, where $C(x)$ is given by (10) for uniform $N_{\rm bt}$. Equation (13) gives

$$\left(\frac{\Delta V_g}{\Delta\Psi_s}\right)_{\rm trap} = \frac{(C_s + C_0)\exp(2C_0/C_{\rm ox}) - (C_s - C_0)}{2C_0} \times \exp\left(-\frac{C_0}{C_{\rm ox}}\right) \tag{14}$$

where $C_0 = q\sqrt{\epsilon_{\rm ox}N_{\rm bt}}$. Note that this dc expression assumes that bulk-oxide traps all the way to the metal gate are engaged. In reality, the sweep rate of V_g is finite and we expect an experimental $\Delta V_g/\Delta\Psi_s$ value lower than the dc value of (14).

For the Al_2O_3 MOS biased at $V_g = 0.3$ V, $\Delta V_g/\Delta\Psi_s$ is computed from (13) using the uniform $N_{\rm bt}$ and other parameters extracted from the high-frequency C and G dispersions in Fig. 8. Fig. 9 plots the real part of $\Delta V_g/\Delta\Psi_s$ with respect to $(\Delta V_g/\Delta\Psi_s)_{\rm no\,trap} = (C_{\rm ox} + C_s)/C_{\rm ox}$ as a function of frequency (solid curve). The imaginary part is negligible. The experimental stretch-out of V_g can be estimated by comparing the $dC_{\rm tot}/dV_g$ slope of the 1-MHz C–V data to that of the ideal simulated C–V in Fig. 2. However, since both $N_{\rm bt}$ and τ_0 are bias dependent, the contribution of bulk-oxide traps to

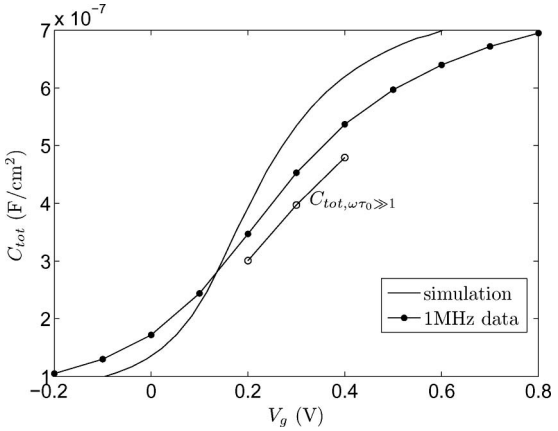


Fig. 10. Total capacitance of ideal simulation (solid line) and 1-MHz data (solid dots), as well as $C_{\mathrm{tot},\omega\tau_0\gg1}$ from model fitting (open circles) versus V_g near the flatband bias condition. The extracted N_{bt} at $V_g = 0.2$ V and $V_g = 0.4$ V are 1.81×10^{19} cm^{-3}eV^{-1} and 2.6×10^{19} cm^{-3}eV^{-1}, respectively. The extracted τ_0 at $V_g = 0.2$ V and $V_g = 0.4$ V are 2.5×10^{-7} s and 6×10^{-8}, respectively.

the 1-MHz C_{tot} is also bias dependent and has the effect of steepening the slope. A more accurate way is to determine the slope using the extracted $C_{\mathrm{tot},\omega\tau_0\gg1} = C_s C_{\mathrm{ox}}/(C_s + C_{\mathrm{ox}})$ near $V_g = 0.3$ V. This is shown in Fig. 10, which gives a $dC_{\mathrm{tot},\omega\tau_0\gg1}/dV_g$ slope a factor of 1.74 smaller than that of the simulated $C-V$ with no stretch-out. This factor is compared to the computed ratio in Fig. 9 since $dC_{\mathrm{tot},\omega\tau_0\gg1}/dV_g$ is inversely proportional to $\Delta V_g/\Delta\Psi_s$ for the same $C_{\mathrm{tot},\omega\tau_0\gg1}$ and Ψ_s between the two curves. While the calculated dc ratio of 1.71 is comparable to the measured inverse-slope ratio, the calculated ratio of 1.31 at ~0.1 Hz corresponding to the sweep rate is lower. The model-to-data match can be fine tuned by allowing nonuniform N_{bt}, with higher trap densities toward the gate. The dashed curve in Fig. 9 shows the model calculation with a nonuniform N_{bt} that reproduces the experimental $(\Delta V_g/\Delta\Psi_s)_{\mathrm{trap}}$ at 0.1 Hz without affecting the dispersions at higher frequencies.

