An Acoustic Charge Transport Imager for High Definition Television Applications

NASA Grant #NAGW-2753

First Semiannual Report for FY '94

W.D. Hunt and K.F. Brennan
School of Electrical and Computer Engineering
and
C.J. Summers
Georgia Tech Research Institute
Georgia Institute of Technology
Atlanta, Georgia 30332-0250

April, 1994

TABLE OF CONTENTS

1.0 Program Objectives and Approach

The primary goal of this research is to develop a solid-state high definition television (HDTV) imager chip operating at a frame rate of about 170 frames/sec at 2 Megapixels/frame. This imager will offer an order of magnitude improvement in speed over CCD designs and will allow for monolithic imagers operating from the IR to UV.

The technical approach of the project focuses on the development of the three basic components of the imager and their subsequent integration. The camera chip can be divided into three distinct functions: 1) image capture via an array of avalanche photodiodes (APDs), 2) charge collection, storage and overflow control via a charge transfer transistor device (CTD) and 3) charge readout via an array of acoustic charge transport (ACT) channels. The use of APDs allows for front end gain at low noise and low operating voltages while the ACT readout enables concomitant high speed and high charge transfer efficiency. Currently work is progressing towards the optimization of each of these component devices.

In addition to the development of each of the three distinct components, work towards their integration and manufacturability is also progressing. The component designs are considered not only to meet individual specifications but to provide overall system level performance suitable for HDTV operation upon integration. The ultimate manufacturability and reliability of the chip constrains the design as well.

The progress made during this period is described in detail in Sections 2-4.

2.0 Heterojunction Acoustic Charge Transport (HACT) Device

2.1 HACT Design and Sensitivity Analysis

The purpose of this section is to introduce and analyze an alternative HACT structure which is relatively insensitive to changes in residual surface concentration and fluctuations in the doping concentration of the epitaxial layers. The proposed device structure utilizes a pn junction at the surface to provide isolation from the surface charge and to create the top vertical confinement. A lightly doped n-type channel layer is then separated from the substrate by a p-type AlGaAs layer. The p-type AlGaAs serves two purposes: First, the band edge discontinuity provides the bottom part of the vertical confinement. Second, the p-type doping serves to complete the depletion of the transport layer. By using dopants to deplete the transport layer, channel thicknesses an order of magnitude larger than those obtained in current HACT devices can be achieved, sensitivity to the surface states is reduced, and processing variations can be avoided. A sensitivity analysis of HACT performance to layer thickness and doping variations was also accomplished.

2.1.1 HACT Design

Comparison to previous HACT designs shows that the design reported here has much less sensitivity to the surface states. Conduction band contour plots for the two different HACT designs at one specific acoustic power level will then be compared for three different residual surface charge concentrations showing the insensitivity of the p-HACT structure. The electron packet shape of the p-HACT structure is analyzed by injecting a fixed number of electrons into the channel and allowing them to equilibrate with the acoustic potential. Finally, transient simulations using several values of carrier lifetime in the channel layer provide a measure of the maximum charge transfer efficiency of the proposed structure.

Figures 2.1 and 2.2 display the device designs for a conventional n-HACT design and the p-HACT design analyzed for this report. One dimensional simulations of the conduction band of the structures with varying residual surface state concentrations are shown in Figures 2.3 and 2.4 for the n-HACT and p-HACT structures, respectively. Notice that the conduction band experiences a large change in the case of the n-HACT structure when the residual surface state density is increased. This is not the case for the conduction band of the p-HACT structure which shows nearly a complete insensitivity to changes in the surface state density. It is also important to note that the channel thickness in the case of the p-HACT structure is nearly an order of magnitude wider than the transport channel for the n-HACT device structure.

Figures 2.5 and 2.6 show energy surface plots of the steady state conduction band for the n-HACT and p-HACT structures with a surface state density of -1.5×10^{12} cm^{-2} and -1.0×10^{12} cm^{-2} respectively. The acoustic power scaling factor, P_s, for both structures was 4. These figures are useful for illustrating the potential well created by the acoustic potential. Hereafter, the conduction band will be illustrated by using contour plots of the energy surface. Comparison of Figures 2.1.5 and 2.1.6 shows that the well created in the p-HACT structure (Figure 2.1.6) is approximately an order of magnitude larger than the n-HACT structure, as expected from the one dimensional simulations. Notice in the n-HACT structure the top and bottom vertical confinement is provided by the two heterojunction barriers, while in the p-HACT structure there is only one abrupt change in the conduction band diagram. The steady state conduction band contours for the p-HACT structure with surface concentrations of -5×10^{11}, -1.0×10^{12}, and -5.0×10^{12} cm^{-2} are displayed in Figures 2.7 through 2.9. As seen from these figures the residual surface state concentration does not significantly change the shape of the potential well in the p-HACT structure.

To determine the maximum charge capacity of the potential well excess electrons are injected into the channel at the conduction band minimum. The structures are simulated until the electrons reach steady state with the acoustic potential. The number of injected electrons is increased until it appears that the

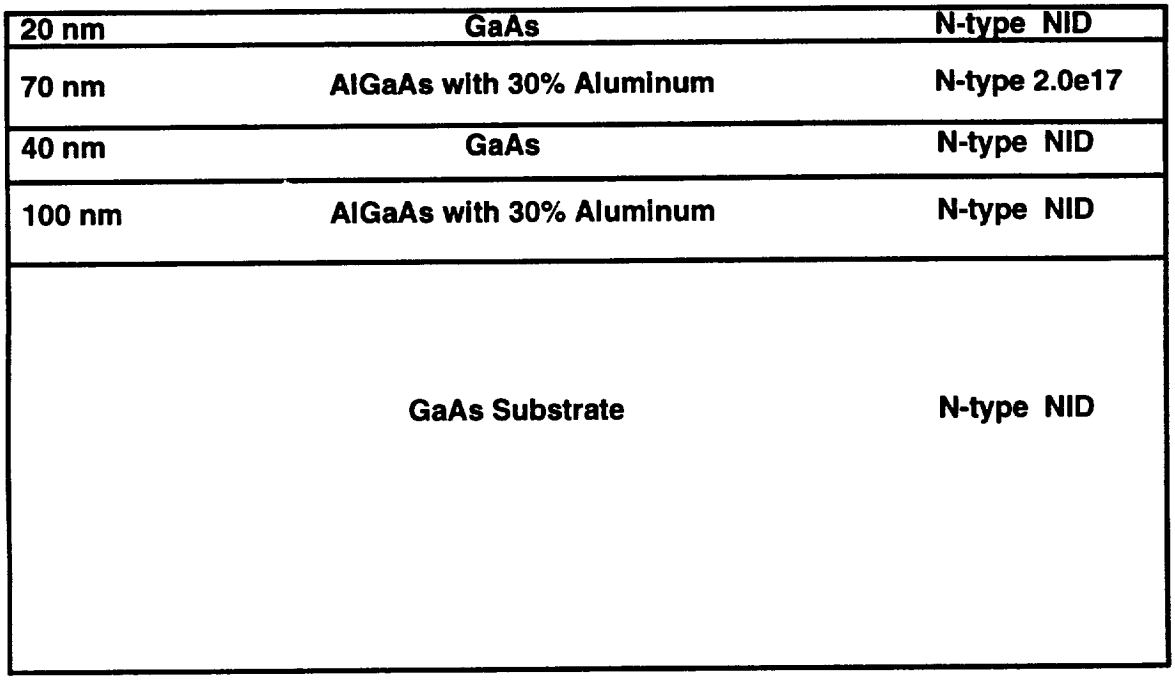


Figure 2.1 Epitaxial layer structure of a conventional n-type heterojunction acoustic charge transport device (n-HACT).

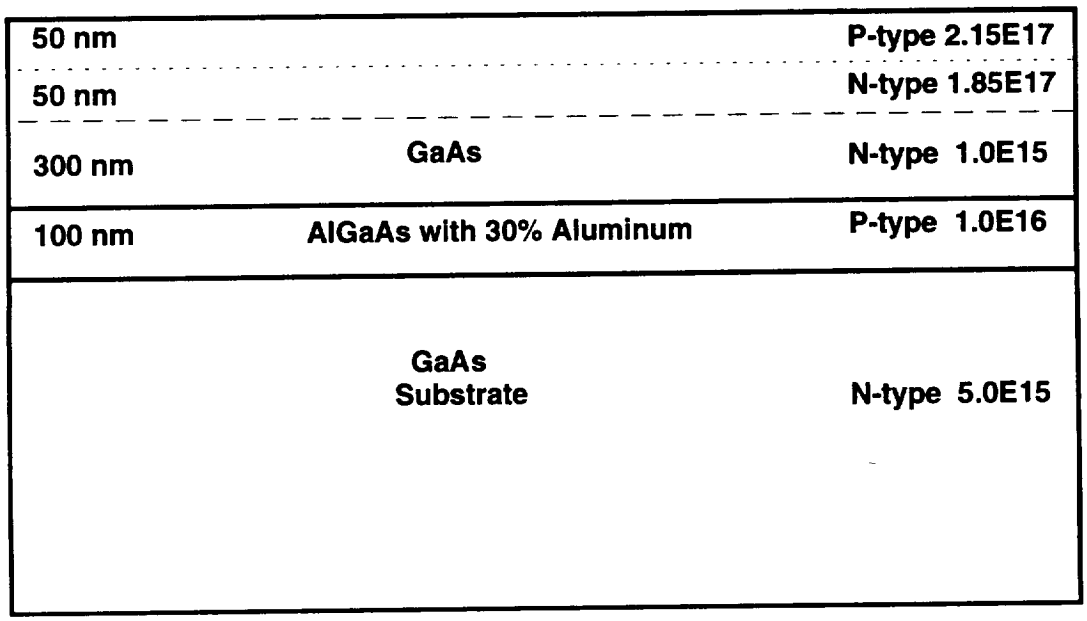


Figure 2.2 Epitaxial layer structure of the alternative p-type heterojunction acoustic charge transport device (p-HACT).

