Electron capture kinetics at AlF₃/SiO₂ interfaces

I. Thurzo, T.U. Kampen, D.R.T. Zahn, D. König

Abstract

Charge deep-level transient spectroscopy (Q-DLTS) was applied to study the mechanism of capturing electrons at the AlF₃/SiO₂ interfaces of AlF₃-coated oxidized p-type Si. The process may be viewed as comprising three consecutive steps: (i) formation of an inversion layer in p-Si; (ii) electron tunneling through the ultra-thin SiO₂ interlayer; and (iii) capture of electrons at F-vacancies in AlF₃ after surmounting a potential barrier \( E_b \). A semi-empirical model of non-exponential kinetics of the charge transients has been developed and compared with experiment. Three discrete electron traps of a quasi-discrete spectrum of traps in AlF₃ could be resolved. The thermal activation energy of a Q-DLTS peak (0.6–0.7 eV) seems to reflect \( E_b \), while the full-width at half-maximum (FWHM) corresponds to thermal activation of minority carriers over the band gap of p-Si (≈1.1 eV). Coulomb-blockade effects may be responsible for the temperature dependence of the capture cross-section \( \sigma_n = \sigma_{n0} \exp(-E_b/kT) \).

Keywords: AlF₃; Electron states; Spectroscopy

1. Introduction

The solar cells based on the field effect induced by a large sheet density of fixed positive charges in p-type Si were first treated by Green et al. [1]. Then there was a gap of more than two decades until the proposal to design alternative devices with a high density of fixed negative charges appeared [2,3]. At least two discrete electron energy levels possibly originating from fluorine vacancies (\( V_F \)) at the AlF₃/SiO₂ interface were deduced from mercury probe capacitance–voltage (C–V) measurements. Evidence for sub-stoichiometric fluoride content at the AlF₃/SiO₂ interface came from (unpublished) inclined-angle Rutherford backscattering measurements (measurement carried out by Schrempel [4]).

The above-mentioned C–V results found a strong support from preliminary, recently conducted isothermal charge deep-level transient spectroscopy (Q-DLTS) experiments [5]. Two predicted distinct energy levels could be detected, yet, when interpreting the Q-DLTS results in terms of trap depths, the corresponding thermal activation energies of related Q-DLTS peaks were inconsistent with expected (calculated) energy level positions above the valence band edge of AlF₃. Moreover, the full-width at half-maximum (FWHM) of either of the two isothermal Q-DLTS peaks appeared much smaller than that calculated for the charge emission from a discrete energy level.

The present contribution aims at removing at least the main inconsistencies concerning the handling of
the preliminary data, and presenting original thermal-scan Q-DLTS spectra which prompted for formulating a new model of the response. Before doing so, a brief description of the Q-DLTS implementation, as well as of the closely related feedback charge capacitance method (FCM) will be provided. Finally, possible ways of improving the model and extending its applicability will be outlined as concluding remarks.

2. Basics of the Q-DLTS and FCM techniques

The Q-DLTS is actually a valuable modification of the capacitance DLTS developed by Lang [6]. There is fundamental work by Miller et al. [7], where important physical and instrumental aspects of the classical DLTS are reviewed. When designing the Q-DLTS, the methodology presented by Farmer et al. [8] under the acronym QTS (charge transient spectroscopy) was substantially modified towards achieving a better immunity to leakage (steady-state) currents, as well as improving energy resolution [9]. The FCM technique was introduced by Mego [10] as a tool for performing quasi-static C–V measurements in the time domain. The transition from the Q-DLTS to the FCM can be done using the same hardware while reprogramming the time base [9].

