AN ABSTRACT OF THE THESIS OF

Mehran G. Aminzadeh for the degree of Doctor of Philosophy in Electrical and Computer Engineering presented on April 25, 1988.

Title: Recombination and Generation Lifetime Characterization of p/p^+ Epitaxial silicon Wafers

Abstract approved: Redacted for privacy

Professor Leonard Forbes

The refresh times of all dynamic charge storage devices, best characterized by the generation lifetime at room temperature and the recombination lifetime at higher device operating temperatures ($T>70^{\circ}C$), strongly influence the efficient and successful operation of dynamic charge storage devices such as DRAM's and CCD's. Both recombination and generation lifetime characteristics of p/p^+ epitaxial silicon wafers are investigated. Extremely high generation lifetimes on the order of 20-30 msec are observed on p-type epitaxial layers which are correlated by DLTS measurements where no significant trap levels with concentrations higher than $3x10^{11}$ cm^{-3} are observed.

Schroder's [1] simple technique for the determination of recombination lifetimes of materials using a pulsed MOS capacitor technique at elevated temperatures ignores the lateral quasi-neutral bulk generation for short base width devices such as thin epitaxial layers. Consequently, calculations using Schroder's technique indicate that the recombination lifetime of a given material is a function of MOS capacitor diameter. A simple one dimensional approach

was developed [2] for the measurement of recombination lifetimes in which quasi-neutral bulk generation in the lateral area of MOS capacitors and the time dependence of the width of space-charge region are taken into consideration for short-base-width devices.

The apparent recombination lifetime of thin p-type epitaxial layers were limited to 2 μsec by the lower lifetime of the epitaxial layer - p^+ substrate interface. Investigation of wafers with various epitaxial layer thickness revealed that the actual recombination lifetime of the bulk epitaxial material is an order of magnitude higher than the apparent values. Precipitates at the epitaxial layer - substrate interface were suspected as the main limiting factor causing a localized degradation of recombination lifetimes at the epitaxial layer - substrate interface in p/p^+ epitaxial wafers. Application of an intrinsic gettering cycle and CMOS simulation heat treatments suggests up to three fold improvement of apparent recombination lifetimes.

Recombination and Generation Lifetime Characterization of p/p^+ Epitaxial Silicon Wafers

by

Mehran G. Aminzadeh

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of
Doctor of Philosophy

Completed April 25,1988

Commencement June 1988

APPROVED:

Redacted for privacy

Professor of Electrical and Computer Engineering in charge of major

Redacted for privacy

Head of department of Electrical and Computer Engineering

Redacted for privacy

Dean of Graduate School

Date thesis is presented: April 25, 1988

Typed by researcher for: Mehran G. Aminzadeh

to my parents

ACKNOWLEDGEMENT

I express my sincere appreciation to professor Len. Forbes for his guidance, and support during this investigation. I am grateful to Dr. John Wager and Dr. Steve Goodnick for their timely and critical review of this manuscript and helpful discussions. Also to be thanked is Dr. Babak Sabi for his continuous encouragement and his helpful discussions.

The author is grateful to Dr. Howard Mollenkopf, Dr. Witawat Wijaranakula, and Dr. Albert Derheimer of SEH America, Inc. for their valuable discussions and support throughout this investigation.

This work was sponsored in part by SEH America, Inc. and the Intel Research Council.

TABLE OF CONTENTS

TABLE OF CONTENTS — Continued

TABLE OF CONTENTS — Continued

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF TABLES

RECOMBINATION AND GENERATION LIFETIME CHARACTERIZATION OF p/p^+ EPITAXIAL SILICON WAFERS

I. INTRODUCTION

Since the invention of integrated circuits the technology trends have always been focused on size and especially toward faster speeds of electronic devices. With the recent development of 32 bit microprocessors, clearly the issue of speed remains the main goal of the future electronic industry. These systems clearly require large amounts of information stored in dynamic memories, commonly used in various electronic devices, to be handled more efficiently. The refresh time, which is the minimum time interval required for refreshing a dynamic storage device without loss of information, is very important in dynamic memories. The advantages of simple operation and small geometry of dynamic memories have to be compromised with intelligent and expensive circuits to refresh the memory circuits. Refreshing of memory circuits is a time consuming process causing a loss of computing time, so it is desired to have very long refresh times to increase computing time and have faster circuits. Imperfections or defects present in silicon material, best characterized by carrier lifetimes, strongly influence the refresh time of dynamic memories.

Prior to the invention of field-effect transistors, most semiconductor devices were minority carrier operated. The gain of bipolar transistors and their switching times are dependent on minority carrier transport across the base. The leakage current of

p-n junctions is also dependent on minority carrier lifetime. Solar cells rely on efficient transport of minority carriers. Although metal oxide semiconductor (MOS) devices and Schottky diodes are majority carrier devices, the minority carrier lifetime plays an important role in the leakage current of these devices. The refresh time of dynamic random access memories (DRAM's) and efficient operation of charge coupled devices (CCD's) are also strongly dependent on lifetime of the material.

In recent years, a thin layer of lightly doped silicon deposited on a heavily doped substrate by a chemical vapor deposition process known as epitaxial deposition has attracted advanced bipolar and MOS technologies. Epitaxial silicon offers an improved active layer quality, free from the defects normally observed in bulk Czochralski silicon, and results in increase in yield and device performance. Epitaxial silicon wafers have superior lifetimes since a heavily doped substrate limits the diffusion current collected by a space charge region in the lightly doped layer. Naturally, dynamic charge storage devices such as DRAM's and CCD's benefit by substituting the initial material from polished wafers to epitaxial wafers. However, technology trends are constantly pushing toward reduced impurities in the material or higher generation and recombination lifetimes. It is necessary to establish an understanding for the material limitations of epitaxial silicon by means of electrical characterization for very large scale integration (VLSI) and ultra large scale integration (ULSI) applications.

The main goal of this study is characterization of generation and recombination lifetimes of p/p^+ epitaxial wafers and a basic

understanding of factors which limit the lifetimes in epitaxial material. It is shown that surfaces and interfaces are primarily responsible for generating minority carriers causing a significant reduction in apparent or measured generation and recombination lifetimes. Deep level Transient spectroscopy (DLTS) measurements indicates no significant deep levels present in the p-type epitaxial layers and proper accounting of surface generation using modified Zerbst technique [3] reveals extremely long actual generation lifetimes for the p-type epitaxial layer. From the study of epitaxial wafers with various epitaxial thicknesses the epi-layer - substrate interface is identified as a source of generation of a significant number of minority carriers causing low apparent recombination lifetimes. In addition, properties of recombination lifetime of p/p^+ epitaxial wafers such as temperature dependence of τ_r and possible sources causing the reduction of apparent lifetimes are investigated in this work.

The generation lifetime in silicon is determined from the thermal emission rate of generated carriers within a space charge of a device due to the trap levels present within the bandgap. The generation lifetime measured at room temperature, is very sensitive to the energy of impurities or defects within the silicon bandgap. The generation lifetime samples a volume limited to the space charge region which is usually of the order of a few microns from the surface. This measurement is commonly used for characterization of a material or process. In contrast, the recombination lifetime is relatively insensitive to the position of energy levels of impurities or defects. The depth sampled by the recombination lifetime is a

material property and is typically tens to hundreds of microns. At elevated temperatures (60-100°C) or higher which is a more realistic representative of device operating conditions, the diffusion process, characterized by a recombination lifetime, dominates over the generation within the space charge region. In contrast to various reports in the literature on generation lifetimes, recombination lifetimes are seldom reported in the evaluation of epitaxial wafers. Although, as a first order evaluation of the material quality the properties of the active region of a device are of main concern, the effects of recombination lifetime in any dynamic charge storage device especially at normal device operating temperatures is inevitable. The current understanding of lifetime limitations, especially recombination lifetime on epitaxial wafers, is very limited.

Even though generation and recombination lifetimes in silicon are inverse processes, the physical conditions under which they apply are quite different, resulting in values which differ by a large amount. The pulsed MOS capacitor technique or its variations have been well established for generation lifetime measurements. On the other hand, numerous techniques are available to characterize the recombination lifetime of a material or a device. The photoconductive decay method, surface photo voltage, reverse recovery technique, open circuit voltage decay, pulse MOS capacitor, and admittance techniques are some of many existing techniques for recombination lifetime evaluation. When one wishes to study lifetime it appears logical to adapt the measurement technique and test structure similar to the device and application.

A brief review of available lifetime measurement techniques and effects of oxygen precipitation behavior and intrinsic gettering on lifetimes is discussed in the next chapter. In order to appropriately evaluate the lifetimes on high quality epitaxial wafers the existing lifetime techniques have been replaced or modified (Chapter II and III). The measurement technique and test structures used in this study were kept simple to avoid complications or introduction of secondary defects. MOS capacitors for investigation of generation and recombination lifetime properties of p/p^+ epitaxial silicon wafers and Schottky barrier diodes for deep level transient spectroscopy (DLTS) measurements were fabricated. The test capacitors and Schottky diodes were fully characterized to meet the acceptance criteria prior to further lifetime or DLTS study as described in Chapter IV. The recombination lifetime characterization is discussed separately in Chapter V. The main emphasis in this study is the recombination lifetime properties of starting p/p^+ silicon material and identification of factors limiting lifetime values. Contribution of interface between p^+-substrate and the lightly doped epitaxial layer is examined by recombination lifetime measurements. For the first time, the size dependence of the test structures and effects of epitaxial thickness on the recombination lifetime of p-type epitaxial material has been studied.

II. BACKGROUND AND LIFETIME MEASUREMENT TECHNIQUES

The continuous trend towards smaller geometry and higher performance of semiconductor devices demands improved process control and more importantly, starting material free from defects and imperfections. The lifetime in semiconductors is a measure of material quality. Generation lifetime characterizes the active region of a device while recombination lifetime which is different in principle, is a measure of the bulk property of a material. Various techniques have been used to characterize lifetimes prior or subsequent to high temperature processing during device fabrication. The work described herein is mainly restricted to evaluation of p/p^+ epitaxial silicon wafers used as starting material for many technologies such as CMOS or bipolar.

In this chapter, the theoretical background used for the measurements and characterization in the following chapters is developed. The first section deals with the general recombination - generation processes in semiconductors, the generation lifetime measurment techniques, the Zerbst formulation for evaluation of generation lifetime using pulsed MOS capacitors, and a discussion of the maximum possible recombination and generation lifetimes in silicon. Assessment of the recombination lifetime using the pulsed MOS (Metal Oxide Semiconductor) capacitor technique is discussed in the second section, followed by a brief development of the Deep Level Transient Spectroscopy (DLTS) technique. In addition, a brief discussion on effects of intrinsic gettering on lifetimes are

described in the last section.

A. RECOMBINATION - GENERATION PROCESSES

There are three main recombination - generation mechanisms in semiconductors namely, the Shockley-Read-Hall (SRH), radiative, and Auger. The recombination rate R for minority carrier electrons in a p-type semiconductor depends on the number of excess carriers δn which can be written as [4]

$$R = A(\delta n) + B(\delta n)^2 + C(\delta n)^3 \qquad (1)$$

where A, B, C are constants which depend on the individual process. The recombination lifetime, τ_r, is defined to be the number of excess carriers divided by the recombination rate ($\delta n/R$)

$$1/\tau_r = A + B(\delta n) + C(\delta n)^2 = 1/\tau_{srh} + 1/\tau_{rad} + 1/\tau_{Auger} \qquad (2)$$

In a SRH process, recombination is completed via intermediate energy levels within the bandgap which act as "stepping stones" [5] in the transition of electrons and holes between the conduction and valence bands. Radiative recombination is a band to band process ; therefore, both electrons and holes must be present simultaneously for the recombination to take place. Thus the radiative recombination is inversely proportional to the number of excess carriers, δn. In Auger recombination, energy is given to a third carrier (either an electron or a hole) and the lifetime is inversely dependent on $(\delta n)^2$.

The Auger process is most important at high temperatures [6] and only under high level injection conditions. Radiative recombination is normally the dominating process in the direct band gap semiconductors like GaAs. SRH recombination is the dominant recombination-generation process in indirect bandgap semiconductors such as silicon. Generation of carriers at room temperature is due to the inverse SRH process shown in Fig. 1. Since both electron and hole emission processes must take place to complete the inverse SRH process, intermediate trap levels located around the midgap are the most effective recombination-generation centers within the bandgap. Trap levels close to the valence or conduction bands are inefficient recombination-generation centers since the probability of either an electron emission or hole emission is very low.

The free carrier lifetime in a semiconductor is the time required for the free carrier concentration in the bulk semiconductor to return to thermal equilibrium after a disturbance from equilibrium has occurred. When the system is driven out of equilibrium both the majority and minority carrier concentrations changes from their thermal equilibrium values. Depending on the reduction or increase of the carrier concentration after a transient, equilibrium will be reached by generation or recombination process, respectively. If the minority carriers dominate the majority carriers in establishing the return to equilibrium, then the lifetime is determined by the lifetime of the minority carriers. At room temperature, the minority carriers actually determine the lifetime of silicon wafers by a thermal emission process which is sensitive to the position of the energy level of defects or impurities present in the material. The

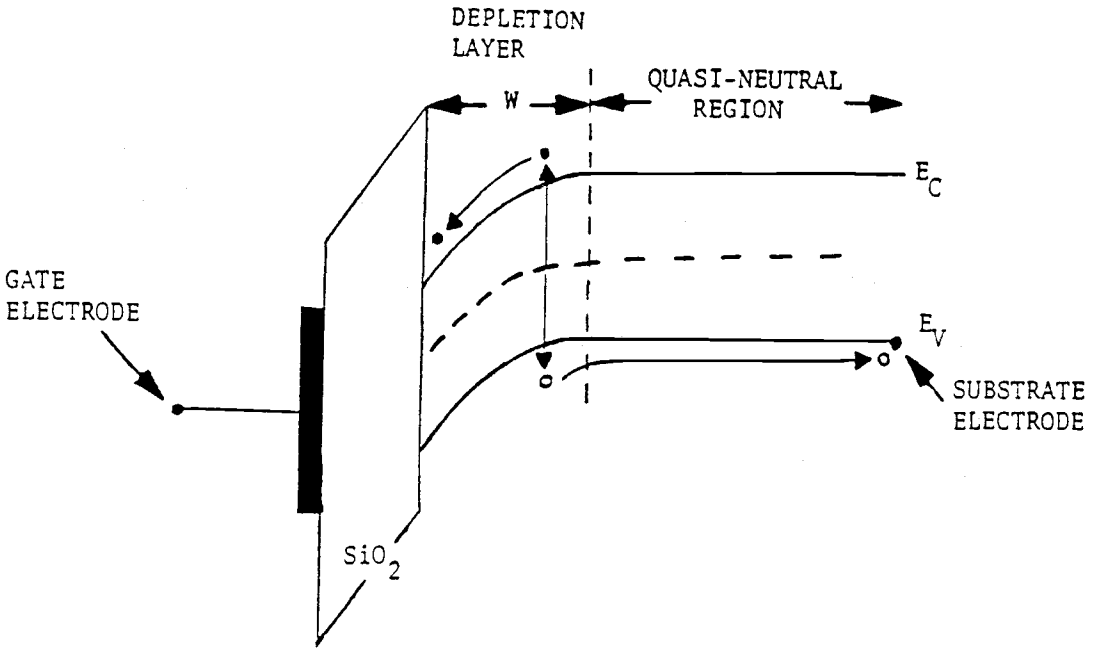


Figure 1) Generation within the space charge region corresponding to inverse SRH recombination process for p-type silicon.

term lifetime implies that an entity ceases to exit. However, in the generation process carriers are being continuously generated instead of being destroyed. It may seem surprising that this process is described by a lifetime, but none the less it is described as a generation lifetime and is a very useful parameter in device behavior.

1. Recombination vs. generation lifetime

Of the three main recombination - generation processes discussed in Section A1, the Shockley - Read - Hall process is the dominant

process in silicon. From the theory of the SRH process [7], the recombination rate of electron - hole pairs is given by

$$R = \frac{(pn - n_i^2)}{\tau_{po}(n + n1) + \tau_{no}(p + p1)} \tag{3}$$

$$\tau_{no} \equiv 1/(\sigma_n v_{th} N_T) \quad , \quad \tau_{po} \equiv 1/(\sigma_p v_{th} N_T) \tag{4}$$

where n1 and p1 are the number of electrons and holes in conduction and valence bands when the Fermi level coincides with trap energy, n and p are free carrier concentrations of electrons and holes, τ_{no} and τ_{po} are the electron and hole lifetimes respectively, n_i is the intrinsic carrier concentration, v_{th} is the carrier thermal velocity, σ_n and σ_p are electron and hole capture cross sections, and N_T is the trap concentration. In the following, the recombination and generation lifetimes, τ_r and τ_g are discussed using the pulsed MOS capacitor technique within the SRH framework.

At elevated temperatures diffusion from the bulk silicon is the dominant process in relaxation to equilibrium in pulsed MOS capacitors. The reason is simply because the diffusion process (shown in Fig. 2 is a function of the number of minority carriers (or α n_i^2) while the generation within the space charge region (Fig. 1) is proportional to n_i. The diffusion process is characterized by a diffusion length or a recombination lifetime, τ_r. For quasi - neutral bulk generation (diffusion process) in a pulsed MOS capacitor the minority carrier concentration changes significantly, but the majority carrier concentration changes by a negligible amount. In general, the concentration of electrons under the non-equilibrium

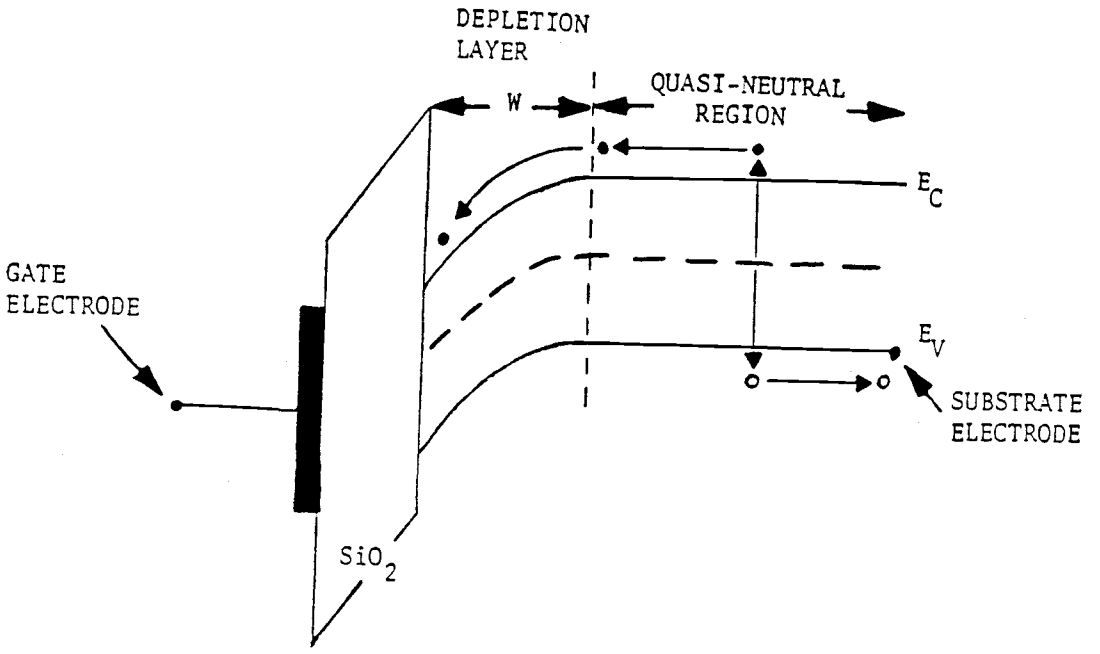


Figure 2) Diffusion from the bulk at elevated temperature dominates the generation within the space charge region.

condition is given by $n=n_o-\delta n$ and for holes by $p=p_o-\delta p$, where n_o and p_o are the equilibrium concentration of electrons and holes and δn and δp are the deficient concentrations due to the pulsed condition. The number of reduced carriers is assumed to be the same ($\delta n=\delta p$) when the number of deep level traps is much smaller than majority carrier concentration. The concentration of majority carriers does not change significantly, therefore, $p\approx p_o$ and for a p-type sample $p_o>>\delta n, n_o$. Furthermore, for the pulsed capacitance condition, the free carrier electron concentration becomes negligible ($n\approx 0$). Therefore, the recombination rate using Eqn. (3) results in

$$R \approx \frac{- p_o \delta n}{\tau_{po}(n_o - \delta n + n1) + \tau_{no}(p_o + p1)} \quad . \tag{5}$$

For trap centers whose energy level is close to E_i, $p_o >> p1, n1$ so the recombination rate reduces to

$$R \approx \frac{- \delta n}{\tau_{no}} = - \left(\frac{n_i^2}{N_a} \right) \frac{1}{\tau_{no}} \equiv - \left(\frac{n_i^2}{N_a} \right) \frac{1}{\tau_r} \tag{6}$$

The negative sign in Eqn. (6) indicates a generation process due to quasi - neutral bulk generation (diffusion). Consequently, the recombination lifetime of the material is equal to the minority carrier (electron) lifetime ($\tau_r = \tau_{no}$).

For evaluation of the generation lifetime, τ_g, consider Eqn. (3) with the concentration of electrons, n, and holes, p, to be essentially zero within the space charge region. Writing n1 and p1 in term of the trap energy level, Eqn. (3) becomes

$$R = \frac{- n_i}{\tau_{po} \exp[-(E_i - E_T)/KT] + \tau_{no} \exp[-(E_T - E_i)/KT]} \equiv \frac{- n_i}{\tau_g} \quad . \tag{7}$$

The generation lifetime, τ_g, defined in Eqn. (7) can be written as [4]

$$\tau_g = 2\tau_r \, (\sigma_n/\sigma_p)^{0.5} \, \text{COSH} \left[\, (E_T - E_i)/KT + \ln \, (\sigma_n/\sigma_p)^{0.5} \, \right] . \tag{8}$$

Equation (8) clearly shows that the generation lifetime is a strong function of the location of the trap level and also depends on the capture cross section of electrons and holes. Note that τ_g can be

much larger than τ_r ($=\tau_{no}$) if the trap energy level is not at the midgap or if the electron capture cross sections are much larger than holes.

The volume sampled by the recombination lifetime is quite different than the generation lifetime. For the recombination lifetime the sampled volume is determined by the minority carrier diffusion length, which is a material property. However, in the generation lifetime the sampled volume is limited by the space charge region. In high-quality silicon materials the minority carrier diffusion length is normally greater than 50 or 100 μm while the depletion region sampled in τ_g measurements is on order of a few microns. The generation lifetime is an important parameter in charge coupled devices, and in establishing the refresh time of dynamic RAM's. The recombination lifetime is an important factor in determining the switching time of bipolar devices, the leakage current of p-n junctions, and I^2L devices in bipolar technology. τ_r is also very important at normal operating temperatures for any dynamic charge storage device.

2. Generation lifetime measurement techniques

Many papers have been published dealing with the non-equilibrium pulsing a MOS capacitor from accumulation to deep depletion. The capacitance-time (C-t) transient technique involving a voltage step applied to a metal-insulator-semiconductor structure was introduced by Zerbst [3] and was further analyzed in more detail by Heiman [8]. The generation lifetime of the material is obtained from a time

derivative of the C-t response. Inversion to deep depletion pulsing was investigated by Hofstein [9]. Later Schroder [10,11] identified a significant component of C-t transient current due to lateral surface effects or lateral surface generation velocity, resulting in a reduction of the effective or measured generation lifetime. The Zerbst technique and a simple model developed by Schroder which incorporates the surface effects are discussed in detail in the next section. Pierret [12,13] has described and refined a technique using a linear voltage ramp to maintain a constant capacitance for measurement of the generation lifetime, τ_g. This method permits direct evaluation of τ_g from the slope of a plot of various applied voltage ramp rates versus the measured depletion layer width. Contrary to the time derivative calculation of the C-t response required for τ_g evaluation in the Zerbst technique, Pierret's procedure provides a simple means for the evaluation of τ_g. However, the experimental setup for this procedure is more complex compared to the Zerbst technique and it suffers from the breakdown in thin oxides due to the high voltages required during the voltage ramp. Lin [14] introduced a double-sweep C-V technique for determination of the generation rates and the doping concentration in non-uniform doping profile devices.

The retention time, defined as the time for return to the inversion capacitance after application of a voltage step, is extremely long [15] in high quality silicon materials. Care must be taken in the interpretation of the results [16]. Numerous modifications to the original Zerbst technique have been reported [15,17,18,19] to reduce the measurement time of generation lifetimes.

It has been shown that by pulsing from inversion to deep depletion and taking the C-t data over a small portion of transient response, τ_g can be evaluated in much shorter times without any information loss. This is discussed in detail in Chapter III. Recently, Keller [18] reported a rapid measurement based on establishing intermediate capacitance values by means of voltage pulses and calculating the time derivative in an identical fashion to the Zerbst technique, but only for short times. This method is similar to that discussed and used for the τ_g measurement in Chapter III [15], except various voltage pulses are employed rather than a single voltage step. Although this method may be more complex for routine evaluations, τ_g values can be obtained in about 10% of the total retention times. Fahrner et al. [19] investigated a technique which involves an increase in the measurement temperature to speed up the measurements. As pointed out by Schroder [11], for high quality silicon wafers the quasi-neutral bulk generation starts to become important at lower temperatures when compared to the space charge generation. Therefore, while increasing the temperature may reduce the measurement time, the accuracy of τ_g using such a technique is questionable. Furthermore, they assumed that the bulk generation is independent of the depletion layer thickness. This assumption does not hold for short base width devices [2] namely p/p^+ epitaxial wafers used in this study. Recently, Radzimski et al. [20] reported analysis of MOS capacitor C-t data that enables a plot of the generation lifetime as a function of depletion layer width which is useful for non-uniform lifetime characterization.

3. Generation lifetime using a pulsed MOS capacitor

The MOS capacitor is the most widely used structure for assessing the $Si-SiO_2$ interface and the quality of the bulk silicon. Measurements of the generation lifetime is possible by operating the MOS capacitor in a non-equilibrium mode (known as a pulsed MOS capacitor) introduced by Zerbst [3]. The Zerbst technique or its various modifications [15,17,18,19] have been commonly used for generation lifetime measurement. The generation lifetime is a very important parameter in monitoring deep level defects caused by interstitial or substitutional impurities in silicon which cause a degradation in performance of DRAM's and CCD's.

In the pulsed MOS capacitor measurement, the capacitor is pulsed from an accumulation to a deep depletion while monitoring the capacitance as a function time. The observed transient in the gate voltage results from formation of a deep depletion layer under the gate electrode causing the capacitor to be in a non - equilibrium condition. The gate voltage is kept at a fixed value until the capacitance returns to its equilibrium or inversion capacitance, C_{inv}. Figure 3 shows four processes contributing to relaxation to equilibrium: 1) generation of carriers within the space charge region with a generation lifetime, τ_g, 2) surface generation, S, of the surface under the gate , 3) lateral surface generation, S_0, from the surface of depletion region, and 4) diffusion process from the bulk silicon characterized by a recombination lifetime, τ_r. The diffusion contribution is normally negligible at room temperature (diffusion at elevated temperatures is discussed in section B).

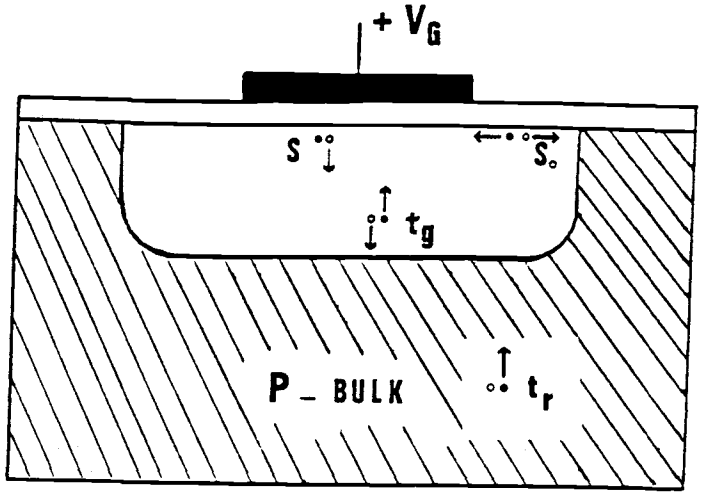


Figure 3) Schematic representation of a pulsed MOS capacitor showing the four carrier generation mechanisms involved in relaxation to equilibrium.

In the pulsed MOS capacitor technique, the surface generation, S, is a maximum immediately after the transition from accumulation to deep depletion. Because the surface under the gate electrode quickly becomes inverted while the lateral surface remains depleted immediately after the pulse. Once an inversion layer is formed, traps at the SiO_2 - Si interface become filled, thereby, reducing surface generation under the gate electrode. After the gate bias transient, S decreases rapidly while the lateral surface generation remains constant. As time progresses, the depletion width contracts and an equilibrium inversion layer is eventually formed. The period in which the capacitance relaxes back to C_{inv} is known as retention time.

The Zerbst's relation, commonly used for the evaluation of τ_g,

is derived here. From Gauss's law, the oxide field is proportional to the net charge per unit area [21]

$$C_{ox}[\ V_g - \phi_s(t)\ ] = q\ [\ N_I(t) + \int_0^{W(t)} N(x)\ d(x)\ ] \qquad (9)$$

where $\phi_s(t)$ is the instantaneous band bending, $W(t)$ is the depletion width, $N_I(t)$ is the instantaneous inversion layer density, $N(x)$ is the doping density at position x, and C_{ox} the oxide capacitance in F/cm^2. Differentiating Eqn. (9) with respect to time for a constant gate voltage after the pulse and assuming uniform doping concentration results in

$$\frac{dN_I}{dt} = -\frac{C_{ox}}{q}\frac{d\phi_s}{dt} - N_a\frac{dW}{dt} \quad . \qquad (10)$$

The space charge layer, in terms of the measured capacitance, is given by:

$$W(t) = \epsilon_{si}\ \text{Area}\ [\ 1/C(t) - 1/C_{ox}\ ] \qquad (11)$$

where $C(t)$ is the measured high frequency capacitance, and ϵ_{si} the permitivity of silicon. Neglecting the voltage drop across the inversion layer [21], the rate of change of the band bending becomes

$$d\phi_s/dt = q/\epsilon_{si}\ W(t)\ N_a\ dW/dt \qquad (12)$$

and with the use of Eqn. (10,11, and 12) the rate of change of the inversion layer density becomes

$$\frac{dN_I}{dt} = -\frac{C_{ox}}{C(t)} N_a \epsilon_{si} \frac{d}{dt}\left(\frac{1}{C(t)} - \frac{1}{C_{ox}}\right) = -\frac{N_a}{2}\frac{\epsilon_{si}}{\epsilon_{ox}} T_{ox} \frac{d}{dt}\left(\frac{C_{ox}}{C(t)}\right)^2 . \quad (13)$$

Equation (13) is a general relation between the rate of change of the inversion layer density and the rate of change of depletion layer width. Zerbst's analysis assumes that the rate of change of the inversion layer density or generation rate is linearly proportional to the depletion layer width,

$$dN_I/dt = n_i/\tau_g \ (W - W_f) \quad (14)$$

where W_f is the final (equilibrium) depletion width. The rate n_i/τ_g is the value used in the steady state analysis of reverse biased p-n junctions. Schroder and Nathanson [10] have shown that the rate of change of inversion layer density, taking the surface generation and the extended space charge region into account, gives

$$dN_I/dt = n_i/\tau_g \ (W - W_f) + n_i \ S \ A_G/A + n_i \ S_o \ (W - W_f) \ P/A \quad (15)$$

where A_G is the area of the gate, A is the area of the gate plus the lateral surface area shown in Fig. 3, P is the perimeter of the gate, and W_f is the final depletion width. For simplicity, Schroder assumes that the lateral and longitudinal width of the space charge regions are the same. This is a fairly good first order approximation. For a circular gate of diameter D, Eqn. (15) simplifies to

$$\frac{dN_I}{dt} \approx n_i\ (W - Wf)\ \left[\frac{1}{\tau_g} - \frac{4S}{D} + \frac{4S_o}{D}\right] + n_i\ S \qquad (16)$$

for $t>0^+$ on the C-t curve, S is much smaller than S_o, so the (4S/D) term in Eqn. (16) may be neglected. Using the Eqn. (13 and 16) we obtain

$$-\frac{d}{dt}\left[\frac{C_{ox}}{C(t)}\right]^2 = 2\frac{n_i}{N_a}\left[\frac{C_{ox}}{C_{inv}}\left(\frac{C_{inv}}{C_{ox}} - 1\right)\frac{1}{\tau'_g} + \frac{\epsilon_{ox}}{\epsilon_{si}}\frac{S}{T_{ox}}\right] \qquad (17)$$

where τ'_g is the effective generation lifetime defined by

$$1/\tau'_g = 1/\tau_g + 4S_o/D\ . \qquad (18)$$

Equation (17) is the Zerbst relation which relates the capacitance from the C-t curve to the generation lifetime and surface generation velocity. If the left hand side of Eqn. (17) is plotted versus $C_{inv}/C_{ox} - 1$, the slope of the linear portion of this plot is inversely proportional to the effective generation lifetime and the intercept to an average surface generation velocity, S. By plotting the transient capacitance as shown in Fig. 4, the bulk generation may be separated from surface generation and diffusion from the bulk silicon. The Zerbst plot is equivalent to a plot of current density, J, versus the depletion layer change, $W-W_f$. The following relations show how current density, J, and $W-W_f$ are related to C(t) data.

$$J = -q\frac{N_a}{2}\frac{\epsilon_{si}}{\epsilon_{ox}}T_{ox}\frac{d}{dt}\left[\frac{C_{ox}}{C(t)}\right]^2 \qquad (19)$$

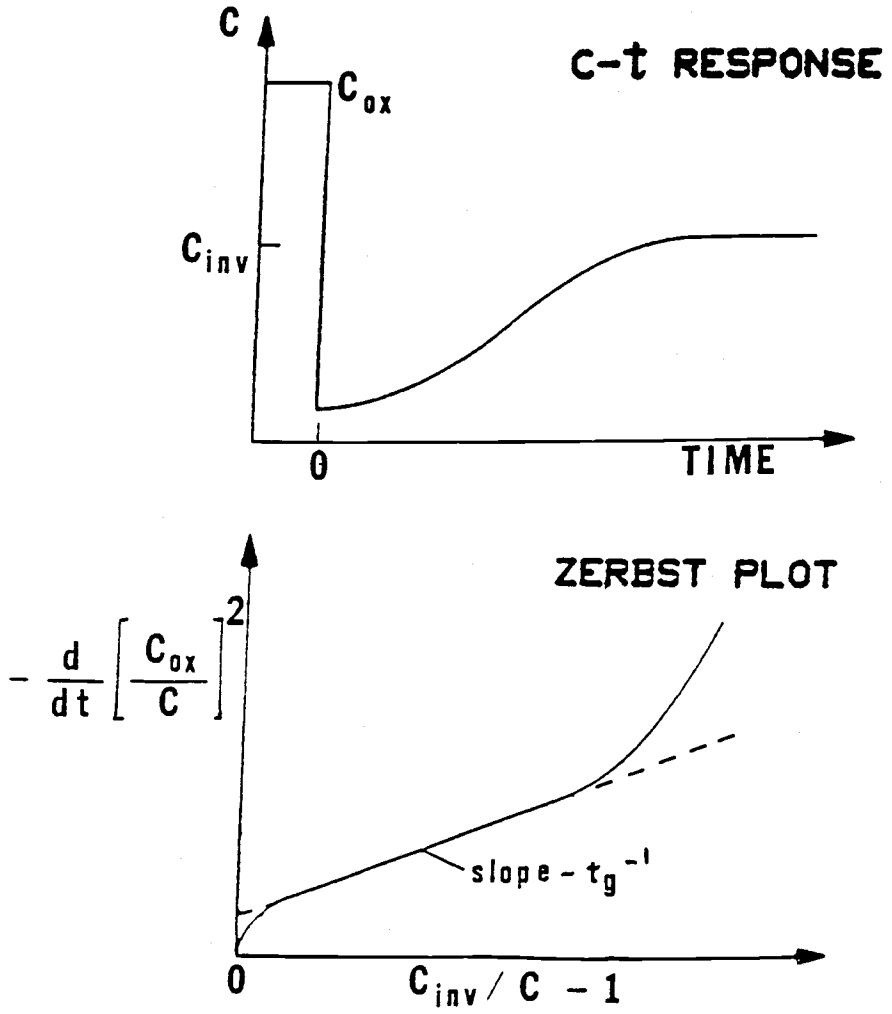


Figure 4 a) Pulsed MOS capacitor C-t response, and b) Zerbst plot. The slope of linear portion of such a plot results in the generation lifetime and its intercept in an average surface generation velocity.

$$W - W_f = (1/C - 1/C_{inv})\ \epsilon_o\ \epsilon_{si}\ \text{Area} \ . \qquad (1)$$

Collins and Churchill [22] have shown that the Eqn. (14) is not completely adequate for the non-equilibrium analysis of pulsed MOS capacitors since the rate is taken from a steady state analysis. Their exact analysis reveals that the generation region is initially

linearly dependent on $W-W_f$, but as time progresses it becomes smaller than W_f. Figure 5 is a plot of J versus $W-W_f$ (or a Zerbst plot) showing that $q\, dN_I/dt$ drops faster as $W-W_f$ reaches zero. Collins's explanation for a more rapid drop of dN_I/dt is that mobile carriers encroach the depletion layer at small $W-W_f$. In general, a Zerbst plot provides an approximate straight line, provided that the doping profile is uniform and the time interval for the measurement capacitance is chosen appropriately for evaluation of time derivative of $d/dt\ (C_{ox}/C(t))^2$.

Although, theoretically it is possible to extract the generation lifetime from the linear portion of the Zerbst plot, even when there is significant bulk diffusion or surface generation, in reality the

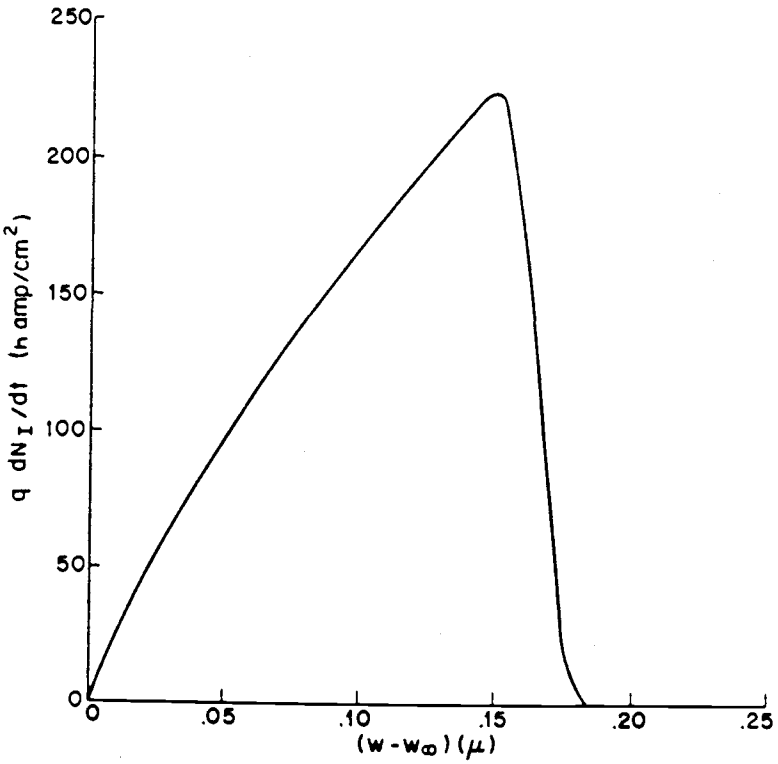


Figure 5) Rate of change of inversion layer density per area versus $W-W_f$, after Collins et al. [22].

interpretation of the transient response in such a case is cumbersome. To examine the bulk generation lifetime of a device, the surface generation S, and especially the lateral surface generation, S_o, must be minimized (Eqn. (18)). It was demonstrated [9] that the surface generation can be reduced significantly if instead of an accumulation-inversion voltage pulse, an inversion to strong inversion voltage pulse is used. With this technique the contribution of interface traps is minimized by pulsing from inversion to stronger inversion since the majority of interface traps have already been filled throughout the pulse, resulting in smaller surface generation contribution.