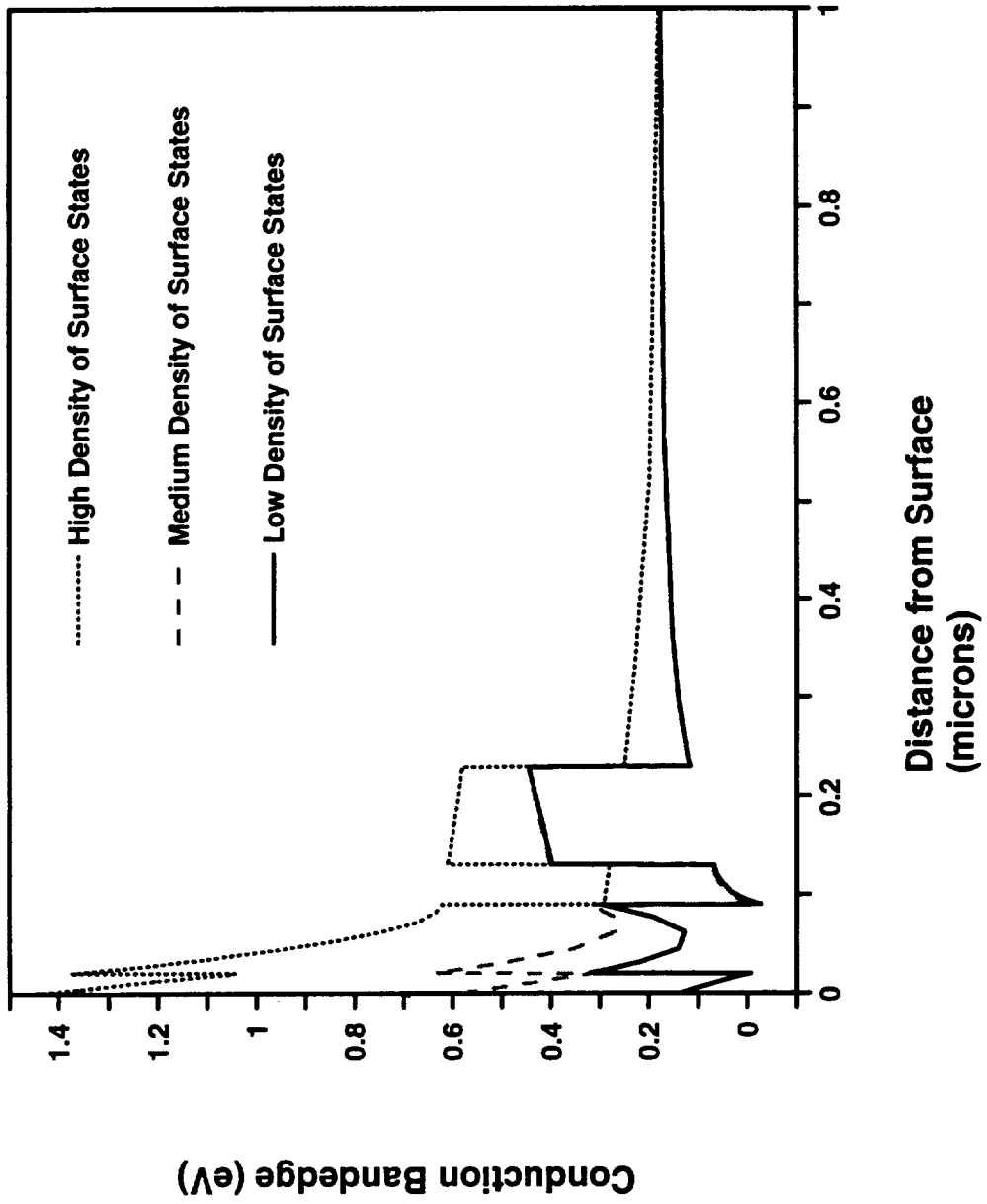


Figure 2.3 One dimensional conduction band diagrams of the n-type HACT device with three different residual surface state concentrations (-$5x10^{12}$, -$1x10^{12}$, and <$5x10^{11}$ cm^{-2}), no acoustic potential.

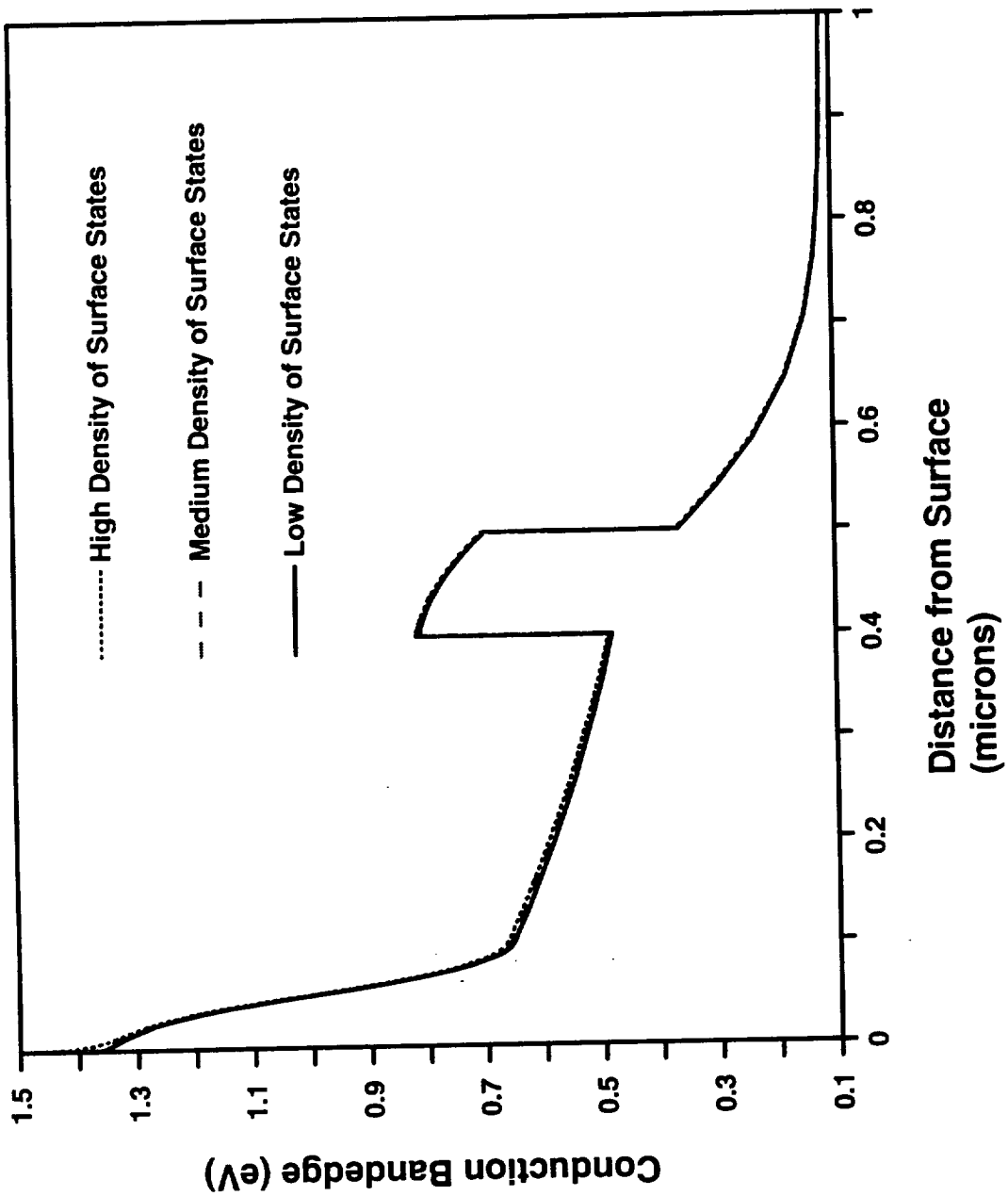


Figure 2.4 One dimensional conduction band diagrams of the p-HACT device with three different residual surface state concentrations ($-5x10^{12}$, $-1x10^{12}$, and $<5x10^{11}$ cm^{-2}), no acoustic potential.