While our work has shown that charging and discharging of bulk-oxide traps at 1 kHz–1 MHz can be satisfactorily explained by an elastic-tunneling-based model, recent research on reliability of thin dielectrics by bias-temperature stress using random telegraph noise and time-dependent defect spectroscopy techniques has revealed inconsistency of trap capture and emission with the elastic tunneling model [14]–[17]. It is not clear whether the difference is due to different types of traps in the oxide or due to different stress and characterization techniques.

VI. Conclusion

In conclusion, a distributed bulk-oxide trap model based on tunneling between the semiconductor surface and trap states in the gate insulator has been developed. It fundamentally differs from the conventional interface state model in that there is a wide frequency spectrum of bulk-oxide trap response at a given gate bias. It is more physical than previously published lumped-circuit models in the literature. The model is validated with the $Pt/Al_2O_3/n\text{-}In_{0.53}Ga_{0.47}As$ dispersion data in strong accumulation and near the flatband. Unlike surface states, which are in units of areal density, bulk-oxide traps are characterized by a volume density, extracted from fitting of the capacitance and conductance data.

It is further shown that the commonly employed method of extracting the interface state density from $C-V$ stretch-out could yield unphysical numbers inconsistent with the high-frequency dispersion data. On the other hand, the bulk-oxide trap model, in particular, with nonuniform trap density in the oxide film, can explain $C-V$ stretch-out independent of the dispersion at higher frequencies.

References

[1] M. Passlack, R. Droopad, and G. Brammertz, "Suitability study of oxide/gallium arsenide interfaces for MOSFET applications," *IEEE Trans. Electron Devices*, vol. 57, no. 11, pp. 2944–2956, Nov. 2010.

[2] Y. Xuan, Y. Q. Wu, T. Shen, T. Yang, and P. D. Ye, "High performance submicron inversion-type enhancement-mode InGaAs MOSFETs with ALD Al_2O_3, HfO_2 and HfAlO as gate dielectrics," in *IEDM Tech. Dig.*, 2007, pp. 637–640.

[3] N. Goel, P. Majhi, C. O. Chui, W. Tsai, D. Choi, and J. S. Harris, "InGaAs metal–oxide–semiconductor capacitors with HfO_2 gate dielectric grown by atomic-layer deposition," *Appl. Phys. Lett.*, vol. 89, no. 16, pp. 163 517-1–163 517-3, Oct. 2006.

[4] E. J. Kim, E. Chagarov, J. Cagnon, Y. Yuan, A. C. Kummel, P. M. Asbeck, S. Stemmer, K. C. Saraswat, and P. C. McIntyre, "Atomically abrupt and unpinned Al_2O_3/ $In_{0.53}Ga_{0.47}As$ interfaces: experiment and simulation," *J. Appl. Phys.*, vol. 106, no. 12, pp. 124 508-1–124 508-8, Dec. 2009.

[5] G. Brammertz, K. Martens, S. Sioncke, A. Delabie, M. Caymax, M. Meuris, and M. Heyns, "Characteristic trapping lifetime and capacitance–voltage measurements of GaAs metal–oxide–semiconductor structures," *Appl. Phys. Lett.*, vol. 91, no. 13, pp. 133 510-1–133 510-3, Sep. 2007.

[6] Y. Hwang, R. Engel-Herbert, N. G. Rudawski, and S. Stemmer, "Effect of postdeposition anneals on the Fermi level response of HfO_2/ $In_{0.53}Ga_{0.47}As$ gate stacks," *J. Appl. Phys.*, vol. 108, no. 3, pp. 034111-1–034111-4, Aug. 2010.

[7] E. H. Nicollian and J. R. Brews, *MOS (Metal Oxide Semiconductor) Physics and Technology*. New York: Wiley, 1982.

[8] W. Shockley and W. T. Read, Jr., "Statistics of the recombinations of holes and electrons," *Phys. Rev.*, vol. 87, no. 5, pp. 835–842, Sep. 1952.

[9] F. P. Heiman and G. Warfield, "The effects of oxide traps on the MOS capacitance," *IEEE Trans. Electron Devices*, vol. ED-12, no. 4, pp. 167–178, Apr. 1965.