4. Maximum possible generation - recombination lifetimes in silicon

The recombination and generation lifetimes of pure silicon, completely free of defects and impurities provide an upper limit of what may be expected for a high-quality material. For a perfect silicon sample without any deep levels, the dominant process of recombination is via the radiative process. The maximum possible lifetime for the doping densities used in this study is approximated in the following using the radiative process. The Van Roosbroeck-Shockley relation [23] states that the rate of optical generation of electron - hole pairs equals the radiative recombination rate at equilibrium. Therefore, the emission rate at a frequency v in an interval dv is given by

$$R(v)dv = P(v)\Omega(v)dv \qquad (21)$$

where P(v) is the probability per unit time of absorbing a photon of energy hv, and $\Omega(v)d(v)$ is the density of photons. From Planck's radiation law [24] the density of photons is written as

$$\Omega(v)dv = 8\pi v^2 (n/c)^3 \, [\exp(hv/KT) - 1]^{-1} \, dv \, . \tag{22}$$

The absorption probability is equal to inverse of mean lifetime $\tau(v)$ of the photon

$$P(v) = 1/\tau(v) \tag{23}$$

Since the mean lifetime is proportional to the mean free path, $1/\alpha(v)$, and the velocity of photon Eqn. (21) becomes

$$R(v)d(v) = \frac{8\pi v^2 n^3 \, dv}{c^3 \, [\exp(hv/KT) - 1]} \, . \tag{24}$$

The total number of recombinations per second per unit volume becomes

$$R = \frac{8\pi n^2}{c^2 \, h^3} \int \frac{\alpha \, (hv)^2 \, d(hv)}{\exp \, (hv/KT) - 1} \, . \tag{25}$$

In order to estimate the integral in Eqn. (25) the absorption coefficient as a function of frequency, hv, must be known. Since only an order of magnitude of maximum lifetime is of interest, a simple approach was considered for evaluation of Eqn. (25) as follows. The

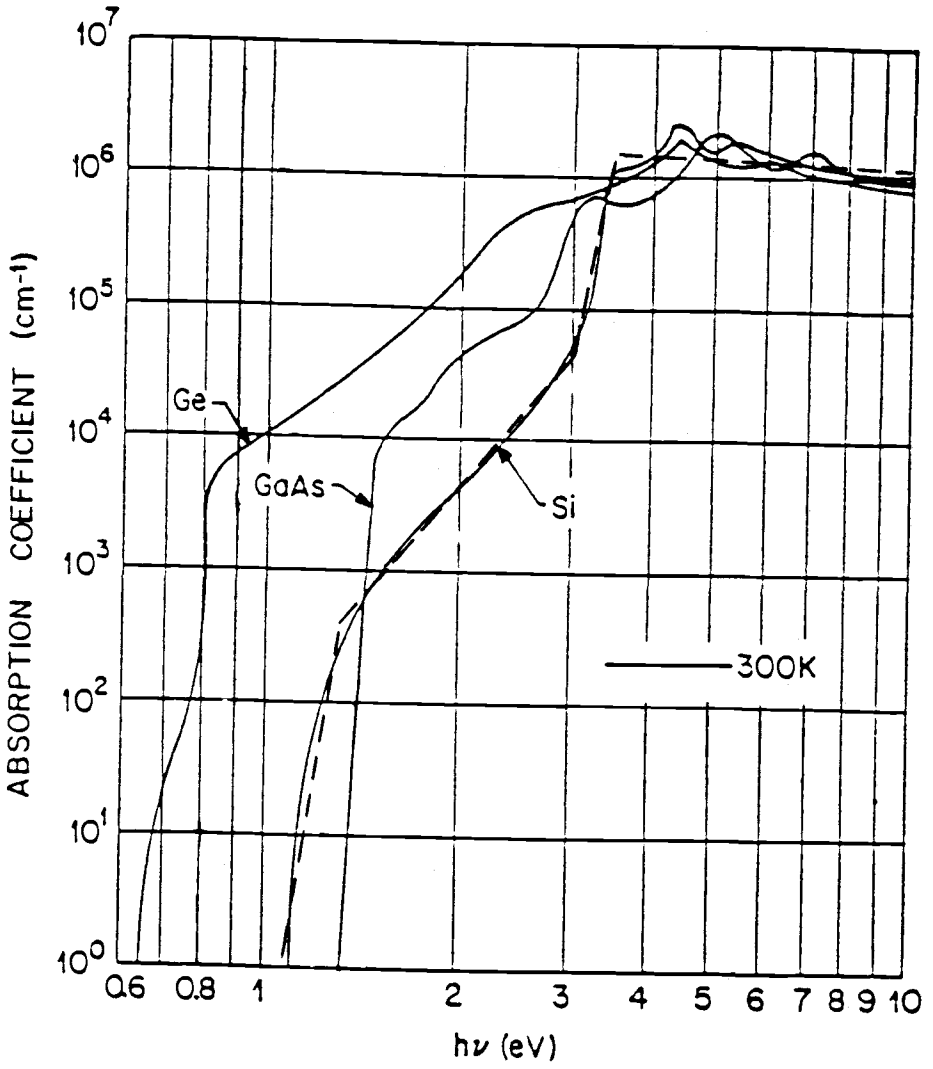


Figure 6) Measured absorption coefficient for pure Ge, Si, and GaAs, after S.M. Sze [25].

absorption coefficient curve of silicon as a function of hv was broken into four sections and a power law fit was made to each section. Figure 6 shows the measured absorption coefficients for pure Ge, Si, GaAs , and the simple power law fit for Si. Using this approximation for $\alpha(hv)$, the integral of Eqn. (25) was evaluated numerically using Simpson's rule in the range of 1.1 eV to 10 eV at room temperature which resulted in a recombination rate of about

$1x10^5$ $cm^{-3}sec^{-1}$. For a doped silicon sample the maximum generation or recombination lifetime is given by [26]

$$\tau = n_i^2/(R\ N_a) \tag{26}$$

where N_a is the doping concentration of the p-type material. For the doping density of the epitaxial layer used in this study ($3\text{-}5x10^{14}$ cm^{-3}), the maximum generation or recombination lifetime was calculated to be about 0.45 sec. A recombination lifetime of 0.47 sec has been calculated by Dumke [27] for spontaneous radiative recombination of silicon for a doping density of 10^{15} cm^{-3}. Blakemore [26] and Pankove [28] have reported 0.3 sec and 0.5 sec lifetimes for ideal silicon with concentrations of 10^{15} cm^{-3}.

This is an extreme upper limit of the lifetime expected in silicon material. In reality, either deep level traps or surface effects would dominate the radiative recombination process resulting in orders of magnitude lower measured lifetimes.

B. RECOMBINATION LIFETIME

Different recombination mechanisms were briefly discussed in Section A. In this section, subsequent to the recombination lifetime measurement techniques, evaluation of the recombination lifetime using the pulsed MOS capacitor technique is discussed.

1. Recombination lifetime measurement techniques

A comprehensive review of recombination mechanisms is given by Bruevich [29] with a summary of electron and hole capture cross sections and recombination lifetime values for various impurities in silicon and germanium. As mentioned in the previous chapter, numerous measurement methods for evaluation of the minority carrier recombination lifetime, τ_r, exist and only a brief introduction is given here.

Table 1. gives a summary of some common minority carrier recombination lifetime measurement techniques. The reverse recovery time method [30,31] and open circuit voltage decay are measured on a p-n junction diode structures. Injection level and trapping problems are fairly common for these techniques. Use of a p-n junction is avoided in this work to eliminate the possibility of introducing another impurity. Optical methods on bulk material such as the photoconductive decay measurement [32,33,34] and the surface photovoltage technique [35] are also common. Proper optical setup of these measurements could be rather involved. In the photoconductivity

Method	test structure	measurement parameter	disadvantage	Ease of measurement	Ref.
Reverse recovery time	pn Junction	I-t	Injection level & trapping problems	simple	[30,31]
Open circuit voltage decay	pn Junction	voltage decay	Injection level & trapping problems	simple	
Photo conduct-ivity decay	bulk material	conductance decay	DC light must saturate traps or injection problems occur	involved optical setup	[32-34]
Surface photo voltage	capacitance probe or Schottky diode	voltage vs. $1/\alpha$	voltage & $1/\alpha$ may have non-linear relation	complex setup	[35]
Pulsed MOS capacit-ance	MOS capacitor	capacitance vs. time	elevated temp., two dimensional effects for short base width device	simple	[1,2]

Table 1. Common minority carrier (recombination) lifetime measurement techniques.

decay technique, the recombination lifetime is evaluated from the conductance decay and in the surface photovoltage technique from a plot of the voltage versus inverse absorption coefficient ($1/\alpha$). Also, the small signal admittance method is used on pn junctions and npn or pnp transistors for measurement of the diffusion length and surface recombination velocity in thin semiconductor layers [36,37]. In addition, from the amplitude or phase shift of the frequency dependence of the short-circuit photocurrent in a Schottky contact illuminated by intensity modulated light, the minority carrier lifetime, τ_r, can be evaluated [38]. Recently, Spirito et al. [39]

reported a conductivity modulated technique using a lateral pn diode for τ_r profile in thin epitaxial layers. Schroder et al. [1,40] showed that τ_r and hence the diffusion length of a material may also be obtained by using a pulsed MOS capacitor at elevated temperatures. Even though an MOS capacitor is a zero injection device and the recombination of carriers does not take place, τ_r is evaluated from the quasi-neutral bulk generation (or diffusion process) which is the dominant mechanism in C-t relaxation to equilibrium at elevated temperatures. This technique uses the same MOS capacitor test structure and measurement setup as the Zerbst method for the τ_g measurements; therefore, it is more favorable than other techniques. The pair of τ_g and τ_r values using the pulsed MOS capacitor technique reflect both near surface (active region of a device) and bulk properties of the substrate using a common test structure and processing conditions. The recombination lifetime measurement technique used in this study [2] is a modification of Schroder's method accounting for lateral and time dependent quasi-neutral bulk generation for short base width devices. This is discussed in detail in next section and Chapter III.

2. Recombination lifetime of short base width devices using pulsed MOS capacitance technique

Schroder et al. [1] have described a simple method for the determination of the recombination lifetime on silicon wafers using the pulsed MOS capacitor technique at elevated temperatures of 70 - 100°C. This technique relies on the dominance of the quasi-neutral

bulk generation (diffusion current) over the space charge and surface generation at elevated temperatures. The technique is summarized in the following. The rate of change of the space charge region W(t) in a pulsed MOS capacitor is given by [3,8]

$$\left(1 + \frac{\epsilon_{ox}}{\epsilon_{si}} \frac{W}{T_{ox}}\right) \frac{dW}{dt} + \frac{1}{N_a} \frac{J_{total}}{q} = 0 \qquad (27)$$

where ϵ_{ox} and ϵ_s are the dielectric constants of oxide and silicon, T_{ox} the oxide thickness, N_a the doping concentration, J_{total} the total current density contributing to relaxation to equilibrium and q the charge of the electron. At elevated temperatures, quasi-neutral bulk generation dominates and a plot of $1-(C_i/C)^2$ versus time gives a straight line [1]. The slope is related to an effective diffusion length, L'_n as follows:

$$\text{Slope} = \left(\frac{C_i}{C_{ox}}\right)^2 \left(\frac{n_i}{N_a}\right)^2 \left(\frac{\epsilon_{ox}}{\epsilon_s}\right) \left(\frac{2}{T_{ox}}\right) \left(\frac{D_n}{L'_n}\right) \qquad (28)$$

where C_i is the initial capacitance after the pulse, C_{ox} the oxide capacitance, n_i the intrinsic concentration at the measured temperature, D_n the diffusion coefficient of electrons (for a p-type sample).

Schroder's technique of measuring the recombination lifetime works well for long base width polished silicon wafers, but it is shown below that it is not appropriate for short base width samples. This simple approach ignores the lateral quasi-bulk generation and

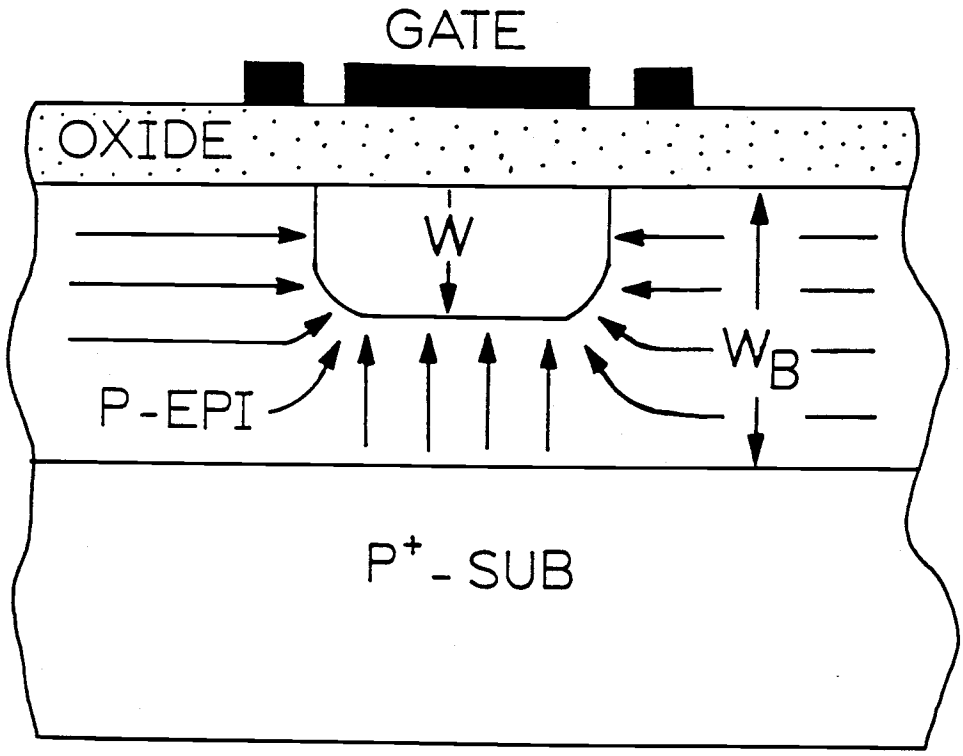


Figure 7) Cross sectional view of a short base width device on p/p^+ epitaxial silicon. Minority carriers generated within one diffusion length in the lateral direction contribute to the lateral current. The volume of generated carriers (diffusion) under the gate is limited by the epitaxial thickness and the time-dependent space charge width.

the time dependence of the width of the space charge region, W, in short base width devices. Consequently, calculations using Schroder's technique indicate that the recombination lifetime is a function of device diameter. These factors are discussed in the following. Figure 7 shows a cross section of a p/p^+ epitaxial wafer where a quasi- neutral bulk current is generated one diffusion length away from the depletion region edge laterally and is bounded by the epitaxial thickness. This can be best treated as a two dimensional problem (assuming a uniform recombination lifetime within the

epitaxial layer). However, for simplicity a simple one dimensional approach is proposed [2] in which bulk generation in the lateral area of the device is taken into consideration, resulting in a fairly uniform recombination lifetime which is independent of device diameter for short base width devices.

The total current contributing in the relaxation to equilibrium due to quasi-neutral bulk generation in a pulsed MOS capacitor at elevated temperatures for a circular structure may be approximated by simply adding the lateral current contribution to the bulk diffusion under the space charge region as follows

$$J_{total} = q \frac{n_i^2}{N_a} \frac{D_n}{L_n^2} (W_B - W(t)) + q \frac{n_i^2}{N_a} \frac{D_n}{L_n} \left(\frac{4W_B}{D} \right) \qquad (29)$$

where W_B is the base width, $W(t)$ the depletion region width, L_n the diffusion length, and D the device diameter. The first term in Eqn. (29) is due to the carriers generated within region defined by the epitaxial layer - p^+ substrate interface and the edge of the depletion region under the gate electrode. The time dependence of W is included in this term. The second term is the lateral current component, assuming the minority carriers are generated within one diffusion length, L_n, from the space charge edge. The term $(4W_B/D)$ in Eqn. (29) is a correction factor which results from normalizing the cross sectional area of the lateral current component $(\pi D W_B)$ to the area of the MOS capacitor $(\pi D^2/4)$. Due to the p/p^+ built in field, generation at this interface is normally assumed to be zero $(S_{pp+}=0)$. If this generation is not zero, the recombination lifetime assuming

no interface generation would be an apparent rather than the actual lifetime of the epi-layer. Substituting Eqn. (29) in (27) results in the following solution:

$$\left(1 - \frac{\epsilon_{ox}}{\epsilon_{si}} \frac{U}{T_{ox}}\right) \ln \left| \frac{W_i+U}{W+U} \right| + \frac{\epsilon_{ox}}{\epsilon_{si}} \frac{1}{T_{ox}} (W_i - W) = -\left(\frac{n_i^2}{N_a^2} \frac{D_n}{L_n^2} \right) t \qquad (30)$$

$$U \equiv - W_B \left(1 + 4L_n/D\right)$$

where W_i is the initial depletion width just after the voltage pulse. Equation (30) predicts how the depletion region changes as a function of time for a pulsed MOS capacitor at elevated temperatures. All of the parameters in Eqn. (30) can be measured or evaluated for a given transient response except for the diffusion length, L_n, which may be found as follows. The initial and final values of the space charge region W_i and W_f and the final time (retention time), t_f, may be obtained from the C-t response at elevated temperatures. All of the capacitor parameters in addition to its diameter size are substituted into Eqn. (30) and L_n is solved iteratively using the secant method. Subsequently, the value of L_n is substituted back in Eqn. (30) and a W-t plot is created for comparison to the measured transient response (this is shown later).

C. DEEP LEVEL TRANSIENT SPECTROSCOPY (DLTS)

1. Capacitance DLTS

Capacitance DLTS is a transient technique [41] which is used for characterization of trap levels present within the bandgap of a semiconductor. The capacitance transient technique [42] is widely used for studying traps within the depletion region of a Schottky barrier or a p-n junction diode. Analysis using DLTS normally allows evaluation of the trap type (majority or minority carrier trap), energy level of the trap, capture cross section, and the trap concentration.

Figure 8 [41] is a schematic diagram of emission and capture process of an arbitrary trap in a p-type material. e1, e2, c1, and c2 represent the electron emission, hole emission, electron capture, and hole capture, respectively. The observable traps are within the depletion region of the device. Therefore, the capture rates are zero since the concentration of electrons and holes are essentially zero in the space charge region. Thus, all deep levels act as traps since the recombination cannot occur in the space charge region. The simple rate equation of Fig. 8 shows the transient to be an exponential function of time with a rate e1 + e2. Normally one of these rates dominates. An increase in trapped minority carriers causes an increase in the junction capacitance. Thus, the capacitance transient is always positive for a minority carrier trap while a majority carrier has a negative capacitance transient. Figure 9 [41]

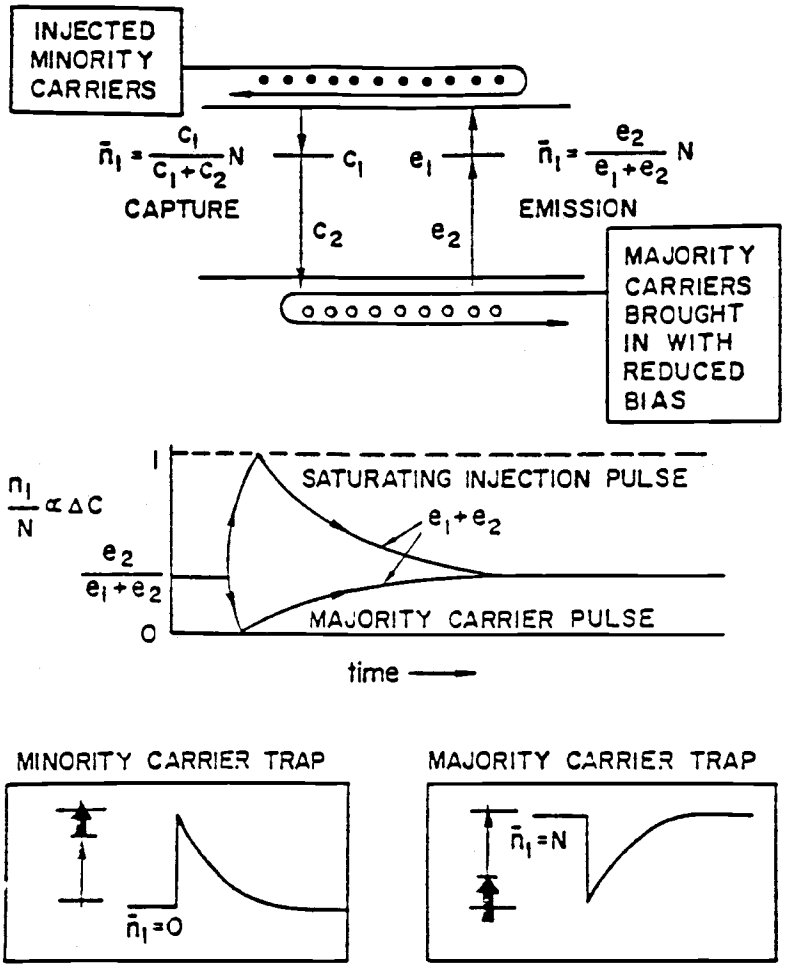


Figure 8) Schematic diagram of the emission and capture processes for an arbitrary trap level and capacitance transient of that trap in p-type material, after D.V. Lang [41].

indicates the capacitance transient and the required pulse sequence used to produce the transient for a) minority carrier trap and b) majority carrier trap in a $p\text{-}n^+$ junction diode. Band bending is omitted for simplicity.

The emission rate of a trapped carrier is determined by Boltzmann statistics and is an exponential function of the lattice temperature, T, and the activation energy of the trap, E_A. For a p-

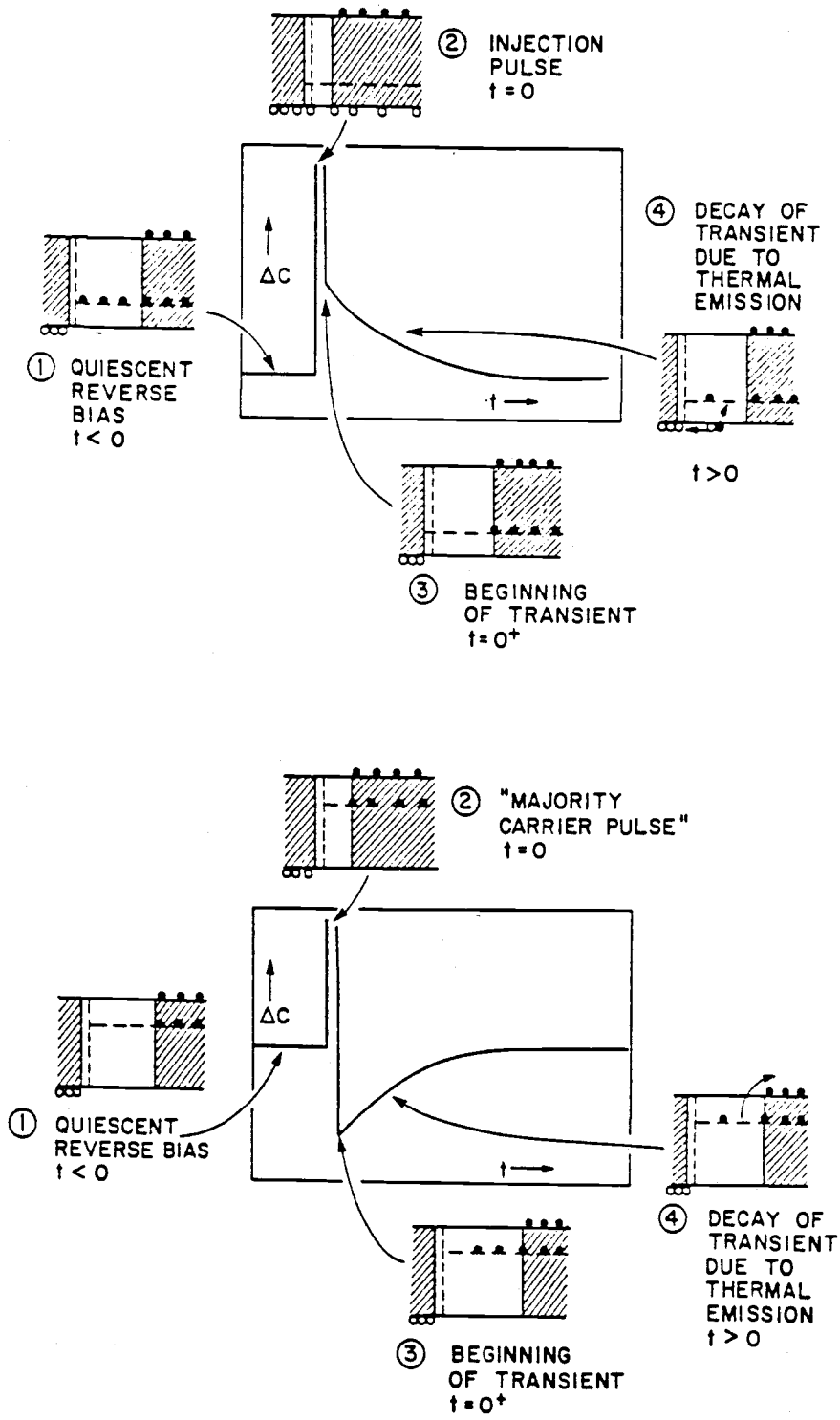


Figure 9) Pulse sequence which is used to produce a capacitance transient for a) minority carrier trap and b) majority carrier trap for a pn^+ junction, after D.V. Lang [43].

type semiconductor the emission rate of electrons, e_n, is given by

$$e_n = \sigma_n \, v_{th} \, N_C \exp(-E_A/KT) \tag{31}$$

where K is Boltzmann constant, σ_n is the minority carrier capture cross section, v_{th} is the electron thermal velocity, and N_C is the effective density of states of conduction band. The activation energy is defined as the energy separation between the conduction band and the trap level. A semi-log plot of emission rate normalized by the square of the temperature (T^2) versus the inverse temperature is called an Arrhenius plot. The product of thermal velocity and density of states in Eqn. (31) has a temperature dependence of T^2. Therefore, the reason for division by T^2 in the Arrhenius plots is to remove the temperature dependence of the pre-exponential term. In general, the capture cross section is also temperature dependent, and care must be taken when interpreting deep levels with known strong temperature dependent of capture cross sections. The activation energy of a trap can be obtained from the slope of this plot and the capture cross section from the intercept.

One way of analyzing the exponential capacitance is by sampling the transient at two different times t_1, and t_2 using a boxcar integrator. Figure 10 [41] shows a schematic diagram of how a capacitance transient changes as a function of temperature. On the right hand side of Fig. 10, the DLTS signal is shown which results from the difference between the capacitance at time t_1 and t_2 plotted as a function of temperature. The normalized capacitance change (normalized DLTS signal) at these two times is written as

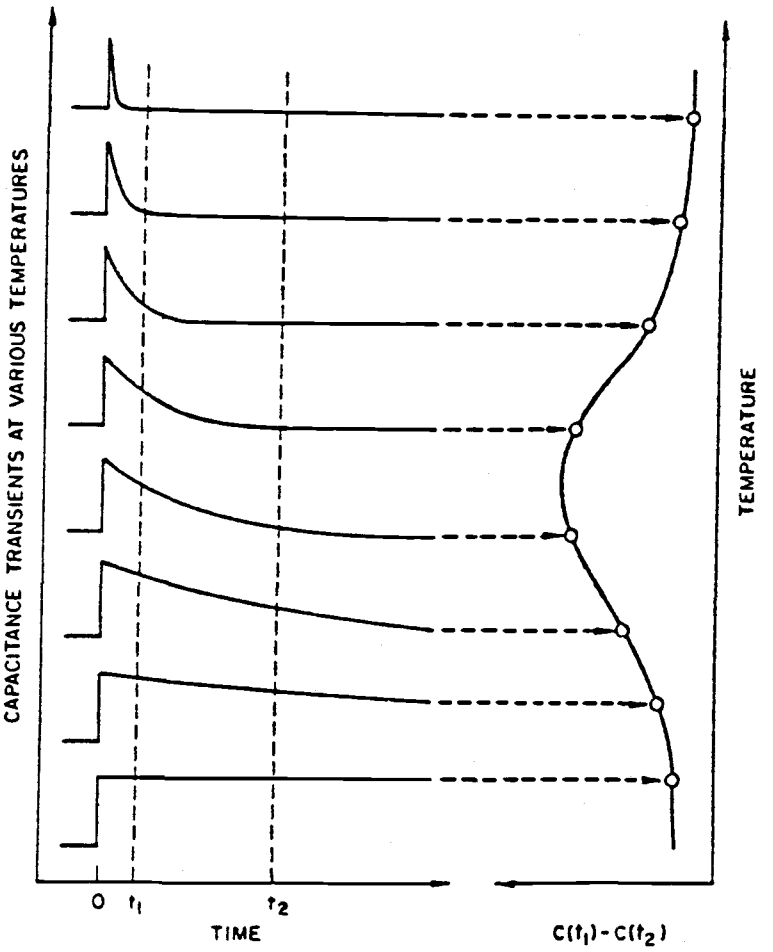


Figure 10) Schematic representation of various temperatures which corresponds to DLTS signal resulting from a double boxcar integrator, after D.V. Lang [41].

$$S(T) = (C(t_1) - C(t_2))/\delta C(0) \tag{32}$$

Where $\delta C(0)$ is the capacitance change at the beginning of the pulse (t=0). For an exponential transient the normalized DLTS signal is $S(T) = \exp(-t_1/\tau_c) - \exp(-t_2/\tau_c)$. Where, τ_c, is the time constant defined as the reciprocal of the sum of electron and hole emission rates ($\tau_c = 1/(e_n+e_p)$. The maximum DLTS signal occurs when

$$\tau_c = (t_2 - t_1)/\ln(t_2/t_1) \ . \qquad (33)$$

Since t_1 and t_2 are fixed parameters of the boxcar integrator the corresponding τ_c or emission rate can be calculated using Eqn. (33). The emission rate corresponding to a peak in the DLTS thermal scan gives a single data point for an Arrhenius plot. The thermal DLTS scan must be repeated for various time constants, τ_c, or the emission rates by changing the rate window t_1 and t_2 to be able to obtain an adequate Arrhenius plot for evaluation of the activation energy and capture cross section of a trap. The concentration of a trap is related to the DLTS peak height by the following relation [41]

$$N_T = 2 \frac{\delta C(0)}{C} N_A = 2 \frac{C(t_1) - C(t_2)}{S_{max}(T)\ C} N_A \ . \qquad (34)$$

The above scheme is valid only for exponential transients. The non-exponential case [44] is common for intermediate and shallow traps. Therefore, care must be taken when analyzing this situation. Kirchner et al. [45] have considered non-exponential and multi-exponential transient decays.

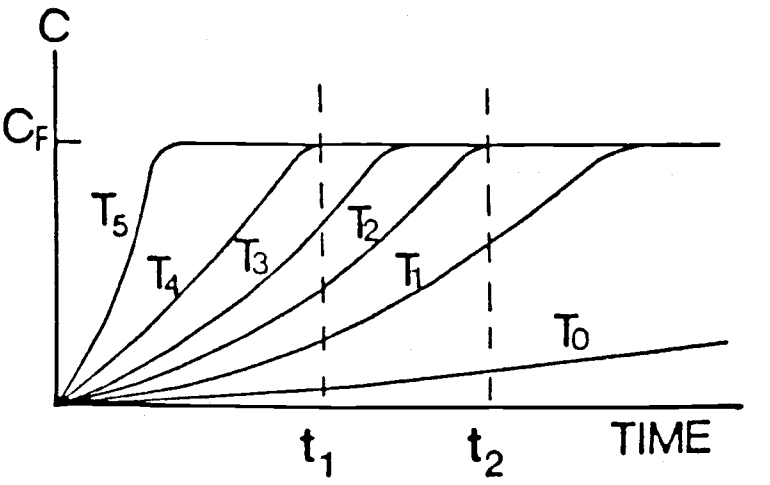
One of the main limitations of using a Schottky barrier instead of a p-n junction for DLTS measurement is that minority carrier traps cannot be detected. This is because the forward current in Schottky barriers is dominated by the majority carrier rather than minority carriers. However, "optical DLTS" may be used to introduce the necessary minority carriers into the sample by using an appropriate

wavelength light in one of two ways. One approach is by using a majority carrier pulse while at the same time illuminating the sample. The other is by pulsing the light on and off and having a fixed reverse bias applied to the diode. This introduces the required minority carriers for detection of a minority carrier trap in a Schottky barrier, provided that the appropriate wavelength and intensity are applied. The size of the capacitance transient in this case depends on the concentration of the trap as well as light intensity and wavelength.

2. Application of DLTS to MOS relaxation transients

N. Pearce et al. [46] have shown an application of DLTS to MOS capacitor relaxation transients, discussed briefly in the following. This technique is useful for direct evaluation of the temperature dependence of the diffusion process from the bulk of the sample or the generation within the depletion region for retention times within the capability of the DLTS setup.

Figure 11 is a schematic representation of C-t response of a MOS capacitor at various temperatures. Application of a double boxcar integrator with a window of t_1 and t_2 results in a capacitive signal similar to a DLTS scan of a p-n junction or a Schottky diode. The peak occurs at a retention time t_2 as shown in Fig. 11b. Since any increase in temperature beyond T_2 increases the value of C at t_1, while the capacitance at t_2 remains at the same inversion value. The peak temperature T_2 and the retention time t_2 define a set of points directly measured from the experimental C-t. By changing the DLTS



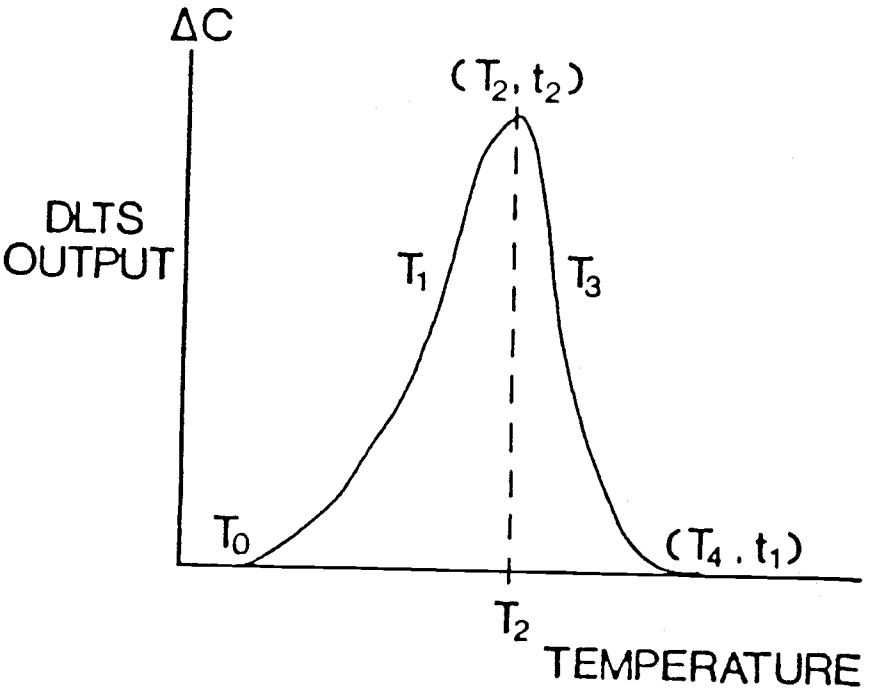


Figure 11 a) Schematic representation of C-t response of a MOS capacitor at various temperatures, b) Illustrates how an application of DLTS results in a peak corresponding to retention time, t2 at the peak temperature, after Pearce [46].

rate window (t_1 and t_2), the temperature dependence of the retention time can be obtained. The shape of the transient is not exponential, however this is immaterial for the temperature dependence of the retention time. The DLTS peak occurs when the emission rate from the deep level coincides with the rate window (Eqn. (33)). The distinction between generation within the depletion region or bulk generation (diffusion process from the bulk) can be easily observed by examining the activation energy. The temperature dependence of the dominant process causing the return to equilibrium can be evaluated from an Arrhenius plot.

D. EFFECTS OF INTRINSIC GETTERING ON LIFETIMES

In Czochralski grown silicon wafers point defects, vacancies, interstitial oxygen and other impurities play an important role in establishing the electrical characteristics of the material. During crystal pulling of the silicon ingot, oxygen atoms are incorporated in the Si melt from the quartz crucible. Oxygen concentrations normally vary between 25 and 33 ppma (of the order of 10^{18} atoms/cm^3). Oxygen atoms precipitate [47,48] during the cool down period of the crystal growth or during annealing after the crystal growth. These oxide precipitates, when properly controlled, act as nuclei of secondary defects, and can getter heavy metal impurities away from the active region of devices reducing or eliminating their electrical activity around the active region.

Gettering techniques attempt to control unintentional contamination introduced into silicon wafers prior to or during wafer processing. Transition metals which may have been introduced during initial wafer preparation or during device fabrication, are the main contaminations which degrade the lifetime in silicon. Clearly, the ideal method of improving the material characteristics is to completely eliminate the contaminations. However, at present elimination of all types of unintentional contamination and impurities is not possible during crystal growth or device fabrication. An alterative way is using gettering schemes to reduce or remove transition metal contamination during silicon wafer processing. Some gettering techniques are as follows:

1) Extrinsic gettering - dislocation generation at the wafer's back surface.
2) Chlorine oxidation - Chemical reaction at the wafer's front surface.
3) Intrinsic gettering - micro-defect formation in the wafer by controlled oxygen precipitation.

Beneficial effects of intrinsic gettering such as lifetime improvement have been established in the literature. Proper intrinsic gettering [49] using oxygen precipitation can getter metallic impurities from the active layer of the device leaving a denuded zone at the surface [50,51], improving generation lifetimes. In addition to intrinsic gettering by oxygen precipitation, an intrinsic gettering process involving silicon point defects has also been reported [52]. Udea et al. [53] identified Cu and Ni as the predominant metal components of the precipitates involved in silicon point defects in both Czochralski and oxygen free float zone wafers. In the present study, only intrinsic gettering by oxygen precipitation is considered.

The fundamental heat treatment steps of intrinsic gettering are shown in Fig. 12 [54]. A high temperature heat treatment prior to nucleation is typically used in bulk wafers to out-diffuse the oxygen from the surface and form a denuded zone in the active region of a device. Subsequent to high temperature denudation, a nucleation process of oxygen normally performed at temperatures ranging from 650 to 800°C results in formation of oxygen precipitates. Subsequent to nucleation, high temperature annealing is used to grow oxygen precipitates. While some of the precipitates grow, the rest dissolve

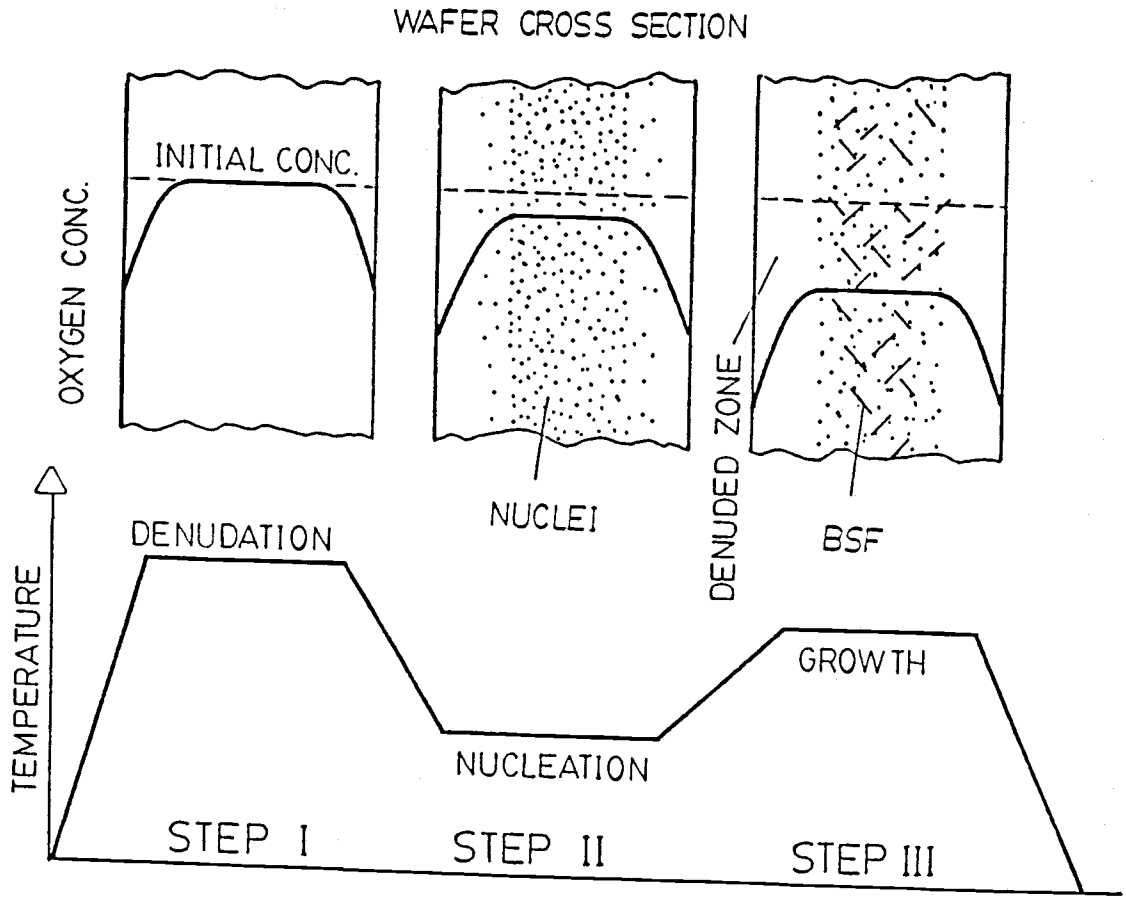


Figure 12) The fundamental heat treatment steps for intrinsic gettering, after W. Wijaranakula [54].

or shrink [55]. There seems to be a threshold value [56] of approximately $6x10^{17}$ cm^{-3} of oxygen below which precipitation may not occur. The oxygen precipitates induce bulk stacking faults (BSF) and dislocation loops which act as gettering centers for transition metals, removing them from the active region.