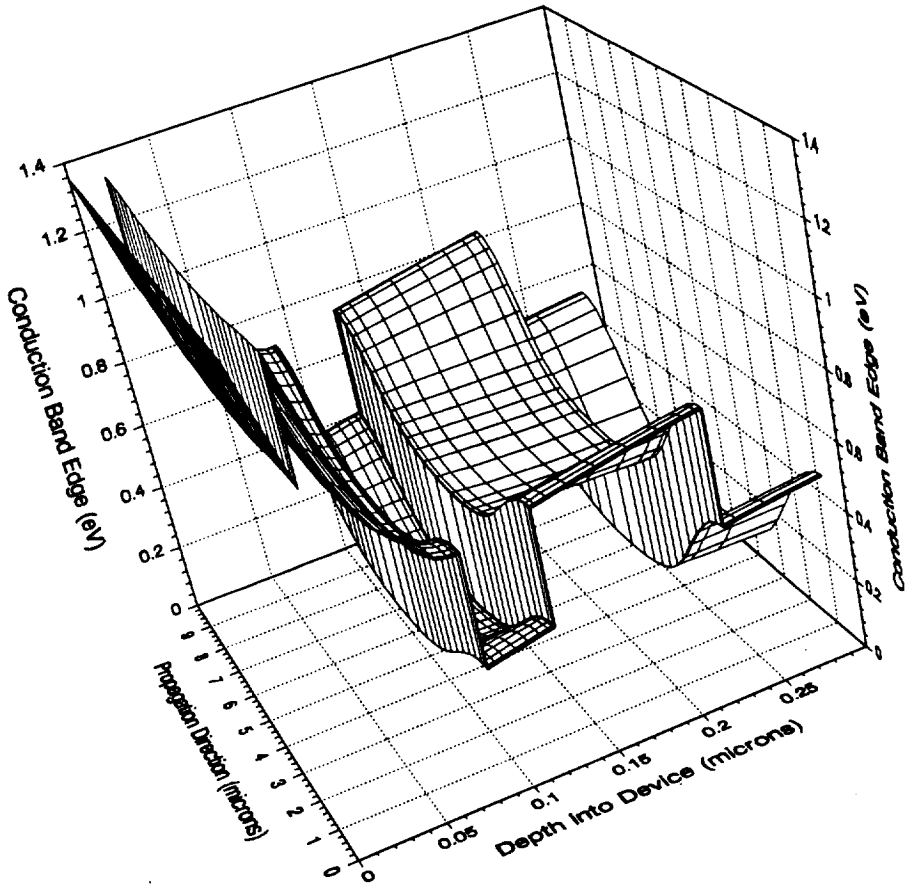


Figure 2.5 Two dimensional conduction band diagram for the n-HACT structure with an applied surface acoustic wave. Only the epitaxial layers are displayed.

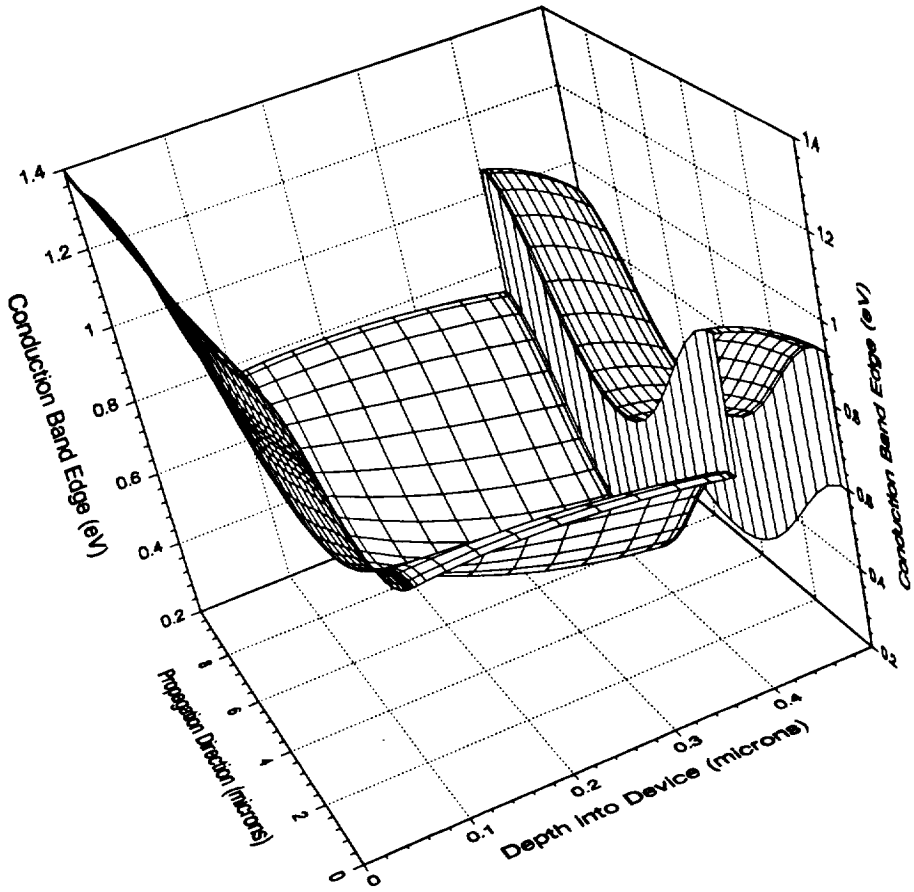


Figure 2.6 Two dimensional conduction band diagram for the p-HACT structure with an applied surface acoustic wave. Only the epitaxial layers are displayed.

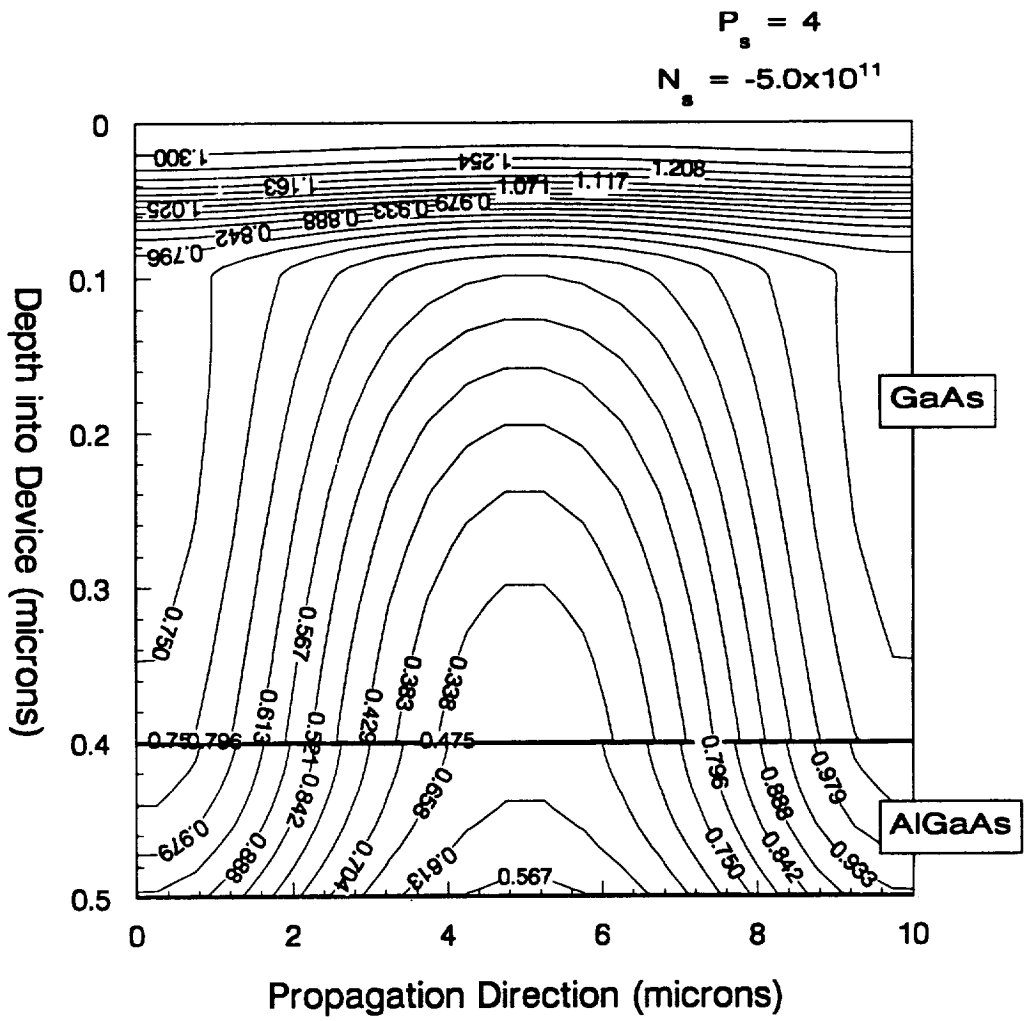


Figure 2.7 Equi-energy contour plot of the p-HACT structure with an applied surface acoustic wave and a residual surface state concentration of -5.0×10^{11} cm^{-2}.