Starting with the Q-DLTS mode, after an excitation pulse, i.e. in the rest of the measurement period, three values of the transient charge are sampled. After a predefined number of cycles the three acquired values \( Q(t_1) \), \( Q(t_2) \) and \( Q(t_3) \) \( (t_2 = 2t_1, t_3 = 4t_1) \) are combined to obtain a weighted sum:

\[
\Delta Q = Q(t_1) - \frac{1}{2} Q(2t_1) + \frac{1}{4} Q(4t_1).
\]

(1)

Introducing the emission time constant as \( \tau_p = e_p^{-1} \), \( e_p \) being the emission rate, one can express the kinetics of the emitted charge \( Q(t) \) through:

\[
Q(t) = Q_0 \left[ 1 - \exp \left( -\frac{t}{\tau_p} \right) \right].
\]

(2)

The amplitude of any DLTS peak from a discrete energy level amounts to only 17% of the total emitted charge \( Q_0 \), nevertheless the rate-window concept [6,7] holds and the trap parameters can be assessed in the standard way, taking \( \tau_{\text{max}} = t_1 \) for the peak. There is the possibility of finding the time constant \( \tau_{\text{max}} \) by either sweeping \( t_1 \) at a constant temperature \( T \) (isothermal Q-DLTS), or scanning temperature while keeping \( t_1 \) fixed (thermal Q-DLTS).

The evolution of the charge \( Q(t) \) in the FCM mode is to some extent similar to the one in the Q-DLTS mode, still now there are two essential special features. First the channel number one is sampling the baseline \( Q(t_0) \) at the point in time \( t_0 \) with respect to the leading edge of the pulse. The trailing edge of the pulse is immediately followed by the incremental charge \( \Delta Q_i \) related to the instantaneous (geometrical) capacitance \( C(0) = \Delta Q_i / \Delta U \). The time-dependent portion of the charge \( Q(t) \) is assigned to the charges released from the trap level. Then the overall capacitance composed of \( C(0) \) and the “excess” capacitance reads:

\[
C(t_2) = Q(t_0) - \left( \frac{3}{2} Q(t_2) + \frac{1}{2} Q(3t_2) \right) / \Delta U.
\]

(3)

The excess charge acquired in the FCM mode is equal to the full charge \( Q_0 \) released from traps \( (t_2 \gg 5\tau_p) \).

Restricting our consideration about the Q-DLTS to large-pulse excitation \( (\Delta U \gg kT/q) \), the description of the kinetics of emission (Eq. (2)) is relatively simple compared to what is going on during the pulse. The capture rate \( c_p \) is a function of the free-carrier density \( p(x) \) at the position \( x \), where the capture takes place, as well as of the capture cross-section \( \sigma_p \) and thermal velocity \( v(T) \):

\[
c_p = \sigma_p v(T) p(x).
\]

(4)

An accurate treatment of the capture at the Debye tail of free carriers within a space charge region was provided by Pons [11].

3. Experimental results

3.1. FCM capacitance–voltage characteristics

The representative C–V plot in Fig. 1 belongs to Au/AlF₃/SiO₂/p-Si devices comprising 1.5 nm SiO₂ coated by 20 nm thick AlF₃ layer. The diameter of the Au dots (gate electrode) was \( \phi = 1 \) mm, corresponding to an area of \( 7.85 \times 10^{-7} \text{ m}^2 \). Evidently there is an excess capacitance at gate biases ranging from 0.5 to 2 V, which is dependent on the biasing and timing conditions. This is the range of biases to be explored by Q-DLTS.
3.2. Thermal-scan Q-DLTS spectra

It might appear at first sight that there is a complete equivalence between isothermal and complementary thermal Q-DLTS measurements. The equivalence is restricted to the uniquely defined rate-window only. In the thermal-scan Q-DLTS mode the FWHM of a peak of the signal is carrying information about the thermal activation energy $D_E$ and consequently also about $\sigma_p$. Then it is in principle possible to obtain both parameters from a single temperature scan.

To exemplify the thermal-scan Q-DLTS response of the devices under test we appeal to Fig. 2, showing the evolution of the spectra upon proceeding from depletion at the SiO$_2$/p-Si interface ($U_g = 0.5$ V) to strong inversion ($U_g = 2.5$ V) while keeping the interface at accumulation during the pulse ($U_g + \Delta U = -0.5$ V). Evidently, there is a discrete spectrum of at least three energy levels involved in the response, followed by a transition to a quasi-continuum of electron states in strong inversion. Now two distinct values of thermal activation energies $D_E$ can be considered, namely one determined from the corresponding Arrhenius’ plot ($D_{E_{Arr}}$) and another one ($D_{E_{FWHM}}$) obtained from FWHM, respectively.