In spite of the fact that proper gettering can improve the lifetime, too much precipitation causes an insufficient denuded zone depth as well as excessive surface defects, volume mismatch, and warpage due to growth of oxygen in the bulk silicon. Experimental

evidence [57-61] indicates that in p-type bulk silicon wafers, oxide precipitates cause degradation of the lifetime when located within the active region of the device. Hwang et al. [62] studied oxygen precipitation in n-type and p-type bulk silicon wafers and modeled the oxide precipitates as having an interface with the bulk silicon similar to the $Si-SiO_2$ interface. The recombination lifetime of p-type bulk silicon was more severely degraded due to oxygen precipitation than n-type silicon. To explain the lifetime difference between p-type and n-type silicon, Hwang et al. have postulated the existence of positive charge due to oxide precipitates similar to the $Si-SiO_2$ interface causing an attractive field for electrons in p-type material and a repulsive field for holes in n-type silicon.

Intrinsic gettering schemes are an effective means of reducing contamination effects in MOS devices. However, many problems arise in conjunction with the implication of these schemes. At lower temperatures preferred in MOS device processing for shallower junctions, the oxygen is removed from the interstitial supersaturated solid solution as clusters of SiO_x so it behaves differently from the precipitates observed at higher temperatures [63]. Therefore, the role of SiO_x as nucleation sites is not as effective at lower temperatures. Furthermore, little denudation of oxygen occurs at lower temperatures.

As far as the electrical properties of a material is concerned, it is well known that proper intrinsic gettering on bulk (polished) wafers improves the lifetimes. Pre-epitaxial intrinsic gettering of the substrate of epitaxial material results in an improvement of generation lifetime of epitaxial layers [64-66]. However, there are

no reports on recombination lifetime of p/p^+ epitaxial layers. The question that needs to be answered is whether the limitation on recombination lifetime is due to the epitaxial layer itself, the substrate material or the epi-layer - substrate interface. The emphasis in this study is on characterization of τ_r on p/p^+ wafers and identification of sources which limit τ_r. Attempts are made to overcome these limitation to improve material quality.

III. EXPERIMENTAL SETUP AND PROCEDURE

In this chapter the fabrication of MOS capacitors and Schottky diodes as well as the lifetime and DLTS setup and procedure are described. Fabrication of MOS capacitors, the C-V characterization setup, and measurement of generation and recombination lifetimes are discussed in the first section. This is followed by description of the fabrication of Schottky diodes, the DLTS setup and measurement technique of majority carrier and optical DLTS.

Various wafers were fabricated to form MOS capacitors and Schottky diodes to better quantifying the p/p^+ epitaxial wafers and the factors dominating the recombination and generation lifetimes on these wafers. The properties of the wafers used in this study are summarized in table 2. Control wafers which were 13.5 μm p/p^+ epitaxial wafers were always processed along with other samples for control purposes. Electrical characterization of these wafers were routinely examined to insure identical process conditions. When control wafers in a given process exhibited poor electrical properties, such as abnormal C-V characteristics, excessive surface generation, or degraded lifetimes all of the wafers in that process were eliminated from further considerations. To better quantify the epi-layer - substrate interface, various epi-layer thickness wafers were prepared and tested for lifetime and DLTS measurements. The 6 μm epitaxial sample listed in Table 2 was prepared by using the planar etch on 13.5 μm epi-layers (planar etch is described later). 110 μm epi-layers were grown by vendor A under identical growth conditions as 13.5 μm layers. All of the wafers in this study were provided by

Sample	Epitaxial thickness (μm)	Res. (Ω-cm)	Remarks
p/p+	110	35	epitaxial layer deposited by CVD
p/p+	70	35	final polished from 110 μm
Control p/p+	13.5	35	epitaxial layer deposited by CVD
p/p+	5-6	35	planar etched from 13.5 μm
p/p+ from Vendor B	12	20	
p-type polished	---		
Float zone	---	100-150	
n/n+	50	40	
p-type epi Fe implanted	5	40	Fe was implanted at 10^{11} cm^{-2} dose prior to epi deposition

Table 2. A summary of main wafers used for lifetime and DLTS study. All wafers were provided by vendor A unless specified.

vendor A except for a set of p/p^+ wafers from vendor B for comparison. For obtaining an intermediate epi-layer thickness (between 13.5 and 110 μm) some of the 110 μm layers were exposed to the planar etch for long times, resulting in high surface generation velocities and high surface currents on MOS capacitors which were not appropriate for lifetime measurements. The 70 μm thick epi-layers were obtained after several final polishes of the 110 μm epi-layers (this was done by vendor A). For comparison of recombination lifetime values, wafers with 50 μm n/n^+, p/p^+ from a different vendor, and p-type Float zone wafers were also fabricated and tested. Also a set of

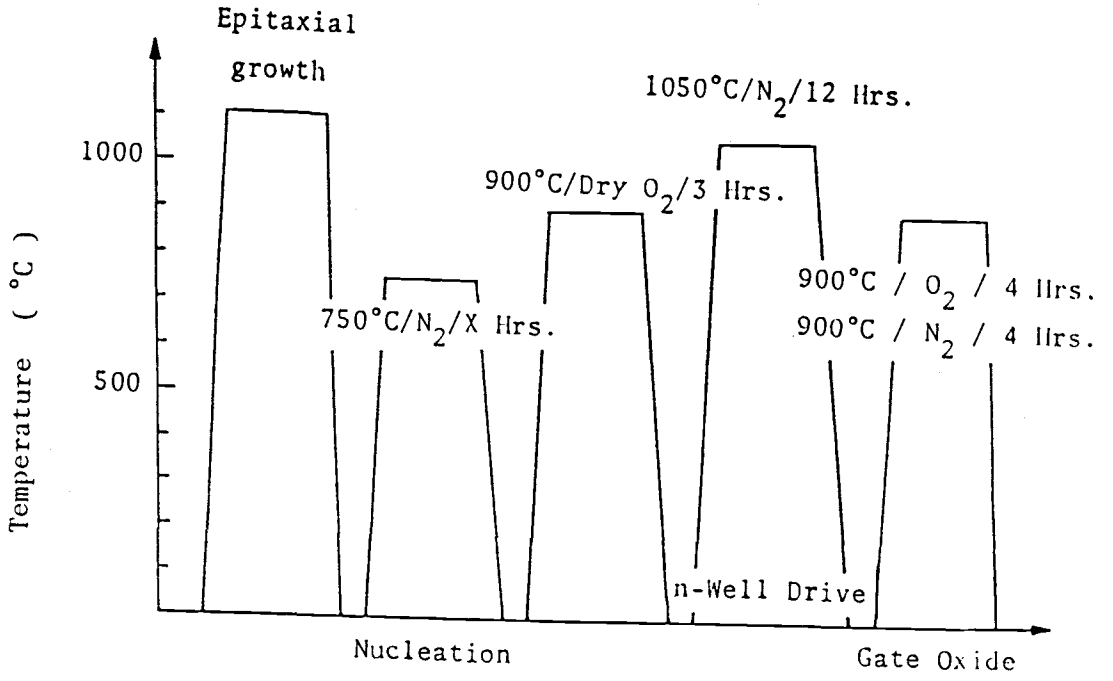


Figure 13a) Post- epitaxial nucleation and CMOS simulation heat treatment steps.

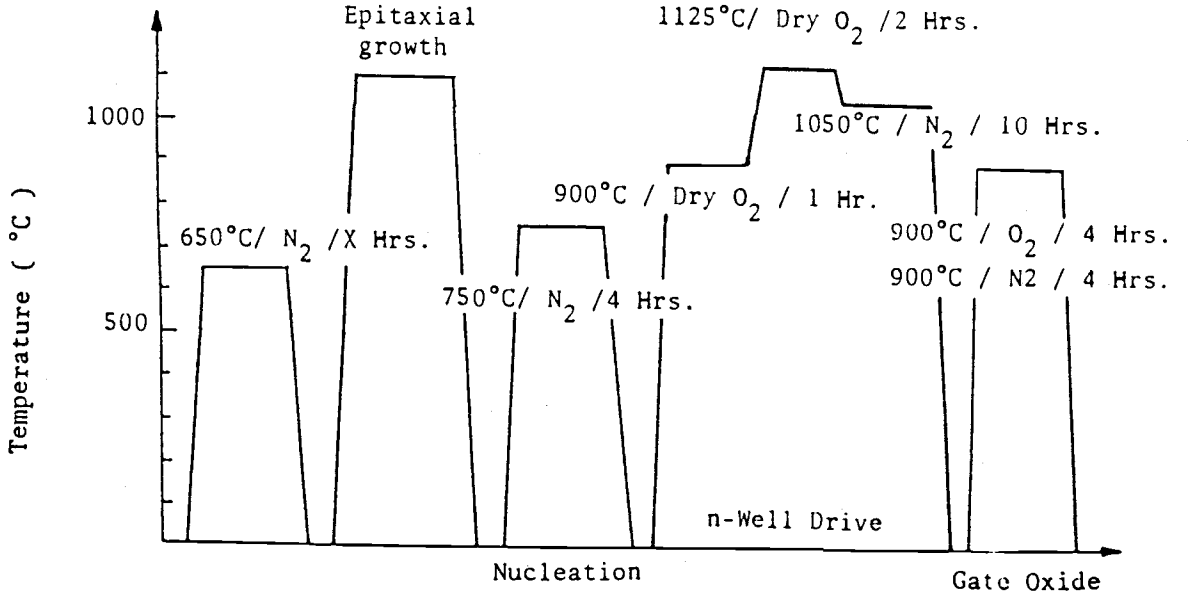


Figure 13b) Pre-epitaxial anneal, post-epitaxial nucleation and CMOS simulation heat treatment steps.

samples with p-type substrates which had Fe implanted at a dose of 10^{11} cm^{-2} and an energy of 10 keV prior to 5 μm of epitaxial deposition were fabricated and their DLTS and lifetimes characteristics are discussed in the next chapter for illustrative purpose. Figures 13a and 13b show two heat treatment cycles used for evaluation of the effect of intrinsic gettering and a CMOS simulation heat treatment on generation and recombination lifetimes.

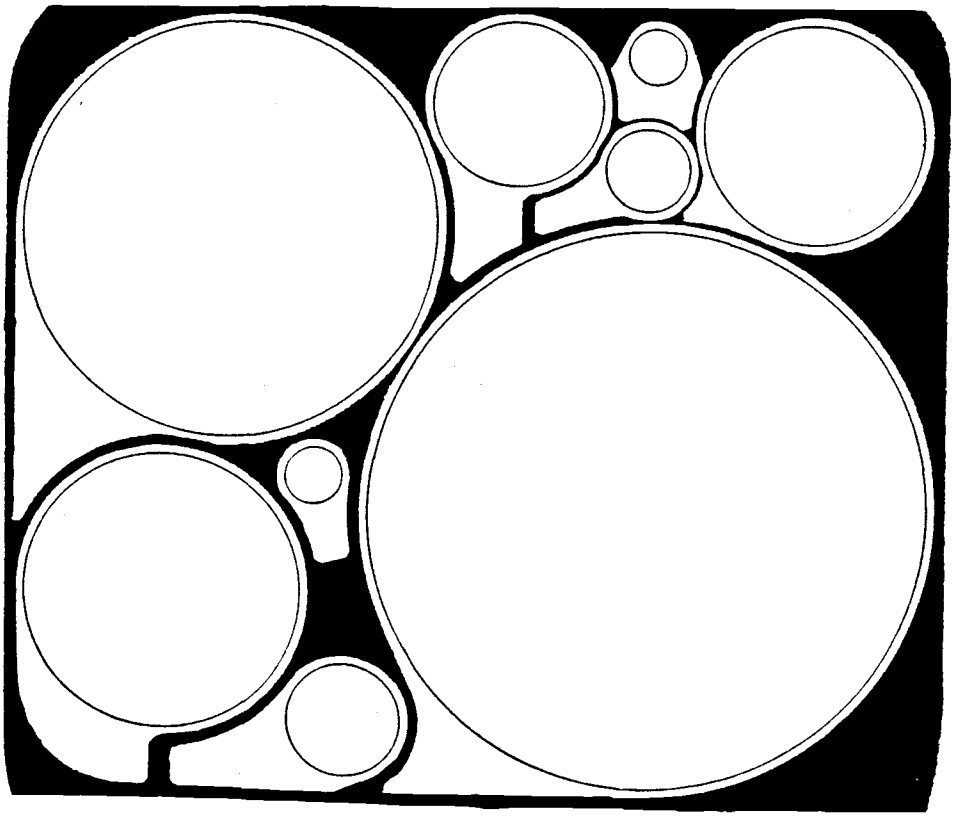


Figure 14) Mask set with various size diameters ranging from 123 μm to 1272 μm with a 6 μm guard ring spacing, after W.I. Sze [68].

A. MOS CAPACITORS AND LIFETIME SETUP

1. Fabrication of MOS capacitors

A large number of MOS capacitors were fabricated for studying generation and recombination lifetimes. p/p^+ epitaxial silicon wafers with, 30-50 Ω-cm / 0.01-0.02 Ω-cm, (100) orientation, 100 mm in diameter were the main material used in this study. In order to make a direct comparison, the 100 mm wafers were cleaved into four quarters and some of the quarters were used for fabrication of MOS capacitors for generation and recombination lifetime measurements while the remaining ones were used to form Schottky diodes for DLTS study.

It is common to etch the native oxide from a silicon wafer with

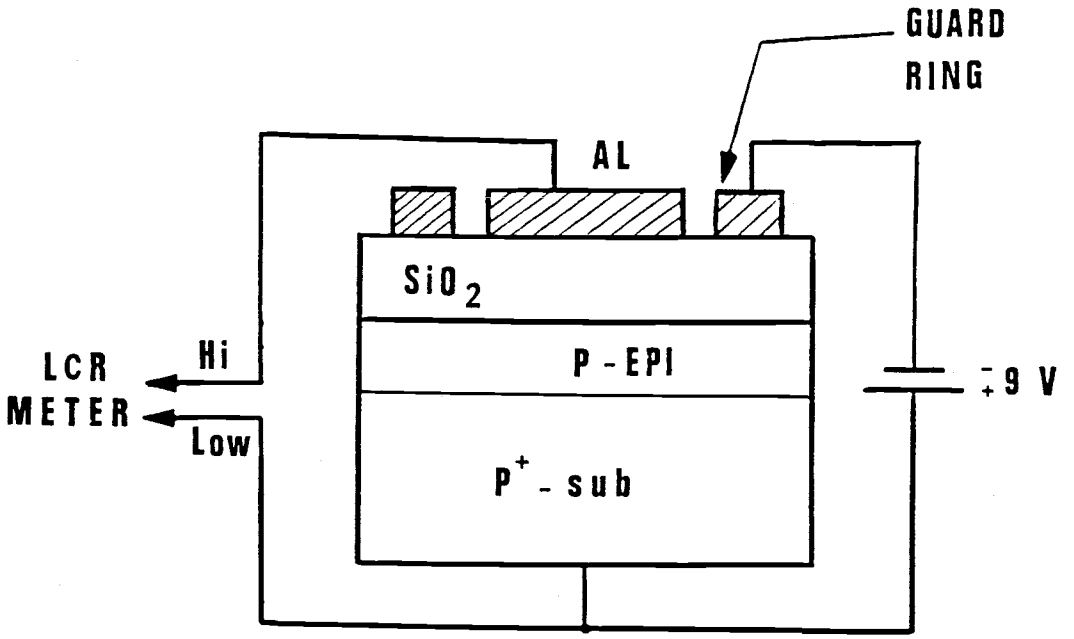


Figure 15) Cross section of an MOS test capacitor.

a 10% HF etch (hydrofluoric acid) prior to thermal oxidation. Chemical cleaning of the sample such as RCA cleaning would leave residual impurities on the sample causing various problems in the lifetime test. Experience has shown a slightly better quality $Si-SiO_2$ interface for MOS capacitors were obtained when no chemical treatment was applied prior to oxide growth. All of the thermal oxides were grown at 900°C in dry O_2 for about 4 hours to give oxide thicknesses in the range of 400 to 500 Angstrom. The oxidation was followed by an anneal at 900°C in N_2 to reduce the fixed charge density in the oxide (see the "Deal triangle" [21]). Aluminum was evaporated on the wafers using a conventional evaporator. Various size circular dots ranging from about 120 μm to 1270 μm in diameter with a guard ring structure were fabricated using photolithography. Figure 14 shows the mask set

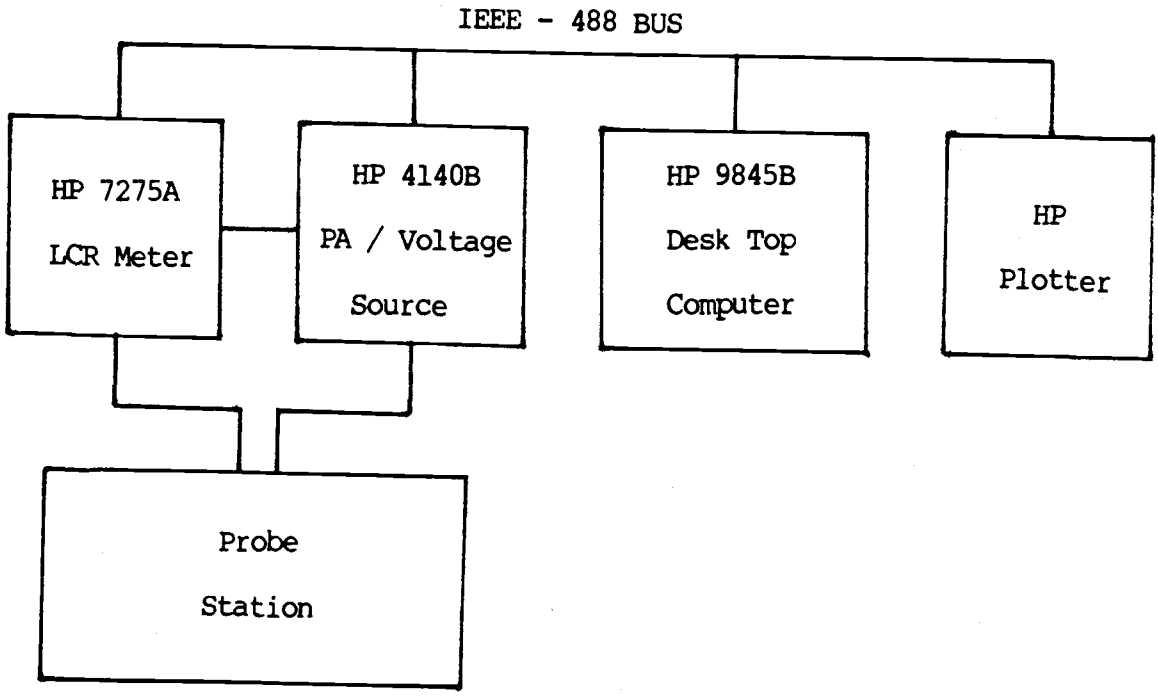


Figure 16) Block diagram of the device characterization setup.

used. The gap spacing between the guard ring and the aluminum dot was 6 μm. Subsequent to metallization, the samples were normally annealed in forming gas (10% H_2, 90% N_2) for 20 to 60 minutes at 400 to 425°C to minimize the Si-SiO_2 interface state densities [21]. The guard rings were biased to an accumulation voltage (-9V for p-type and +9V for n-type) with a battery to prevent surface charge from altering the C-V, and generation and recombination lifetime measurements. A cross sectional view of an MOS capacitor on a p/p^+ wafer is shown in Fig. 15.

2. C-V characterization setup and procedure

High frequency and quasi-static capacitance vs. voltage (C-V) measurements were performed by using an HP4275A LCR meter and an HP 4140B pico-amp meter, respectively. Data acquisition was performed with an HP 9845B mini-computer, which was interfaced to the measurement instruments via an IEEE-488 bus. A block diagram of the characterization setup is shown in Fig. 16. The probe station was placed in a black metal box which was maintained at ground potential and shielded from light to be able to measure quasi static C-V.

A typical C-V plot of a p/p^+ wafer is shown in Fig. 17. The high frequency C-V, measured at 1MHz is initially swept from accumulation to deep depletion (curve 1) and then after an application of light which causes inversion the bias was swept from inversion back to accumulation (curve 2). A computer program was written to calculate all the relevant data such as the oxide thickness, doping concentration, flatband and threshold voltages from the C-V data as

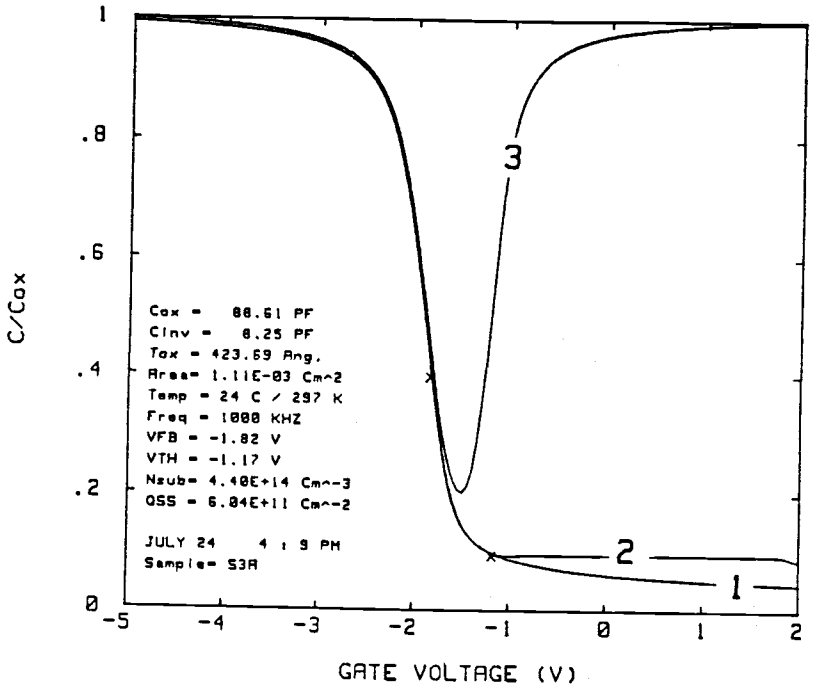


Figure 17) Typical C-V characteristics of a p-type MOS capacitor.

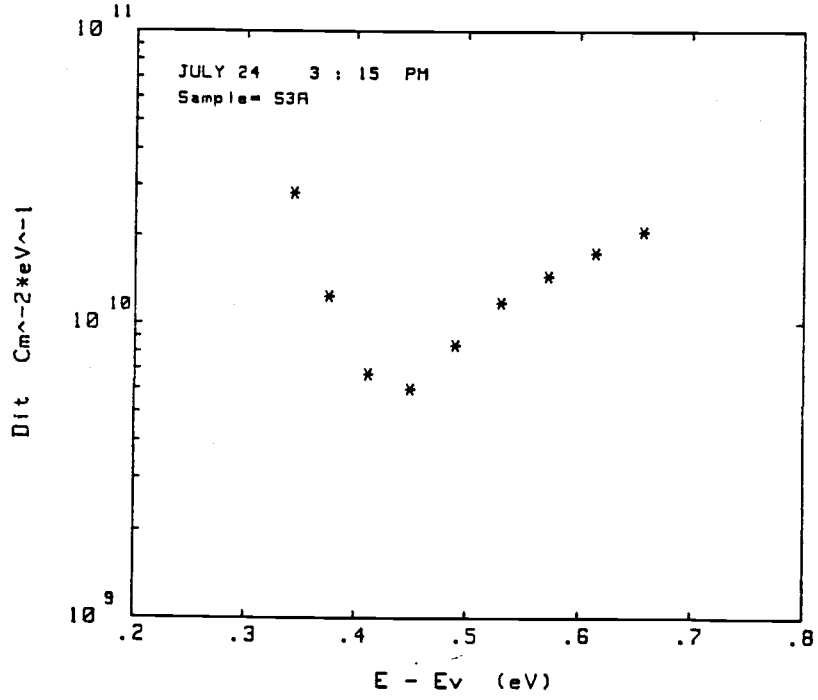


Figure 18) Typical Si-SiO_2 density of interface traps (Dits) distribution within silicon band gap.

shown in Fig. 17. The quasi static C-V plot is indicated by curve 3 in Fig. 17. The equations involved in calculating the above parameters are discussed in Appendix A. The combination of high frequency and quasi static C-V enables evaluation of the density of interface states as a function of energy (discussed in Appendix B). Using the quasi static and 1MHz combined C-V data, $Si-SiO_2$ interface state densities in the low 10^{10} to 10^{11} $cm^{-2}eV^{-1}$ were typically measured for the samples fabricated, as shown in Fig. 18.

3. Generation lifetime setup and procedure

The capacitance transient response was monitored using an LCR meter at 1MHz as a function of time after application of a voltage step. The time measurement was measured with a real time clock (HP 98035A). Since the Zerbst analysis involves a time derivative of the capacitance, in practice the time interval between the measurements is very important. Time intervals which are too short or too long result in either a noisy time derivative or loss of information.

One of the principal difficulties in applying the Zerbst analysis to the MOS capacitor transient recovery is the extremely long retention time of the MOS capacitor in high lifetime materials. Typically, the retention time is thousands of seconds at room temperature, so long times are required to obtain the required number of data points for any statistical analysis. A modified C-t transient response [15] using a step from accumulation to depletion followed by light and then a small voltage step from inversion to stronger inversion speeds up the generation lifetime measurements and

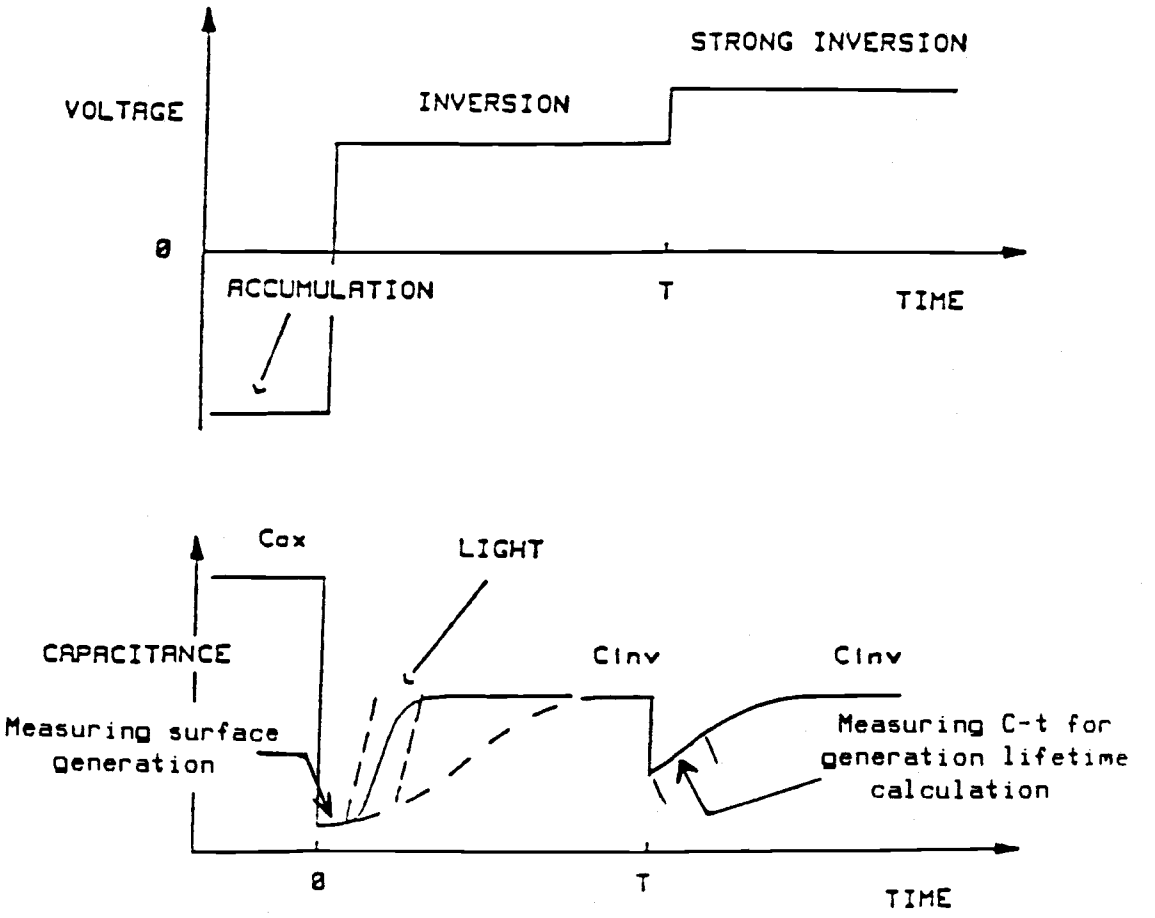


Figure 19) Schematic diagram of the modified voltage steps and the C-t transient response for faster τ_g evaluation.

separates the surface generation from the bulk generation. A schematic diagram of the modified pulsed MOS capacitor for fast Zerbst analysis of p-type wafers is shown in Fig. 19. First the capacitance is pulsed from accumulation to deep depletion. Then from the initial slope of capacitance-time plot the surface generation, S_o is estimated [10]. In order to speed up the measurement, light is applied causing the capacitor to reach inversion very rapidly. The light is then turned off and another voltage step is applied, driving the capacitance from inversion to deep depletion. The presence of the

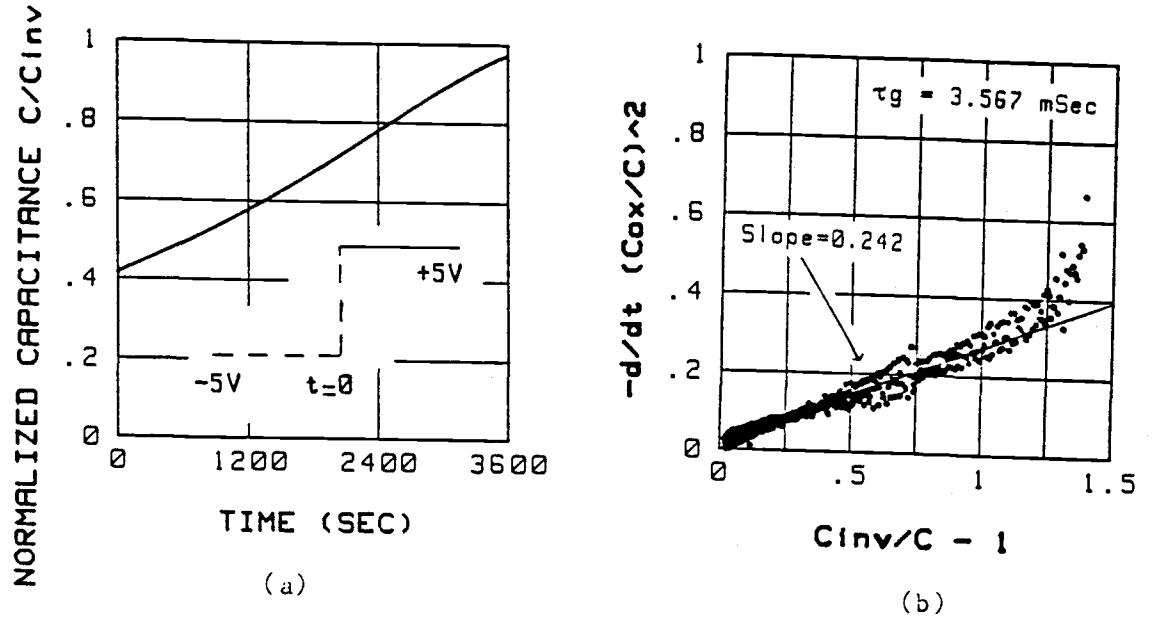


Figure 20 a) C-t response of a typical MOS capacitor, b) Zerbst plot of part a).

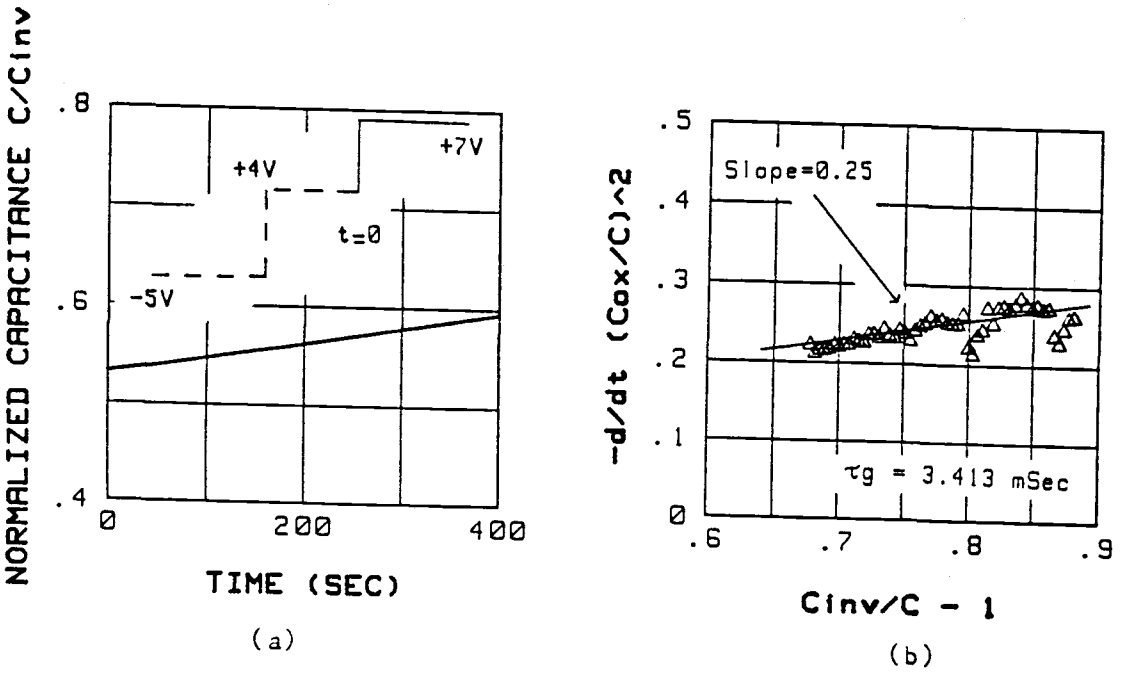


Figure 21 a) Partial C-t response of the same device as Fig. 16 using two voltage steps, b) Zerbst plot of part a).

inversion layer under the gate electrode after the second voltage step minimizes the surface generation, S. At room temperature the quasi-neutral bulk generation (diffusion from the bulk) is negligible. If generation within the depletion region is larger than the combined surface generation contributions, only a small region of the C-t transient response is necessary for evaluation of the generation lifetime. The generation lifetime, τ_g, is inversely proportional to the slope of the Zerbst plot ($-d/dt(C_{ox}/C)^2$ vs. $(C_{inv}/C - 1)$) as evident from Eqn. (17). However, if the MOS capacitor response is dominated by surface generation, an abnormal non-linear Zerbst plot results. Therefore, care must be taken if a partial Zerbst plot is used. Figure 20 [15] shows a typical C-t response of a p/p^+ epitaxial wafer (normalized to inversion capacitance) for a single accumulation to depletion voltage step. Its Zerbst plot in shown in Fig. 20b. The effective generation lifetime for this sample is about 3.5 msec. Figure 21a shows a partial C-t response of the same device using the two step voltages shown in the inset of the figure. From the C-t response of the first step a surface generation velocity of 0.8 cm/sec is obtained. The partial C-t response after the second voltage step results in a generation lifetime of about 3.4 msec. The percentage change in the effective generation lifetimes measured by the two Zerbst techniques is less than five percent, but the modified Zerbst technique reduces the measurement time about 80% compared to the complete C-t. This fast Zerbst method was used for the determination of the generation lifetimes in the work described hereafter.

4. Recombination lifetime setup

The elevated temperatures in the range of 60-100°C mentioned in Chapter II for the C-t transient response measurement for the recombination lifetime test are achieved using a thermal chuck which is controlled by a DC power supply. Use of an AC power supply was avoided since it causes large noise on the instrument (LCR meter). The temperature of the wafer is accurately monitored by a copper-constantan thermocouple which is attached to a dummy sample sitting on the hot chuck next to device under test. The reference temperature is ice water. Subsequent to an application of an accumulation bias of -5V on p/p^+ MOS capacitors, the voltage was stepped to a deep depletion bias of +4V. Since the C-t transient response is much faster at elevated temperatures, the LCR meter was used in a fast mode with a response time of about 200 msec.

B. SCHOTTKY BARRIER DIODES AND DLTS

1. Fabrication of Schottky diodes

Starting materials were p/p^+ epitaxial wafers with 30-50 Ω-cm/ 0.01-0.02 Ω-cm, (100) orientation, 100 mm wafers. Various techniques and metals were used to obtain low leakage Schottky diodes on p-type epitaxial samples. The surface preparation was simple, but very critical. Subsequent to native oxide removal using dilute HF the sample is rinsed in deionized water for 10 to 15 minutes to minimize residual chemical concentration. The sample is then immediately taken

to a vacuum chamber for metal deposition. Evaporation of aluminum on to the surface was not a reliable way for formation of Schottky diodes on p/p^+ samples. This normally resulted in excessive leakages current in the diode which could not be tolerated in DLTS measurements. Titanium has a barrier height of 0.61 eV [25] which is slightly higher than Aluminum's 0.58 eV. In general, the Ti diodes had much lower leakage currents, but they had to be annealed for about 15 minutes at 400°C either in N_2 or forming gas ambient to have reasonable C-V and built-in voltage characteristics. However, other remaining problems remained such as high series resistance due to oxidation of Ti at room temperature and more importantly, adhesion when accomplishing wire bonding. As a consequence, the following scheme was adopted for the formation of Schottky diodes. Subsequent to sputtering of Ti, Al is sputtered on the sample through a metal mask without bringing the sample out of vacuum chamber. This results in low leakage Schottky diodes. Only after the Ti/Al Schottky diodes were annealed the C-V characteristics and built-in voltage of diodes adequately resembled properties of the p-type epitaxial layer. The device was easily packaged for DLTS measurements since gold bonds easily to annealed Al. Sputtering of Ti and Al was performed at 1keV using argon ions.

In order to be able to probe deeper in the epitaxial layer with the depletion layer of the diodes some wafers had portions of the epitaxial layer slowly etched by a planar etch. The planar etch is a mixture of 20 ml of HF, 150 ml of nitric acid, and 50 ml of acetic acid. The etching rate is approximately 2 μm per minute. Once the diodes were prepared, I-V and C-V measurements were performed to

characterize their electrical behavior prior to DLTS measurements. A doping profile and the built-in voltage were obtained using the C-V data (see Appendix C).

2. DLTS setup

A schematic diagram of the DLTS setup is shown in Fig. 22. It consists of a Boonton capacitance meter and a dual gate boxcar averager which feeds the differential capacitance measured at specified intervals (t1 and t2) in the Y-axis of an X-Y plotter. The temperature is measured by a thermocouple which is fed to the X-axis of the plotter. The X-Y plotter in Fig. 22 was replaced by a dual channel multi-meter (Keithley 619) for temperature and differential capacitance measurements, an HP 9845A computer, and an HP plotter for more flexibility in storage and plotting of the DLTS data.

3. Majority carrier DLTS and optical DLTS

In conventional DLTS, filling and emptying of deep level traps is obtained by modulating the depletion region which is done by means of varying the applied reverse bias. Since the Schottky diode is a majority carrier device, the optical DLTS measurements were performed to enable injection of minority carriers. For the optical DLTS test, the Schottky diode was biased at a fixed voltage. The carriers (minority and majority) were injected into the silicon surface using an infra-red GaAs LED placed next to the Schottky diode and triggered by a square wave generator. Injection of carriers by the GaAs LED

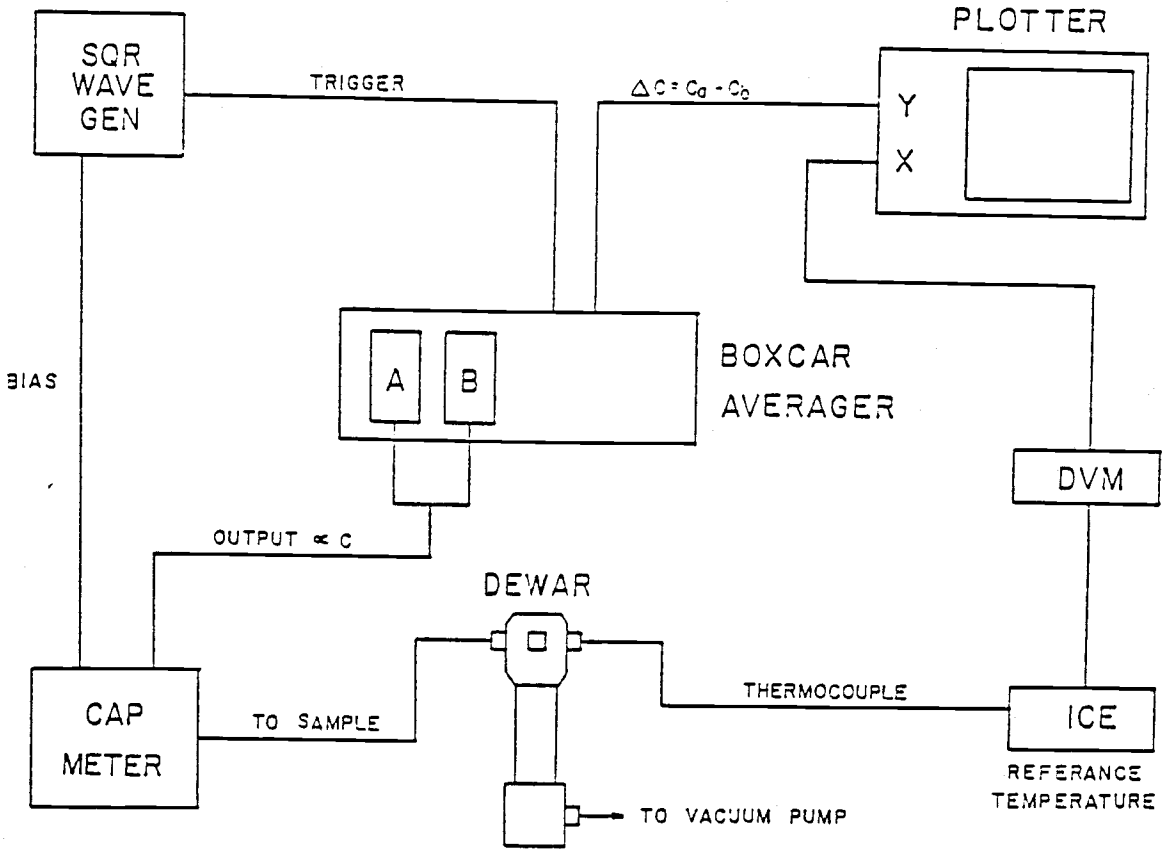


Figure 22) Schematic diagram of the DLTS setup.

causes an increase in the capacitance. As the LED is turned off the boxcar averager which is triggered by the square wave measures the capacitance transient at a specified rate window for detecting emission of any minority carriers through a large temperature range. Using this technique the minority carriers can be detected only if enough carriers are injected into the silicon and significant minority carrier traps are present. A helium refrigeration was used to cool the samples to temperatures of around 40 K. The DLTS scan was performed during the slow warm up to room temperature. The warm up rates were about 2-3 K per minute.