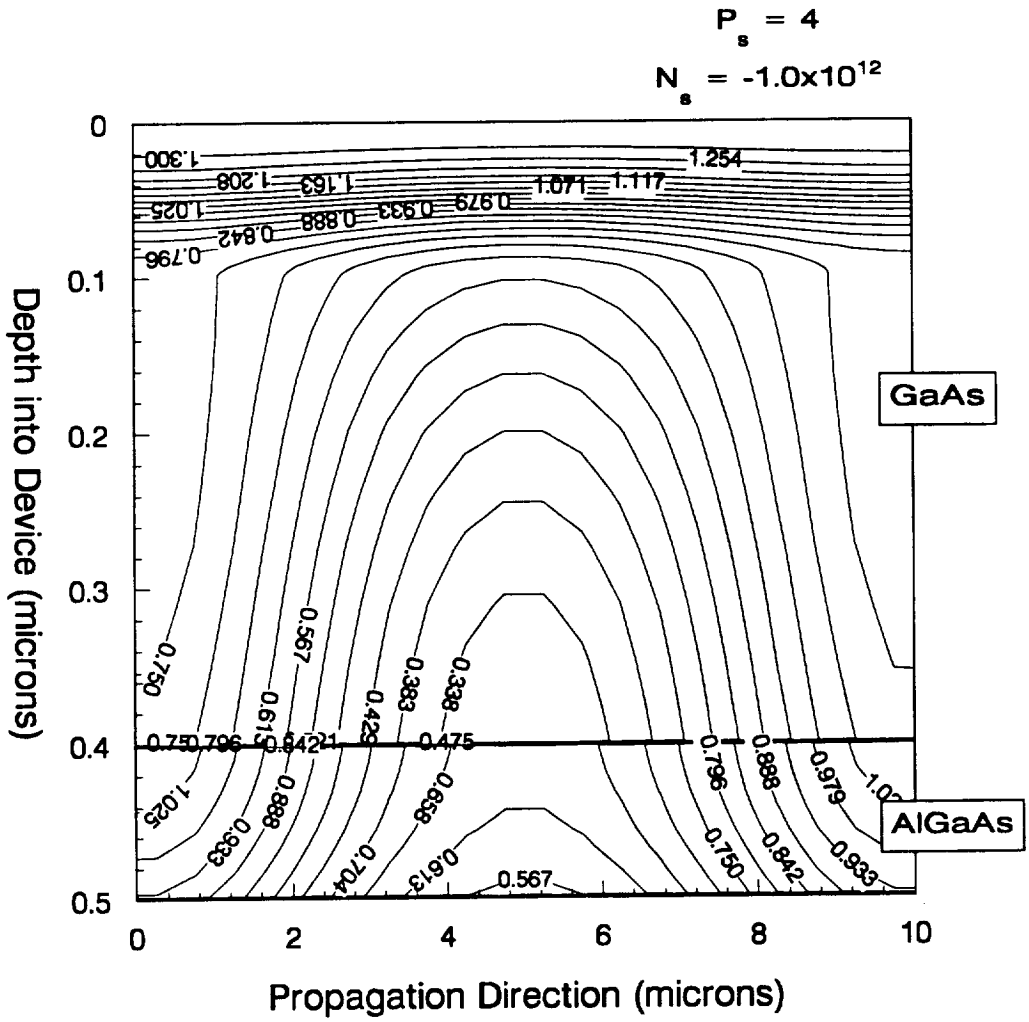


Figure 2.8 Equi-energy contour plot of the p-HACT structure with an applied surface acoustic wave and a residual surface state concentration of -1.0×10^{12} cm^{-2}.

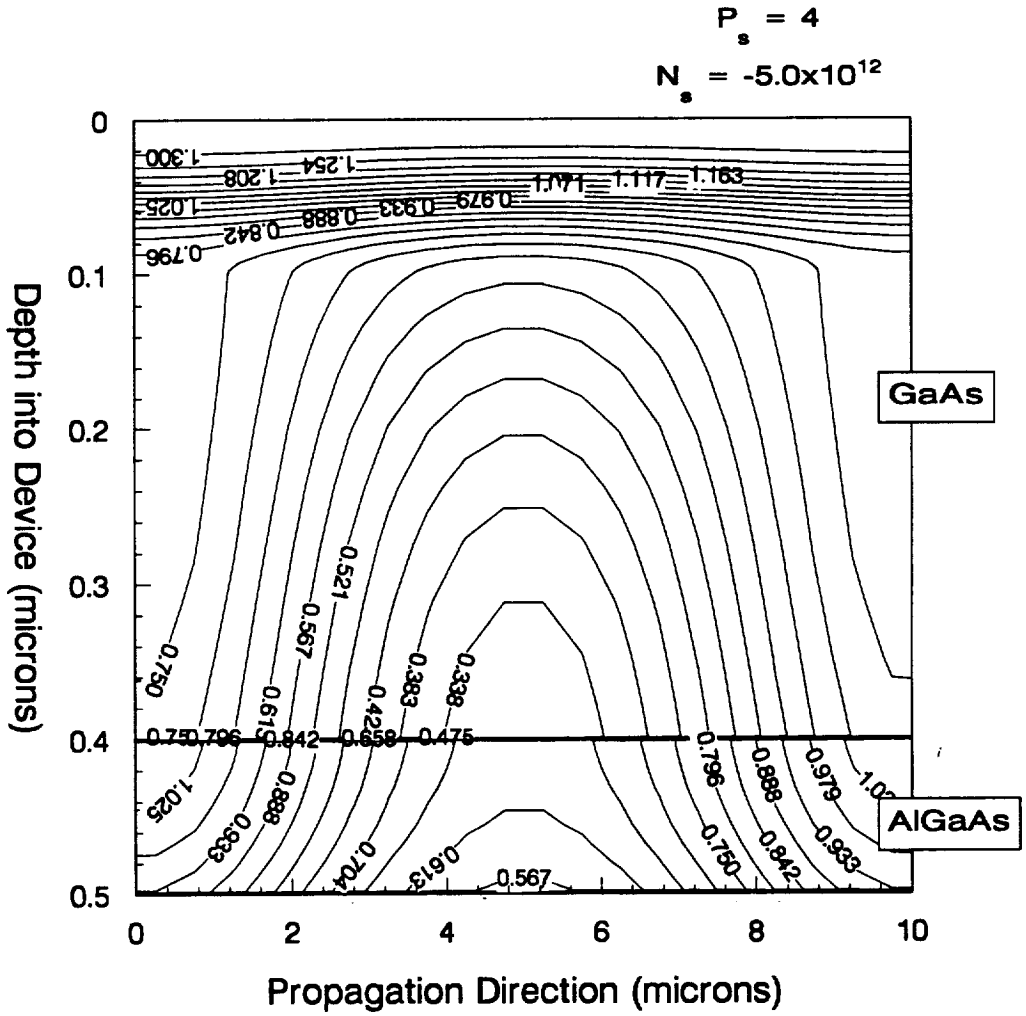


Figure 2.9 Equi-energy contour plot of the p-HACT structure with an applied surface acoustic wave and a residual surface state concentration of -5.0×10^{12} cm^{-2}.

electrons are not being confined within a single wavelength. Figure 2.10 is an example of the electron charge packet shape when 10^6 electrons/cm are injected into the middle of the simulation domain for the p-HACT structure. To determine the electron packet shape the electron density simulated at steady state was subtracted from the electron density simulated 0.6 nanoseconds after the charge was injected. Notice that the electrons are confined to the center of the simulation domain, which is the lowest part of the conduction band in Figure 2.9. The width of the packet is approximately 2 μm wide in the propagation direction. As will be seen in the figures to be presented the width of the packet increases as the amount of injected charge increases. Also note that an image charge, removal of electrons, has been created in the substrate just below the confining heterojunction. This image charge will not be plotted in later figures, but one should be aware that it exists.

The electron packet shape obtained using the p-HACT structure with a residual surface state concentration of $-1.0\text{x}10^{12}$ cm^{-2} and three different injected electron concentrations is shown in Figures 2.11 through 2.13. For an injected electron concentration of 10^7 electrons/cm, Figure 2.11, the packet is confined within 5 μm of the simulation domain. Even at 10^8 electrons/cm, Figure 2.13, the packet is still fairly well confined to the simulation domain. Due to the similarities in the characteristics of the p-HACT structure at other residual surface state densities charge packet shapes at other surface state concentrations will not be presented. Simulations at surface state densities of $1.0\text{x}10^{11}$ and $1.0\text{x}10^{13}$ cm^{-2} also show similar insensitivity.

To determine the maximum charge transfer efficiency of the p-HACT structure electrons are injected into the channel layer in the same manner as for the packet shape calculations. The structure is simulated for 1000 wavelengths, 4 μsec, at different time intervals. The amount of electrons remaining in the packet was recorded to determine the transfer efficiency. The fraction of electrons remaining in the packet as a function of time is shown in Figure 2.14 for three different values of SRH lifetime in the channel layer. At high SRH lifetimes the fraction of remaining electrons is very close to unity, indicating

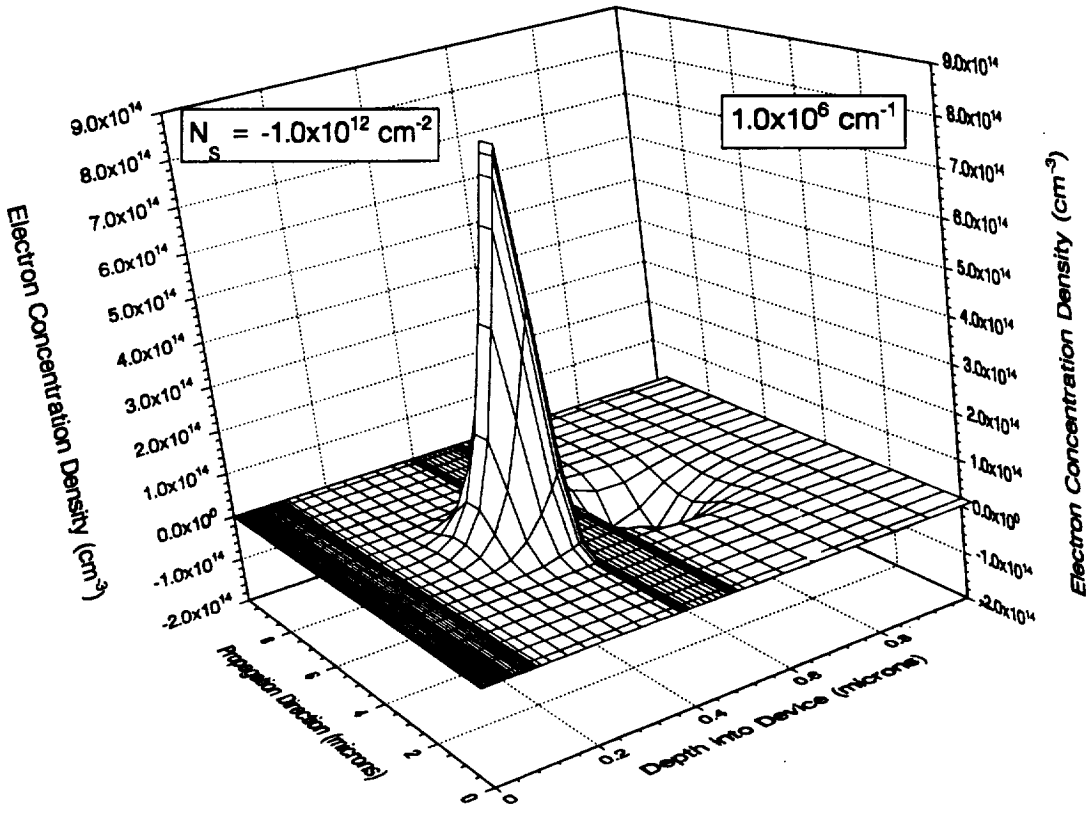


Figure 2.10 Electron packet shape for the p-HACT structure with an applied surface acoustic wave and 10^6 electrons/cm injected into the channel. The loss of electrons from 0.6 to 0.8 microns deep is the image charge.