Arrhenius’ plots of peaks derived from thermal-scan Q-DLTS data provided values of thermal activation energies $D_E$ within the interval from 0.6 to 0.7 eV, these are consistent with those extracted earlier from isothermal Q-DLTS data [5]. The analysis of the FWHMs of the peaks in the thermal-scan Q-DLTS spectra led to a controversy with the calculated equivalent half-widths, i.e. the FWHM of each of the experimentally observed $\Delta Q$ versus $T$ peaks was always much lower than the one computed for emission of charge from a discrete energy level, while taking the $D_E$ from the corresponding Arrhenius’ plot. To match the experimental FWHM of any thermal Q-DLTS peak as due to a discrete level, one had to consider $D_E$ values falling to a narrow interval around $D_{E_{FWHM}} = 1.1$ eV. This applies for all energy levels (peaks) resolved by the second-order filter, as nicely illustrated in Fig. 3, showing the simultaneous Q-DLTS response of well resolved levels II and III, respectively. We have realized that the non-random, almost unique $D_{E_{FWHM}}$ of about 1.1 eV may nearly correspond with the band gap $E_g = E_C - E_V$ of Si.

![Fig. 1. Capacitance–voltage (FCM – $\Delta U = 0.06$ V) characteristics of an Au/AlF$_3$/SiO$_2$/p-Si device exhibit a remarkable dispersion of capacitance at positive (reverse) biases.](image-url)
rather than to the depths of the energy levels. This hot track has been followed when formulating the new model of the kinetics of the \( Q \)-DLTS response of the Au/AlF\(_3\)/SiO\(_2\)/p-Si diodes.

### 3.3. Modeling kinetics of \( Q \)-DLTS spectra

After having recognized the failure of the standard concept, we have abandoned the notion of the unique excited volume and considered a three-step mechanism of charge transfer comprising all the solids involved:

1. After driving the SiO\(_2\)/p-Si interface from weak accumulation (at \( U_g = 0.5 \text{~V} \)) toward inversion (at \( U_g > 0 \)) after the pulse, generation of a sheet of minority carriers (electrons) at the interface is taking place. At the same time the generated majority carriers (holes) are leaving the space charge region under the impact of the electric field.

2. The accumulated electrons behave like a reservoir of charges to cross the ultra-thin SiO\(_2\) layer via tunneling.

3. The electrons reaching the AlF\(_3\) side of the structure are forced to surmount a potential barrier \( E_b \) before being captured at deep energy levels (fluorine vacancies \( V_F \)).
On the time scale of our experiments \( t_1 = 2 \times 10^{-6} \) to 0.6 s, we neglect the time electrons need for tunneling through the SiO\(_2\) layer, concentrating on steps 1 and 3, respectively. The finite durations of both pulse and period have been ignored as well.

As a first approximation, we describe the evolution of the inversion layer by the time constant \( \tau_g \) characteristic of the generation of electron/hole pairs at the SiO\(_2\)/Si interface:

\[
\tau_g = \tau_{g0} \exp\left(\frac{E_g}{kT}\right) \tag{5}
\]