IV. RESULTS AND DISCUSSION (I)

The results and discussion are broken into two chapters with the recombination lifetime characterization to follow in the next chapter. In this chapter the results of measurement of the generation lifetime for testing of the active region quality, and DLTS measurements for detection of deep levels are discussed. It is shown that p/p^+ epitaxial silicon wafers have a very high generation lifetime (>20-30 msec) when surface generation is properly accounted for, and the density of deep levels as estimated by DLTS is negligible. The characteristics of MOS capacitors and Schottky barrier diodes which were carefully examined prior to the lifetime and DLTS measurements are discussed in detail. In addition, a simple criterion for sample acceptance in terms of I-V and C-V measurments used prior to lifetime and DLTS measurement is described here.

A. C-V CHARACTERIZATION OF MOS CAPACITORS

Subsequent to fabrication of MOS capacitors, C-V measurements were routinely performed to characterize the test capacitor characteristics prior to generation or recombination lifetime measurements. Doping density, oxide thickness, density of $Si-SiO_2$ interface traps, fixed charge within the oxide, and threshold and flat band voltages were the main parameters obtained from C-V tests. Various p/p^+ epitaxial wafers with different epitaxial thicknesses as well as p-type polished wafers and n/n^+ wafers were examined in this study. Figure 23a shows C-V characteristics of a p/p^+ epitaxial wafer

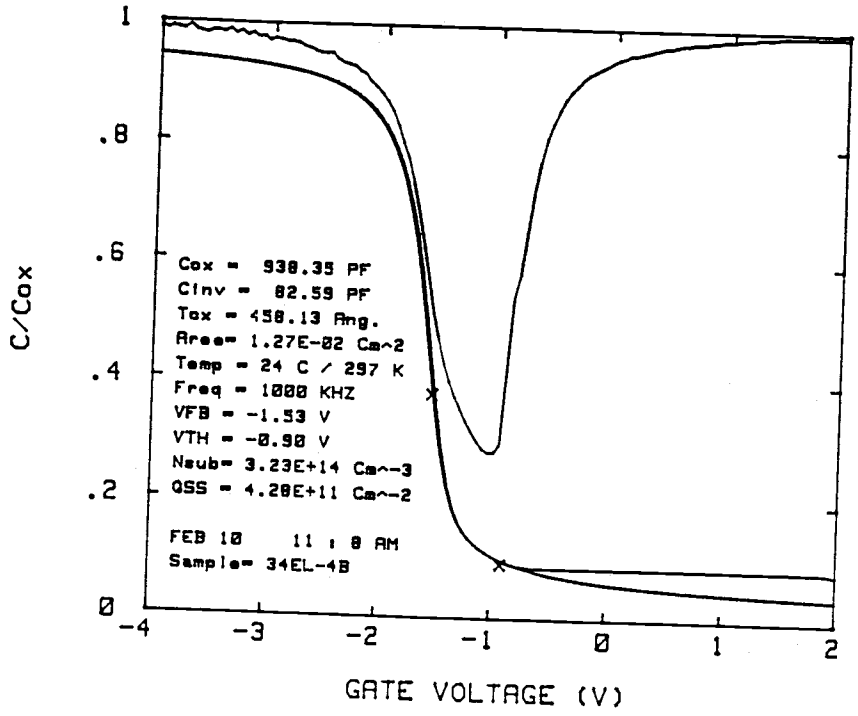


Figure 23a) High frequency and quasi-static C-V plots of a p/p^+ epitaxial wafer without post-metallization anneal.

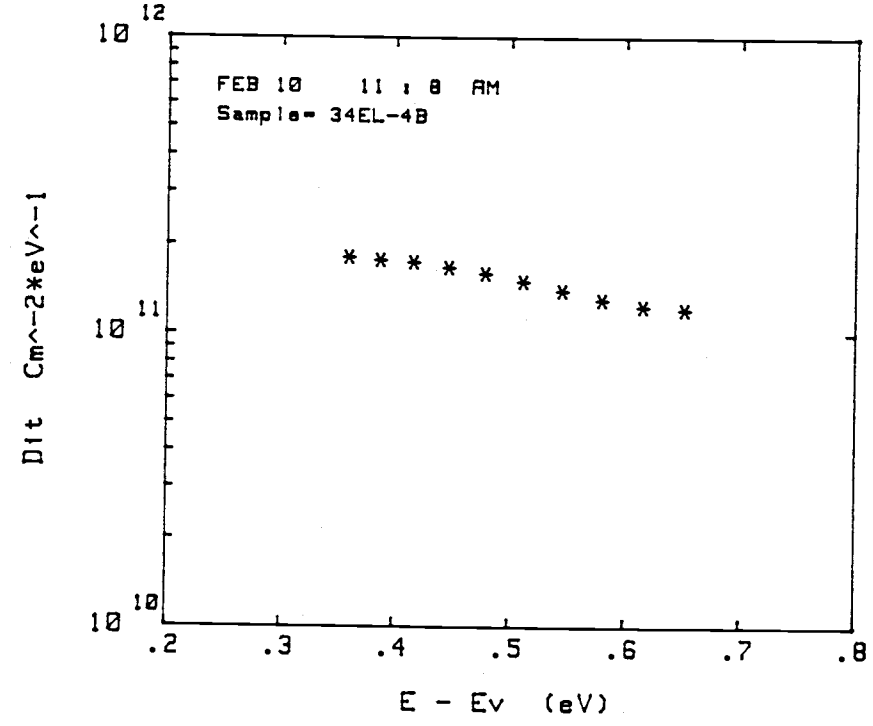


Figure 23b) Distribution of density of interface traps of Fig. 23a.

without a post-metallization anneal. High frequency C-V measurement reveals reasonable flatband and threshold voltages which are marked with (X's) in the figure. The density of interface traps (Fig. 23b) is rather high, about 10^{11} $cm^{-2}eV^{-1}$. This is due to dangling bonds at $Si-SiO_2$ interface. It is well known that a post metallization anneal at about 400-450°C in hydrogen ambient reduces the interface trap density. Therefore, all of the samples were annealed in forming gas (90% N_2, 10% H_2) at 400°C for 20-60 minutes subsequent to aluminum metallization and prior to C-V and lifetime measurements.

The acceptance criteria of a MOS capacitor for lifetime measurements was set at $|V_{th}| < |-1.0\ V|$ for a small density of fixed charge in the SiO_2 and a density of interface traps of less than $5x10^{10}$ $cm^{-2}eV^{-1}$ to ensure low surface generation due to the $Si-SiO_2$ interface. The oxide thickness criteria was 450-550 Angstrom.

C-V plots and the density of interface traps of p/p^+ epitaxial wafers with 6 μm, 13.5 μm, 70 μm, and 110 μm epitaxial thickness are shown in Fig. 24-27, respectively. Flatband voltages are in the range of -1.5 to -1.25 V. The threshold voltages are greater than -1.0 V, and the density of interface traps (Dits) are less than $3x10^{10}$ $cm^{-2}eV^{-1}$ for all of p/p^+ MOS capacitors shown in figures 24-27. Therefore, the surface generation velocities are expected to be small and the capacitor characteristics meet the acceptance criteria. The well-known U shaped density distribution is observed for forming gas annealed MOS capacitors. The discrepancy between the high frequency and quasi-static C-V curves in the accumulation region is due to series resistance effects at high frequency. Since the MOS capacitors are very large (diameter = 1270 μm) and the measurement were

performed at 1MHz, the series resistance causes lower measured capacitance readings in the accumulation region. In the depletion or inversion regions the capacitance is substantially smaller, therefore the series resistance does not introduce significant error in measured capacitance allowing accurate determination of MOS capacitor parameters. The oxide thickness, C_{ox}, was measured at 10kHz to avoid series resistance problem in determination of oxide thickness from the C_{ox} value. The C-V plot and Dits of a 50 μm n/n^+ epitaxial wafer are shown in Fig. 28. The n/n^+ wafers were used for comparison of the lifetimes to p/p^+ epitaxial wafers. In general, the C-V characteristics of all annealed samples indicate very good MOS capacitor behavior.

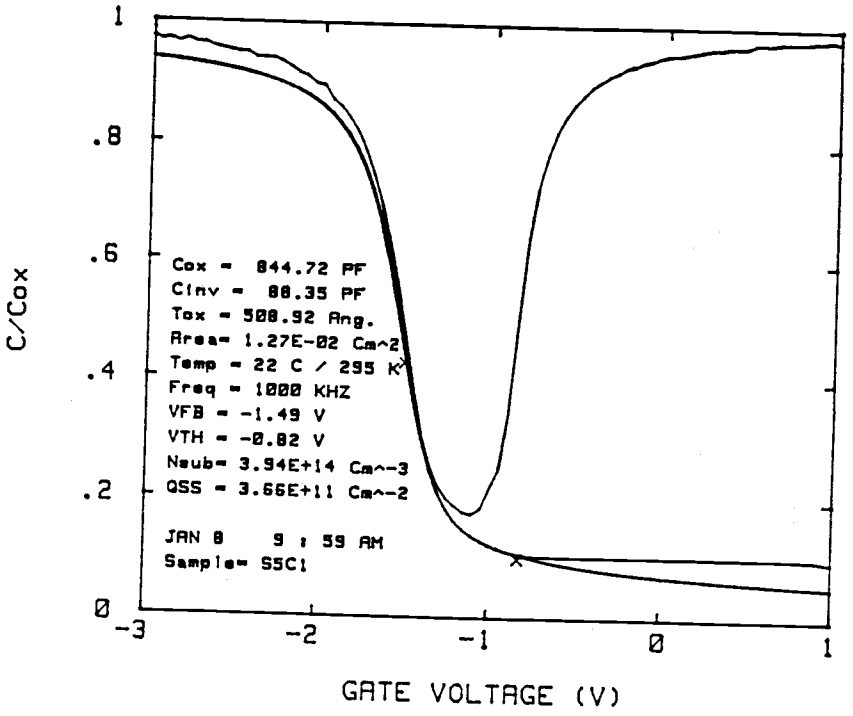


Figure 24a) High frequency and quasi-static C-V plots of a 6 μm p/p$^+$ epitaxial layer.

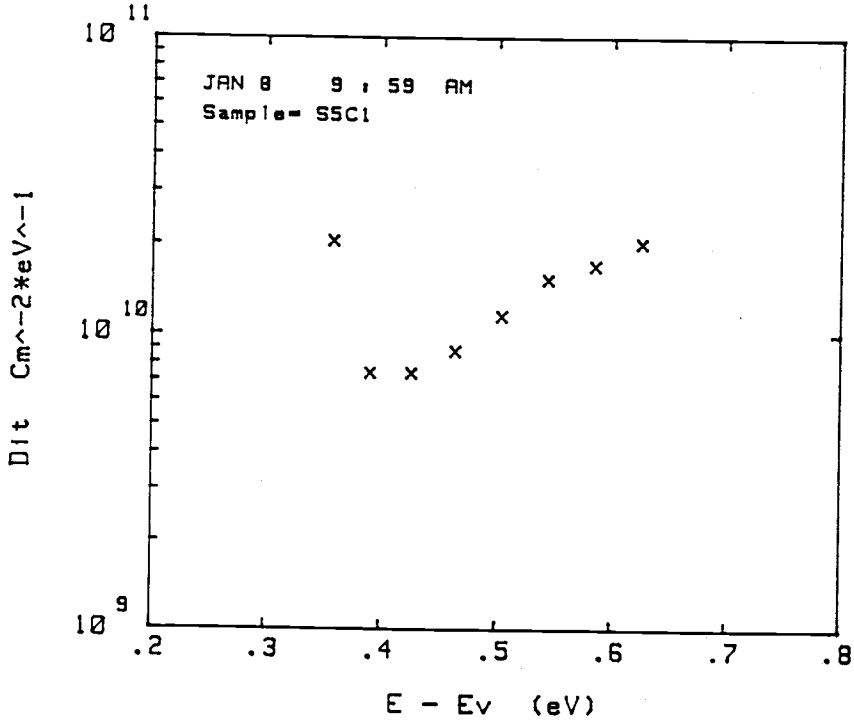


Figure 24b) Distribution of density of interface traps of Fig. 24a.

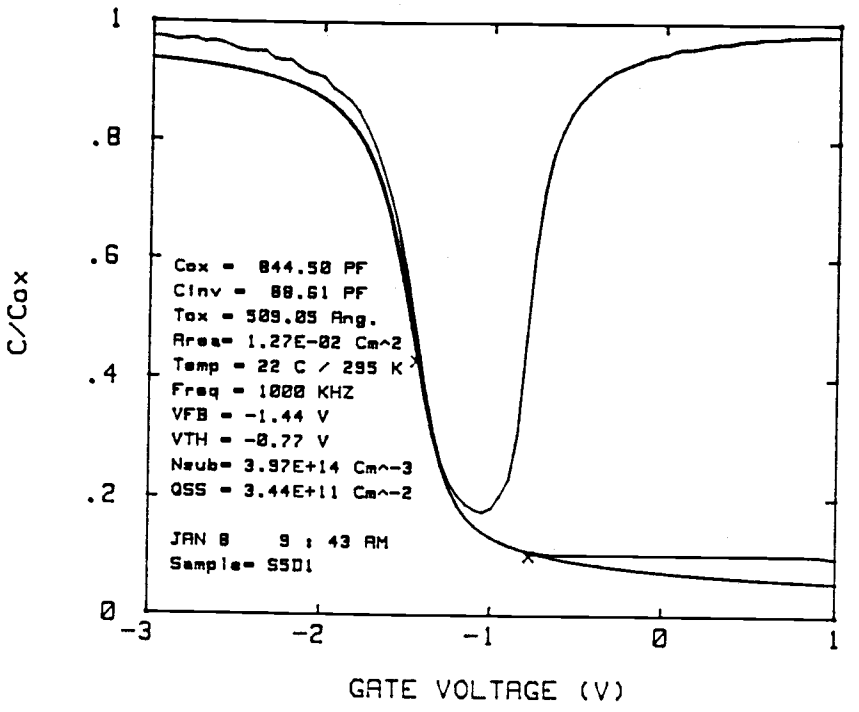


Figure 25a) High frequency and quasi-static C-V plots of a 13.6 μm p/p^+ epitaxial layer.

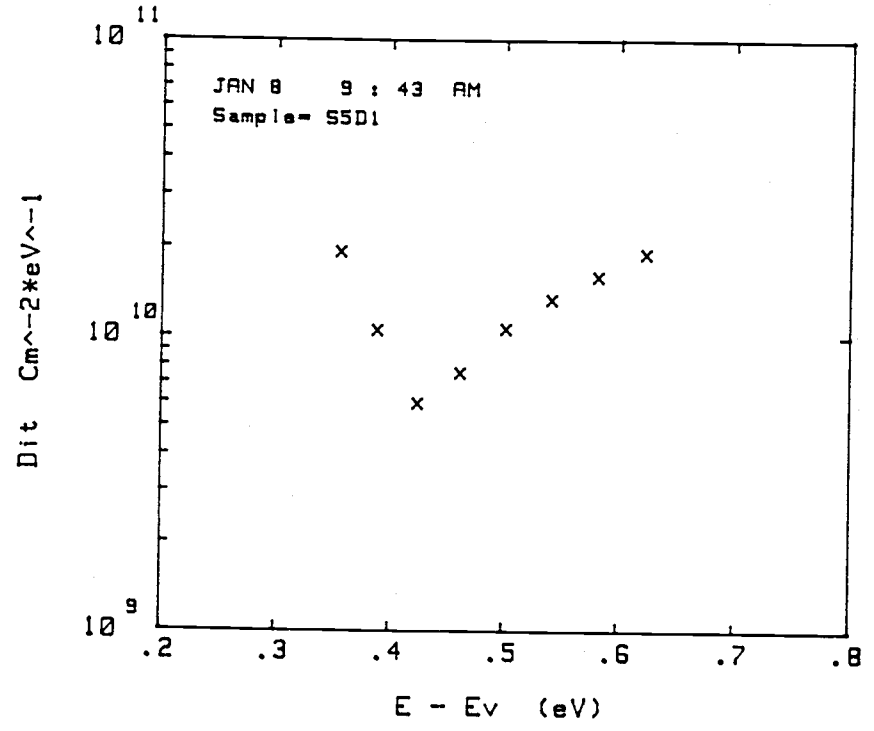


Figure 25b) Distribution of density of interface traps of Fig. 25a.

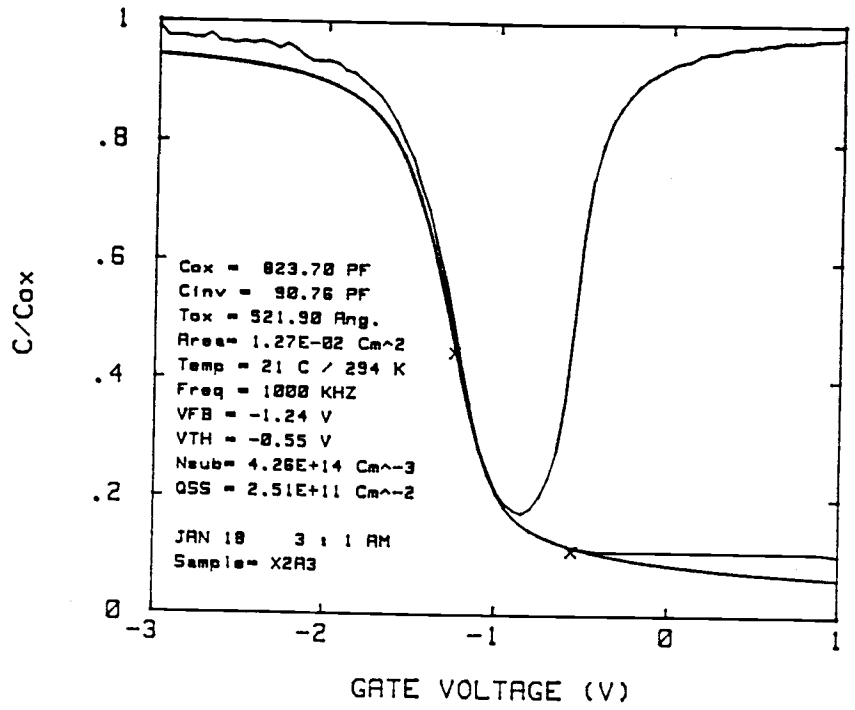


Figure 26a) High frequency and quasi-static C-V plots of a 70 μm p/p$^+$ epitaxial layer.

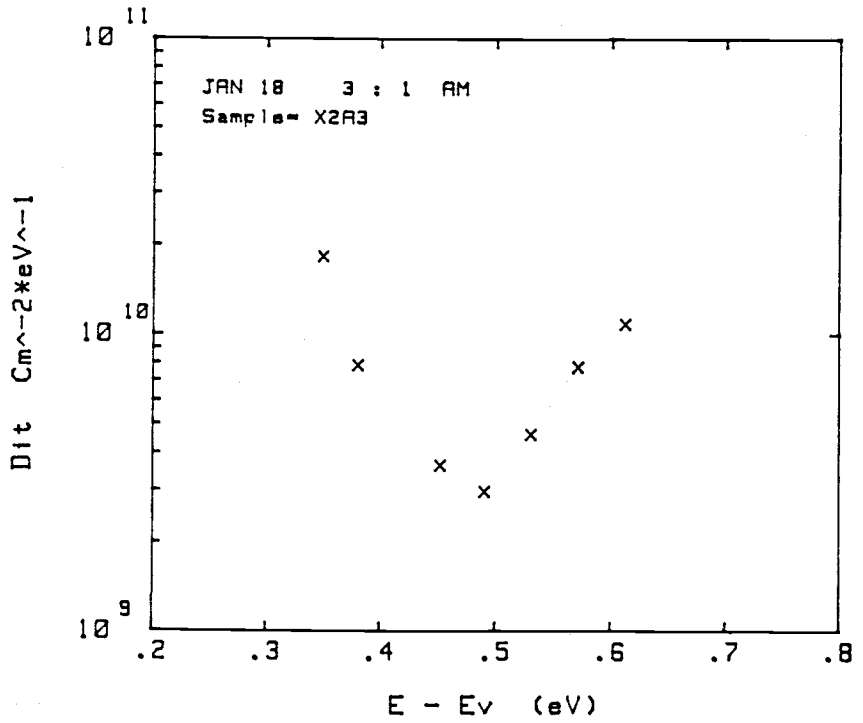


Figure 26b) Distribution of density of interface traps of Fig. 26a.

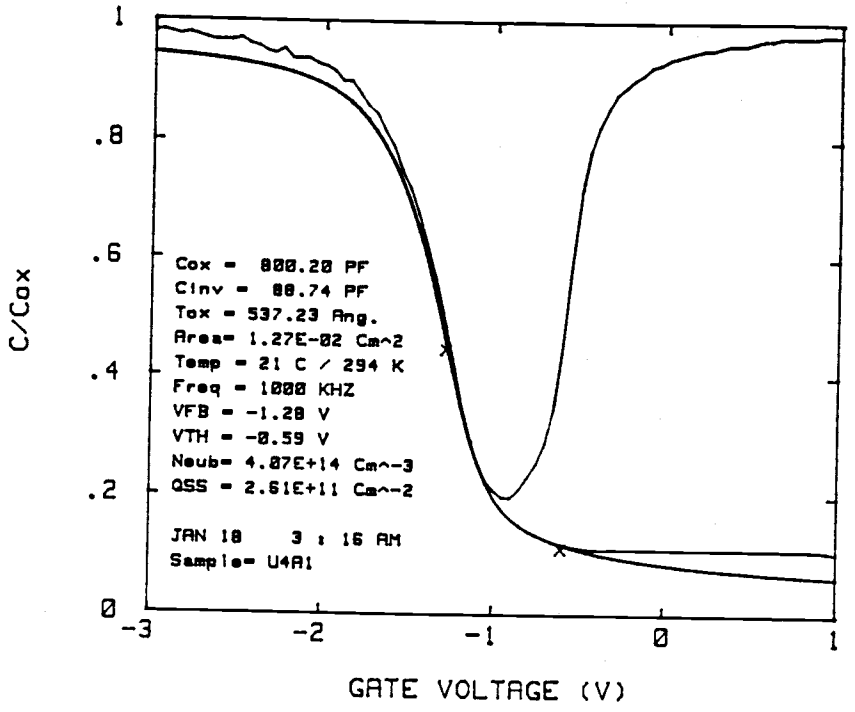


Figure 27a) High frequency and quasi-static C-V plots of a 110 μm p/p$^+$ epitaxial layer.

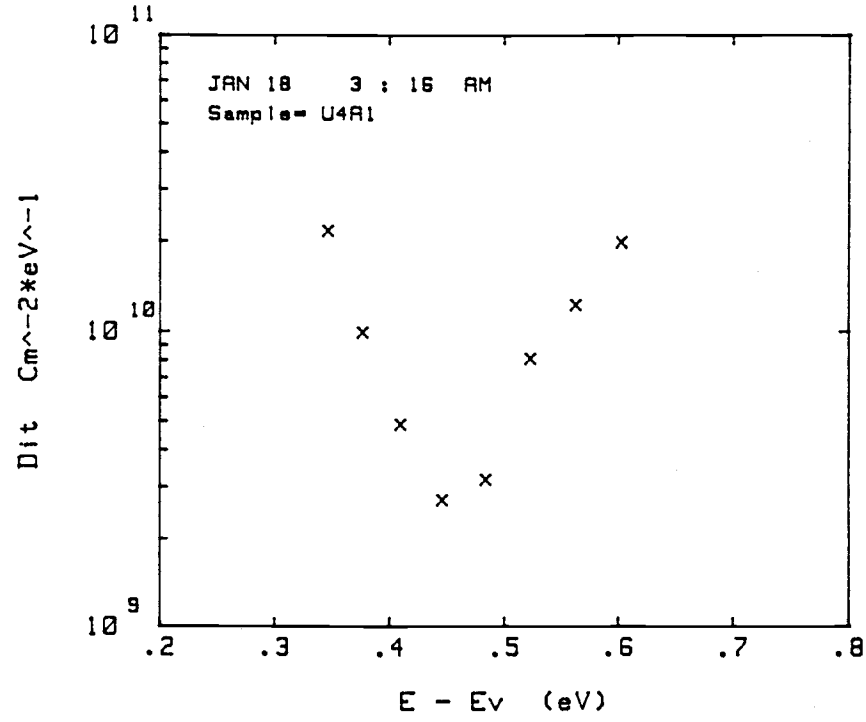


Figure 27b) Distribution of density of interface traps of Fig. 27a.

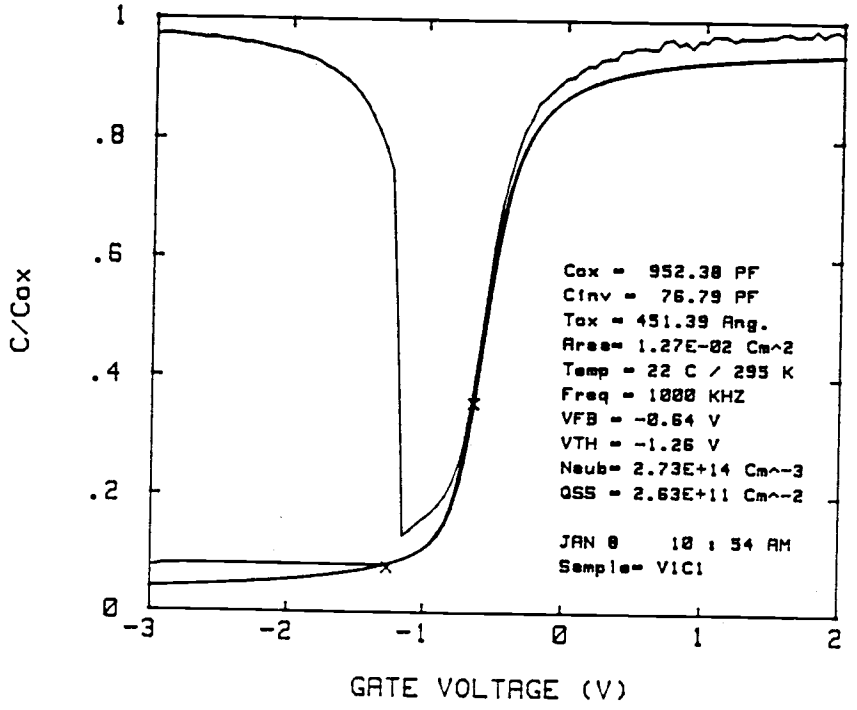


Figure 28a) High frequency and quasi-static C-V plots of a 50 μm n/n$^+$ epitaxial layer.

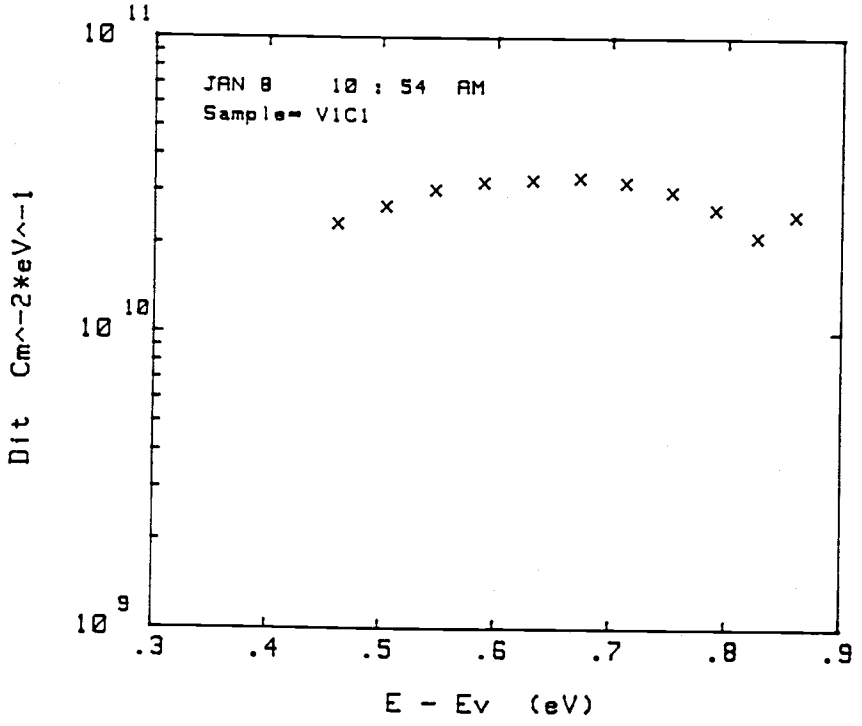


Figure 28b) Distribution of density of interface traps of Fig. 28a.

B. GENERATION LIFETIME

A typical C-t transient response of a MOS capacitor at room temperature using an accumulation-depletion voltage step is shown in Fig. 29. Initially the capacitance increases rapidly due to the surface generation of carriers under the gate electrode. However, once an inversion layer is established it blocks surface generation under the gate causing a slower capacitance change. Generation of carriers within the space charge region dominates over the surface generation for the rest of the C-t transient response. The generation lifetime of the sample shown in Fig. 29 was measured using the Zerbst technique to be 3.2 msec. A typical C-t response of an MOS capacitor

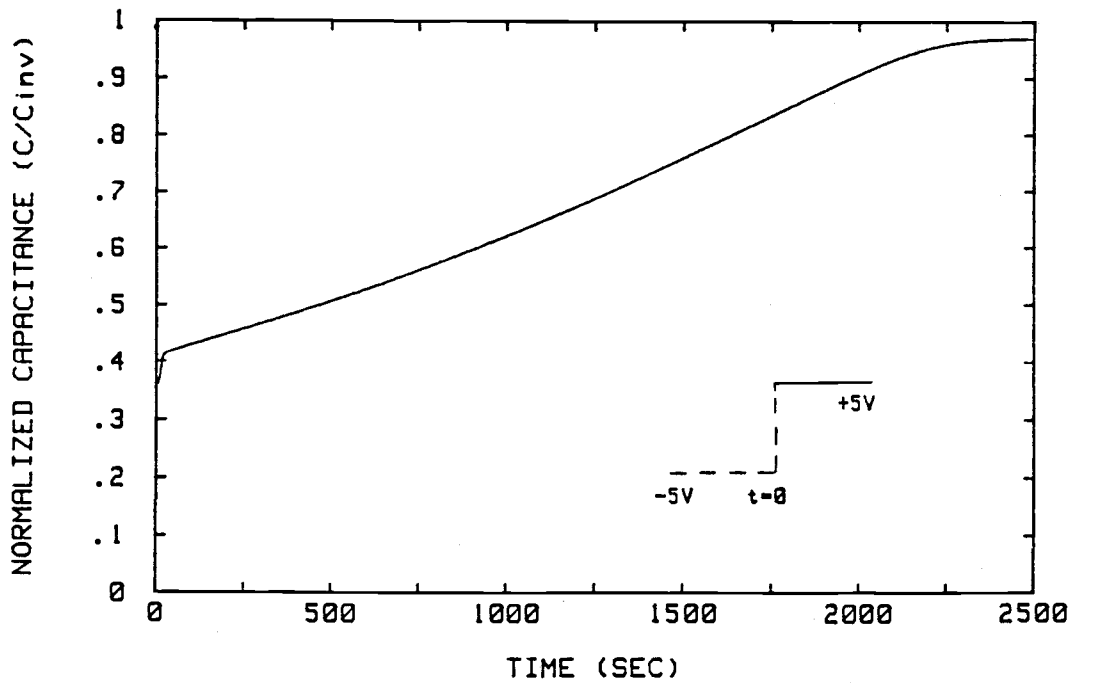


Figure 29) Typical C-t response of an MOS capacitor at T=22°C using an accumulation - depletion voltage step.

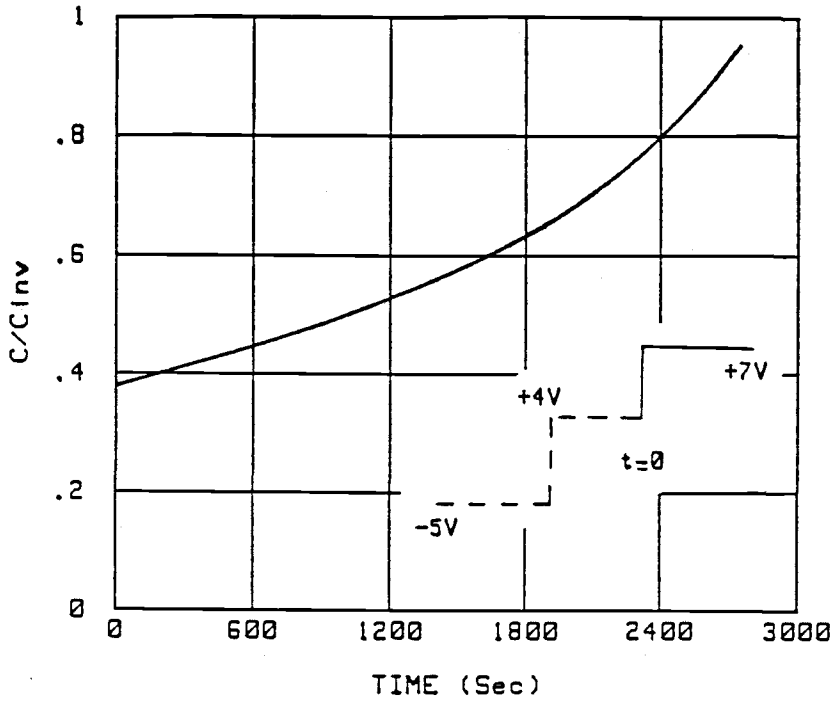


Figure 30a) Typical C-t response of an MOS capacitor at T=22°C using a modified voltage step.

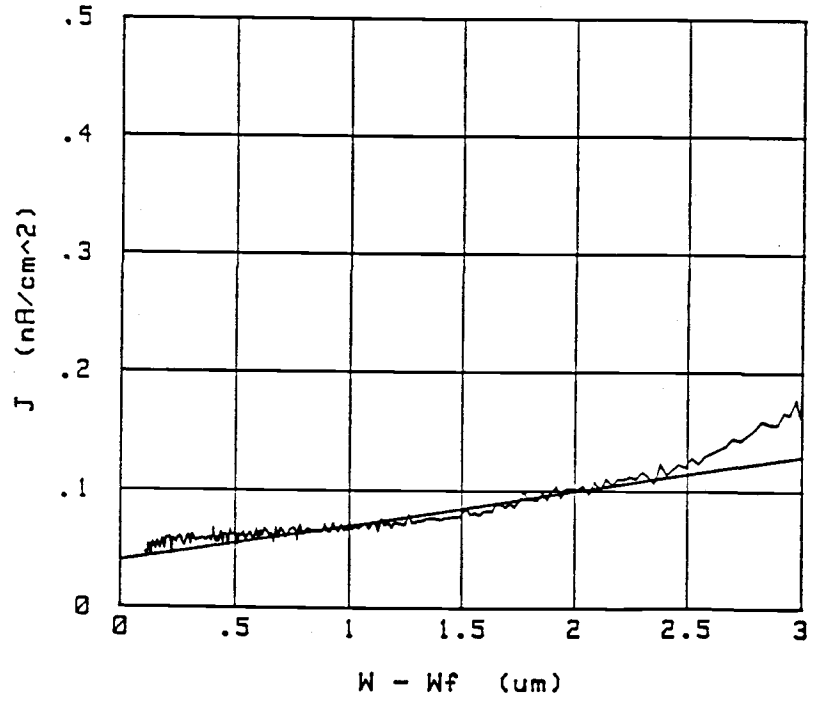


Figure 30b) Zerbst plot of Fig. 30a) in terms of J vs. $W-W_f$.

measured at T=22°C using the two voltage step scheme discussed in the previous chapter is shown in Fig. 30a. This C-t response does not show any rapid capacitance change at the beginning of the transient since the inversion layer was established before the C-t was measured. Figure 30b is an alternative way of plotting the Zerbst relation using Eqns. (19) and (20). The generation lifetime for J versus $W-W_f$ is inversely proportional to the slope of such a plot,

$$\tau'_g = qn_i[\ \delta(W-W_f)/\delta J\ ]\ . \qquad (35)$$

In the next section, surface generation effects on high quality silicon materials are discussed and a simple method for proper evaluation of the real generation lifetime and lateral surface generation is examined. Generation lifetime results for various thickness p/p^+ epitaxial wafers as well as for n/n^+ wafers and p-type bulk wafers are presented in section 2.

1. Surface generation effects for high quality silicon materials

The Zerbst technique has become widely used in the characterization of silicon materials and in the determination of the generation lifetime of the active region of the device. More recent adaptations of this technique [10,11,16] made first order approximations to correct and account for lateral surface effects for accurate generation lifetime measurements. While these simple approximations are sufficient for low lifetime materials in the range of microseconds, it is shown here that these approximations lead to

significant errors for long lifetime materials.

According to the technique of Schroder et al. [10] the measured generation lifetime, τ'_g, should be corrected to account for surface generation in the peripheral areas as given by Eqn. (18) in Chapter II. Equation (12) is repeated here for convenience:

$$1/\tau'_g = 1/\tau_g + 4S_o/D \qquad (18)$$

where τ_g and τ'_g are the actual and measured generation lifetimes respectively. Assuming Eqn. (18) is valid, an accurate value of S_o is required for the determination of long generation lifetimes in high quality materials. The Zerbst technique of evaluation of surface generation velocity, S, from the intercept of Zerbst plot gives an average value of S during the C-t transient response. The main difficulty in using the Zerbst technique is that the lateral surface generation, S_o, is normally different from the surface generation under the gate electrode, S. This is because the Si-SiO_2 interface immediately under the gate anneals differently from the gap area out side the gate electrode due to the presence of aluminum on the gate electrode. Presumably, because of the high concentration of hydrogen in the Al [67], the Si-SiO_2 interface under the gate area is easily annealed, resulting in lower S values. Schroder's technique [11] of evaluation of S_o from the initial C-t response is adequate for samples with small generation lifetimes, in the range of microseconds, but it is not accurate enough for long lifetime samples.

Unfortunately, no simple and accurate technique exists for the

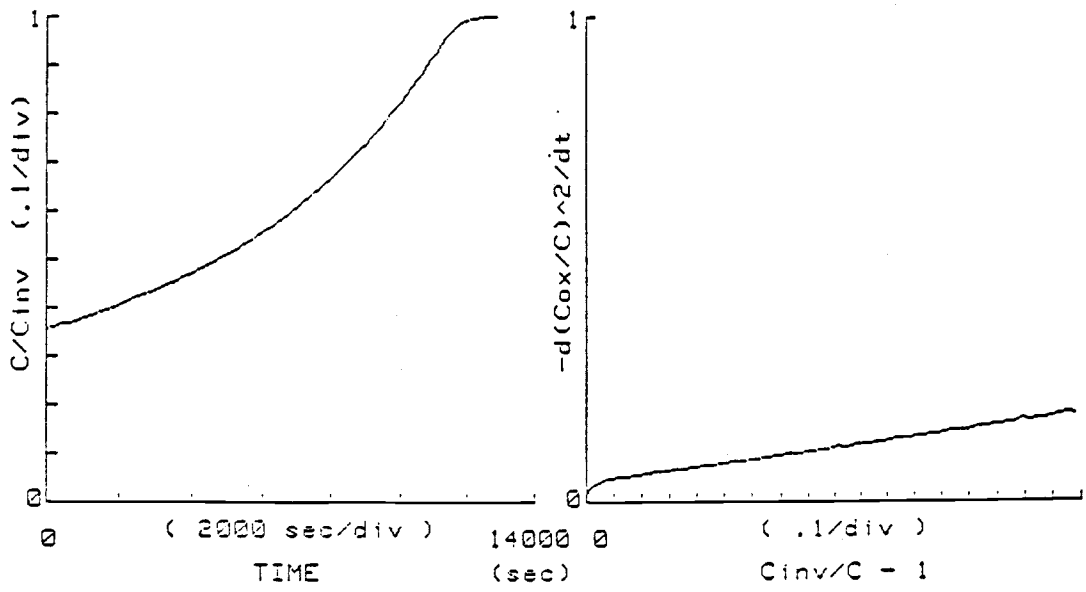


Figure 31) Typical C-t and Zerbst plot of a p/p^+ epitaxial wafer, after W.I. Sze [68].

correct evaluation of the surface generation. The only reasonable way of properly determining the actual generation lifetime of a high quality silicon material is by comparing devices of different diameter sizes. To better quantify the surface effects and generation lifetimes, different diameter MOS capacitors were fabricated (see Fig. 14). For studying surface effects, three samples which received the heat treatment of Fig. 13a are discussed here. Subsequent to Al metallization samples #1 and #2 were annealed in forming gas while sample #3 was not. Samples #2 and #3 were cleaved quarters of the same wafer. The summary of heat treatments on the three samples discussed in this section are shown in Table 3.