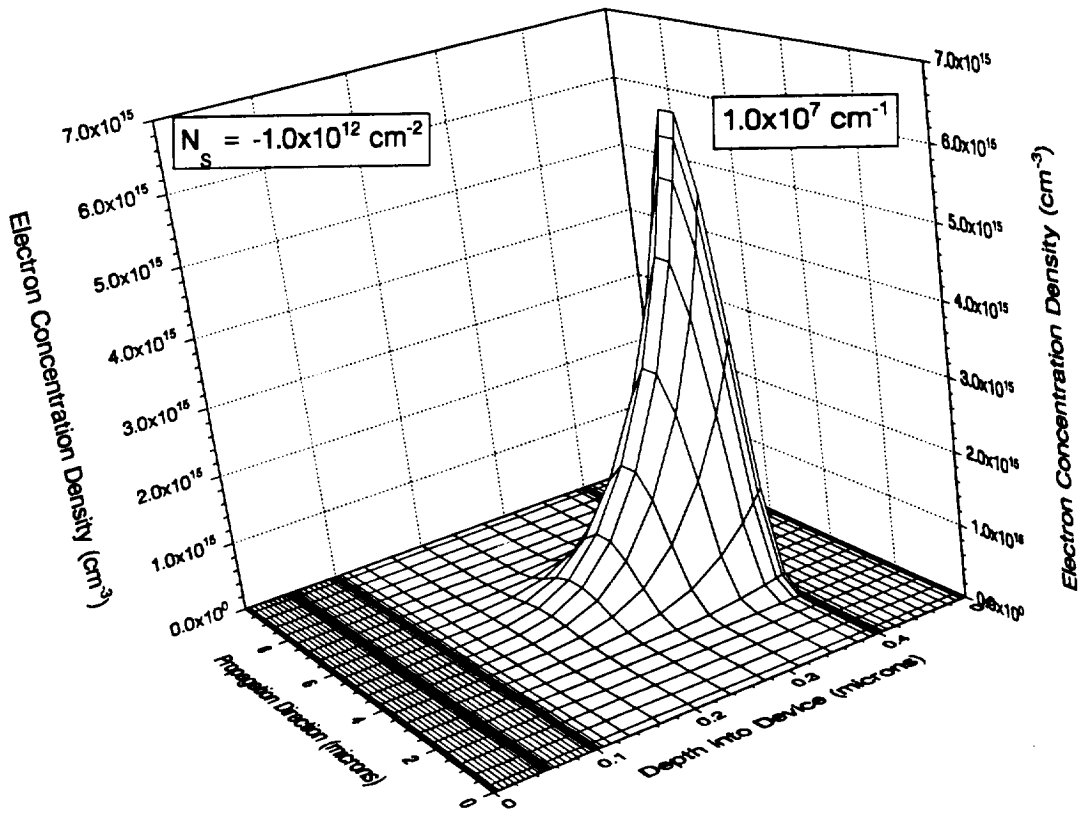


Figure 2.11 Electron packet shape for the p-HACT structure with 10^7 electrons/cm injected into the transport channel.

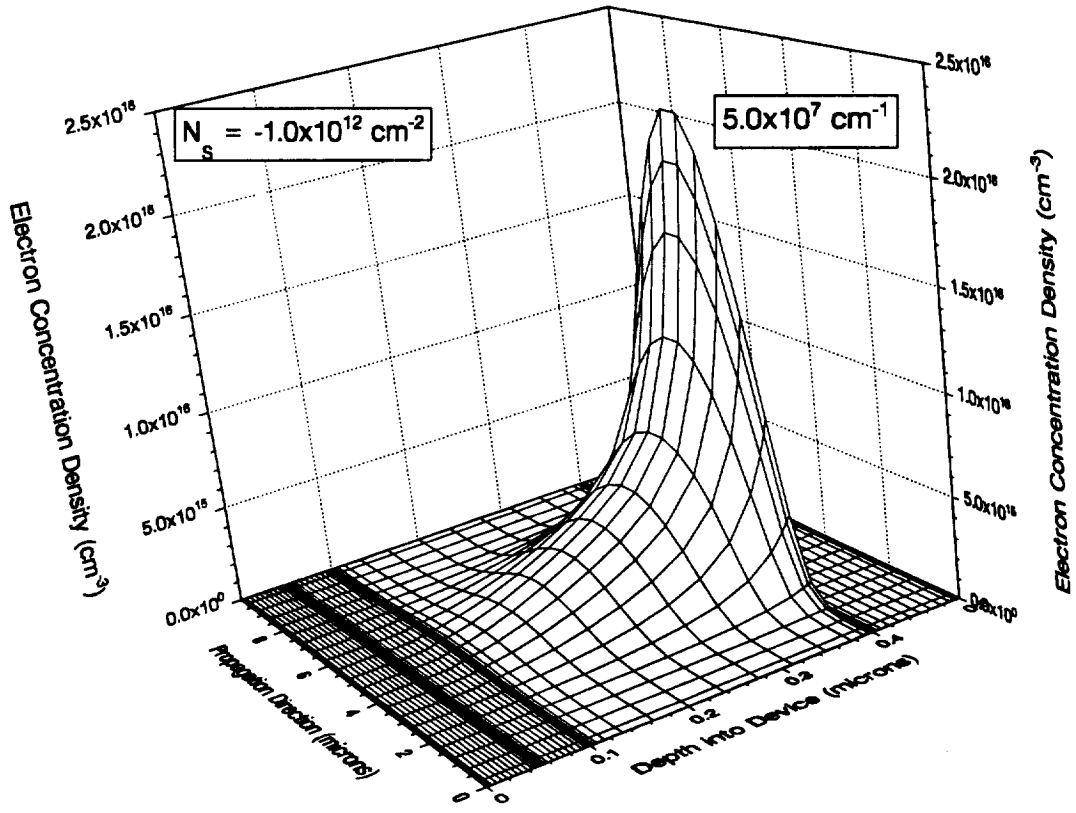


Figure 2.12 Electron packet shape for the p-HACT structure with $5x10^7$ electrons/cm injected into the transport channel.

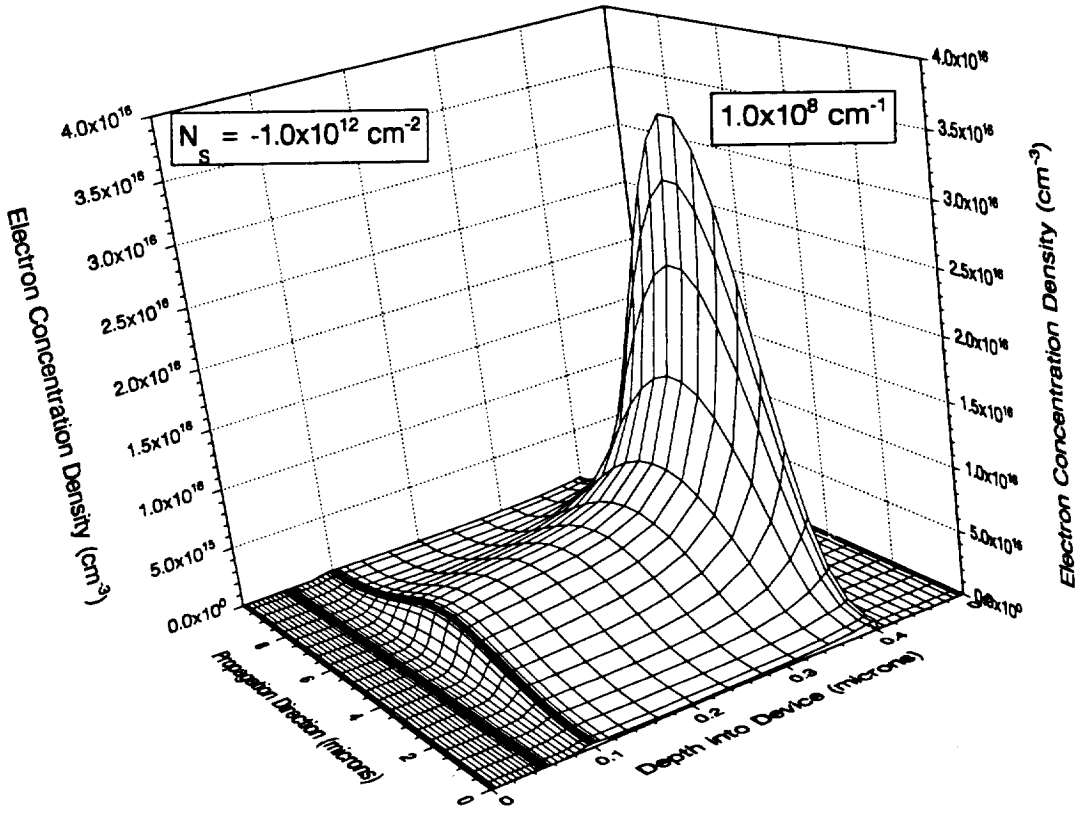


Figure 2.13 Electron packet shape for the p-HACT structure with 10^8 electrons/cm injected into the transport channel.