The appearance of \( E_g \), resembling the value of \( 1.1 \) eV observed for \( \Delta E_{\text{FWHM}} \), instead of \( E_g/2 \) may be surprising with regard to the Shockley–Read–Hall statistics [12, 13] and the widely accepted notion of the generation/recombination processes of electrons and holes, as a process mediated by some mid-gap defect levels in Si. Adopting Eq. (5), the time dependence of the charge per unit area, \( Q_g(t) \), in the inversion layer is expressed in analogy with Eq. (2) as:

\[
Q_g(t, T) = Q_{g0}(T) \left[ 1 - \exp\left( -\frac{t}{\tau_g(T)} \right) \right] \tag{6}
\]

where the steady-state charge \( Q_{g0}(T) \) in general is a function of temperature. Neglecting possible charge capture events within the SiO\(_2\) layer, the number of free electrons per unit area available for being captured at the AlF\(_3\)/SiO\(_2\) interface is \( n_s(t, T) \approx Q_g(t, T)/q \). The last step of the response represents capture of the free electrons at the fluorine vacancies \( (V_F) \) at a time-dependent rate (s\(^{-1}\)):

\[
c_n(t) = c_0 \frac{n_s(t)}{L_c} \tag{7}
\]

where \( c_0 = \sigma_n v \) is a constant and \( L_c \) is the effective width of the spatial confinement of \( V_F \) in AlF\(_3\). Postulating that there is also the barrier-to-capture \( E_b \), the capture cross-section \( \sigma_n \) is also thermally activated:

\[
\sigma_n = \sigma_{n0} \exp\left( \frac{-E_b}{kT} \right) \tag{8}
\]

Denoting the sheet density of initially empty \( V_F \) by \( N_s \) and realizing that the capture probability is proportional to the number of free electrons \( n_s(t) \) and to the one of initially empty states \( N_s \), we arrive at an expression for the sheet density of captured electrons \( Q_s(t) \):

\[
Q_s(t) = q[N_s n_s(t)]^{1/2} \left[ 1 - \exp(-c_n(t)) \right] \tag{9}
\]

The quantity \( N_s \) is expected to decay with time to some equilibrium value \( N_s(\infty) \), so Eq. (9) is only correct as long as the fraction of occupied states is much less than \( N_s \). Finally, the correlated charge \( \Delta Q_s \) is expressed via Eq. (1) after replacing \( Q(t) \) by \( Q_s(t) \).

Now let us verify our model concerning the thermal-scan \( Q \)-DLTS spectra as originating from capture at
traps in AlF$_3$. In Fig. 4, the result of simulating the level II response is reproduced along with the list of parameters involved. A comparison of the Arrhenius’ plots originating from both the experimental data on the thermal peak II, and those from the simulation depicted in Fig. 4, is presented in Fig. 5. The FWHMs of the simulated peaks (Fig. 4) are consistent with what has been observed experimentally.

A few guidelines are provided on how to improve the new model performance. Apart from including the finite durations of both pulse and period, we are primarily interested in the energy level positions $E_t$ in the gap. With regard to the expected extremely high barrier for transitions of electrons from the V$_F$-related traps back to the p-Si substrate, it is not feasible to determine the energy positions of the trapping levels using DLTS with bias pulse excitation in dark. One is also interested in estimating the (sheet) densities of individual trapping levels resolved by the $Q$-DLTS. For doing this one has to update Eq. (9), inserting a term which would reflect the change $\Delta I$ in occupancy of either level $E_t$ as a consequence of applying the pulse $\Delta U$. Since we need again the quantity $E_n$, the energy level positions should be estimated in an independent way. Still another way of improving the model could be the introduction of a correct temperature dependence $Q_{\text{eff}}(T)$, as well as for $v(T)$ on the AlF$_3$ side of the devices.

4. Conclusions

A new model of the response to bias pulse excitation in dark of Au/AlF$_3$/SiO$_2$/p-Si devices has been suggested. The rate-determining step is the formation of a fluorine vacancy, followed by tunneling and capture of the minority carriers (electrons) at fluorine vacancies. A discrete energy spectrum of deep traps has been resolved. The thermal activation energies of the peaks from Arrhenius’ plots seem to reflect the height of the barrier-to-capture of electrons by fluorine vacancies, while the half-widths of the peaks are limited by the thermal energy of the electron–hole pair generation rate in Si ($\approx 1.1$ eV). To achieve the full spectroscopic property, a combination with optically excited $Q$-DLTS filtering remains a challenge for future initiatives.

References