A typical capacitance versus time transient and the resulting Zerbst plot are shown in Fig. 31. A stream of nitrogen gas was applied on this sample during the measurement to reduce moisture

	sample #1	sample #2	sample #3
750°C in N_2	---	4 Hrs.	4 Hrs.
CMOS simulation	---	Yes	Yes
900°C in dry O_2	4 Hrs.	4 Hrs.	4 Hrs.
900°C in N_2	4 Hrs.	4 Hrs.	4 Hrs.
425°C in forming gas (10% H_2)	4 Hrs.	4 Hrs.	---
Actual τ_g (msec)	30	20	20
S_o (cm/sec)	1.15	0.95	6.8

Table 3. Summary of heat treatments for the three samples used for the generation lifetime comparison.

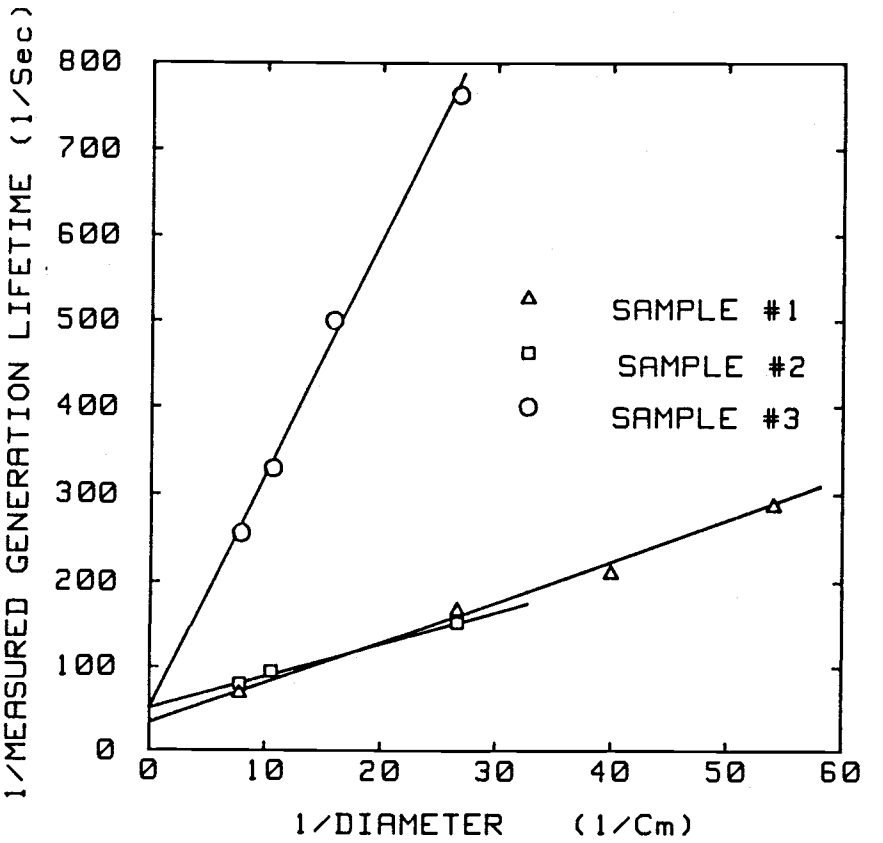


Figure 32) Reciprocal of measured generation lifetime versus reciprocal MOS capacitor diameter.

induced surface leakage. Note the retention time is extremely long when compared to a similar MOS capacitor without the application of nitrogen (Fig. 30). The C-t response of various diameter size devices was measured. The surface generation velocities obtained using Schroder's technique [11] from the initial slope of the C-t response were inconsistent using different diameter devices on the same wafer. Figure 32 shows the reciprocal of the measured lifetime, $1/\tau'_g$, versus the reciprocal of the device diameter, 1/D, for the three samples. Good linearity between $1/\tau'_g$ versus 1/D indicates simple assumptions made in the derivation of Eqn. (18) are valid. The actual generation lifetime can be extracted from the intercept of Fig. 32. Significant differences exist between the apparent or measured lifetime and the actual generation lifetime of moderate or small size samples as seen in Fig. 32. Direct determination of the real generation lifetime from a single sample without any correction due to surface generation would require capacitors with 1.0 cm diameter. The surface generation velocity can be evaluated from the slope of Fig. 32 using Eqn. (18). Extracting the surface generation from the slope of Fig. 32 does not rely on a specific measurement technique except for assumptions made by Zerbst and Schroder discussed in Chapter II.

Sample #1 had a surface generation velocity of 1.15 cm/sec and generation lifetime of 30 msec as determined from the slope and intercept of Fig. 32. Both samples 1 and 2 had identical actual generation lifetime of about 20 msec. This is expected since both samples had identical furnace heat treatments except for the post-metallization anneal. Sample 3, without the forming gas anneal, had a

lateral surface generation velocity of 6.8 cm/sec while sample 2 which was annealed had S_o of 0.95 cm/sec. Even though the measured generation lifetime values of sample #3 without the post metallization anneal are low, the actual value is the same as sample #2 with the post metallization anneal. Slightly lower actual τ_g value for samples 2 and 3 were not significant for high τ_g values in the range of tens of milli-seconds.

From a comparison of various diameter MOS capacitors, the actual generation lifetime of the material is obtained without depending on a specific surface generation measurement technique. The reported results have important implications in determining the factors limiting the maximum observable generation lifetimes of silicon materials. Contrary to common belief, this finding indicates that CMOS dynamic circuits on epitaxial wafers are limited by surface effects and not by the bulk generation lifetime within the space charge region.

2. Generation lifetimes of p/p^+ epitaxial wafers

Various p/p^+ epitaxial wafers from different locations of crystal ingot experienced the heat treatment described in Fig. 13b and samples for generation lifetime measurements were prepared. The calculated generation lifetimes (using the method discussed in the previous section) for wafers from different ingot locations with various pre-epitaxial annealing times are shown in Fig. 33. There appears to be a very slight increase in τ_g for samples which experienced a pre-epitaxial anneal and the CMOS simulation cycle.

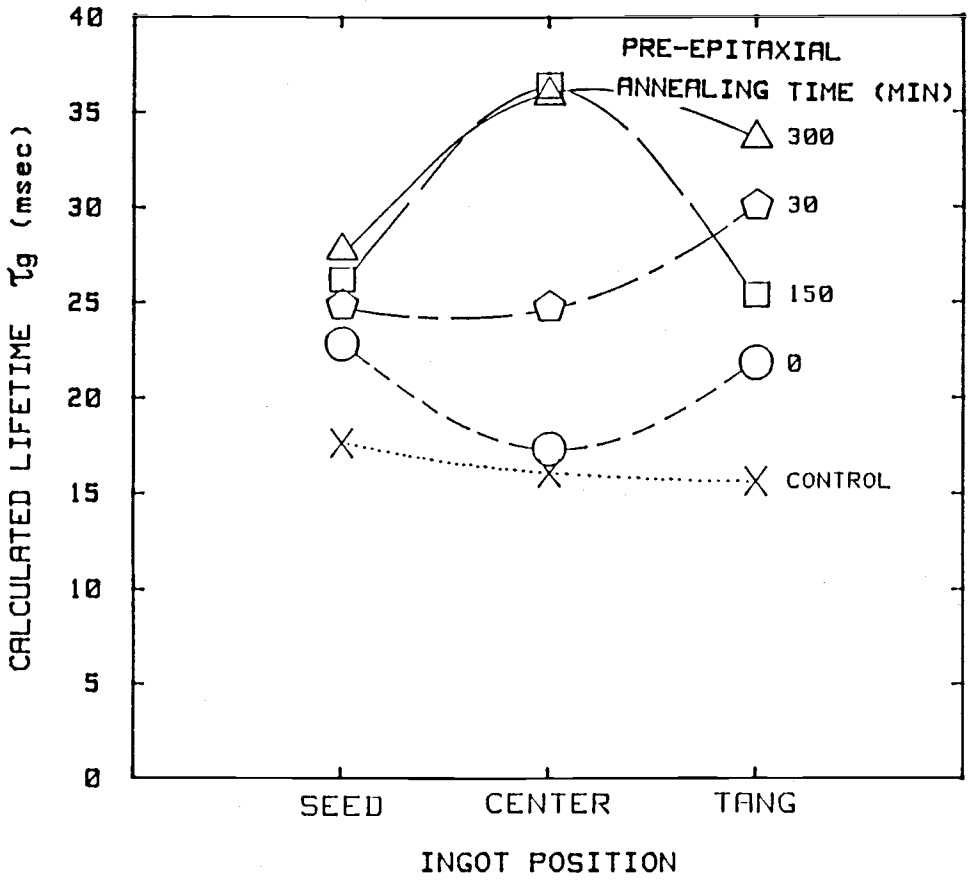


Figure 33) Calculated generation lifetime versus position in ingot for various pre-epitaxial annealing times, after W.I. Sze [68].

Even after exposure to the high temperature CMOS cycle for long times, the τ_g values are still very high. However, there is no correlation between annealing time and τ_g. In general, the generation lifetimes of all of the samples in Fig. 33 are very high in the range of tens of milli-seconds and the main limitation on these epitaxial wafers are primarily due to the surface effects.

For the study and comparison of recombination lifetimes, only the measured or effective generation lifetimes, τ'_g, on the largest size MOS capacitors were tested since evaluation of the exact τ_g

value as discussed in previous section was very time consuming. A summary of the measured generation lifetimes, τ'_g, of various samples also used in recombination lifetime study later is given in Table 4.

In general, the τ'_g values of various epitaxial layers including the n/n^+ are in the range of 3-7 msec. However, τ'_g of p/p^+ wafer from vendor B is slightly lower. This difference is expected since the epitaxial qualities from different vendors are not likely to be identical. The polished p-type wafers had considerably lower lifetimes, perhaps due to various defects such as transition metals or oxygen related defects within the active layer of the wafer. Generation lifetimes of Fe implanted sample are more than two orders

sample	Epitaxial thickness (μm)	Res. (Ω-cm)	Doping (cm^{-3})	τ'_g (msec)
p/p+	110	35	$4.0x10^{14}$	3 - 5
p/p+	70	35	$4.0x10^{14}$	3 - 5
p/p+	13.5	35	$4.0x10^{14}$	3 - 5
p/p+	5-6	35	$4.0x10^{14}$	3 - 5
p/p+ from vendor B	12	20	$6.0x10^{14}$	1 - 1.5
p-type polished	---	40	$3.0x10^{14}$	0.15-0.5
n/n+	50	40	$2.5x10^{14}$	5 - 7
p-type epi Fe implanted	5	40	$3.0x10^{14}$	0.01-0.02

Table 4. Summary of measured generation lifetimes.

of magnitude lower than p/p^+ epitaxial wafers. This is expected since Fe is a deep level in silicon as shown by DLTS later in this chapter.

C. CHARACTERIZATION OF SCHOTTKY DIODES

The Schottky barrier diodes were fabricated using sputtered Ti and Al, as described in the previous chapter, for detecting any deep levels on p/p^+ epitaxial wafers using DLTS. Various Schottky barrier diodes were fabricated and their C-V and I-V characteristics were routinely checked to ensure acceptable diode characteristics before DLTS measurements were undertaken. Typical I-V and C-V characteristics of two samples are discussed in detail this section.

1. I-V Characteristics

Forward and reverse I-V characteristics of a 110 μm epitaxial layer and a 5 μm planar etched epi-layer Schottky diodes are shown in Fig. 34a and 35a, respectively. The reverse bias characteristics and leakage currents of both diodes are adequate for DLTS measurements. The diode on the 5 μm epi-layer has a higher reverse leakage current since the depletion region of this device has punched through the low resistivity p^+ substrate (as seen in Fig. 35c, 35d). The turn on voltages of the Schottky diodes are 0.1 to 0.2 volts. The current density of Schottky diodes are given by [25,69]

$$J = J_s \exp(qV/nKT) \qquad \text{for } V >> KT/q \qquad (36)$$

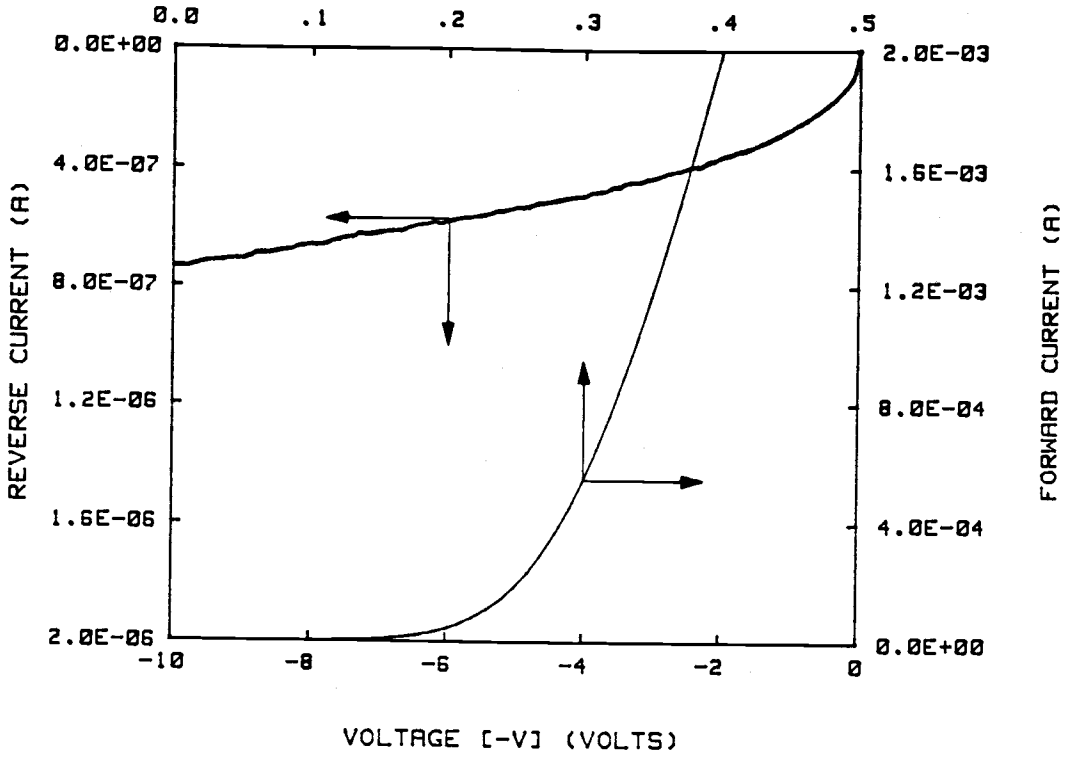


Figure 34a) I-V plot of a 110 μm epi-layer Schottky diode.

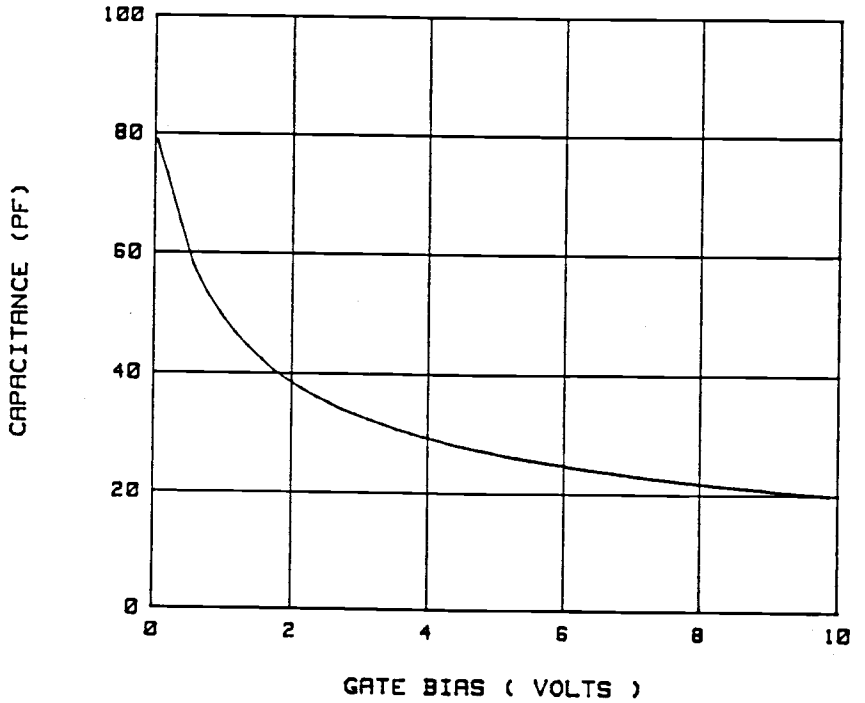


Figure 34b) C-V plot of a 110 μm epi-layer Schottky diode.

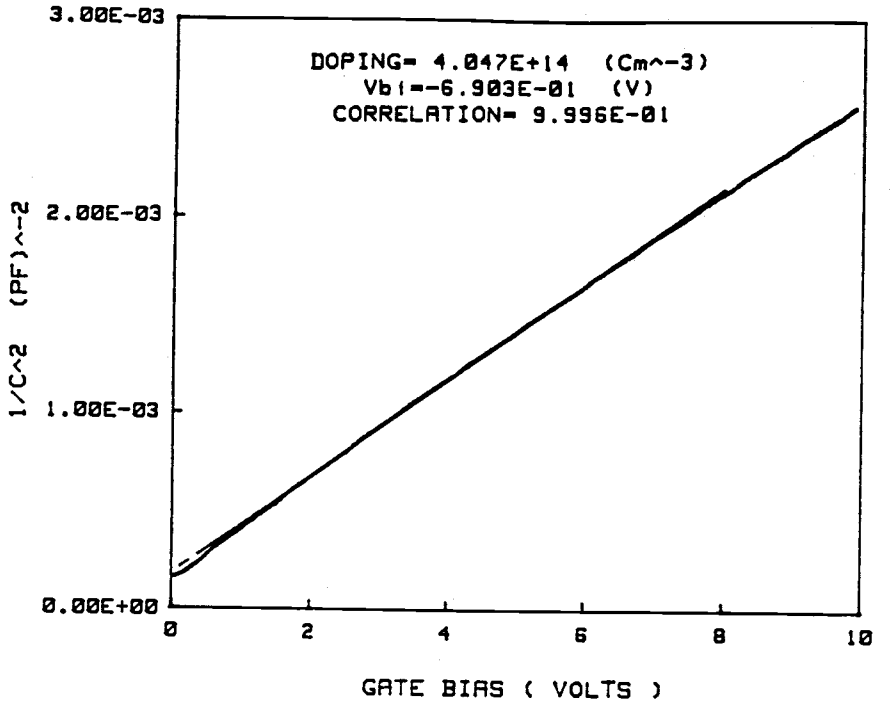


Figure 34c) $1/C^2$ - V plot of a 110 μm epi-layer Schottky diode.

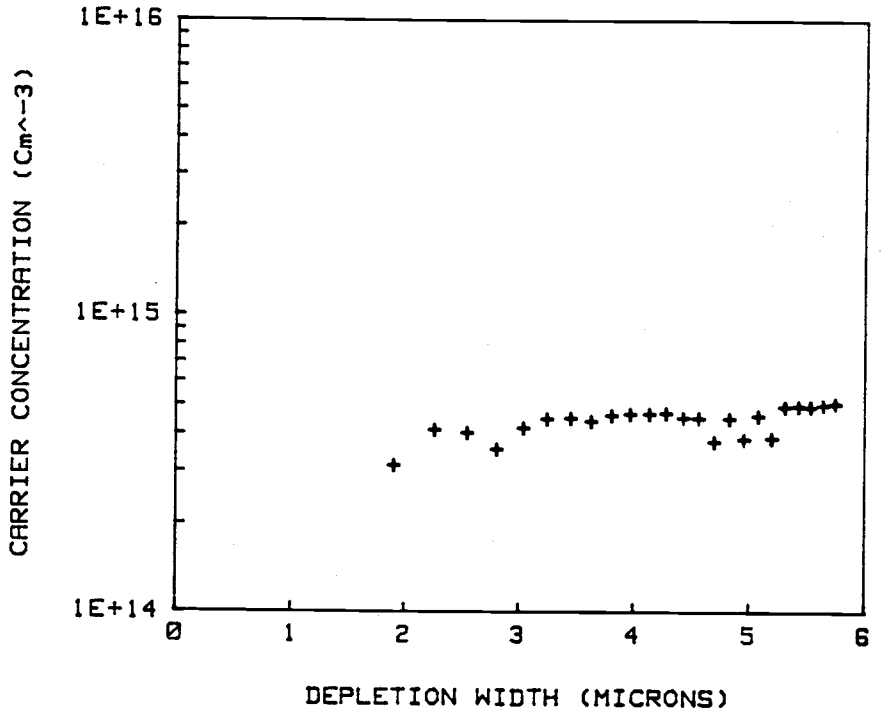


Figure 34d) Free carrier profile plot of a 110 μm epi Schottky diode.

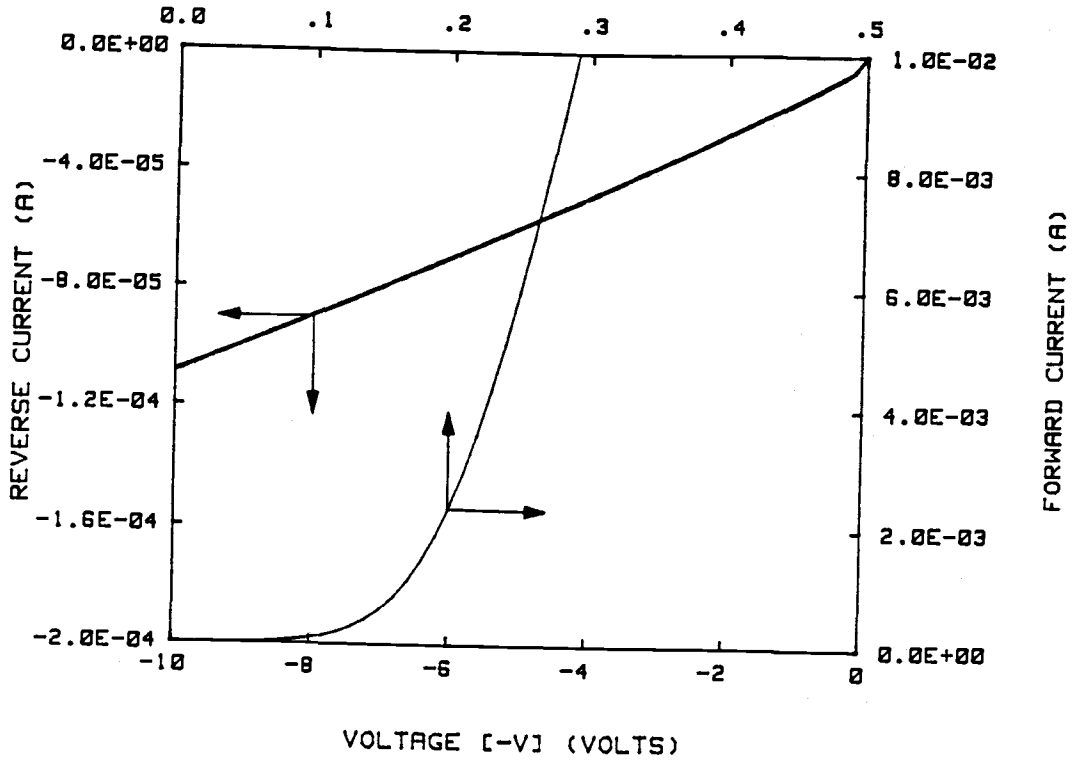


Figure 35a) I-V plot of a 13.5 μm epi-layer (control) Schottky diode.

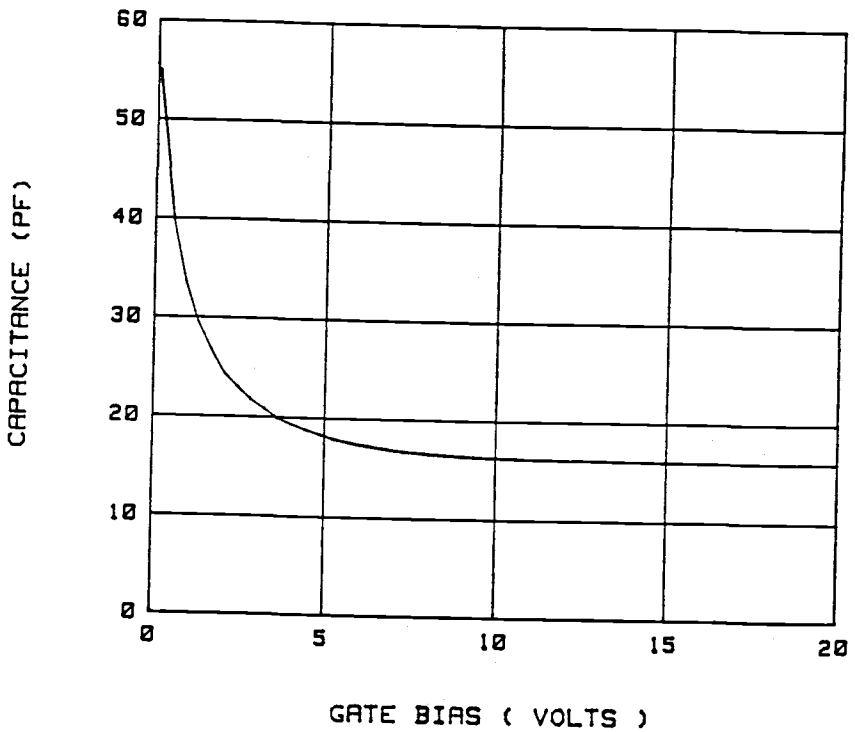


Figure 35b) C-V plot of a 13.5 μm epi-layer (control) Schottky diode.

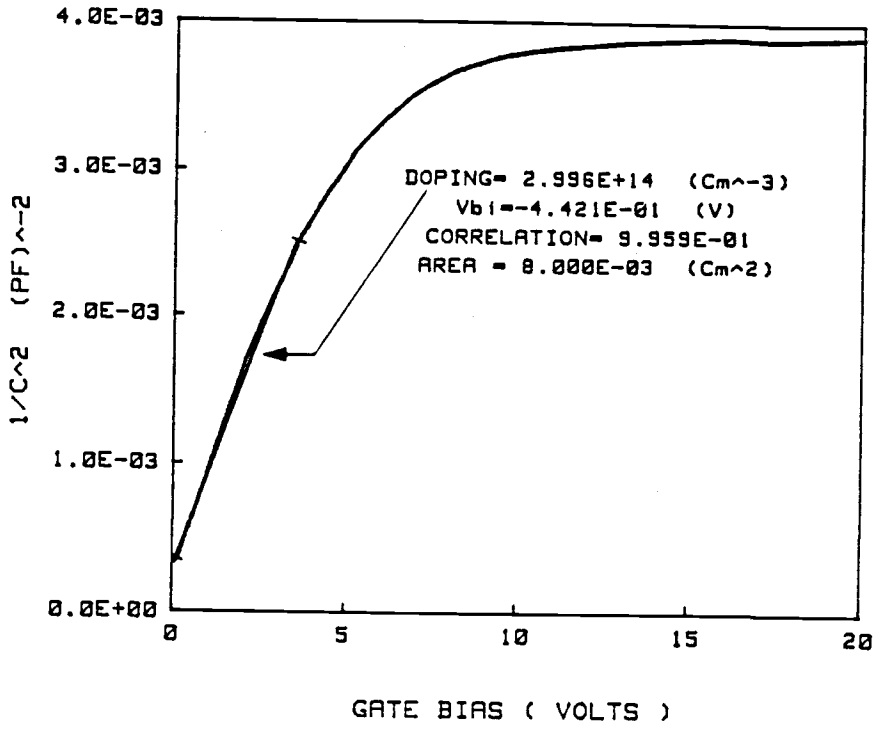


Figure 35c) $1/C^2$ - V plot of a 13.5 μm epi (control) Schottky diode.

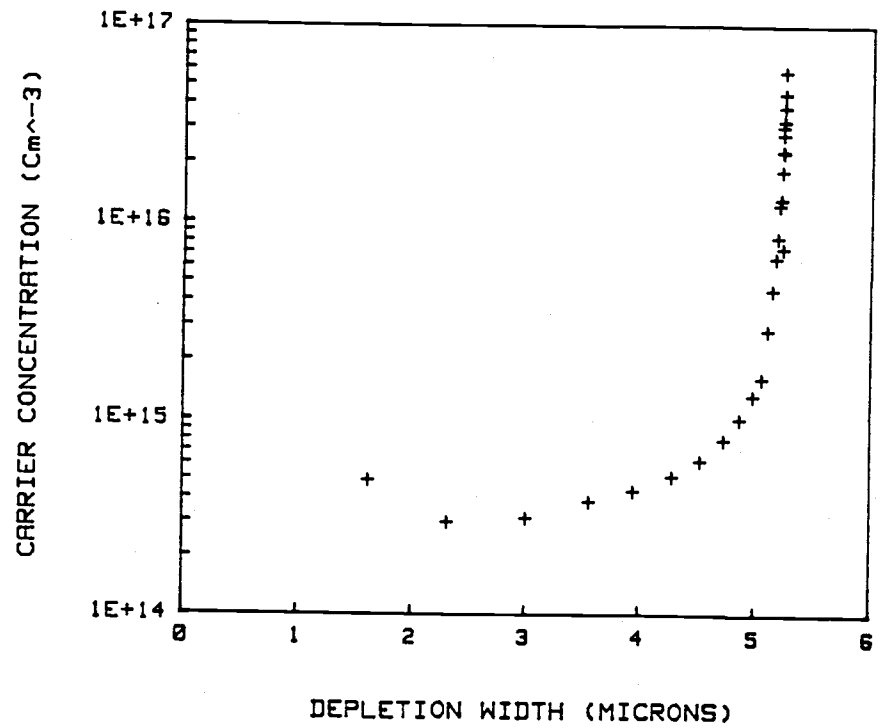


Figure 35d) Free carrier profile of a 13.5 μm epi (control) diode.

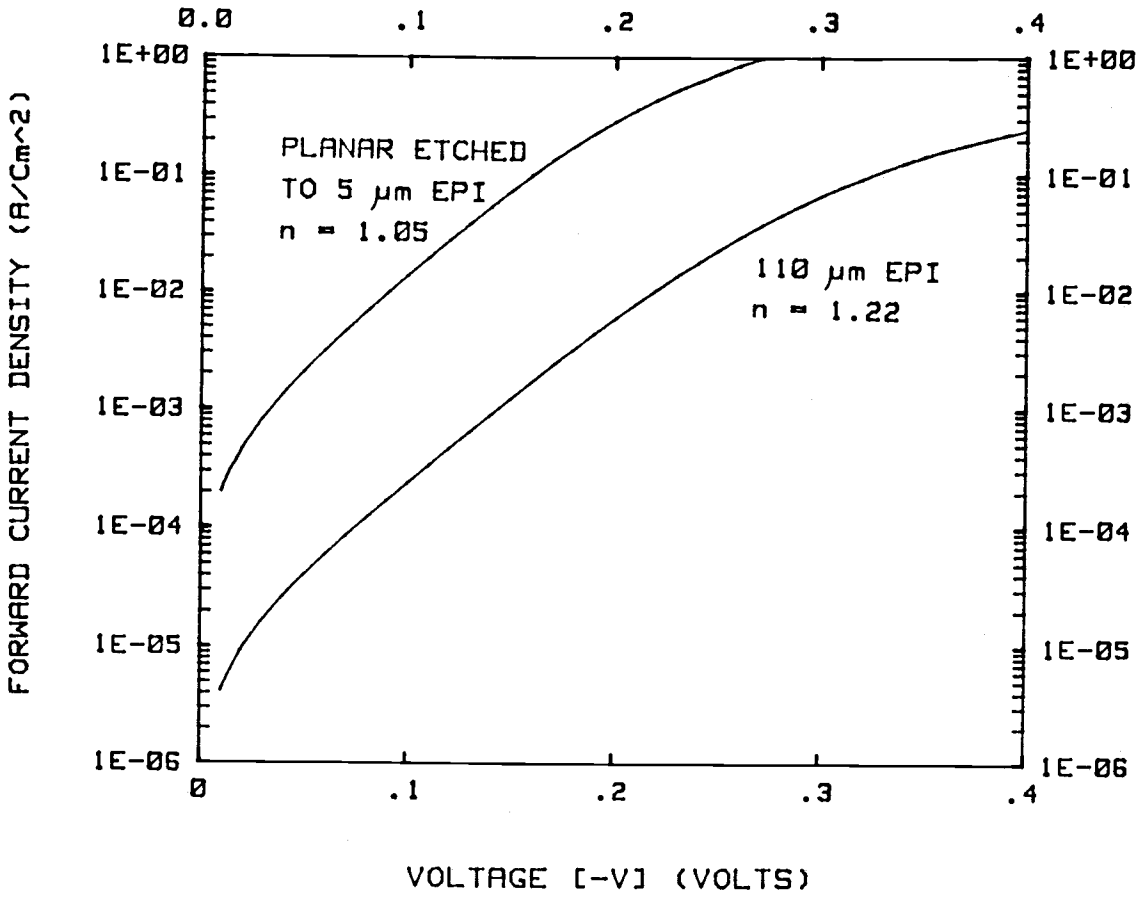


Figure 36) Plot of ideality factor of Schottky diodes.

where J_s is the saturation current density and n is the ideality factor. The current density versus forward voltage of the two diodes are shown in Fig. 36. The ideality factor was calculated from the following relation.

$$n = q/KT \left[\delta V/\delta (\ln J) \right] . \qquad (37)$$

The ideality factor for the 5 μm epitaxial sample is about 1.05 and that of the 110 μm epi-layer about 1.22. The saturation current density, J_s, is obtained by extrapolating the current density from

the linear region of the semi-log plot to V=0. The saturation current density of the 5 μm epitaxial sample is more than an order of magnitude higher than that of the 110 μm epi-layer diode. This may be explained by the fact that the rough surface of the 5 μm sample which was planar etched from a 13.5 μm epi-layer caused a lower barrier height for the diode. Therefore, J_s which is an exponential function of barrier height, is much higher for the 5 μm epi-layer diode. The acceptance criteria for leakage current was 200 μA (or 25 mA/cm^2) at the highest reverse bias required in DLTS (typically checked at 10 volts).

2. C-V Characterization

C-V, $1/C^2$ versus voltage, and the free carrier concentration profile (see Appendix C) of the 110 μm and 5 μm layers are shown in Fig. 34 and 35. The 110 μm epitaxial sample shows a uniform doping of about 4×10^{14} cm^{-3} and a built-in voltage of -0.7 V. The built-in voltage of the 5 μm epitaxial sample is -0.44 V in agreement with the lower barrier height observed from I-V characteristics. A large built-in voltage criteria ranging between -0.35 V and -0.8 V was set because of different metals and surface conditions used in this work. C-V characterization show reasonable Schottky diode characteristics for DLTS study. Figure 36c shows the epi-layer - p^+ profile of the planar etched sample. This profile is a representative of a epi-layer - substrate prior to any heat treatment. Approximately 7 to 8 μm of the epitaxial layer was removed (as shown in the figure) by the planar etch in an attempt to probe the p/p^+ interface by DLTS.

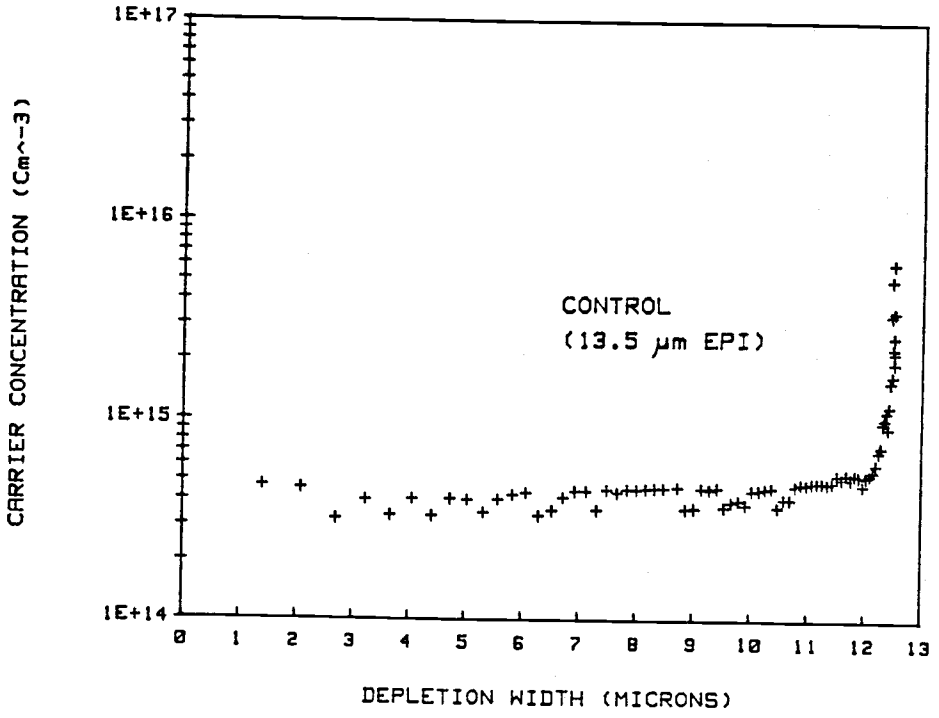


Figure 37a) Carrier profile of a control sample.

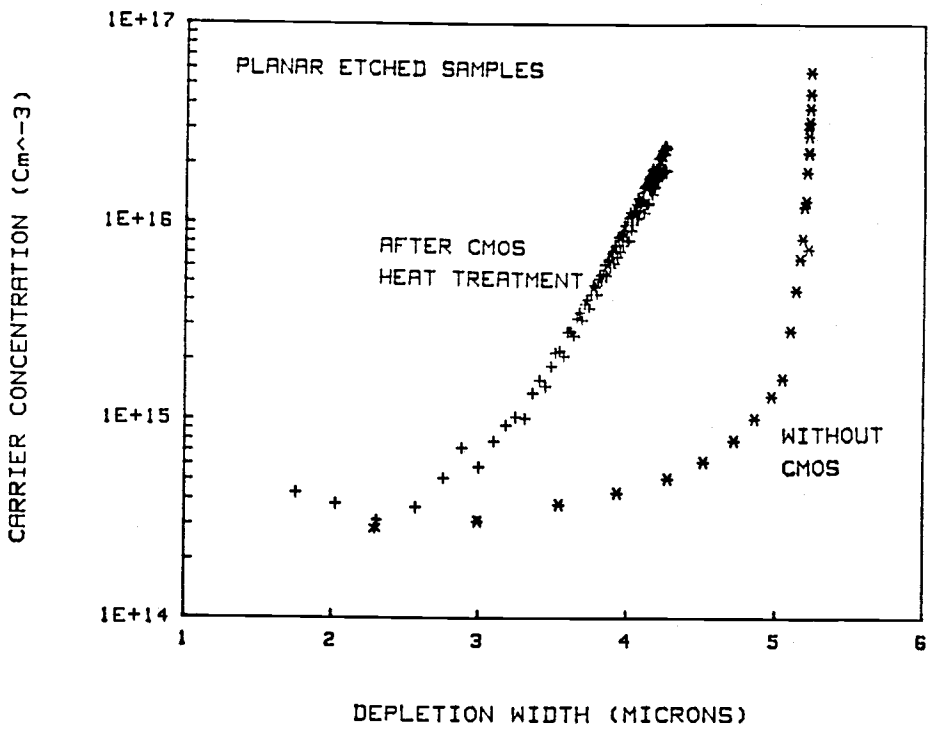


Figure 37b) Carrier profile of planar etched samples with and without the MOS cycle.

However, it was impossible to completely drive the depletion region into the heavily doped p^+ substrate which was doped in the range of $5x10^{18}$ to 10^{19} cm^{-3}.

One simple way of examining the p/p^+ profile is by C-V measurement of Schottky diodes fabricated on an epitaxial wafer. This profile could play an important role in oxygen out-diffusion from the substrate during high temperature processing treatments. The carrier profile of a control sample with 13.5 μm epitaxial thickness (specified by the vendor) is shown in Fig. 37a. A reverse bias of up to 40 V was applied to obtain this profile. The carrier concentration of a sample which experienced the CMOS heat treatment cycle of Fig. 13a is shown in Fig. 37b. Parts of the epitaxial layer were removed in order to probe the p/p^+ interface at a much lower reverse voltage. A profile of the 5 μm epitaxial sample of Fig. 35d without any heat treatment is also plotted for comparison. The two samples in Fig. 37b had different amounts of the epitaxial layer removed. Therefore, the out diffusion of p^+ into the epi-layer cannot be estimated from this figure.