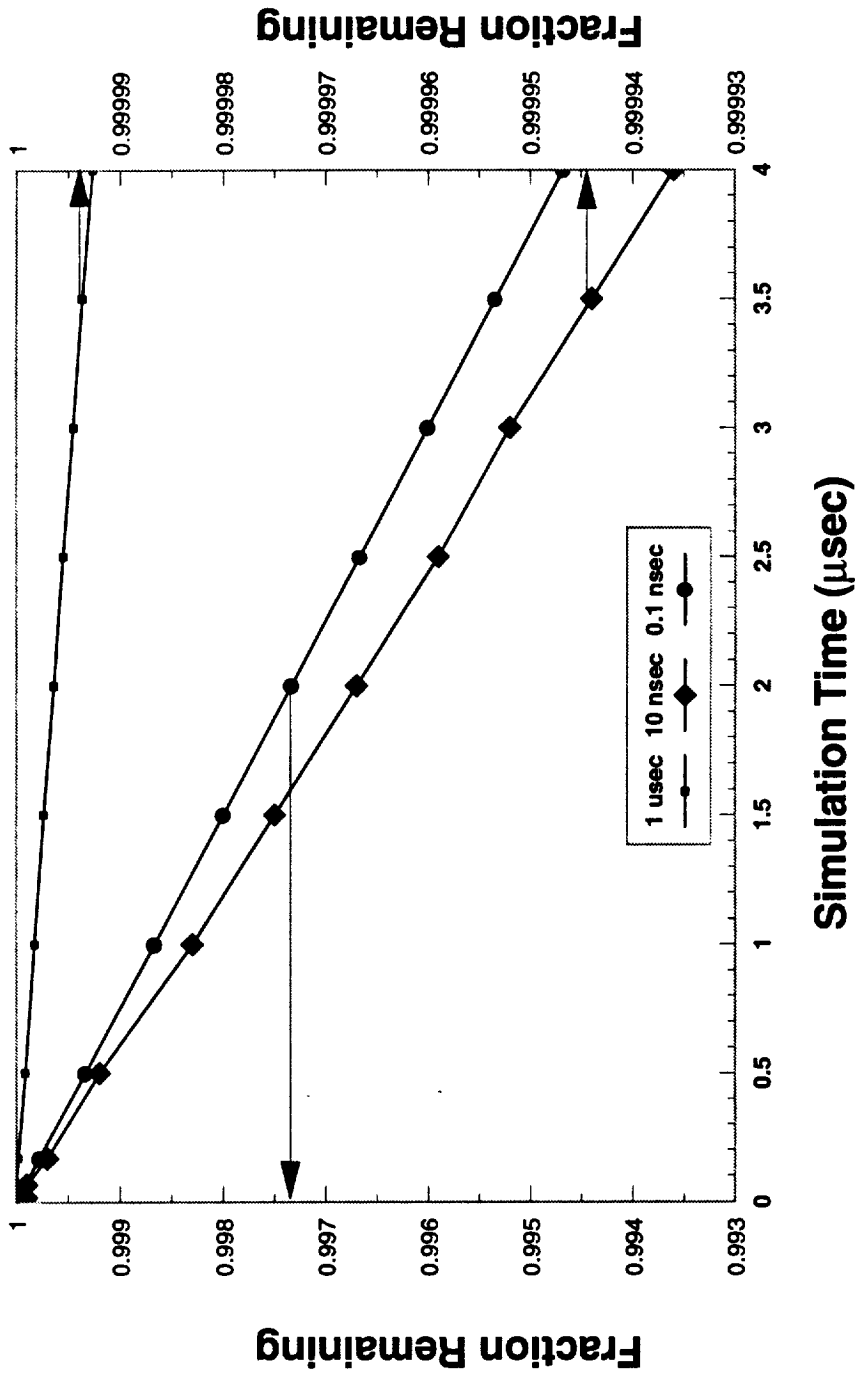


Figure 2.14 Fraction of packet electrons remaining in the p_HACT structure as a function of elapsed time for three different SRH lifetimes: 1 μsec, 10 nsec, and 0.1 nsec.

excellent charge transfer efficiency. The calculated charge transfer efficiency, given by the number of 9's in efficiency, as a function of SRH lifetime is given in Figure 2.15. For a SRH lifetime of 0.1 nsec the structure has 5 9's transfer efficiency and increases to 8 9's for 1 μsec lifetime. However, this analysis does not include any interface recombination at the heterojunction nor does it account for slow traps in the channel layer. Both of these effects will act to lower the charge transfer efficiency.

2.1.2 HACT Material Sensitivity Study.

While the layer thickness tolerance is constrained to less than 100 Å in the MBE process used to grow the HACT material, the layer doping tolerance is controllable only to within one significant figure. Thus, in order to evaluate the sensitivity of the HACT device performance on layer doping, we ran a series of simulation experiments. The figure of merit for the sensitivity evaluation was chosen to be the potential barrier height induced by the SAW. This has a direct effect on both charge capacity and charge transfer efficiency. The method chosen to investigate the sensitivity was "response surface methodology" (RSM). Basically this model assumes a standard form for the output parameter (i.e. the barrier height). In our case, linear terms plus first order interactions were included in the model. Thus we assume that the response surface, i.e. the barrier height is determined solely by these linear and first order interaction terms. The sensitivity is then determined by the gradient of the surface at the nominal settings. The coefficients for each term are determined by best fit from the simulation experiments. We chose to use two levels at 50% variation about the nominal for each layer doping. Surface state effects were also considered by including a nominal and an order of magnitude variation. The six-parameter problem resulted in a 32-experiment matrix. The levels for each layer and the surface state concentration are shown below in Table 2.1.

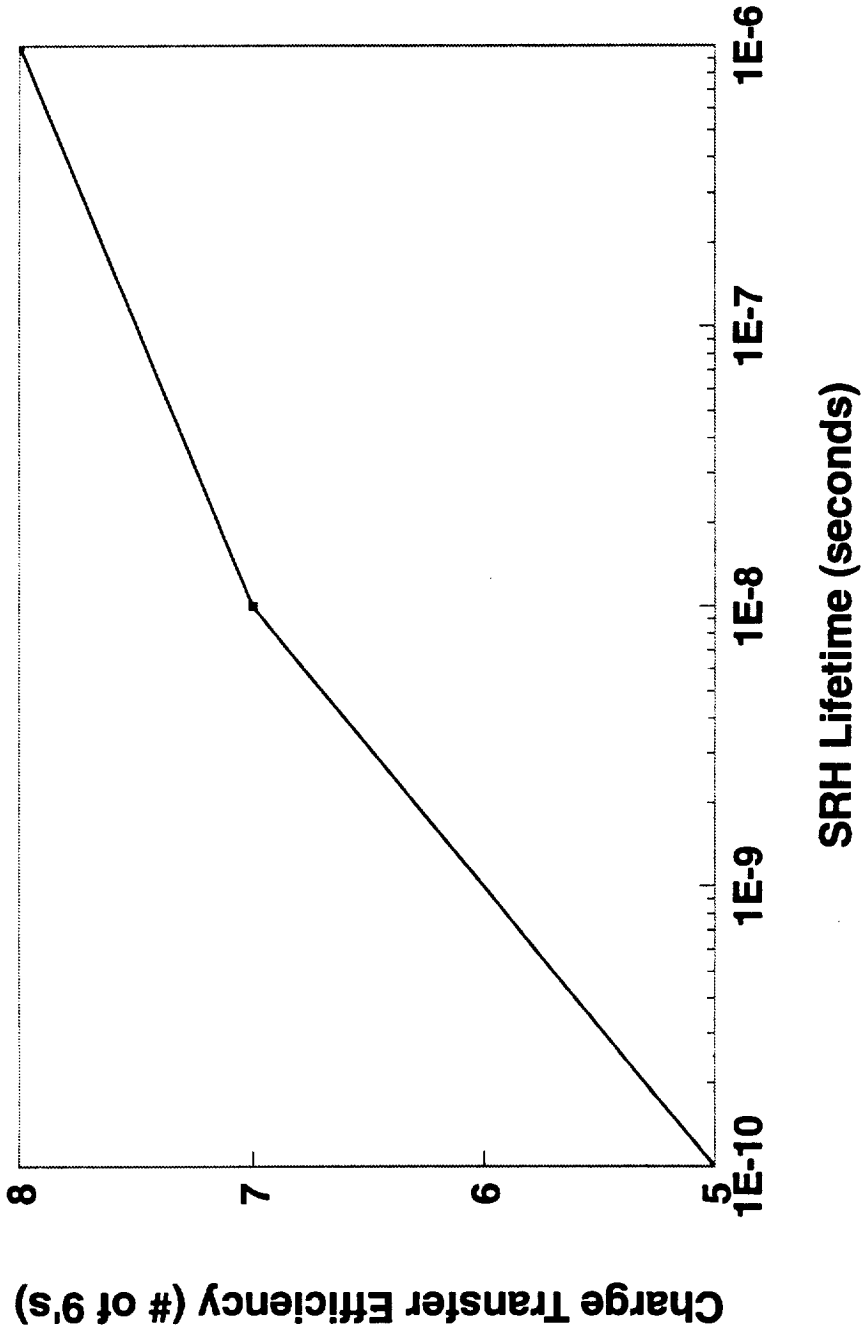


Figure 2.15 Number of 9's in the charge transfer efficiency as a function of the value of the value of the SRH lifetime in the channel region of the p-HACT device structure.