[10] Y. Yuan, L. Wang, B. Yu, B. Shin, J. Ahn, P. C. McIntyre, P. M. Asbeck, M. J. W. Rodwell, and Y. Taur, "A distributed model for border traps in Al_2O_3–InGaAs MOS devices," *IEEE Electron Device Lett.*, vol. 32, no. 4, pp. 485–487, Apr. 2011.

[11] H. Preier, "Contribution of surface states to MOS impedance," *Appl. Phys. Lett.*, vol. 10, no. 12, pp. 361–363, Jun. 1967.

[12] D. S. L. Mui, J. Reed, D. Biswas, and H. Morkoç, "A new circuit model for tunneling related trapping at insulator–semiconductor interfaces in accumulation," *J. Appl. Phys.*, vol. 72, no. 2, pp. 553–558, Jul. 1992.

[13] E. J. Kim, L. Wang, P. M. Asbeck, K. C. Saraswat, and P. C. McIntyre, "Border traps in $Al_2O_3/In_{0.53}Ga_{0.47}As$ (100) gate stacks and their passivation by hydrogen anneals," *Appl. Phys. Lett.*, vol. 96, no. 1, pp. 012906-1–012906-3, Jan. 2010.

[14] J. P. Campbell, J. Qin, K. P. Cheung, L. C. Yu, J. S. Suehle, A. Oates, and K. Sheng, "Random telegraph noise in highly scaled nMOSFETs," in *Proc. IEEE Int. Rel. Phys. Symp. Tech. Dig.*, 2009, pp. 382–388.

[15] H. Reisinger, T. Grasser, W. Gustin, and C. Schlünder, "The statistical analysis of individual defects constituting NBTI and its implications for modeling DC- and AC-stress," in *Proc. IEEE Int. Rel. Phys. Symp. Tech. Dig.*, 2010, pp. 7–15.

[16] T. Grasser, H. Reisinger, P.-J. Wagner, and B. Kaczer, "The time dependent defect spectroscopy for the characterization of border traps in metal–oxide–semiconductor transistors," *Phys. Rev. B*, vol. 82, no. 24, pp. 245 318-1–245 318-10, 2010.

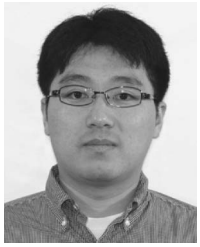
[17] D. Garetto, Y. M. Randriamihaja, A. Zaka, D. Rideau, A. Schmid, H. Jaouen, and Y. Leblebici, "Analysis of defect capture cross sections using non-radiative multiphonon-assisted trapping model," *Solid State Electron.*, vol. 71, pp. 74–79, May 2012.

Yu Yuan received the Ph.D. degree in electrical and computer engineering from the University of California, San Diego, in 2011.

She is with the Process Technology Department, Qualcomm, CA. Her research interests include design, modeling, and characterization of nonclassical nanoscale MOSFETs.

Bo Yu received the Ph.D. degree in electrical and computer engineering from the University of California, San Diego, La Jolla in 2009.

He is currently with the Process Technology Department, Qualcomm, San Diego, CA.

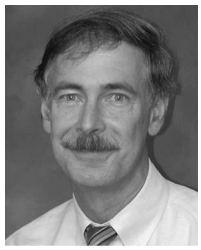


Jaesoo Ahn is currently working toward the Ph.D. degree in the Department of Materials Science and Engineering, Stanford University, Stanford, CA.

His current research interests include atomic layer deposition of high-k gate dielectrics on III–V substrates.

Paul C. McIntyre received the Sc.D. from Massachusetts Institute of Technology, Cambridge.

He is a Professor in materials science and engineering, a Senior Fellow of the Precourt Institute for Energy, and a Director of the Geballe Laboratory for Advanced Materials with Stanford University, Stanford, CA.



Peter M. Asbeck (F'00) received the Ph.D. degree from Massachusetts Institute of Technology, Cambridge in 1975.

Since 1991, he has been with the University of California, San Diego, where he is currently the Skyworks Chair Professor with the Department of Electrical and Computer Engineering.

Mark J. W. Rodwell (F'03), photograph and biography not available at the time of publication.

Yuan Taur (F'98) received Ph.D. from University of California, Berkeley. He worked at IBM T. J. Watson Research Center, New York from 1981 to 2001.

Since 2001, he has been a professor at University of California, San Diego.

Dr. Taur was elected IEEE Fellow in 1998. He has been the Editor-in-Chief, IEEE ELECTRON DEVICE LETTERS from 1999 to 2011.