D. DEEP LEVEL TRANSIENT SPECTROSCOPY (DLTS)

Schottky diodes were fabricated and used for DLTS measurements on p/p^+ epitaxial wafers. They are preferred over pn junctions since they do not introduce additional high temperature steps in the process or impurities in the material which complicates the analysis. Trap levels in the bandgap of silicon act as efficient generation-recombination centers and are detrimental to any dynamic charge storage device. In order to verify that epitaxial layers are free of any deep levels, DLTS measurements were employed on Schottky diodes fabricated on various p/p^+ epitaxial wafers. The main objective of the DLTS measurements was to detect any deep levels that are introduced by oxygen precipitation or bulk stacking faults into the epitaxial layer from the p^+ substrate. A band of energies due to oxygen precipitates in bulk wafers have been reported [62,70] which are quite different from discrete trap levels normally observed due to metals like Au, Cr, or Cu in silicon. It was noticed [70] that these levels were field dependent perhaps due to Poole-Frenkel barrier lowering. Majority and minority carrier DLTS results on p/p^+ epitaxial wafers are discussed in the next two sections, followed by DLTS results on Fe implanted sample.

1. Majority carrier DLTS measurements of p/p^+ epitaxial wafers

No significant DLTS peak was observed on p-type epitaxial layers of any thickness over a large temperature range of 40-300 K. Various

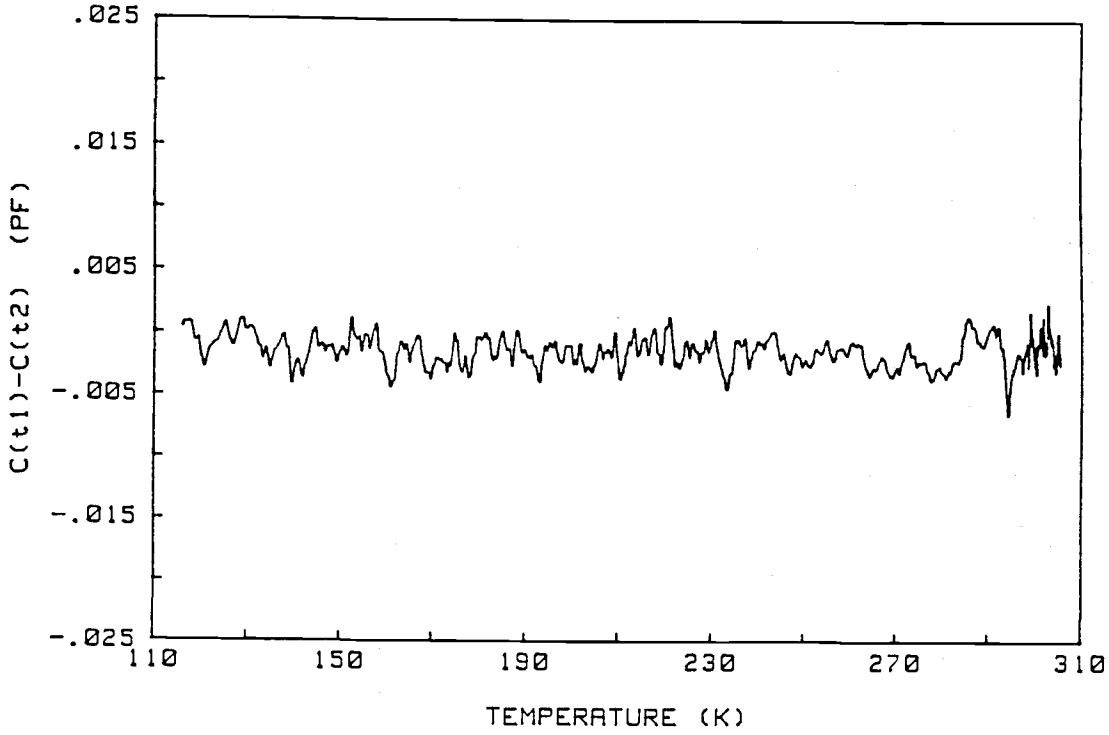


Figure 38) Typical majority carrier DLTS spectra of a 13.5 μm epitaxial Schottky diode.

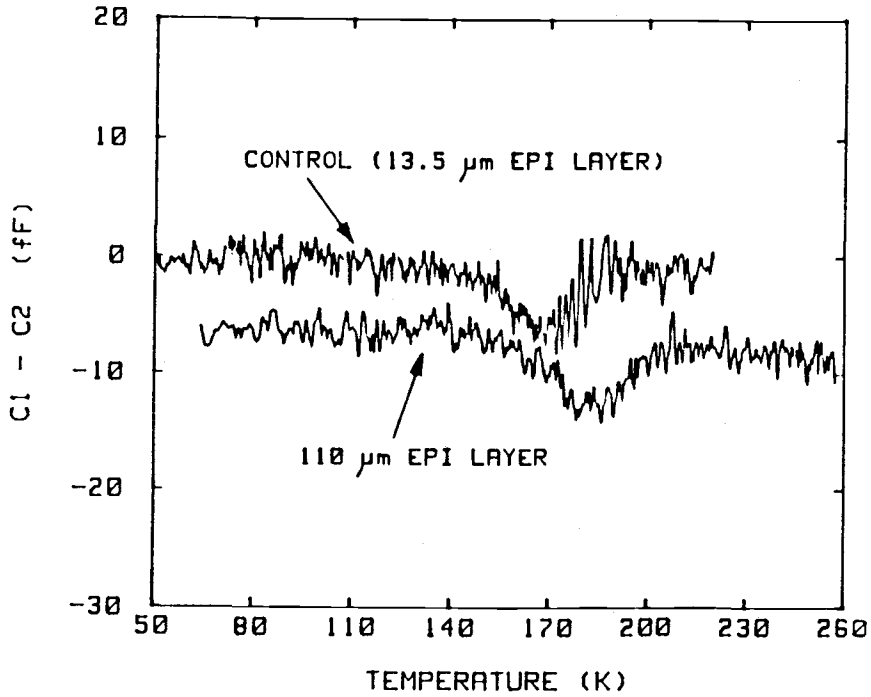


Figure 39) Majority carrier DLTS spectra of a control and a 110 μm epi-layers showing an extremely small and broad peak.

voltage pulses were applied to examine different volumes of the epitaxial layer. Figure 38 shows a typical DLTS spectra for a control (13.5 μm) p/p^+ epitaxial wafer. On some occasions majority carrier DLTS revealed only an extremely small and broad DLTS peak at about 170-190 K. This is shown in Fig. 39 where the t1 and t2 rate window was set at 10 and 40 msec (or an emission rate of 46.2 sec^{-1}). Results of both heat treated and as received samples indicate that the trap concentration within the epi-layer is essentially below the sensitivity limit of our DLTS setup. The sensitivity of the DLTS setup used in this study is on the order of 10^{-4} or approximately 1 part per 10^4 impurity boron atoms. Therefore, DLTS analysis was impossible, since first of all the signal was extremely small and comparable in magnitude to the noise and secondly, the peak temperature was difficult to obtain accurately due to the broad nature of the DLTS signal (see Fig. 39). The DLTS signal of $\delta C/C$ for p/p^+ epitaxial wafers was about $4x10^{-4}$. Assuming a simple exponential transient and a doping concentration of $4x10^{14}$ cm^{-3} results in a majority carrier trap concentration of $3x10^{11}$ cm^{-3} (Eqn. (34)). In other words, the p-type epitaxial layer has a minority carrier trap concentration of less than about $3x10^{11}$ cm^{-3}.

The main purpose of the planar etch treatment of some epitaxial wafers was to be able to punch the depletion region of the diodes into the epi-layer - p^+ interface or p^+ substrate for DLTS evaluation. A significant concentration of deep levels was originally expected to be located at the p^+ interface due to a high concentration of oxygen precipitation and secondary defects (commonly called bulk stacking faults) [55]. Therefore, various majority

carrier voltage pulses were employed for DLTS measurements to probe different depths of the epitaxial material, from a few microns below the surface to the p^+ interface. Unfortunately, the heavily doped p^+ substrate (about 10^{19} cm^{-3}) made it impossible to punch through it. During the epitaxial growth the boron atoms out-diffuse into the epitaxial layer making it impossible to detect the electrical activity of oxygen precipitates or other impurities at the p^+ substrate using DLTS.

2. Optical DLTS of the p/p^+ epitaxial wafers

Since Schottky barrier diodes are limited to majority carriers only, optical means were employed to inject electron-hole pairs into the epitaxial layers as was discussed previously. The optical DLTS signal of a 110 μm epi-layer with different rate windows, as shown in Fig. 40, reveals a minority carrier trap at about 150-160 K. The Schottky diode was biased at zero volts for this DLTS test. The Arrhenius plots of the 110 μm epi-layer and a control sample (13.5 μm layer) are shown in Fig. 41. The control sample was reversed biased at 0.15 V. The capacitance transient due to emission of minority carriers observed on the oscilloscope was not quite exponential. Therefore, the following discussion, based on the assumption of an exponential transient, is only qualitative. Activation energies of 0.21 eV and 0.28 eV below the conduction band for control sample and 110 μm epitaxial samples were obtained from the Arrhenius plots. Since both the 110 μm and the 13.5 μm epitaxial wafers had identical resistivity and growth conditions, it is reasonable to assume that

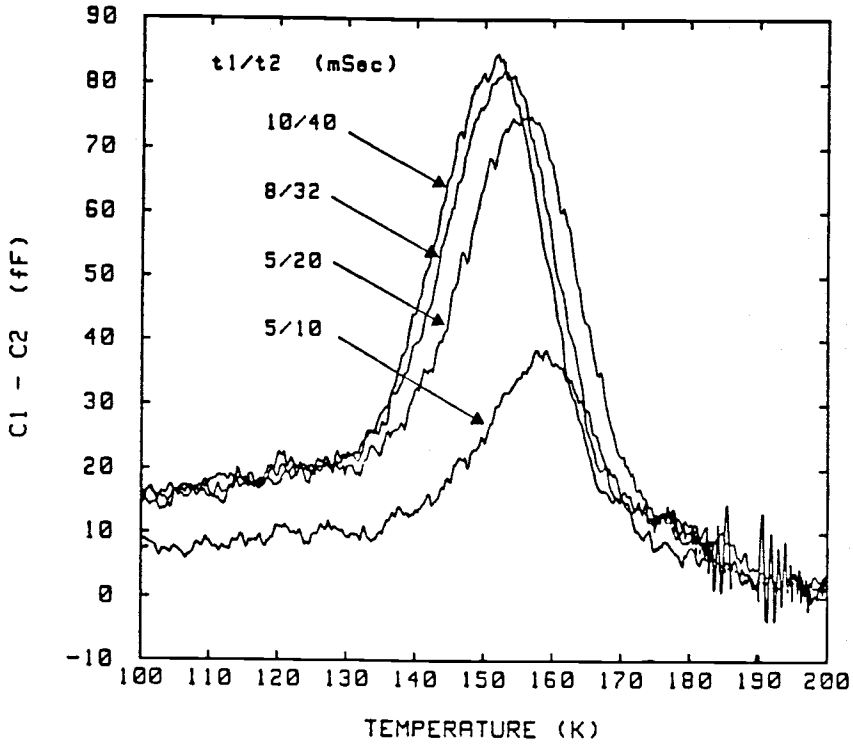


Figure 40) Optical DLTS signal of a 110 μm epi-layer diode.

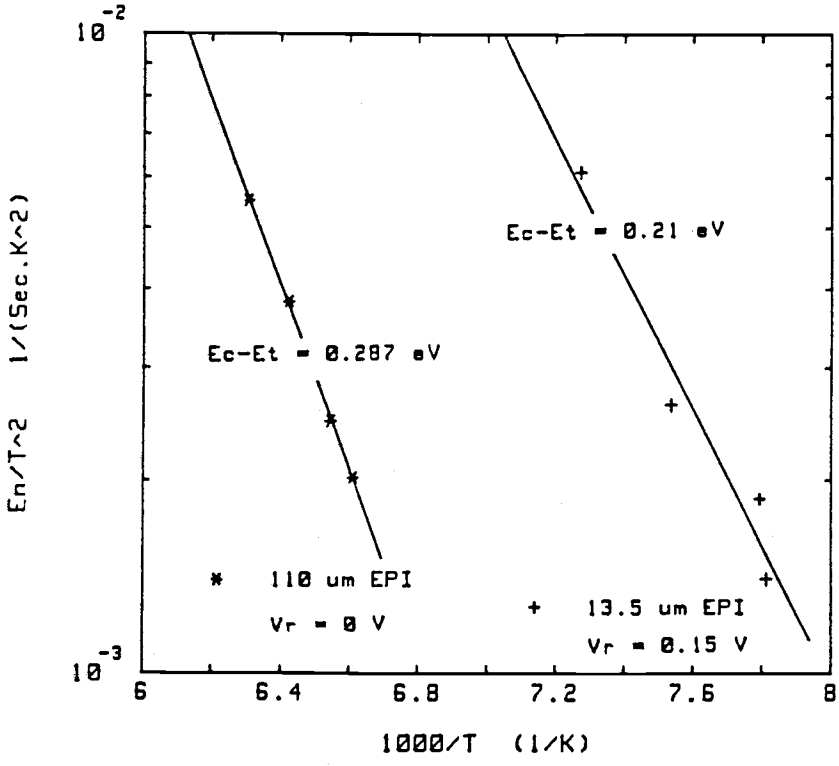


Figure 41) Arrhenius plot of the 110 μm epi-layer diode biased at 0 V and a 13.5 μm control diode reverse biased at 0.15 V.

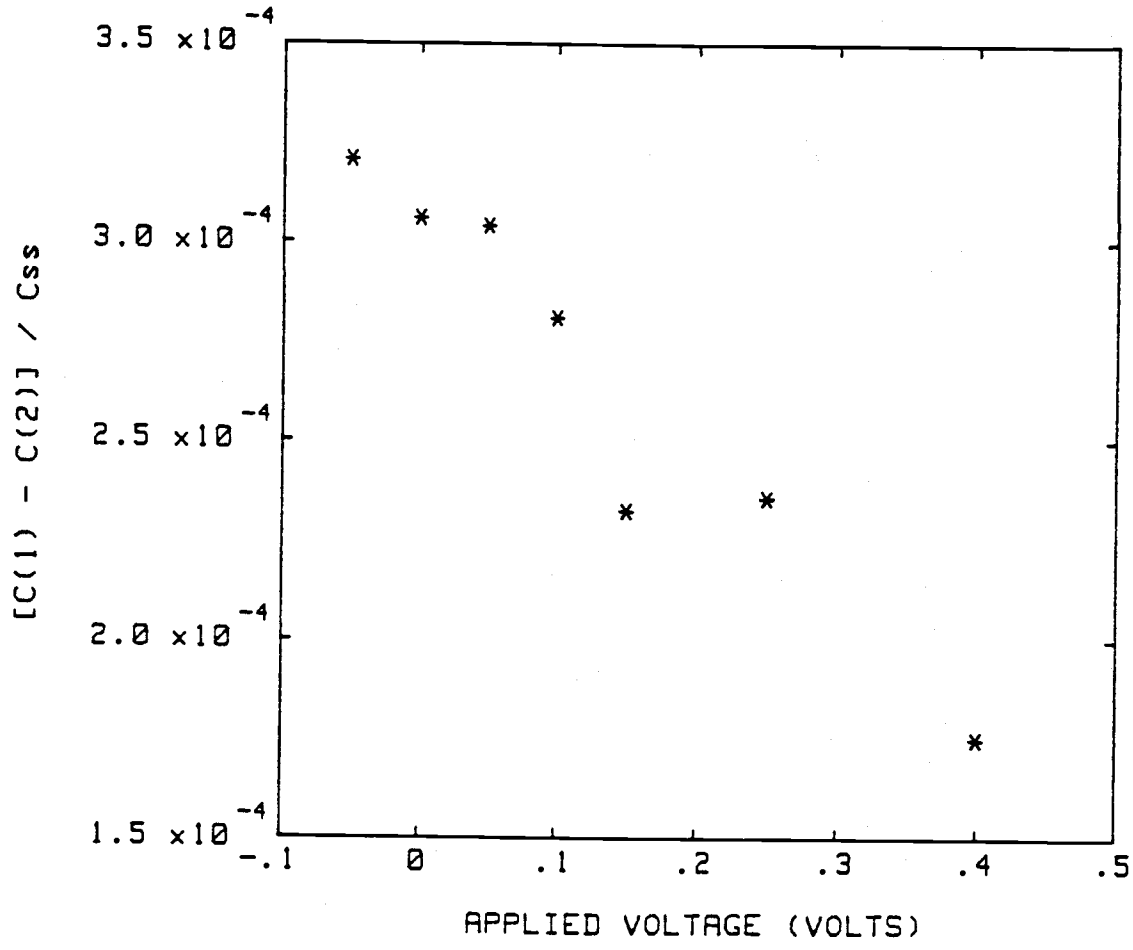


Figure 42) $\delta C/C$ as a function of applied reverse bias on the control Schottky diode.

the observed peaks are due to the same "trap".

For a fixed emission rate the peak height decreased as the reverse bias was increased on the diode. Figure 42 shows the decrease in $\delta C/C$ or "trap" concentration as a function of applied reverse voltage. At a reverse bias of 0.4 V the depletion region is 0.5 μm deeper than for zero volts of Fig. 42. This suggests the presence of a localized trap close to the surface whose concentration decreases as the electric field increases. Poole-Frenkel barrier lowering may be responsible for such a behavior. However, it can be argued that

the higher electric field in the depletion region causes fewer minority carriers to be captured by the localized trap because the minority carriers would be pulled toward the surface. Ti has been reported [25] to have a 0.21 eV energy level below the conduction band of silicon. Ti is suspected to be responsible, since the activation energy for the minority carrier trap observed on p-type epi-layers is very close to the Ti level which was sputtered for Schottky diode formation and secondly, the peak height decreased very rapidly as the reverse bias was increased. In addition, optical DLTS measurements on the few good Al Schottky diodes made on p/p^+ wafers which were processed prior to devising the Ti/Al sputtering process did not reveal any DLTS peaks. The sputtered Ti on epitaxial wafers which was later annealed at 400°C for 30 minutes forms titanium silicide [71] on the surface. The titanium silicide may lead to the deep level observed by optical DLTS. The characteristics of Schottky diodes changes after annealing due to formation of titanium silicide at the surface. The anneal had to be performed to obtain reasonable C-V behavior of the diode. No other minority carrier traps were detected on the epitaxial layers over a wide temperature range. Therefore, the epitaxial layer itself seems to be free of deep levels or oxygen related defects.

3. Trap levels of a Fe implanted sample as an illustration

This section is added as a supplement for an illustrative example of DLTS measurements on an intentionally introduced deep level. Fe was ion implanted at 10^{11} cm^{-2} dose and 10 keV energy on

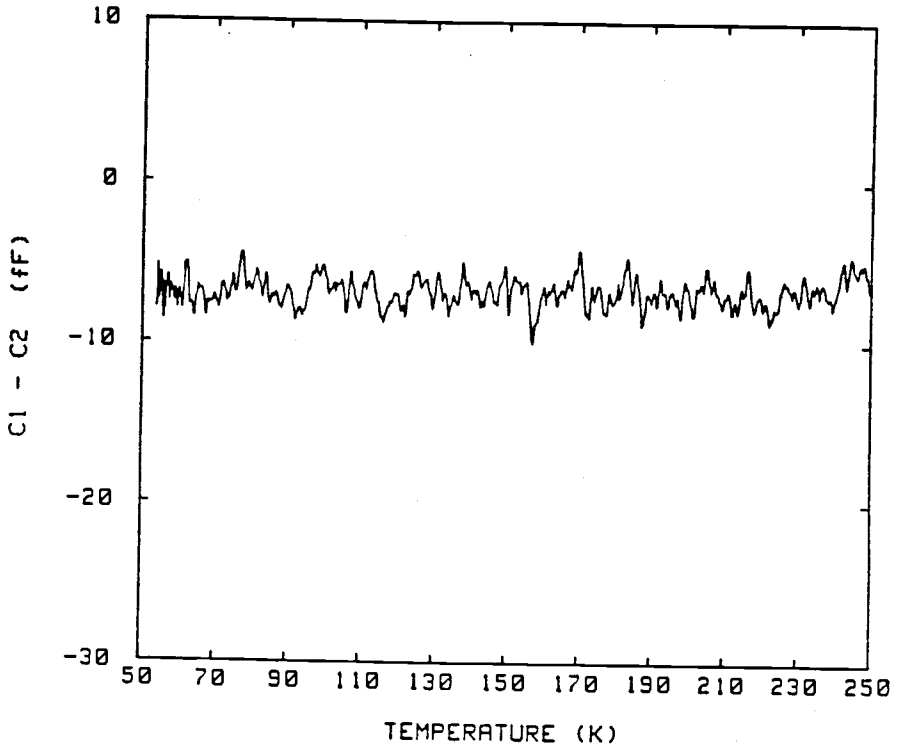


Figure 43) Majority carrier DLTS spectra of a Fe implanted diode.

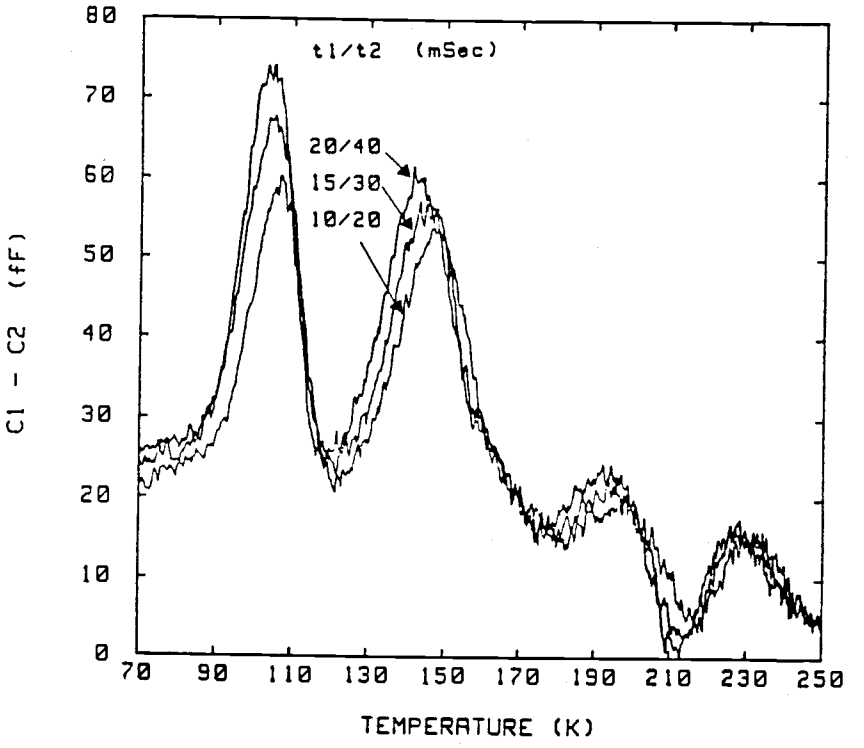


Figure 44) Optical DLTS spectra of a Fe implanted Schottky diode.

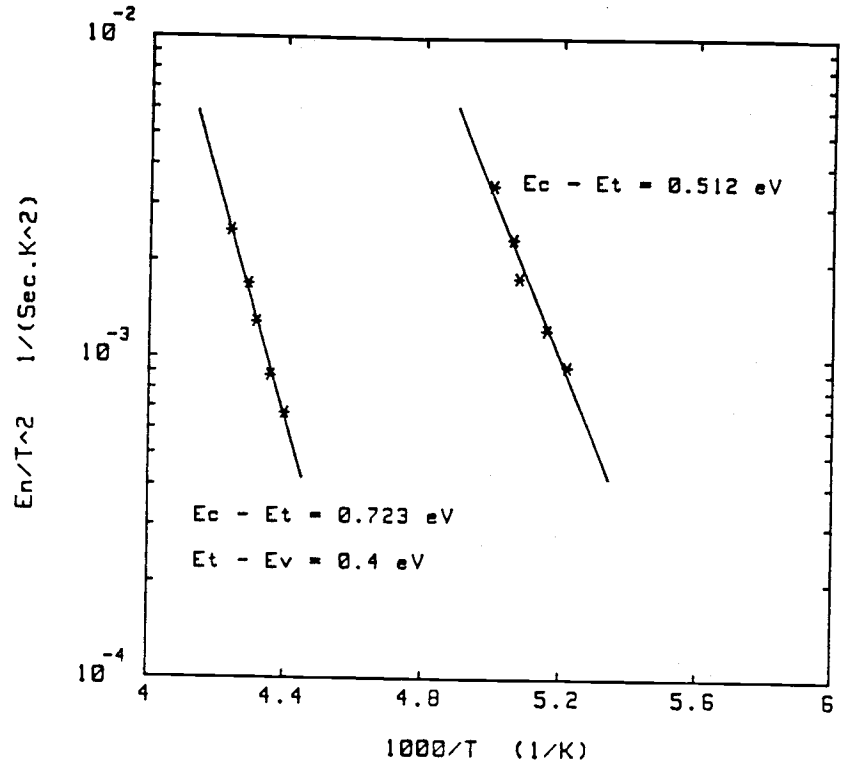


Figure 45a) Arrhenius plot of a Fe implanted diode.

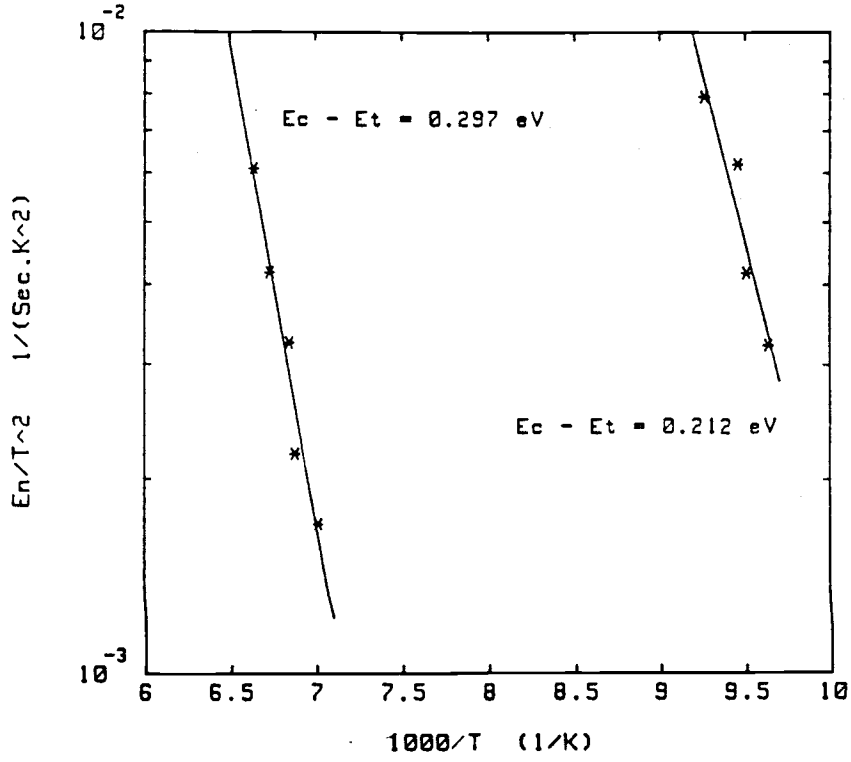


Figure 45b) Arrhenius plot of a Fe implanted diode showing the donor levels detected by optical DLTS.

20-30 Ω-cm p-type substrate. A p-type epitaxial layer of about 5-6 μm thick and the same resistivity as the substrate was grown on the substrate and the Fe. Since Fe is a fast diffuser in silicon with a diffusion coefficient [63] of 1.8×10^{-6} cm^2/sec it can easily diffuse through the epitaxial layer during the epitaxial deposition process which is normally performed at high temperatures in the range of $1100^{\circ}C$ to $1150^{\circ}C$.

The majority carrier DLTS signal of a Schottky diode fabricated on a Fe implanted sample using Ti is shown in Fig. 43. There are no majority carrier traps; however, optical DLTS for the sample shown in Fig. 44 for various rate windows revealed four different peaks which have different peak heights. As was discussed previously, it is difficult to be certain about trap concentrations using optical DLTS because it is not certain what percentage of traps are filled optically. An Arrhenius plot of these peaks is plotted in figures 45a and 45b. Three donor levels 0.21 eV, 0.3 eV, 0.51 eV below the conduction band and a donor level 0.4 eV above the valence band were observed. Donor levels of 0.51 eV below the conduction band and 0.4 eV above the valence band have been reported [25] for Fe in silicon. The 0.21 eV level is suspected to be due to the Ti, as was discussed previously.

Fe implanted samples were used to eliminate any doubts about the capability of detecting deep levels with the DLTS setup used here. It was demonstrated that the DLTS setup used was very sensitive and capable of detecting $\delta C/C$ as low as 10^{-4}. Therefore, it is safe to state that the p-type epitaxial layers are free from oxygen-related or metallic impurity deep levels at least down to a trap level of

$3x10^{11}$ cm^{-3}. This was expected since extremely high generation lifetimes were seen on these epitaxial wafers.

V. RESULTS AND DISCUSSION (II)

Recombination lifetime characterization of p/p^+ epitaxial wafers is discussed in detail in this chapter. Prior to the τ_r characterization of the p/p^+ epitaxial wafers, the recombination lifetime measurement technique of short base width devices is presented. Epi-layer - substrate interface generation, the temperature dependence of diffusion process, and effects of intrinsic gettering heat treatments on τ_r are also discussed in this chapter.

A. RECOMBINATION LIFETIMES OF p/p^+ EPITAXIAL WAFERS

1. Recombination lifetime of short base width devices using the pulsed MOS capacitance technique

For a given material, a relatively uniform recombination lifetime or diffusion length is expected to be independent of the device diameter. A plot of $1 - (C_i/C)^2$ versus time for a 294 μm and a 1256 μm diameter MOS capacitors on a p-type bulk wafer is given in Fig. 46. As expected the recombination lifetime of a polished wafer (long base width) using Schroder's [1] technique was 0.25 μsec regardless of the capacitor diameter size. This technique works very effectively for long base width devices; however, in case of short base width devices this technique is very poor (discussed later). The C-t response of a capacitor with a 630 μm diameter on a p/p^+ epitaxial wafer measured at 70°C is shown in Fig. 47a. The C-t data are plotted as $1-(C_i/C)^2$ versus time in Fig. 47b. From the slope of

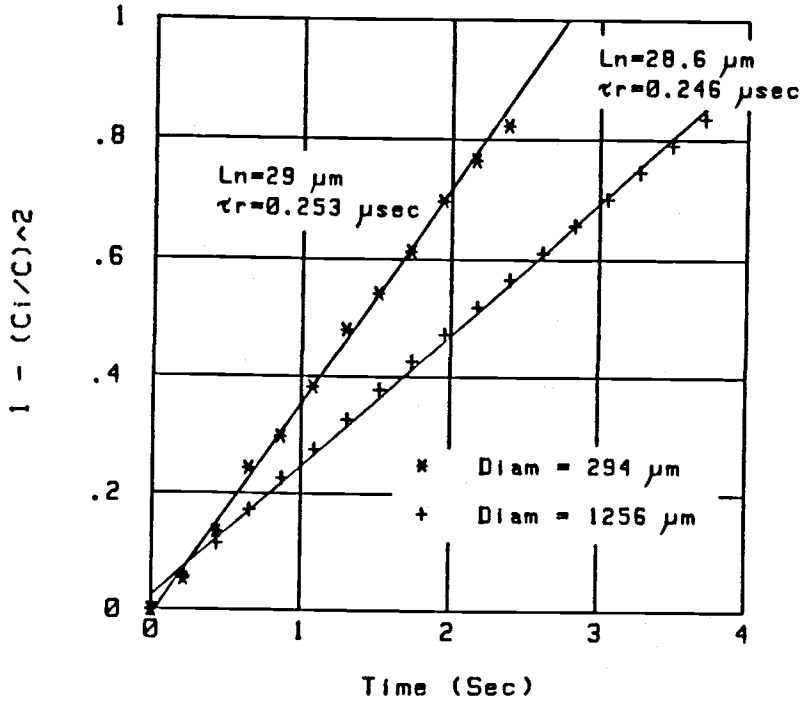


Figure 46) $1 - (C_i/C)^2$ versus time plot of p-type polished wafers for different diameters measured at T=55°C.

this curve an effective diffusion length of 450 μm was obtained using Eqn. (28). Schroder's relation for a short base width device ($L_n^2 = L'_n.W_B$) translates into an actual diffusion length of 78 μm. This technique was repeated on various diameter MOS capacitors and the diameter decreased the measured value of diffusion length also decreased. This implies that the lateral bulk diffusion current becomes significant for smaller samples as was shown in Fig. 7 in Chapter II. Diffusion lengths versus MOS capacitor diameter calculated by both Schroder's technique, and the technique introduced in Chapter II which takes the lateral component of carriers into account are illustrated in Fig. 48. The technique with the lateral bulk correction gives a relatively uniform diffusion length of about

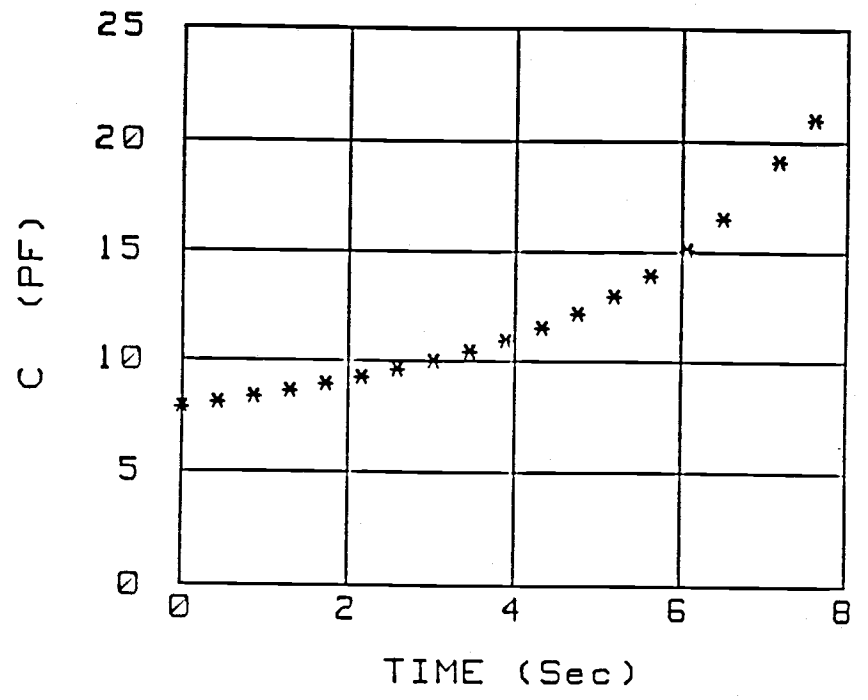


Figure 47a) C-t transient response of a MOS capacitor on p/p^+ epitaxial layer measured at T=70°C. The device diameter is 630 μm.

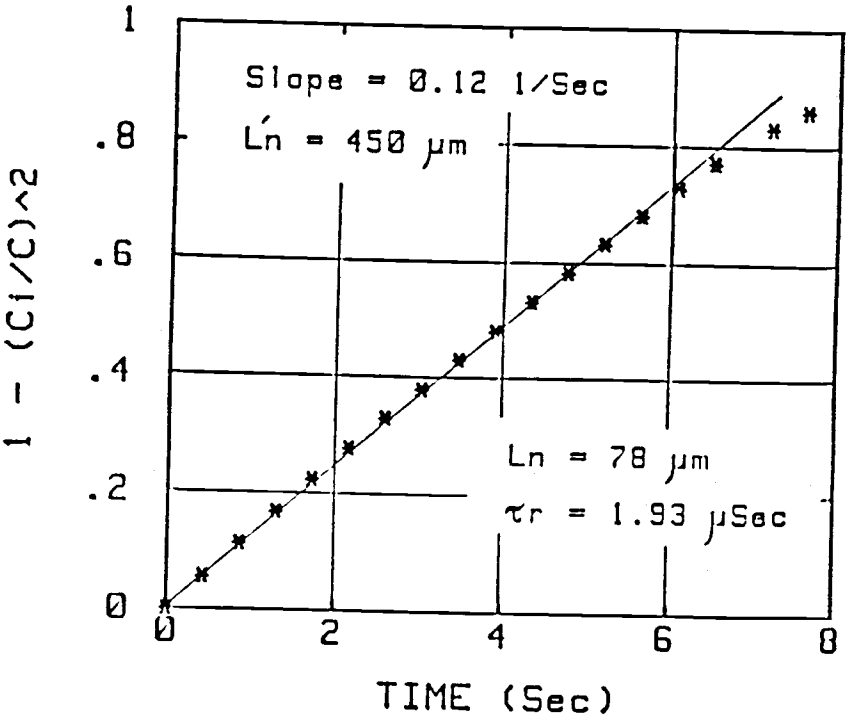


Figure 47b) $1 - (C_i/C)^2$ versus time calculated from C-t response shown in a).

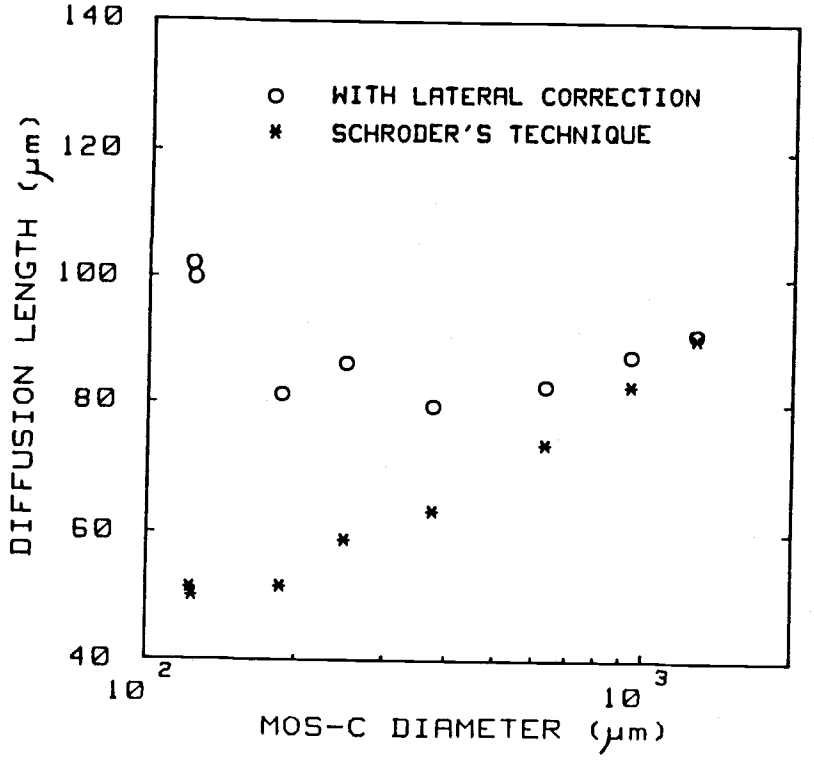


Figure 48) Calculated diffusion length versus device diameter. Schroder's technique indicates that the diffusion length is apparently a function of the device diameter. The technique with lateral bulk correction indicates that the diffusion length is relatively constant.

90 μm independent of the capacitor diameter. The diffusion coefficient of the minority carrier electrons at 70°C is about 31.5 cm^2/sec [72], therefore resulting in a recombination lifetime of 2.6 μsec ($\tau_r = L_n^2/D_n$).

The recombination lifetime of epitaxial wafers is an order of magnitude higher than the bulk wafers. This is expected since it is well known that epitaxial wafers have much less metallic impurities or point defects than the bulk wafers. The metallic impurities and point defects act as effective recombination - generation centers resulting in degraded recombination and generation lifetimes.

The diffusion constant, D_n, and intrinsic carrier density are

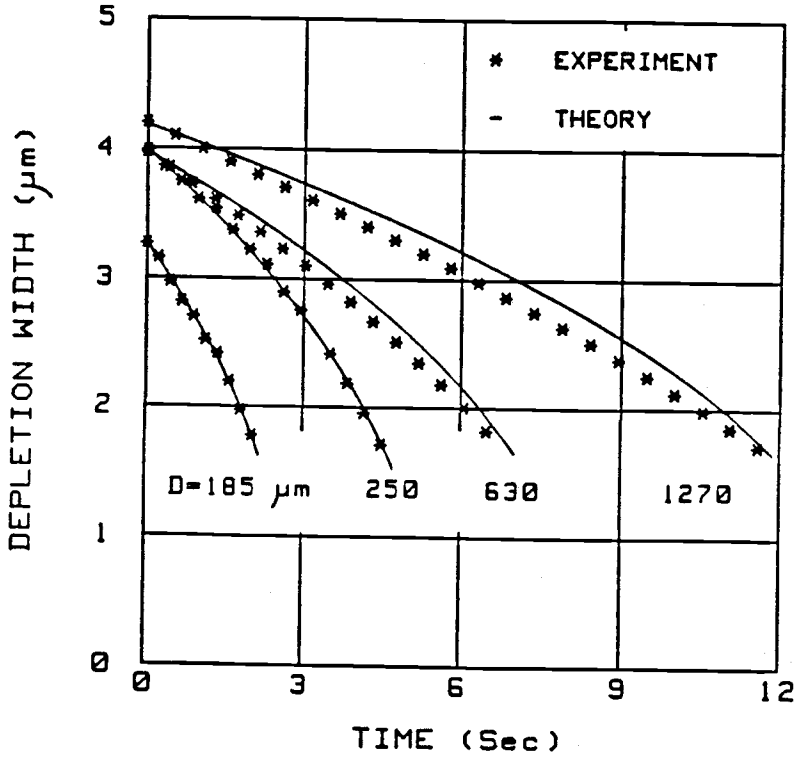


Figure 49) Depletion width versus response time of MOS capacitors of different size devices on the same sample (T=70°C).

calculated from the following semi-empirical expressions [72,73]

$$D_n = \mu_n \, KT/q = 37(300/T)^{1.2} \tag{38}$$

$$n_i = 3.87\text{x}10^{16} \, T^{1.5} \exp(-1.21/2KT) \; . \tag{39}$$

The expression for diffusion constant is valid for low doping concentrations of the order of 10^{14} - 10^{15} cm^{-3}. The diffusion length, L_n, was calculated using Eqn. (30) and the C-t data as follows. The initial and final values of the space charge region W_i and W_f and the final time or retention time, t_f, were obtained from the C-t response at T=70°C. All of the capacitor parameters required for evaluation of L_n in addition to its diameter size were

substituted into Eqn. (30) and L_n was solved iteratively using the secant method. Subsequently, the value of L_n was substituted back in Eqn. (30) and a W-t plot was created for comparison with the measured transient response. Figure 49 shows W-t values calculated from the measured C-t data and W-t values calculated from iterative calculation of L_n for different MOS capacitor diameters. The W-t relation predicted by Eqn. (30) seems to match the experimental data very closely, especially for small diameters where the lateral quasi-neutral bulk generation is significant. The percentage error is less than 6% for the MOS capacitor of 1270 μm in diameter.