TABLE 2.1 DOPING SENSITIVITY SIMULATION EXPERIMENT

Layer	Thickness	Nominal Level	Low Level	High Level
Surface State	-	1×10^{12} cm^{-2}	5×10^{11} cm^{-2}	5×10^{12} cm^{-2}
p-GaAs	500 Å	3.25×10^{17} cm^{-3}	1.63×10^{17} cm^{-3}	4.88×10^{17} cm^{-3}
n-GaAs	500 Å	1.60×10^{17} cm^{-3}	8.00×10^{16} cm^{-3}	2.40×10^{17} cm^{-3}
NID-GaAs	3000 Å	5.0×10^{15} cm^{-3}	7.5×10^{14} cm^{-3}	2.5×10^{15} cm^{-3}
p-AlGaAs	1000 Å	1.0×10^{16} cm^{-3}	5.0×10^{15} cm^{-3}	1.5×10^{16} cm^{-3}
GaAs Buffer	2 μm	3.0×10^{15} cm^{-3}	1.5×10^{15} cm^{-3}	4.5×10^{15} cm^{-3}

The results of the sensitivity analysis are shown in Figure 2.16. As is seen, the material architecture can tolerate a 50% variation in doping of any given layer with only small degradation in barrier height. This is further verified by the histogram of all experiment responses shown in Figure 2.17.

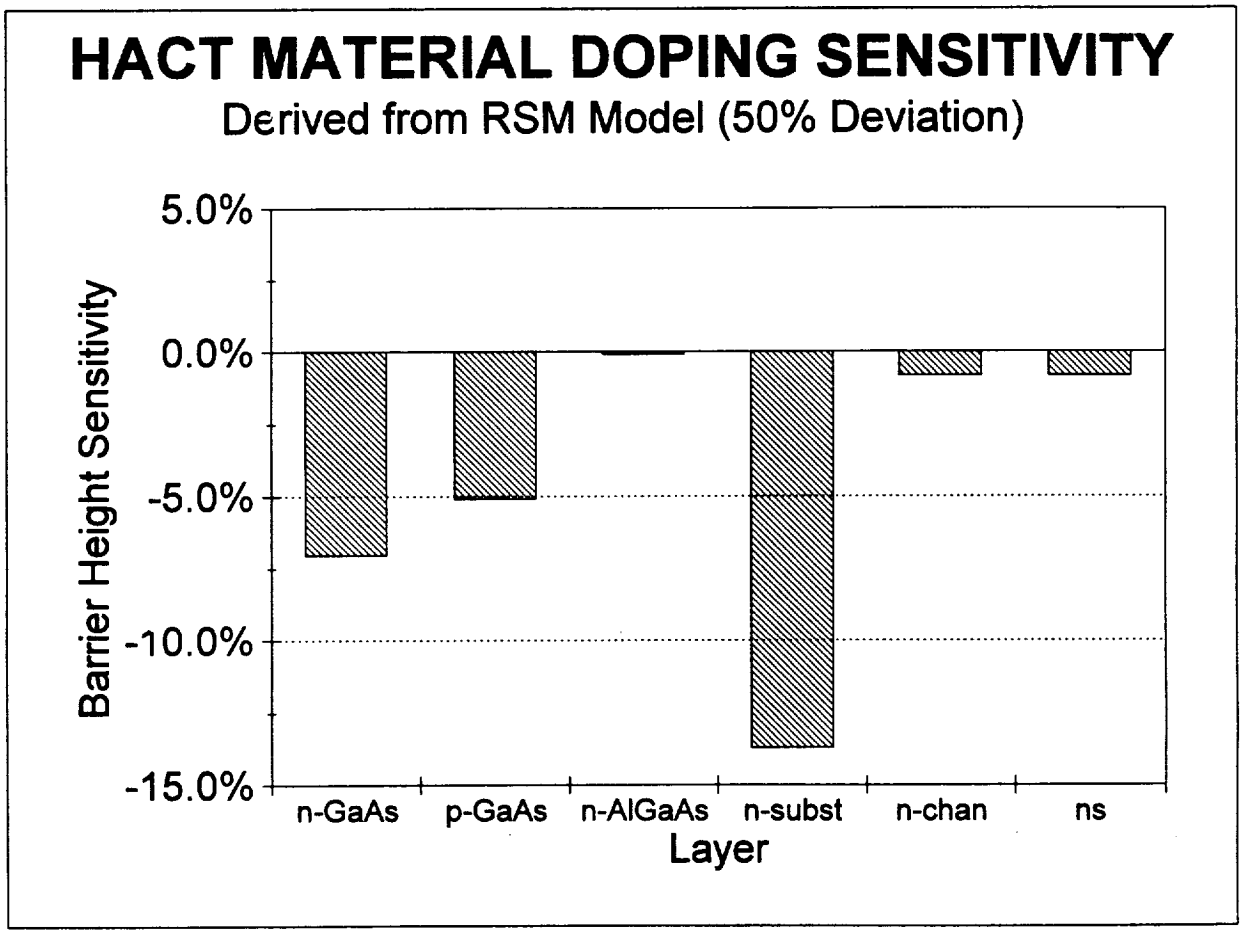


Figure 2.16 Results from HACT material sensitivity analysis.

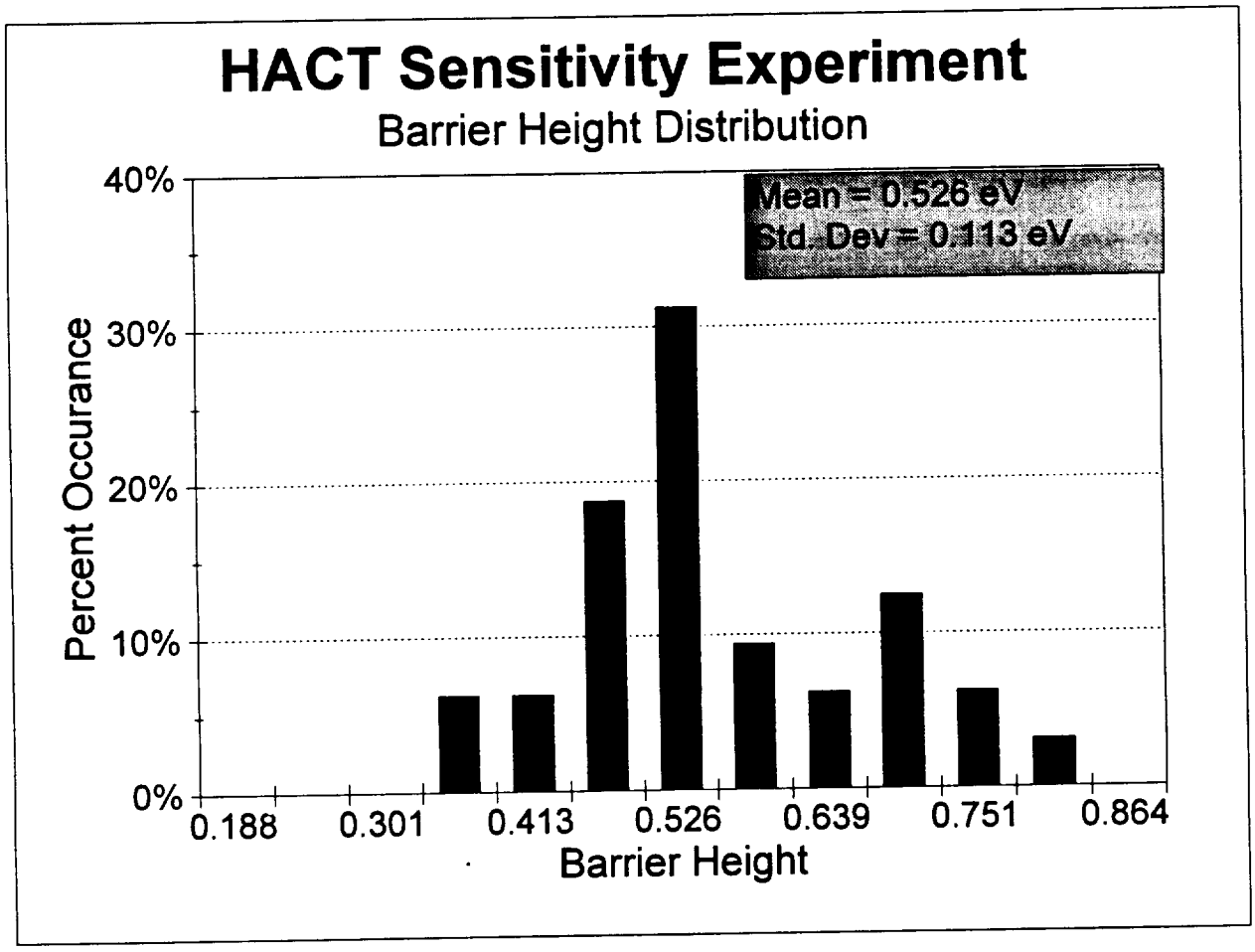


Figure 2.17 Response histogram for the 32-experiment sensitivity analysis.