Although large diameter samples could theoretically be used to measure τ_r for short base width devices ($W_B < L_n$) without a significant lateral correction, for thin epitaxial layers with high diffusion lengths the required diameters would be too large for practical applications. For instance, for the lateral current component to be less than 25% of the total current (see Eqn. (30)) in a 5 μm epitaxial layer with a diffusion length of 200 μm would require a MOS capacitor of 0.4 cm in diameter (assuming an average W(t) value of 2 μm). Therefore, the lateral diffusion of minority carriers cannot be neglected for a proper recombination lifetime measurement of a high quality silicon wafer with a short base width using any practical size MOS capacitors.

While Schroder's technique of recombination lifetime measurement works well for long base width polished silicon wafers, it underestimates the value of τ_r for short base width samples. A simple one dimensional model developed here is more general and includes the lateral quasi-neutral bulk generation and the time dependence of the

space charge region in short base width devices. The technique introduced here results in a relatively uniform recombination lifetime independent of the device diameter. Even though Eqn. (30) is an approximation it predicts the transient response of pulsed MOS capacitors at elevated temperatures closely. All of the recombination lifetime data reported in the following is obtained using the technique introduced here on the largest size MOS capacitor to minimize the lateral contribution.

2. Recombination lifetime of 13.5 μm p/p^+ epitaxial control wafers

The study of recombination lifetime is mainly focused on as received epitaxial wafers which is discussed in this section. Effects of intrinsic gettering and CMOS simulation heat treatments on τ_r are considered in section 6. Numerous control wafers with epitaxial resistivity in the range of 35 to 40 Ω-cm were processed. These control wafers received a dry oxidation at 900°C for 4 hours which was followed by 4 hours anneal in N_2 at the same temperature. Figure 50 shows a histogram of recombination lifetimes measured at T=70°C on 13.5 μm p/p^+ wafers with a resistivity of 40 Ω-cm. The diffusion length for the control samples was about 80 μm. The mean lifetime was 1.8 μsec. Another set of measurements on 35 Ω-cm epitaxial wafers resulted in mean τ_r of 2.3 μsec. The recombination lifetimes were easily reproduced on different process runs, but since DLTS measurements did not reveal any significant traps, much higher recombination lifetimes were expected. In fact, if we assume a capture cross section of 5×10^{-15} cm^{-2}, typical for some discrete deep

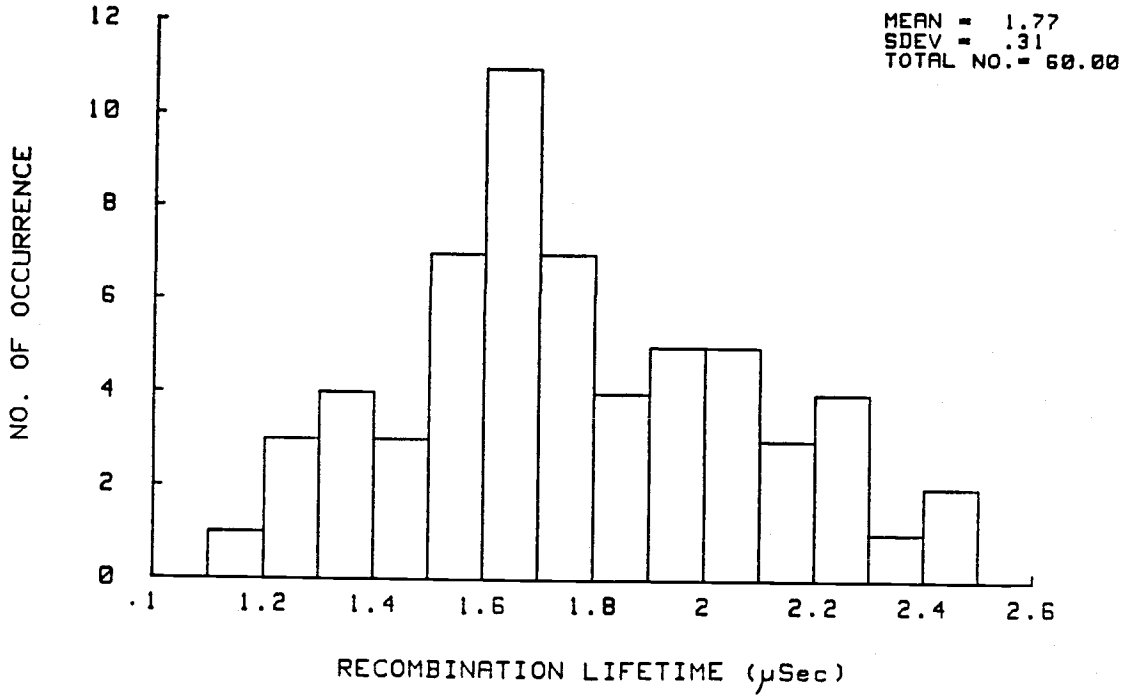


Figure 50) Histogram of the apparent recombination lifetime of a 40 Ω-cm control (13.5 μm) epitaxial sample.

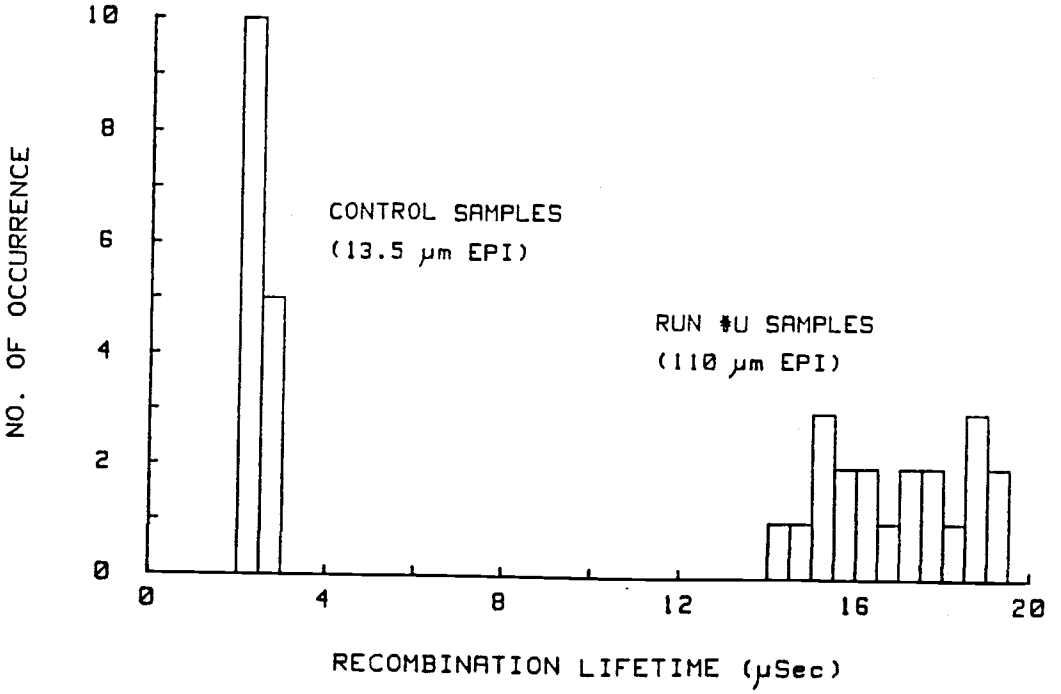


Figure 51) Histogram of the apparent recombination lifetime of a control and a 110 μm epi-layer with 35 Ω-cm resistivity.

levels like gold, and a trap concentration of $3x10^{11}$ cm^{-3} (from DLTS), the recombination lifetime ($\tau_r = 1/(\sigma_n v_t N_t)$) is of order of 50 to 70 of μsec. Since the generation lifetimes are very high and the surface generation velocity is quite small it is suspected that a significant generation of carriers occurs, perhaps at the epi-layer - substrate interface, in addition to the diffusion within the epi-layer causing the lower τ_r values seen on 13.5 μm epitaxial samples.

3. Epitaxial layer - p^+ substrate interface generation

So far the epi-layer - p^+ interface was assumed to be ideal with zero carrier generation or interface generation velocity due to the built-in field of epi-layer - p^+ junction. Since the diffusion coefficient of boron is fairly high a significant redistribution of impurities from the heavily doped substrate occurs during epitaxial growth at 1100-1200°C. The heavily doped boron in the substrate can out diffuse up to 1.5 to 2 μm during epitaxial growth [6,47]. For substrate doping of about $5x10^{18}$ cm^{-3} and epitaxial doping of $4x10^{14}$ cm^{-3} the built-in electric field is about 1.2 kV/cm. Any further heat treatments results in more out-diffusion (see Fig. 37b), reducing the built-in field. Some minority carriers may be swept from the epitaxial layer - substrate interface due to this built-in electric field. In reality a significant current could be generated from this interface for various reasons. There may be contamination prior to the epitaxial deposition, dislocations, vacancies, or oxygen related defects at the epi-layer - p^+ substrate. The epitaxial layer itself is not the limiting factor in the recombination lifetime because very

good quality epi-layer was observed from both generation lifetime and DLTS measurements. Up to now the recombination lifetimes have been calculated under the assumption that no epi-layer - substrate interface generation occurs, so if the epi-layer - p^+ interface generation is significant, the recombination lifetime values are apparent values and not the actual epitaxial recombination lifetimes.

The total current density when diffusion, generation, surface generation, and epi-layer - p^+ interface generation (J_{pp+}) are present is the sum of four processes.

$$J_{total} \approx q \frac{n_i^2}{N_a} \frac{D_n}{L_n^2} W_B \left(1 + \frac{4L_n}{D} \right) + J_{pp+} + J_{gen} + J_s \,. \qquad (40)$$

The time dependence of the depletion region in the diffusion component (Eqn. (29)) is neglected since the epitaxial thickness, W_B, used for the epi-layer - p^+ interface study is much larger than W(t). At elevated temperatures the generation within the space charge region and surface generation (the last two terms) are negligible. According to Eqn. (40) the total current density due to relaxation to equilibrium in a short base width p/p^+ epitaxial wafer is linearly proportional to the epitaxial thickness, W_B, at elevated temperatures. Furthermore, if no significant epi-layer - p^+ interface generation is present, in a linear plot of J versus epitaxial thickness, W_B, the current density must be essentially $J_{gen}+J_s$ when the epitaxial thickness is extrapolated to zero. The main motivation for the τ_r measurements using different epitaxial thicknesses is to clarify whether the epi-layer - p^+ interface is responsible for the

low apparent recombination lifetime values.

Epitaxial layers as thick as 110 μm may be anticipated to have dislocation or mismatch problems which could reduce the generation and recombination lifetimes. However, the generation lifetime and DLTS measurements on thick epi-layers (110 μm) are identical to that of control samples which suggests that the epitaxial quality of thin and thick wafers are the same. A histogram of the apparent recombination lifetimes of 35 Ω-cm control and 110 μm thick epitaxial layers is shown in Fig. 51. The apparent recombination lifetime of thick epi-layer wafers is considerably higher than the those of control wafers. This shows that the thick epitaxial layers are free of mismatch or dislocation. The thick and thin epi-layer are carefully grown under identical processing conditions, so it is expected to have the same actual recombination lifetime for both thin and thick epi-layers.

A summary of measured generation and apparent recombination lifetimes of various samples used for the τ_r study is given in Table 5. The first four samples listed in Table 5 had identical epitaxial growth process except different epitaxial thickness. The generation lifetimes of these samples were all in the range of 3 to 5 msec as discussed earlier. However, the apparent recombination lifetimes increases as the epi-layer - thickness increases. This can be easily explained by Eqn. (40). By increasing W_B the relative magnitude of the diffusion current in the epitaxial layer increases compared to J_{pp+} (J_{gen} and J_s are negligible at elevated temperatures). Figures 52 and 53 are presented to show the contribution of the epi-layer - substrate interface generation. In Fig. 52 the current density

sample	Epitaxial thickness (μm)	Res. (Ω-cm)	Doping (cm^{-3})	τ'_g (msec)	Apparent τ_r (μsec)
p/p^+	110	35	$4.0x10^{14}$	3 - 5	16.9
p/p^+	70	35	$4.0x10^{14}$	3 - 5	9.1
p/p^+	13.5	35	$4.0x10^{14}$	3 - 5	2.41
p/p^+	5-6	35	$4.0x10^{14}$	3 - 5	0.8
p/p^+ from Vendor B	12	20	$6.0x10^{14}$	1 - 1.5	2.06
p-type polished	---	40	$3.0x10^{14}$	0.15-0.5	0.25
n/n^+	50	40	$2.5x10^{14}$	5 - 7	92
p-type epi Fe implanted	5	40	$3.0x10^{14}$	0.01-0.02	0.002
p-type Float Zone	---	100-150	$8x10^{13}$	----	40-60

Table 5. Summary of measured generation and recombination lifetimes.

calculated from the C-t data using Eqn. (19) is plotted as a function of the depletion width for different epitaxial-thicknesses. The total current density versus W_B is illustrated in Fig. 53. Each data point in this figure represents an average of 30 measured C-t data at 70°C. The error bars (minimum to maximum values) are shown in the figure. The apparent recombination lifetime would naturally be smaller than the actual epitaxial lifetime if any generation was present at the epi-layer - substrate interface. As can be seen in Figs. 52 and 53, when W_B is increased from 13.5 μm to 110 μm, J_{total} only increases by a factor of 3. From the slope of Fig. 53 the actual recombination

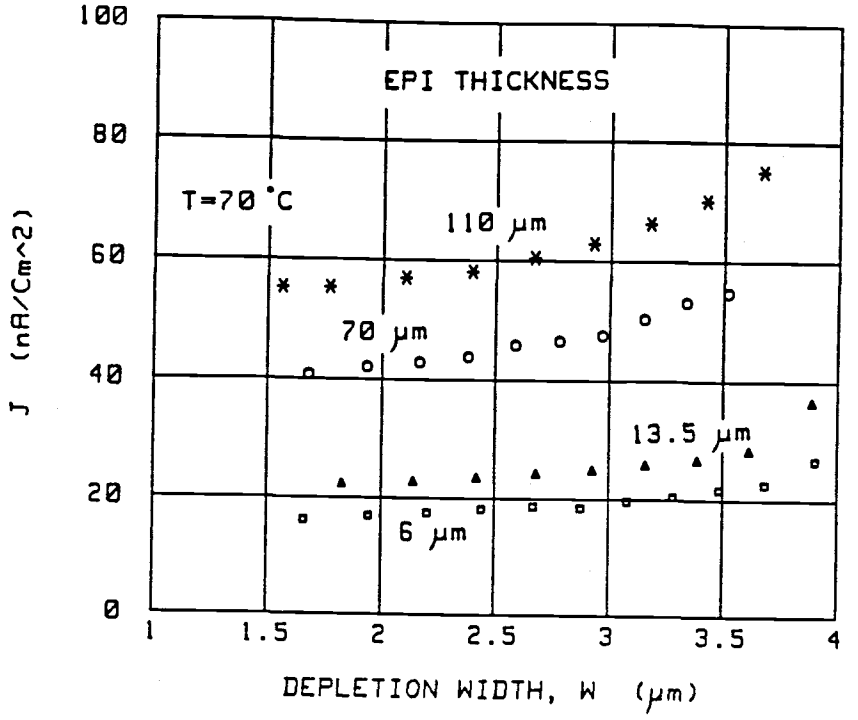


Figure 52) Current density versus depletion width for different epitaxial thicknesses measured at T=70°C.

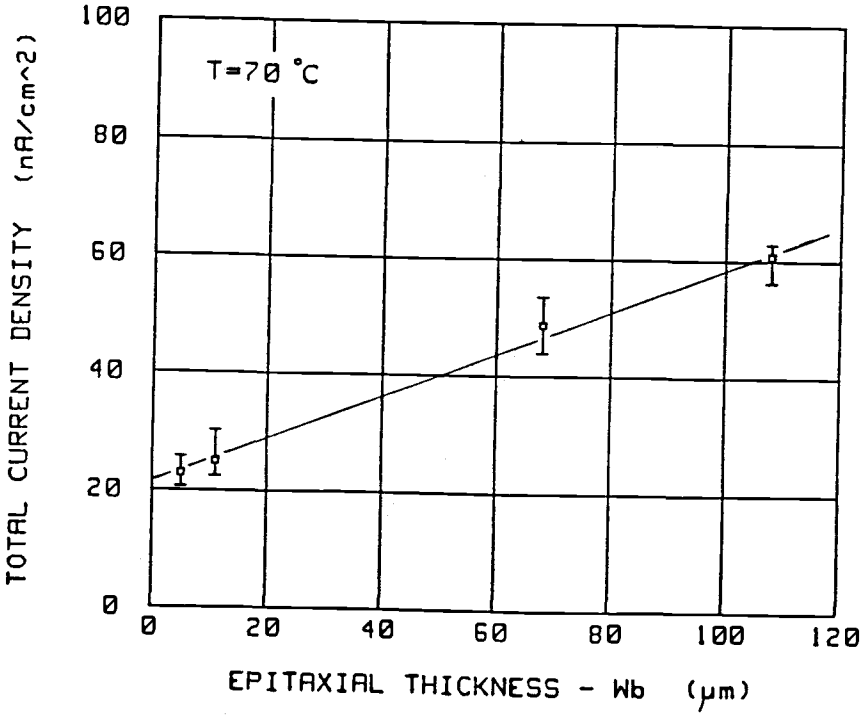


Figure 53) Current density as a function of epitaxial thickness at T=70°C.

lifetime of the epitaxial material may be obtained (Eqn. (40)). A slope of 3.74 $\mu A/cm^3$ from this figure for MOS capacitors of 1270 μm in diameter measured at 70°C results in an actual lifetime of 19 μsec. Note that as W_B is increased the apparent recombination lifetimes measured by a different method starts to approach the actual 19 μsec value. The current density from intercept of Fig. 53 is 21.5 nA/cm^2. The contribution of generation within the depletion region at 70°C is about 2 nA/cm^2 and Si-SiO_2 surface generation for the worst measured value of S_o contributes 2.5 nA/cm^2. Therefore, J_{pp+} is about 17 nA/cm^2. In the case of the 13.5 μm epitaxial layers, the epi-layer - p^+ interface therefore constitutes 70% of the total current and about 75% for the 6μm epi-layer . This suggests that the limiting factor in apparent recombination lifetimes (seen by any charge storage device) for a thin p/p^+ epitaxial wafer is the epi-layer - substrate interface rather than the epitaxial layer itself.

A 20 Ω-cm, 12 μm epi-layer, p/p^+ wafer from a different vendor (Table 5) was studied for comparison. The generation lifetimes of these wafers were about 1 msec which was considerably smaller than τ'_g values of vendor A. Originally it was expected to have τ_r values quite different under the assumption that large variations exists among vendors in epitaxial growth conditions. However, τ_r values measured on 30 devices (from vendor B) averaged a 2.06 μsec which is very close to the τ_r values obtained from the control samples. The fact that τ_g values of different vendors are different but τ_r values are not substantiates the assertion that the epitaxial quality is not the dominant factor in limiting the recombination lifetime of a thin epitaxial wafer. The lifetime of a p-type polished wafer and a Fe

implanted sample are tabulated in Table 5 for comparison.

Oxygen free float zone samples showed τ_r values of approximately 40-60 μsec. Accurate determination of τ_r values for the float zone samples was not possible because of some back side ohmic contact problems due to the high resistivity of these samples. n/n^+ epitaxial wafers show slightly better generation lifetimes than the p/p^+ wafers. However, in contrast to the low τ_r values obtained on p/p^+ wafers, the recombination lifetimes of 30 measured samples averaged 92 μsec on n/n^+ wafers. The reason for such a large difference in n-type and p-type layers is discussed later.

4. The temperature dependence of diffusion process

As mentioned earlier, the C-t response is expected to be dependent on n_i at room temperature and n_i^2 at elevated temperatures. The experimental data for a control and a 110 μm epi-layer are shown in Fig. 54 where the retention times are proportional to inverse n_i at about room temperature and proportional to n_i^2 at elevated temperatures. If no epi-layer - p^+ interface generation is present, the retention time of the 110 μm epi-layer is expected to be about 8 times lower than the 13.5 μm control sample (see Eqn. (40)). The retention time is only about 3 times lower in the 110 μm epi-layer. The reason for the transition of n_i to n_i^2 dependence which occurs at lower temperatures for the thicker epitaxial sample is because the diffusion term which is linearly proportional to the epi-layer thickness is considerably larger for the 110 μm epi-layer.

The epi-layer - p^+ interface generation, which is very

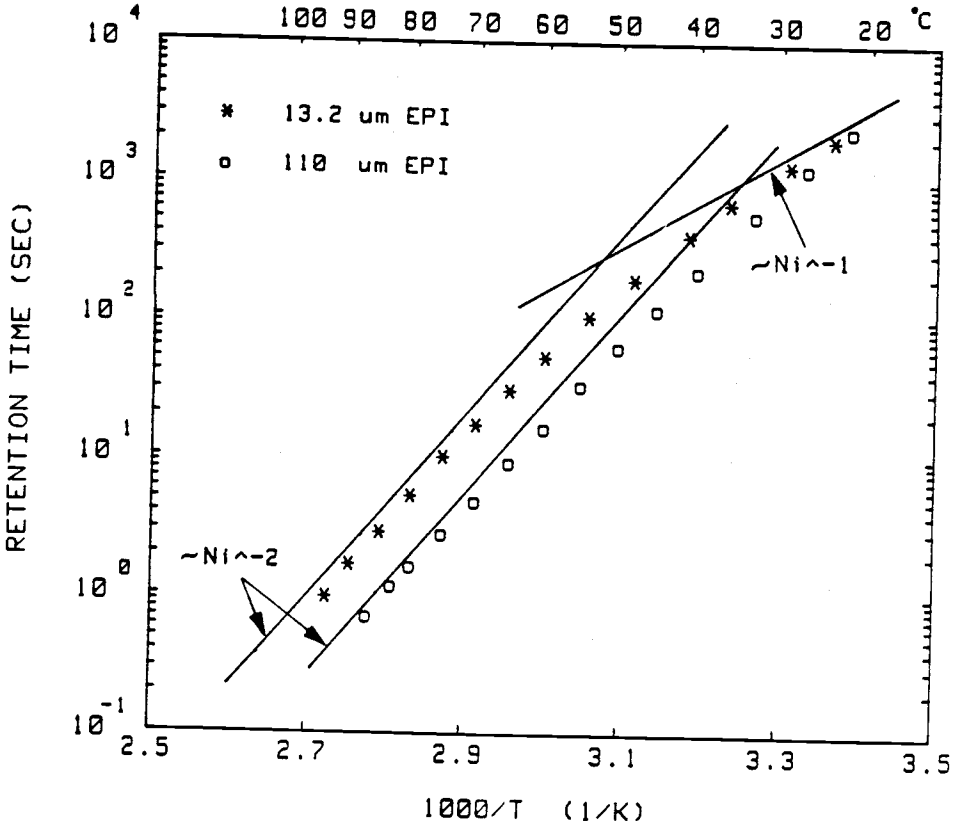


Figure 54) Retention time of C-t transient response as a function of inverse temperature for a control and 110 μm epi-layer.

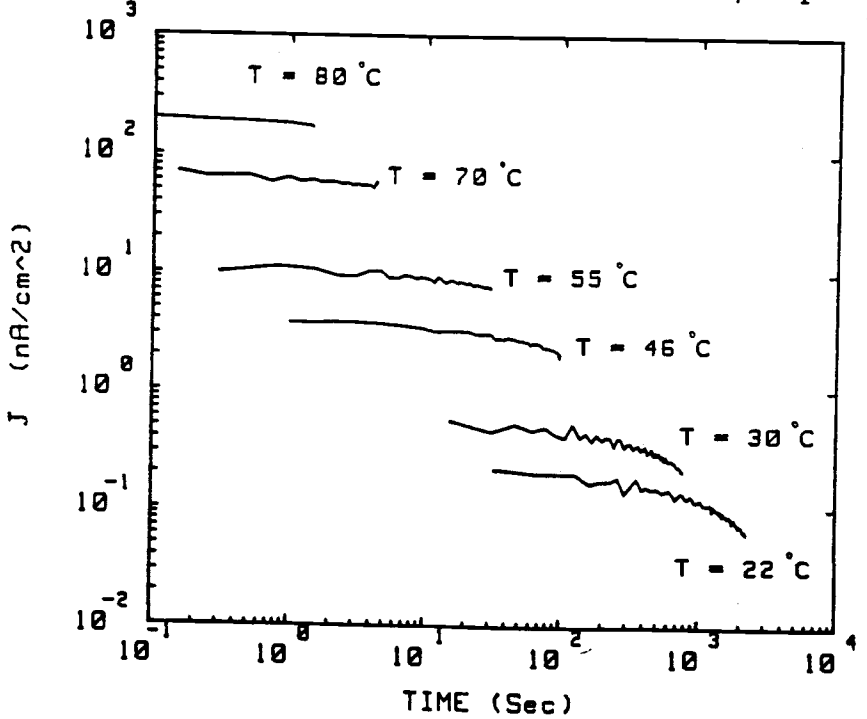


Figure 55) Current density as a function of transient response time for various temperatures.

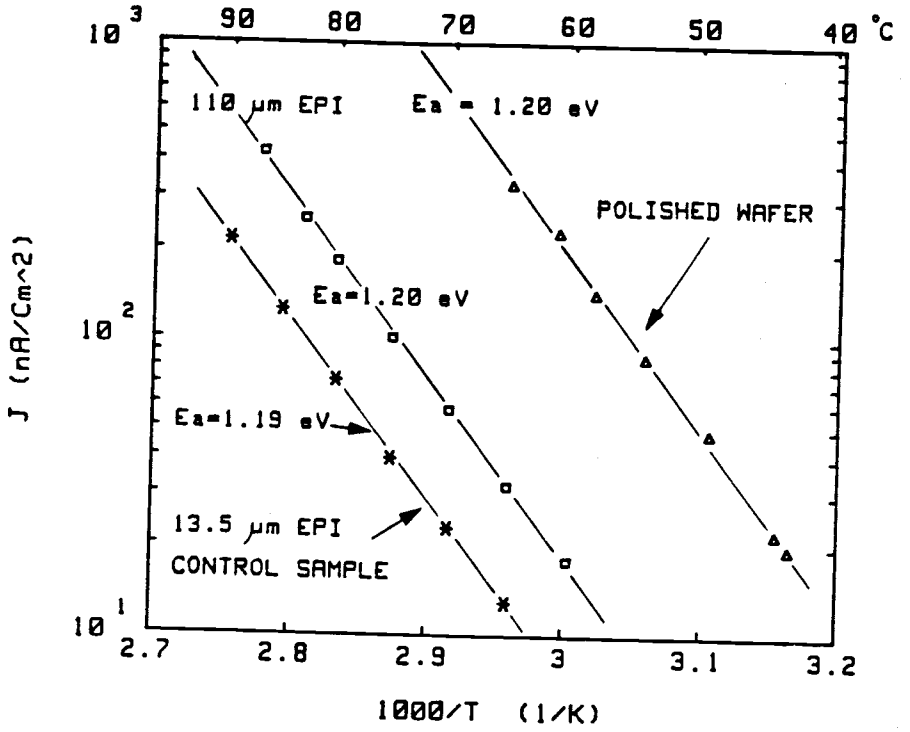


Figure 56) Current density as a function of inverse temperature.

significant in thin epitaxial layers, apparently have an n_i^2 dependence since the ratio of the retention times of the thin and thick epi-layers remains the same at elevated temperatures. Figure 55 is a plot of current density as a function of time after an application of a voltage step to the MOS capacitors for various temperatures on a 110μm epi-layer. At temperatures of about 70°C or higher the diffusion clearly dominates and the diffusion current is constant when plotted versus the response time of the pulsed MOS capacitor. At lower temperatures, since generation within the depletion region is the dominant generation mechanism, the current decreases with time as the depletion region relaxes to equilibrium. A plot of the current density versus inverse temperature in the range

where diffusion is the dominant process is shown in Fig. 56. Due to the exponential nature of the retention times and response time limitations (200 msec in its fast mode) of the LCR meter used in measurement of capacitance transient the measured data is limited to a small temperature range. Ignoring the pre-exponential temperature dependence for a small temperature range of about 25°C, the polished wafers, control samples, and 110 μm epi-layers all had an activation energy corresponding to the silicon bandgap (1.20 eV).

In order to examine the temperature dependence at higher temperatures the DLTS setup was used to obtain an accurate value of retention time of MOS capacitors for a given temperature as discussed in Chapter II section C2. A Boonton capacitance meter was used since the response time for this instrument is much faster than that of the HP LCR meter (about 1 msec). The DLTS signal of a 110 μm epi-layer MOS capacitor in the range of 80 to 130°C is shown in Fig. 57a. The peak heights increase as the temperature increases because the depletion layer shrinks due to an increase in minority carrier density (n_i^2/N_a) at elevated temperatures, causing the inversion capacitance to increase. The retention times obtained from the t2 setting of the rate window of the boxcar averager for a control and a 110 μm epi-layer wafers had an activation energies of 1.213 eV and 1.197 eV, respectively. The ratio of retention times of the 110 μm layer and the control samples from Fig. 57b or current densities in Fig. 56 remains the same. Thus, even though a significant epi-layer - substrate generation was observed in thin epi-layers, its temperature dependence generally remains as n_i^2. If this interface was modeled with a simple interface generation velocity, S_{pp+}, the

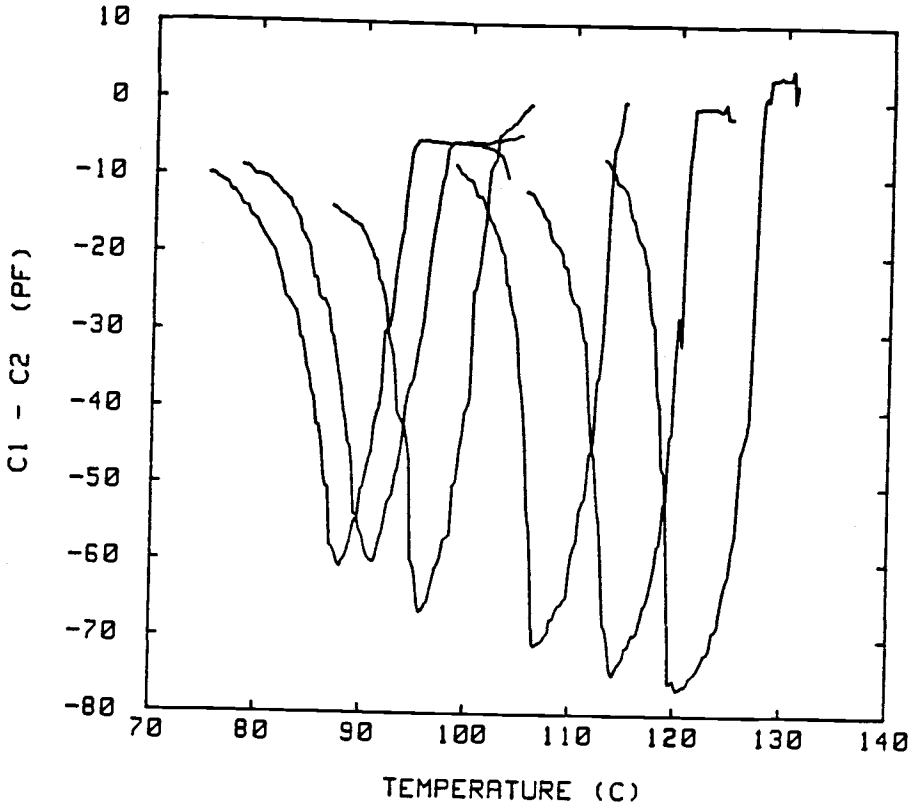


Figure 57a) DLTS signal of 110 μm epi-layer MOS capacitor.

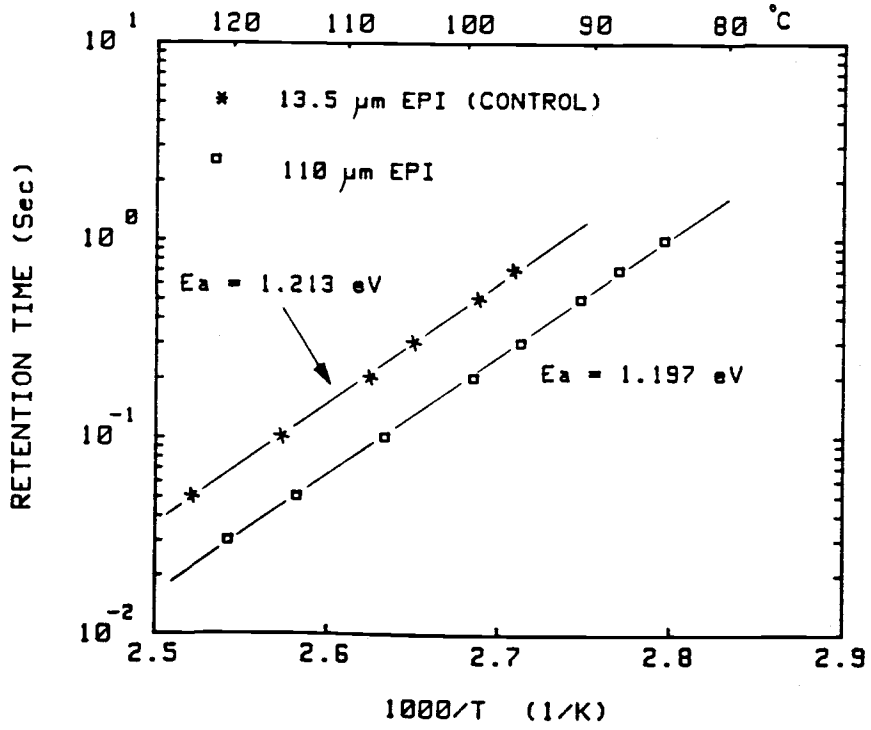


Figure 57b) Retention time obtained from Fig. 57a) for temperature range of 80 - 130°C.

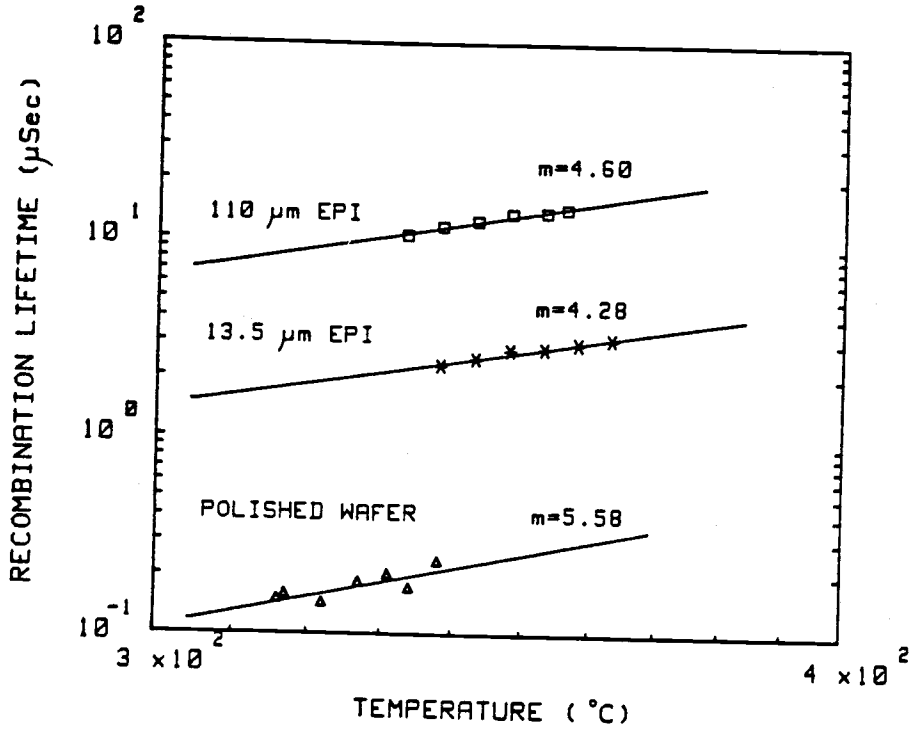


Figure 58) Temperature dependence of recombination lifetimes.

control epitaxial samples would have a close to n_i temperature dependence at lower temperatures of about 70 - 80°C and a n_i^2 dependence at higher temperatures. Therefore, it was concluded that the epi-layer - p^+ interface generation was due to a diffusion process with a poor recombination lifetime at the epi-layer - substrate interface.

An estimate of temperature dependence of capture cross section, σ_n, is obtained from the temperature dependence of the recombination lifetime. τ_r is inversely proportional to the carrier thermal velocity, v_{th}, and σ_n (Eqn. (22)). v_{th} is proportional to $T^{1/2}$ and the temperature dependence of the capture cross section is normally given by T^{-n}, where n depends on trap type. A log-log plot of recombination lifetime over a limited temperature range for a

polished wafer, a control wafer, and a 110 μm epi-layer is shown in Fig. 58. The recombination lifetime varies strongly with temperature. The error could be large since only a very limited temperature range was used; therefore, the n factor is a qualitative value rather than quantitative indicated here. The differences between the capture cross sections of the control and the 110 μm epi-layer samples are not considered to be significant. However, the large temperature dependence of the capture cross sections suggest an attractive type trap [74,75]. The recombination lifetimes of n/n^+ epitaxial wafers were a weak function of temperature ($n \approx 1.6$).

5. Possible sources of epi-layer - p^+ substrate interface current

Some possible sources causing generation of minority carriers at the epi-layer - substrate interface which result in degradation of the measured or apparent recombination lifetime are:

1) Chemical removal of the substrate surface layer prior to growth of epi-layer in the epitaxial reactor.
2) Oxygen precipitation, bulk stacking faults, or residual oxygen forming SiO_x clusters at the substrate.
3) Mismatch between the epitaxial layer and the high boron concentration substrate.

Since the n/n^+ wafers show fairly high recombination lifetimes and that τ_r increases after high temperature processing (discussed in the next section), the possibility that chemical removal of the front surface of the substrate introduces a significant metallic contamination can be eliminated. However, the chemical removal of p^+-

substrate alone does not form a denuded zone at the heavily doped substrate. In fact, during the epitaxial deposition process, oxygen atoms diffuse into the growing epitaxial layer and in turn cause a decrease in the oxygen concentration at the surface region of the substrate wafer. The oxygen concentration of epi-layer is perhaps too small to form oxygen precipitates because the diffusivity of oxygen is high in silicon [75]. On the other hand, it is possible to have a localized trap due to SiO_x clusters, not visible under the optical microscope, in the vicinity of the epi-layer - p^+ substrate interface.

It is generally expected that $\tau_g > \tau_r$ because most generation-recombination centers have energy levels, E_T, which are not equal to E_i. A generation process which is thermally activated has an exponential dependence of E_T/KT. In contrast, the recombination process is not so sensitive to the location of E_T and it is typical to see τ_g much larger than τ_r. Exceptionally good generation lifetime characteristics for p/p^+ epitaxial wafers were observed, but the actual recombination lifetime of 19 μsec is not very long. The ratio of actual generation lifetime to actual recombination lifetime of p/p^+ epitaxial layers studied here, τ_g/τ_r, is about 1050. E_T may be approximated by using Eqn. (8) and a knowledge of the electron to hole capture cross section ratio. E_T-E_i for capture cross section ratio σ_n/σ_p ranging form 1 to 100 results in $E_T-E_i=6.96$ KT to $E_T-E_i=2.35$ KT or $E_C-E_T=0.42$ eV to $E_C-E_T=0.54$ eV at room temperature. The trap concentration of deep levels in the range of 0.42 to 0.54 eV below the conduction band must be $N_T \leq 3\times10^{11}$ cm^{-3} since the DLTS measurements did not show any peaks. This results in an electron

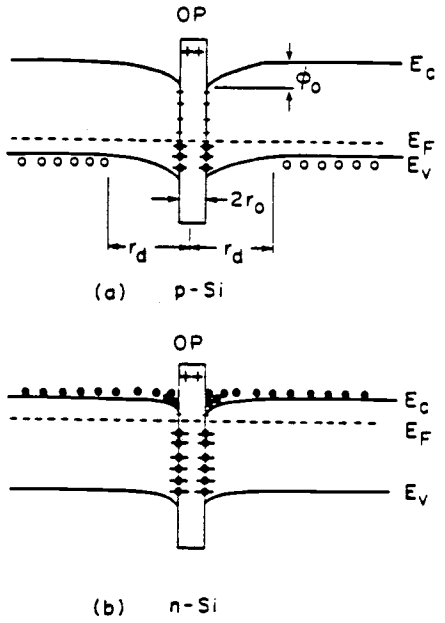


Figure 59) A model for oxide precipitates in n- and p-type silicon wafers, after Hwang et al. [62].

capture cross section of $\sigma_n \geq 2x10^{-14}$ cm^{-2} ($\sigma_n = (\tau_r v_{th} N_T)^{-1}$).

Chan et al. [76] reported electron traps in boron doped Czochralski silicon at E_c-0.41 eV with an electron capture cross section greater than 10^{-14} cm^{-2} due to oxygen precipitation. After comparing E_T and σ_n values obtained on p/p^+ epitaxial wafers with their results it was speculated that oxygen precipitates at the epi-layer - p^+ interface are the main interface generation mechanism. Recently, Hwang et al. [59,62] reported degradation of lifetimes due to small oxygen precipitates (OP), less than 100 angstrom, in bulk wafers where recombination at OP's took place through Si-OP interface states. Similar to the epitaxial wafers in this work, they observed much higher τ_r values in n-type than p-type bulk wafers. They have postulated a band bending around the OP's which is caused by the positive fixed charge of OP - Si similar to the Si-SiO_2 interface

(Fig. 59). The positive fixed charge in the oxide precipitates results in a depletion region in p-type silicon and an accumulation layer around the OP's in n-type silicon. In p-type silicon, once the electrons enter the space charge region, they are swept by the field to the interface to recombine through interface states. In n-type silicon, the same positive fixed charge causes an accumulation region repelling minority carrier holes from the Si-OP interface. The attractive type capture cross sections and temperature dependence observed in the p-type epi-layer supports this model. It may also be argued that a mismatch between epi-layer - substrate creates electron traps degrading the lifetimes in p-type, but not affecting the n-type since the electrons are majority carrier in n-type material.

6. Effects of intrinsic gettering and heat treatment on τ_r

A set of 13.5 μm epitaxial wafers from the seed, center and tang sections of a crystal ingot were selected to examine the effects of intrinsic gettering on the recombination lifetime. One set of samples (control) only experienced the 900°C oxidation and anneal of Fig. 13b. The other set experienced the heat treatment summarized in Fig. 13b with 300 minutes of pre-anneal heat treatment [77] at 650°C. The apparent recombination lifetime versus the location of ingot is shown for both control and intrinsically gettered and heat treated samples in Fig. 60. The control samples regardless of the location of ingot, have τ_r values in the range of 2 to 2.5 μsec. The apparent recombination lifetime for all of the intrinsically gettered wafers improves compared to control samples. For wafers from the seed-end,

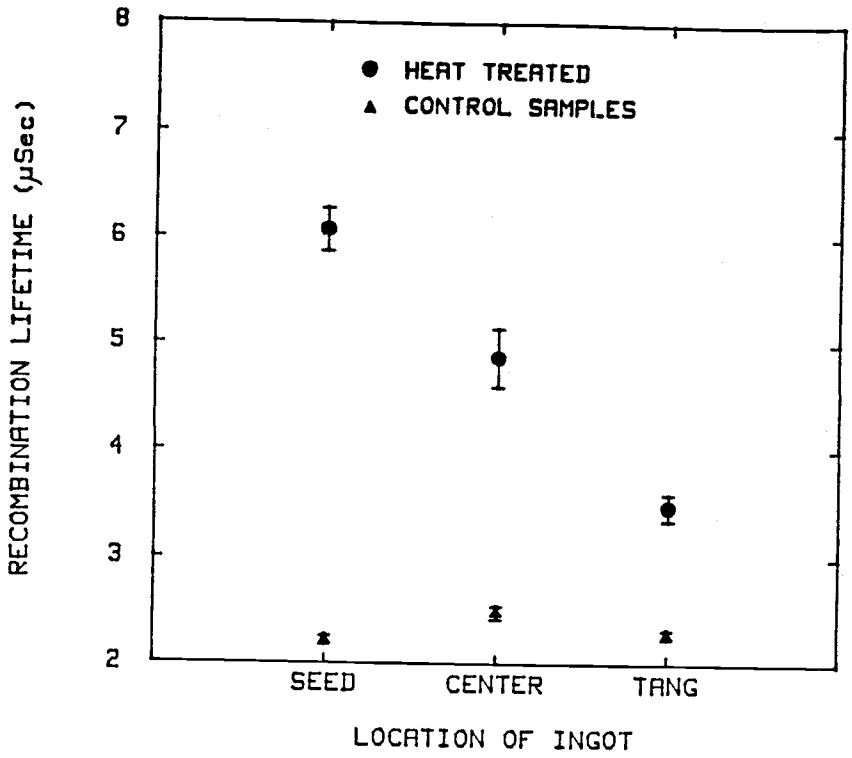


Figure 60) Apparent recombination lifetime versus location of ingot for control and heat treated (see Fig. 13b) samples.

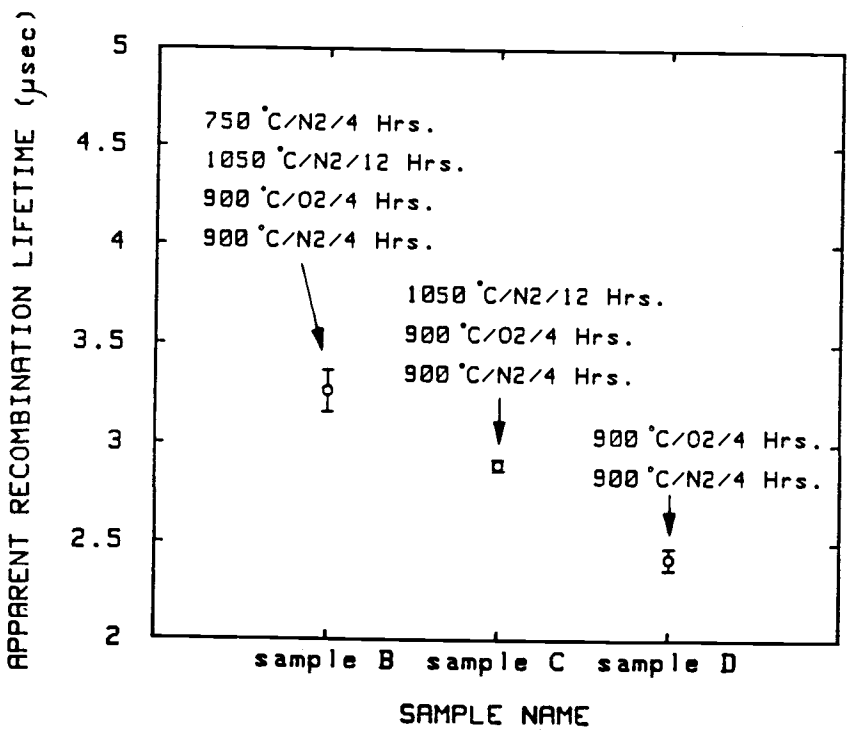


Figure 61) Apparent recombination lifetime of 13.5 μm epi-layers after each heat treatment step.

the apparent recombination lifetime improves up to three times. It is obvious that the observed lifetimes are not primarily dominated by contamination of metallic impurities during the annealing process, since metallic impurities are more easily introduced at higher temperatures which degrades τ_r. The τ_r improvement is believed to be due to out-diffusion of oxide precipitates and formation of a denuded zone in the p^+ substrate [77,78]. Smaller improvement in τ_r for center and tang-end sections of the ingot are due to higher concentration of impurities [47] and lower interstitial oxygen concentration [54] in these sections. The out diffusion of oxygen and formation of a denude zone results in gettering of impurities away from the interface via the oxygen precipitates in the bulk of the p^+ substrate. The effectiveness of this intrinsic gettering depends on a long denuded zone, the concentration of oxygen precipitates, and concentration of impurities at the epi-layer - p^+ substrate interface. In addition, there may be vacancies and dislocations at the interface due to the mismatch between the heavily doped boron substrate and the lightly doped epitaxial layer. The intrinsic gettering is not an effective method for gettering or reducing the electrical activities of vacancies or dislocations. A smaller improvement was observed for the generation lifetimes because the gettering takes place only in the heavily doped substrate and perhaps to some extend at the epi-layer - p^+ interface, not in the active region of the epitaxial layer where no significant oxygen has been observed.

To examine how τ_r changes with different steps of intrinsic gettering cycles, a wafer with 13.5 μm epi-layer was cleaved into

quarter sections and each section received different heat treatment steps. The apparent recombination lifetime of the three quarter sections B, C, and D are shown in Fig. 61. The apparent τ_r values increase for samples C and B with the heat treatment at 1050°C and nucleation at 750°C.

The preliminary investigation of intrinsic gettering indicates that the apparent recombination lifetimes can be improved. However, further investigation is required to identify a more efficient gettering cycle such as a pre-epitaxial intrinsic gettering on the p^+ substrate to eliminate the strong generation of minority carriers at the epi-layer - substrate interface.

7. Suggestions for improving τ_r in thin p/p^+ epitaxial wafers

The epi-layer - p^+ interface has been identified to be the limiting factor in the apparent recombination lifetimes, and it is speculated that oxygen precipitation at that interface is primarily responsible. Obviously, no improvement can result by changing the thickness of the epitaxial layer. Increasing the epitaxial thickness only increases the diffusion current in the epi-layer, resulting in shorter retention times for dynamic charge storage devices. No significant improvement is achieved by just reducing the epi-layer thickness of layers of the order of 10 to 15 μm since the epi-layer - substrate interface is the limiting factor. It was shown that the actual recombination lifetime of the epi-layer was an order of magnitude higher than the apparent values for the thin epi-layers. Therefore, the epi-layer - p^+ substrate interface generation must be

reduced or eliminated in order to benefit from higher actual τ_r values in the epitaxial layer for applications of dynamic charge storage devices. One effective way of eliminating the carrier generation in the epi-layer - p^+ interface from degrading the recombination lifetime characteristics, is to initially grow a heavily doped epi-layer for a few microns on the p^+-substrate followed gradually with a lightly doped layer. This limits the diffusion current collected by the space charge region in the lightly doped layer and undesired effects of p^+-substrate will be eliminated, resulting in an improved apparent recombination lifetime. The gradual change of doping ensures a small built-in electric field so that no significant carriers may be swept out of the substrate. At the present time this solution has some technological limitations, and may not be cost effective. An alternative method is to optimize a heat treatment cycles to form an effective denuded zone in p^+-substrate prior to epitaxial growth. Preliminary study of post-epitaxial intrinsic gettering indicates that the recombination lifetimes improve up to three times. However, a denuded zone on the heavily doped substrate prior to epitaxial deposition can be much more effective than denudation after deposition of epitaxial layer. A three step pre-epitaxial intrinsic gettering cycle similar to Fig. 12 applied to the p^+- substrate was shown [79] to improve the generation lifetimes significantly. Therefore, if the active layer quality of the p/p^+ epitaxial wafers can significantly improve by pre-epitaxial gettering cycle, it is likely that the recombination lifetimes improves also. More investigation is required in this area. Transmission electron microscopy (TEM) study of epi-layer - substrate

interface region can also be very helpful to better quantify the interface effects.

VI. CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK

The main objective of this study was characterization of the lifetime of p/p^+ epitaxial wafers and identification of the factors which limit the lifetime of epitaxial material. Surfaces and interfaces are the dominant factors in generating minority carriers in the p/p^+ epitaxial wafers, causing significantly lower measured or apparent generation and recombination lifetimes. The major points of this research work are summarized in the following.

1) A fast Zerbst C-t transient analysis technique was developed which reduces the generation lifetime measurement time up to 80% in uniformly doped material.

2) Extremely high generation lifetimes of 30-40 msec were observed on p/p^+ epitaxial wafers when the surface generation was properly accounted for. Therefore, the active layer quality of a starting p/p^+ epitaxial material is not limited by the epitaxial quality, but is determined by surface generation as was demonstrated by comparing different size devices.

3) Combined high frequency and quasi-static C-V measurements on MOS capacitors indicated a very good $Si-SiO_2$ interface with trap densities of less than $3x10^{10}$ cm^{-2} eV^{-1}. At room temperature even though the surface generation due to the $Si-SiO_2$ interface is small, the generation lifetimes are extremely large requiring careful account of the surface generation. At elevated temperatures, the

contribution from carriers generated at the Si-SiO_2 interface and within the space charge region is small when compared to diffusion from the bulk epi-layer or epi-layer - substrate interface generation.

4) Majority carrier DLTS measurements revealed no deep levels with concentrations higher than $3x10^{11}$ cm^{-3} within p-type epitaxial layers. This supports the high generation lifetimes observed on p-type epitaxial wafers. Optical DLTS measurements revealed only a minority carrier trap level about 0.21-0.28 eV below the conduction band whose concentration decreased as the reverse bias was increased. Sputtered Ti used in the fabrication of Schottky diodes was responsible for this minority carrier trap. No other significant trap level was observed. Thus, both DLTS and the generation lifetime measurements indicate good epitaxial layer quality.

5) While Schroder's technique for evaluation of the recombination lifetime works well for long base width devices such as polished wafers, it is inadequate and underestimates the recombination lifetime for short base width devices such as thin epitaxial wafers. A simple one dimensional model was developed here including lateral quasi - neutral bulk generation and the time dependence of the space charge region for short base width devices. Apparent recombination lifetimes of about 2 μsec were typically measured on thin (13.5 μm) p/p^+ epitaxial layers.

6) From a study of the transient response of MOS capacitors of

various epitaxial thickness at elevated temperatures, the epi-layer - substrate interface was identified as the source of generation of a significant number of minority carriers. In relaxation to equilibrium, the pulsed MOS capacitors with thin (13.5 μm) epitaxial layers revealed that only 30% of the total current was due to diffusion within epi-layer and the epi-layer - substrate interface was responsible for the rest.

7) The actual recombination lifetime of 19 μsec for the p-type epitaxial layer is an order of magnitude larger than the apparent τ_r values on the thin epi-layers, which translates to an actual diffusion length of 245 μm. Therefore, it is essential to eliminate the generation of minority carriers at the epi-layer - substrate interface to fully utilize the higher lifetimes of epi-layer for applications of dynamic charge storage devices.

8) Assuming Shockley-Read-Hall theory and a discrete trap level, the combined τ_g and τ_r study suggested a trap level of 0.42 to 0.54 eV below the conduction band. The capture cross section for minority carrier electrons at the trap was estimated to be $\sigma_n \geq 2x10^{-14}$ cm^{-2} from the DLTS study. The large capture cross section and the temperature dependence of τ_r suggest an attractive type trap for electrons in the p/p^+ epitaxial material.

9) It is speculated that the positive charge associated with oxygen precipitates and bulk silicon (similar to the positive charge of the $Si-SiO_2$ interface) at the vicinity of epi-layer - substrate interface

cause an attractive electric field for minority carriers in p-type silicon. Therefore, the minority carrier electrons that are attracted to the oxide precipitates recombine at the interface of oxygen precipitates and p-type silicon. The same positive charge of oxygen precipitates in n-type material causes an accumulation layer which repels minority carrier electrons. This model explains the large τ_r difference between n-type and p-type epitaxial wafers. In addition to oxygen, vacancies, dislocations, and other impurities at the epi-layer - substrate interface can play an important role in determining the apparent recombination lifetime. Thus, further investigation is required this area.

10) A preliminary study of intrinsic gettering indicates that the apparent recombination lifetimes improve up to three times. Further investigation is necessary to optimize an efficient gettering cycle to overcome the limitations in the epi-layer - substrate interface.

In contrast to the generation lifetime, the recombination lifetime measured at elevated temperatures is a more realistic representation of device operating conditions. The measurement time of the recombination lifetime is normally less than one minute compared to hours for the generation lifetime measured at room temperature. Furthermore, the recombination lifetime measurement samples a much larger volume of the material. Therefore, the recombination lifetime needs more emphasis.

For the first time, a complete wafer evaluation methodology is described for the characterization of epitaxial material using simple

test structures. Combined generation and recombination lifetime measurements along with the DLTS measurements presented here may be applied for material characterization and a specific acceptance criterion of a given technology or application such as DRAM's and CCD's.

The limitations of p/p^+ epitaxial layers are due to surface and interface effects rather than the epi-layer itself. Precipitates are the main factor in contributing to the interface current observed at the epi-layer - substrate interface. Therefore, it is suggested to grow a heavily doped (oxygen free) epitaxial layer of a few microns on the p^+ substrate and then gradually change the doping concentration to a lightly doped layer with the desired thickness. This should "cover up" any undesired electrical activity at the epi-layer - p^+ substrate interface. An alternative suggestion leading to higher recombination lifetimes is to form a proper denuded zone in the p^+-substrate prior to epitaxial deposition so the epi-layer - substrate interface would be free from oxygen precipitates.

VII. REFERENCES

[1] D.K. Schroder, J.D. Whitfield, C.J. Varker,"Recombination lifetime using the pulsed MOS capacitor," IEEE Trans. Elec. Dev., Vol. ED-31, No. 4, pp. 462, 1984

[2] M. Aminzadeh, L. Forbes,"Recombination lifetime of short base width devices using the pulsed MOS capacitor technique,"IEEE Trans. Elec. Dev., Vol. ED-35, No. 4, pp. 518, 1988

[3] M. Zerbst, Z. angew. Phys., Vol. 22, pp. 30, 1966

[4] D.K. Schroder," The concept of generation and recombination lifetimes in semiconductors",IEEE Trans. Elec. Dev., vol. ED-29, No. 8, pp. 1336, 1982

[5] A.S. Grove,"Physics and technology of semiconductor devices", (J. Wiley & Sons Inc., 1967), pp. 129

[6] R.A. Smith,"Semiconductors",(Cambridge university press, 2nd Ed., 1978), pp.270

[7] W. Shockley, W.T. Read,"Statistics of the recombination of holes and electrons", Phys. Rev., Vol. 87, pp. 835, sept. 1952

[8] F.P. Heiman,"On the determination of minority carrier lifetime from the transient response of an MOS capacitor", IEEE Trans. Elec. Dev. vol. ED-14, No. 11, pp. 781, 1967

[9] S.R. Hofstein,"Minority carrier lifetime determination from inversion layer transient response," IEEE Trans. Elec. Dev., Vol. ED-14, No. 11, pp. 785, 1967

[10] D.K. Schroder, H.C. Nathanson,"On the seperation of bulk and surface components of lifetime using the pulsed MOS capacitor", Solid State Electron., Vol. 13, pp. 577, 1970

[11] D.K. Schroder, J. Guldberg,"Interpretation of surface and bulk effects using the pulsed MIS capacitor," Solid State Electron., Vol. 14, pp. 1285, 1971

[12] R.F. Pierret,"A linear-sweep MOS-C technique for determining minority carrier lifetime,"IEEE Trans. Elec. Dev., Vol. ED-19, No. 7, pp. 869, 1972

[13] R.F. Pierret, D.W. Small,"A modified linear-sweep technique for MOS-C generation rate measurements," IEEE Trans. Elec. Dev., Vol. ED-22, No. 9, pp. 1051, 1975

[14] S.T. Lin ," Simultaneous determination of minority carrier lifetime and deep doping profile using a double-sweep MOS-C technique,"IEEE Trans. Elec. Dev. Vol ED-30, No. 1, pp. 60, 1983

REFERENCES - continued

[15] M. Aminzadeh, L. Forbes,"Fast Zerbst transient analysis and application to intrinsic gettered p-type Epitaxial wafers," in Materials issues in silicon integrated circuit processing, Mat. Res. Soc. Proc., Vol. 71, pp. 59, 1986

[16] D.K. Schroder,"Effective lifetime in high quality silicon devices", Solid State Electron., Vol. 27, No. 3, pp. 247, 1984

[17] K.S. Rabbani, D.R. Lamb,"A quick method for determination of bulk generation lifetime in semiconductor pulsed MOS capacitor measurements", Solid State Electron., Vol. 24, pp. 661, 1981

[18] W.W. Keller,"The rapid measurement of generation lifetime in MOS capacitors with long relaxation times," IEEE Trans. Elec. Dev., Vol. ED-34, No. 5, pp. 1141, 1987

[19] W.R. Fahrner, D. Braeunig, C.P. Schneider, M. Briere,"Reduction of measurement time of lifetime profiles by applying high temperatures," J. Electrochem. Soc., Vol. 134, No. 5, pp. 1291, 1987

[20] Z. Radzimski, J. Honeycutt, G.A. Rozgonyi,"Minority carrier lifetime analysis of silicon Epitaxy and bulk crystals with nonuniformly distributed defects," IEEE Trans. Elec. Dev., Vol. ED-35, No. 1, pp. 80, 1988

[21] E.H. Nicollian, J.R. Brews,"MOS physics and technology", (J. Wiley & Sons Inc., 1982), pp. 409

[22] T.W. Collins, J.N. Churchill,"Exact modeling of the transient response of an MOS capacitor," IEEE Trans. Elec. Dev., Vol. ED-22, pp. 90, 1975

[23] W.Van Roosbroeck and W. Shockley, Phys. Rev., vol. 94, pp. 1558, 1954

[24] R. Eisberg, R. Resnick,"Quantum Physics",(J. Wiley & Sons Inc, 1974)

[25] S.M. Sze,"Physics of semiconductor devices",(J. Wiley & Sons Inc, 2nd Ed., 1981)

[26] J.S. Blakemore,"Radiative and radiationless recombination", (Semiconductor statistics, Dover publication Inc., NY., 1987, republished from 1962, pp. 196)

[27] W.P. Dumke,"Spontaneous radiative recombination in semiconductors", Phys. Rev., Vol. 105, pp. 139, No. 1, 1957

REFERENCES - continued

[28] J.I. Pankove,"Optical processes in semiconductors", (Dover publications, Inc, New york, 1971), pp. 108

[29] V.L.B. Bruevich, E.G. Lansberg,"Recombination mechanisms,"Phys. Status. Solidi, Vol. 29, pp.9, 1968

[30] Y.C. Kao, J.R. Davis,"Correlation between reverse recovery time and lifetime of p-n junction driven by a current ramp," IEEE Trans. Elec. Dev., Vol. ED-17, No. 9, pp. 652, 1970

[31] D.C. Lewis,"On the determination of minority carrier lifetime from the reverse recovery transient on pnR diodes," Solid State Electron., Vol. 18, pp. 87, 1975

[32] K. Venkateswaren, D.J. Roulston,"Recombination dependent characteristics of silicon P^+-N-N^+ Epitaxial diodes," Solid State Electron., Vol. 15, pp. 311, 1972

[33] O.L. Curtis, J.R. Srour,"Steady state recombination in semiconductors containing two or more trapping levels," J. Appl. Phys., Vol. 45, pp. 792, 1974

[34] P.D. Blais, C.F. Seiler,"Measurement and retention of recombination lifetime," in Lifetime Factors in Silicon, ASTM STP 712, pp. 148, 1980

[35] E.O. Johnson,"Measurement of minority carrier lifetime with surface photovoltage," J. Appl. Phys., Vol. 28, pp. 1349, 1957

[36] F.N. Gonzalez, A. Neugroschel,"Measurement of diffusion length, lifetime, and surface recombination velocity in thin semiconductor layers," IEEE Trans. Elec. Dev., Vol. ED-31, No. 4, pp. 413, 1984

[37] A. Neugroschel,"Determination of lifetime and recombination currents in p-n junction solar cells, diodes, and transistors, IEEE Trans. Elec. Dev., Vol. ED-28, No. 1, pp. 108, 1981

[38] H. Reichl, H. Bernt,"Lifetime measurement in silicon epitaxial materials," Solid State Electron., Vol. 18, No. 5, pp. 453, 1975

[39] P. Spirito, G. Cocorullo,"A measurement technique to obtain the recombination lifetime profile layer at any injection level," IEEE Trans. Elec. Dev., Vol. ED-35, No. 12, pp. 2546, 1987

REFERENCES - continued

[40] D.K. Schroder,"Bulk and Optical generation parameters measured with the pulsed MOS capacitor,"IEEE Trans. Elec. Dev., Vol. ED-19, No. 9, pp. 1018, 1972

[41] D.V. Lang,"Deep - Level Transient Spectroscopy: A new method to characterize traps in semiconductors", J. Appl. Phys., Vol. 45, No. 7, pp. 3023, 1974

[42] C.T. Sah, L. Forbes, L.L. Rosier, A.F. Tasch,"Thermal and optical emission and capture rates and cross sections of electrons and holes at imperfection centers in semiconductors from photo and dark junction current and capacitance experiments," Solid State Electron., Vol. 13, pp. 759, 1970

[43] D.V. Lang, in "Thermally stimulated relaxtion in solids", Topics Appl. Phys., Vol. 37, Ed. by P. Braunlich, (Springer, Belin, Heidelberg, Newyork, 1979), chpt. 3

[44] D.V. Lang,"Fast capacitance transient apparatus:Application to ZnO and O centers in GaP p-n junctions", J. Appl. Phys., Vol. 45, No. 7, pp. 3014, July 1974

[45] P.D. Kirchner, W.J. Schaff, G.N. Maracas, L.F. Eastman, T.I. Chappel, C.M. Ransom,"The analysis of exponential and non-exponential transients in deep-level transient spectroscopy", J. Appl. Phys., Vol. 52, No. 11, pp. 6462, 1981

[46] N.O. Pearce, B. Hamilton, A.R. Peaker, R.A. Craven,"Application of deep level transient spectroscopy to MOS relaxation transients", J. Appl. Phys. Vol. 62, No. 2, pp. 576, 1987

[47] S.M. Sze,"VLSI Technology", (McGraw-hill, NewYork 1983)

[48] L. Forbes, F.D. Whitwer, J.D. Peng,"Oxygen precipitation in CMOS wafers," in Impurity diffusion and gettering in silicon, edited by R.B. Fair, C.W. Pearce, and J. Washburn, Proc. of Mat. Res. Soc., Vol. 36, (MRS, Pittsburgh, 1985), pp. 257

[49] J.O. Borland,"The homogenization of midro-defects in the bulk of silicon wafers by a three-step heat treatment," in Defects in Silicon, Electrochem. soc. proc., Vol. 83-9, pp. 236, 1983

[50] W.J. Patrick, S.M. Hu, W.A. Westdorp,"The effect of SiO_2 precipitation in Si on generation currents in MOS capacitors," J. Appl. Phys., Vol. 50, No. 3, pp. 1399, 1979

[51] K. Ikuta, T. Ohara,"Lifetime evaluation of denuded zone quality and intrinsic gettering effects on heavy metals," Japanese J. Appl. Phys., Vol. 23, No. 8, pp. 984, 1984

REFERENCES - continued

[52] K. Nauka, J. Lagowski, H.C. Gatos, O. Ueda,"New intrinsic gettering process in silicon based on interactions of silicon interstitials,"J. Appl. Phys., Vol. 60, No. 2, pp. 615, July 1986

[53] O. Ueda, K. Nauka, J. Lagowski, H.C. Gatos,"Identification of intrinsic gettering centers in oxygen-free silicon crystals," J. Appl. Phys., Vol. 60, No. 2, pp. 622, July 1986

[54] W. Wijaranakula, Ph.D. dissertation, Oregon State University, 1986

[55] W. Wijaranakula, J.H. Matlock, H. Mollenkopf, P. Burke, L. Forbes,"Oxygen precipitation in p/p^{+} (100) Epitaxial silicon material," J. Electrochem. soc., Vol. 134, No. 9, pp. 2310, 1987

[56] C.W. Pearce,"Considerations regarding gettering in integrated circuits", in Semiconductor silicon 1981, edited by H.R. Huff, R.J. Kriegler, and Y. Takeishi, The Electrochem. Soc. Proc., Princeton, Vol. 81-5, pp. 705, 1981

[57] M. Miyagi, K. Wada, J. Osaka, N. Inoue,"Effect of oxide precipitates on minority-carrier lifetime in Czochralski-grown silicon," Apply. Phys. Lett., Vol. 40, No. 8, pp. 719, 1982

[58] C.J. Varker, J.D. Whitfield, P.L. Fejes,"Effects of oxygen precipitation on minority carrier lifetime in silicon crystals," in Silicon processing, ASTM STP 804, Baltimore, 1983, pp. 369

[59] J.M. Hwang, D.K. Schroder, A.M. Goodman,"Recombination lifetime in oxygen-precipitated silicon," IEEE Elec. Dev. Lett., Vol. EDL-7, No. 3, pp. 172, 1986

[60] H. Strack, K.R. Mayer, B.O. Kolbesen,"The detrimental influence of stacking faults on the refresh time of MOS memories," Solid State Electron., Vol. 22, pp. 135, 1979

[61] C.J. Varker, J.D. Whitfield, P.L. Fejes,"Oxygen precipitation in silicon - its effects on minority carrier recombination and generation lifetime," in Defects in Semiconductors II, Mat. Res. Symp. Proc., Vol. 14, pp. 187, 1983

[62] J.M. Hwang, D.K. Schroder,"Recombination properties of oxygen-precipitated silicon," J. Appl. Phys., Vol. 59, No. 7, pp. 2476, 1986

REFERENCES - continued

[63] J.R. Carruthers,"Silicon microdefects and contamination control in MOS device processing," Defects in Silicon, Electrochem. soc. proc., Vol. 83-9, pp. 375, 1983

[64] J.O. Borland, M. Kuo, J. Shibley, B. Roberts, R. Schindler, T. Dalrymple,"Influence of Epi-substrate point defect properties on getter enhanced silicon epitaxial processing for advance CMOS and bipolar technologies," in VLSI Science and Technology, Electrochem. Soc. Proc. Vol. 84-7, pp. 93, 1984

[65] W. Dyson, L. Hellwig, J. Moody, J. rossi,"Gettering of P^+ (100) Si substrates for epitaxial growth," in Defects in Silicon, Electrochem. soc. proc., Vol. 83-9, pp. 246, 1983

[66] W. Dyson, S. O'Grady, J.A. Rossi, L.G. Hellwig, J.W. Moddy," N^+ and P^+ substrate effects on epitaxial silicon properties," in VLSI Science and Technology, Electrochem. Soc. Proc. Vol. 84-7, pp. 107, 1984

[67] P.L. Castro, B.E. Deal,"low-temperature reduction of fast surface states associated with thermally oxidized silicon," J. Electrochem. Soc. Vol. 118, No. 2, pp. 280, 1971

[68] W.I. Sze, Ph.D. Dissertation, Oregon State University, 1988

[69] E.H. Rhoderick,"Metal-semiconductor contacts," Clarendon Press, Oxford, 1978

[70] F.D. Whitwer, H. Haddad, L. Forbes, "DLTS charaterization of precipitation induced microdefects", in Materials issues in silicon integrated circuit processing, Mat. Res. Soc. Proc., Vol. 71, pp. 53, 1986

[71] S.P. Murarka, D.B. Fraser,"Thin film interaction between titanium and poly crystalline silicon," J. Appl. Phys., Vol. 51, No. 1, pp. 342, 1980

[72] J.M. Dorkel, P. Leturcq,"Carrier mobilities in silicon semi-empirically related to temperature, doping and injection level," Solid State Electron., Vol. 24, No.9, pp. 821, 1981

[73] C.D. Thurmond,"The standard thermodynamic functions for the formation of electrons and holes in Ge, Si, GaAs, and GaP," J. Electrochem. Soc. Vol. 122, No. 8, pp. 1133, 1975

[74] M. Lax ,"Cascade capture of electrons in solids," Phys. Rev., Vol. 119, No. 5, pp. 1502, 1960

[75] A.G. Milnes,"Deep impurties in semiconductors",(John Wiley & Sons, 1973), pp. 288

REFERENCES - continued

[76] S.S. Chan, C.J. Varker, J.D. Whitfield, R.W. Carpenter,"Deep levels associated with oxygen precipitation in CZ silicon and correlation with carrier lifetimes," in Microscopic identification of electronic defects in semiconductors, edited by N.M. Johnson, S.G. Bishop, and G.D. Watikins, Mat. Res. Soc. Symp. Proc., Vol. 46, 1985, pp. 281

[77] W. Wijaranakula, P. Burke, L. Forbes, J.H. Matlock,"Effects of preanneal heat treatment on oxygen precipitation in epitaxial silicon," in Materials issues in silicon integrated circuit processing, Mat. Res. Soc. Proc., Vol. 71, pp. 139, 1986

[78] W. Wijaranakula, P. Burke, L. Forbes, J.H. Matlock,"Effects of pre- and postepitaxial deposition annealing on oxygen precipitation in silicon," J. Mater. Res., Vol. 1, No. 5, pp. 698, 1986

[79] J.O. Borland, T. Deacon,"Advanced CMOS epitaxial processing for latch-up hardening and improved epi-layer quality," Solid State Technology, Vol. 27, No. 8, 1984

IX. APPENDICES

Appendix A

MOS capacitor C-V analysis

In the following relations for doping concentration, flatband and threshold voltages, and the doping profile are derived for MOS capacitor C-V data.

Defenition of symbols are as follows:

$\epsilon_s \equiv$ Silicon dielectric constant = 11.7

$\epsilon_{ox} \equiv$ Oxide dielectric constant = 3.82

$\epsilon_o = 8.8542 \times 10^{-14}$ F/cm

$q = 1.60206 \times 10^{-19}$ Coulomb

$K = 8.6178 \times 10^{-5}$ eV/K

$E_g \equiv$ silicon bandgap = 1.21 eV

$A \equiv$ Area of MOS capacitor in cm^{-2}

The intrinsic carrier concentration is calculated from the following expression [73]

$$ni = 3.87 \times 10^{16} T^{1.5} \exp(-E_g/2kT) \quad cm^{-3} \tag{39}$$

For C-V measurements at high frequencies such as 1 MHz the $Si\text{-}SiO_2$ interface charge does not respond to the AC signal of the instrument. Therefore, as far as the high frequency measurement in concerned the total capacitance would be series combination of the oxide capacitance, C_{ox}, and the semiconductor capacitance, C_s, shown in Fig. 62.

$$1/C_{hf} = 1/C_s + 1/C_{ox} \tag{41}$$

The maximum depletion width in thermal equilibrium occurs when $C_{hf}=C_{inv}$. Thus, the semiconductor capacitor and the depletion width for this condition are given by

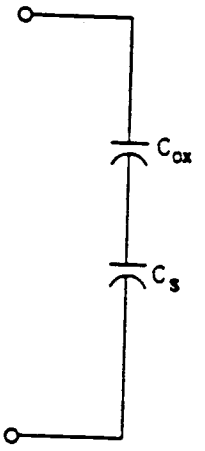


Figure 62) High frequency equivalent circuit of an MOS capacitor.

$$C_{smin} = (C_{ox}.C_{inv})/(C_{ox} - C_{inv}) \quad (42)$$

$$W_{max} = \epsilon_o \epsilon_{si} A/C_{smin} \quad (43)$$

The maximum depletion width, W_{max}, in thermal equilibrium is related to the substrate doping as follows

$$W_{max} = \left| \frac{4\epsilon_o \epsilon_{si} KT \ln(N_{sub}/ni)}{q\, N_{sub}} \right|^{1/2} \quad (44)$$

Solving for N_{sub} we get

$$N_{sub} = (4\epsilon_o \epsilon_{si} KT/q/W_{max}^2) \ln(N_{sub}/ni) \quad (45)$$

Equation (45) is a transcendental equation which may be iteratively solved to determine N_{sub}. The semiconductor flatband capacitance is defined as

$$C_{sfb} = \epsilon_o \epsilon_{si} A/L_{debye} \quad (46)$$

$$L_{debye} \equiv (KT/q\, \epsilon_o \epsilon_{si}/N_{sub})^{1/2} \quad (47)$$

therefore, the measured high frequency flatband capacitance is

$$C_{fb} = (C_{ox}.C_{sfb})/(C_{ox} + C_{sfb}) \quad (48)$$

Flatband voltage is determined by searching for the voltage at C_{fb} in the C-V data file. The threshold voltage of a MOS capacitor is

defined as

$$V_{th} = V_{fb} + 2\phi_f - Q_b.A/C_{ox} \quad (49)$$

where Q_b is the maximum semiconductor charge density due to the depletion charge which is given by

$$Q_b = \pm qN_{sub}W_{max} \quad (+ \text{ for n-type, } - \text{ for p-type}) \quad (50)$$

and the fermi potential, ϕ_f, is defined as

$$\phi_f = \pm KT \ln (N_{sub}/ni) \quad (+ \text{ for p-type, } - \text{ for n-type}) \quad (51)$$

The work function of metal to semiconductor, ϕ_{ms} is equal to the difference between the metal work function, ϕ_m, and the fermi potential, ϕ_f. The total fixed charge in the oxide, Q_{ss} is given by

$$Q_{ss} = C_{ox}/(q.A) \ (\phi_{ms} - V_{fb}) \quad (52)$$

The free carrier concentration profile using MOS capacitor C-V data can be obtained using the following relations [20].

$$N(W) = -2[q\epsilon_o\epsilon_{si}A \ d/dV_G \ (1/C_{hf})^2]^{-1} \quad (53)$$

$$W = \epsilon_o\epsilon_{si}A \ (1/C_{hf} - 1/C_{ox}) \quad (54)$$

Appendix B

Extraction of interface state density as a function of energy

Low frequency equivalent circuit of MOS capacitor is shown in Fig. 63. Where C_{it} represents the capacitance due to the $Si-SiO_2$ interface charge. As mentioned in appendix A in high frequency C-V test the interface states cannot respond to the fast AC signal, so C_{it} is not observed. In a quasi-static C-V test; however, the interface charge does respond to the slowly varing DC bias. Combination of high frequency and quasi-static C-V's allows evaluation of interface states distribution as a function of bandgap energy as discussed in the following. From Fig. 63 we have

$$1/C_{lf} = 1/C_{ox} + 1/(C_s + C_{it}) \qquad (55)$$

where C_{lf} is the measured quasi-static capacitance. The density of interface states is defined as C_{it}/q, which is obtained from Eqn. (41) and (55).

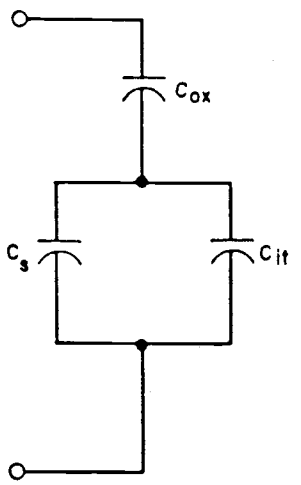


Figure 63) Low frequency equivalent circuit of an MOS capacitor.

$$D_{it} \equiv C_{it}/q = (C_{lf} - C_{hf})/q \; (1 - C_{lf}/C_{ox})^{-1}(1 - C_{lf}/C_{ox})^{-1} \qquad (56)$$

Where D_{it} has units of $cm^{-2}ev^{-1}$. Therefore, only the location of interface traps within the bandgap remains to be found. The incremental charge due to a gate voltage for quasi-static C-V is given by [21]

$$d\Phi_s/dV_G = 1 - C_{lf}/C_{ox} \qquad (57)$$

where Φ_s is the band bending or the barrier height at the surface. Equation (57) is valid only when the semicoductor is in thermal equilibrium. Integration of Eqn. (57) results in Φ_s as a function of gate voltage

$$\Phi_s(V_G) - \Phi_s(V_{fb}) = \int_{V_f}^{V_G} (1 - C_{lf}/C_{ox})\, dV \qquad (58)$$

the band bending at V_{fb} is zero, so the energy level of interface traps is given by

$$E_T - E_V = E_g/2 - \phi_f + \Phi_s \qquad (59)$$

where ϕ_f is the fermi potential (Eqn. (51)) and E_g is the bandgap. Therefore, from quasi-static C-V measurements the band bending as a function of gate voltage may be obtained using Eqn. (58) which can easily be translated to energy of interface traps via Eqn. (59). So, from Eqn. (56) and (59) a plot of density of interface traps as a function of energy level is generated from combination of high frequency and quasi-static C-V's.

Appendix C

Doping profile and built-in voltage from C-V plot of a Schottky diode

The depletion width of a Schottky diode is given by [69]

$$W = [\ 2\epsilon_o\epsilon_s/q/N_{sub}\ (V_{bi} - V_r - KT/q)\ ]^{1/2} \qquad (60)$$

where V_{bi} is the built-in voltage and V_r is the reverse bias on the diode. All other parameters are defined in appendix A. The depletion capacitance is defined as

$$C \equiv \epsilon_o\epsilon_s A/W \qquad (61)$$

Using Eqn. (60) and (61) we get

$$\frac{1}{C^2} = \frac{2(V_{bi} - V_r - KT/q)}{A^2 q\epsilon_o\epsilon_s N_{sub}} \qquad (62)$$

If the doping concentration is uniform throughout the depletion region, a straight line results from a plot of $1/C^2$ vs. V_r. Therefore, for this case the doping concentration is determined from the slope of such a plot

$$N_{sub} = 2(q\epsilon_o\epsilon_s \text{ slope } A^2)^{-1} \qquad (63)$$

The built-in voltage is calculated from intercept of $1/C^2$ vs. V_r plot.

$$V_{bi} = \text{Intercept} + 2KT/q \qquad (64)$$

If the doping concentration is not uniform the differential capacitance determined by differentiating Eqn. (63) with respect to applied reverse bias results in the doping as a function of V_r.

$$N_{sub} = 2/q\epsilon_o\epsilon_s\ [\ d/dV_G(1/C^2).A^2\ ]^{-1} \qquad (65)$$

The depletion layer width, W, is directly calculated from the capacitance measurement (Eqn. (61)). Therefore, combination of Eqns (61) and (65) results in the doping profile.