2.2 Mask Design, Fabrication and Device Evaluation

2.2.1 Mask Design

In order to more fully investigate the enhanced operation made possible by the improvements in the HACT material architecture discussed in the last section, a new mask set was designed. Table 2.2 compares the new mask feature with the previously used mask. The major goals were to improve the processing throughput and the ease of testing. This was accomplished by reducing the die size by roughly a factor of four, allowing more devices to be processed per wafer run. The new layer structure requires that mesas be etched to half-micron or larger. To ensure that photoresist was completely removed in the patterns that overlapped the mesa, small linewidths were increased in width where they ran over the mesa. The feature sizes on the IDT were also increased to improve the DC probe yield. These improvements resulted in a DC probe yield of better that 80% for the new mask, compared to approximately 50% on the previous mask.

In addition to processing improvements, several changes were made to improve the performance to allow easier testing. Diagnostic structures, used to measure CV characteristics, ohmic contact resistance and sheet resistance, were incorporated onto every die. Variations in the uniformity of the material can be identified with these structures, in conjunction with the unique serial numbers on each die. The new design is expected to have an increased output signal level, reduced noise, and lower SAW power. The new mask design is shown in Figure 2.18.

TABLE 2.2 New Mask Features

Feature	Old Mask Design	New Mask Design
Number of Devices per Wafer	14	56
Die Area	18.5 mm × 9.5 mm	8.5 mm × 5.5 mm
Number of Designs	4	2
Test Cell Locations	2 sites per wafer	Located on every die
No. Schottky & Ohmic Bond Pads	1	2
Transducer Line/Gap Width	2 μm	2.5 μm
Maximum Number of NDS Taps	32	100
Min Metal Width over Mesa	1 μm	2.5 μm

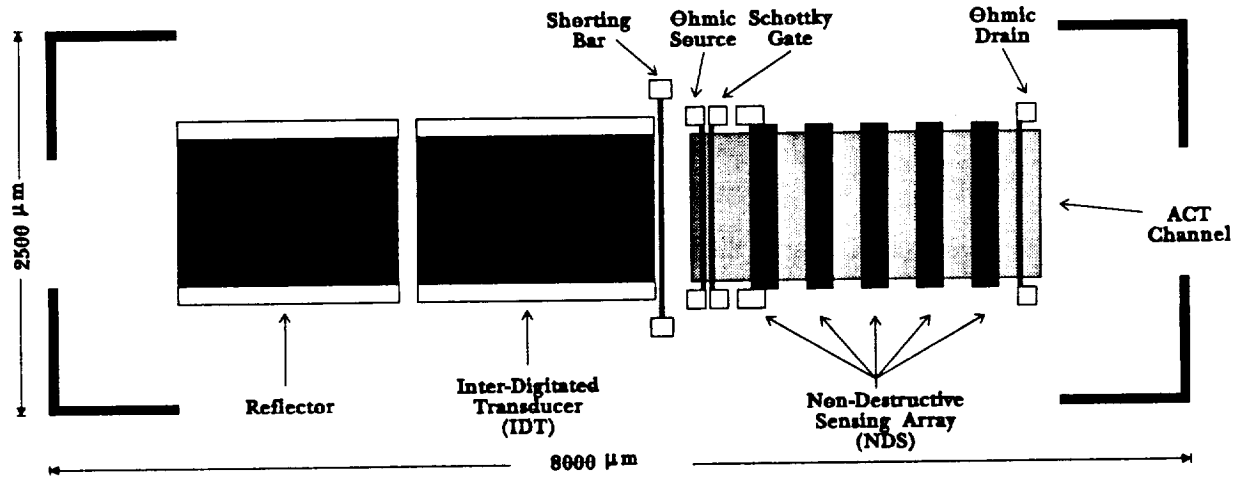


Figure 2.18 Mask Layout for New Test Device.

2.2.2 Test Device Evaluation

As mentioned in the previous section, the SAW transducer design was improved not only to increase the probe yield, but also to allow lower power (or higher performance) operation. This was done in two ways. The reflector design was modified to increase the acoustic potential in the transport area to allow high charge capacity at minimal power levels. Figure 2.19 shows the return loss plot of the new IDT/Reflector design. As is seen, the return loss is better than 8 dB, indicating that more than 80% of the electrical power is converted into acoustic power. This is a vast improvement over the older design, which had a conversion efficiency of only about 30%. A second improvement in the new mask set was the inclusion of an additional mask to allow a piezoelectric overlay material, such as zinc oxide, to be deposited on the wafer. Development is underway to produce high quality ZnO films by sputter deposition. This is expected to reduce the power necessary to operate the HACT device by a factor of ten.

In addition to the transducer design, the test device was also improved to allow easier testing. The channel width was increased by 40% to reduce the insertion loss by approximately 4 dB. The gate-source spacing was also reduced from 2.5 μm to 1.5 μm to improve the injection efficiency into the test device. The number of taps in the to designed was also increased to allow more accurate determination of the charge transfer efficiency and noise level. The development of the HACT material is still in process, so no results from the new mask set are available at this time.

2.2.3 Device Fabrication

Major changes to the HACT device have had a significant impact on wafer processing since the last reporting period. Our first wafer processing run using the new HACT design produced working devices and we are optimistic that further work will lead to increased reliability and yield. Figure 2.20 shows a micrograph of a fabricated HACT. Work was also performed in the area of material validation.

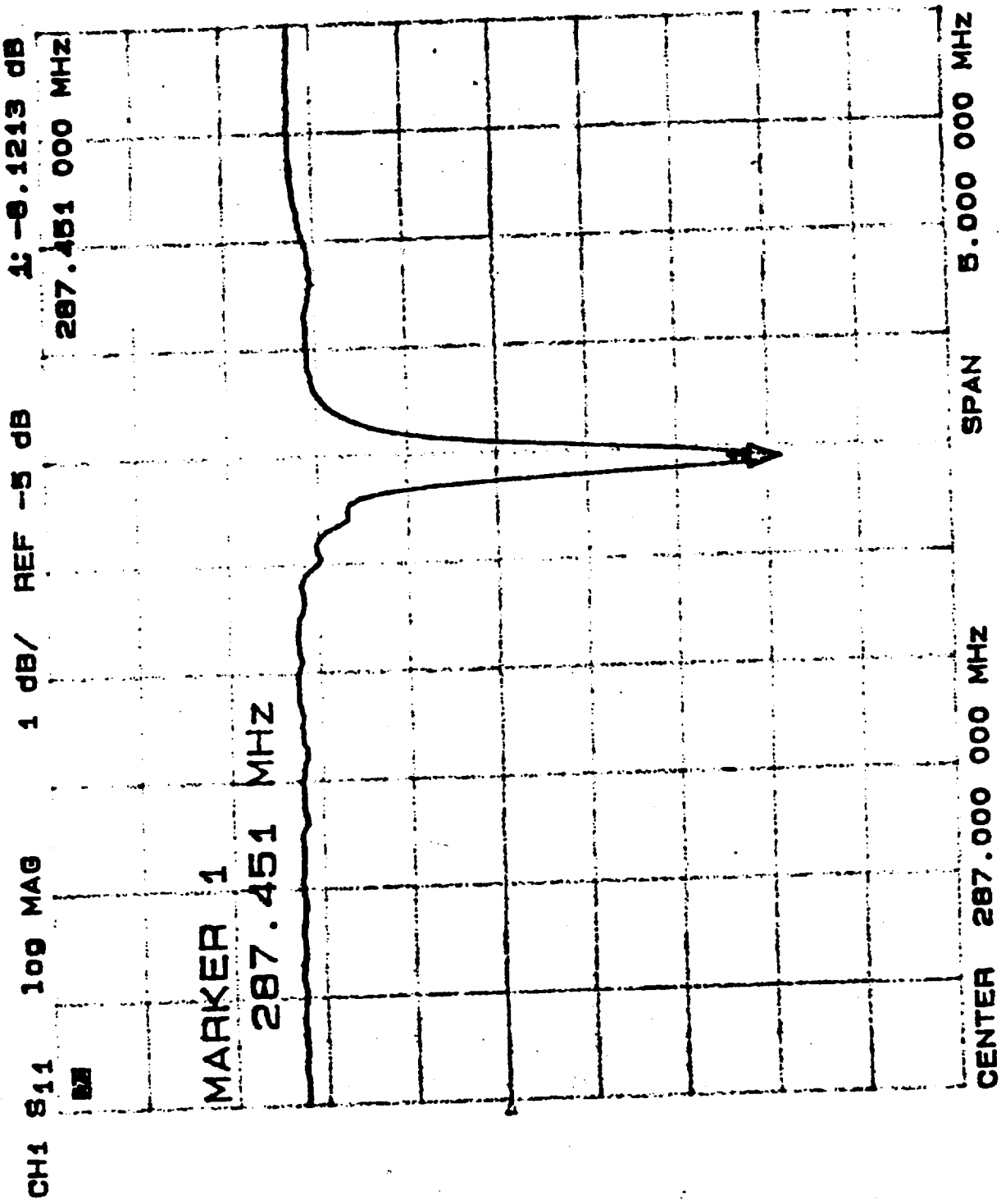


Figure 2.19 Saw transducer return loss.

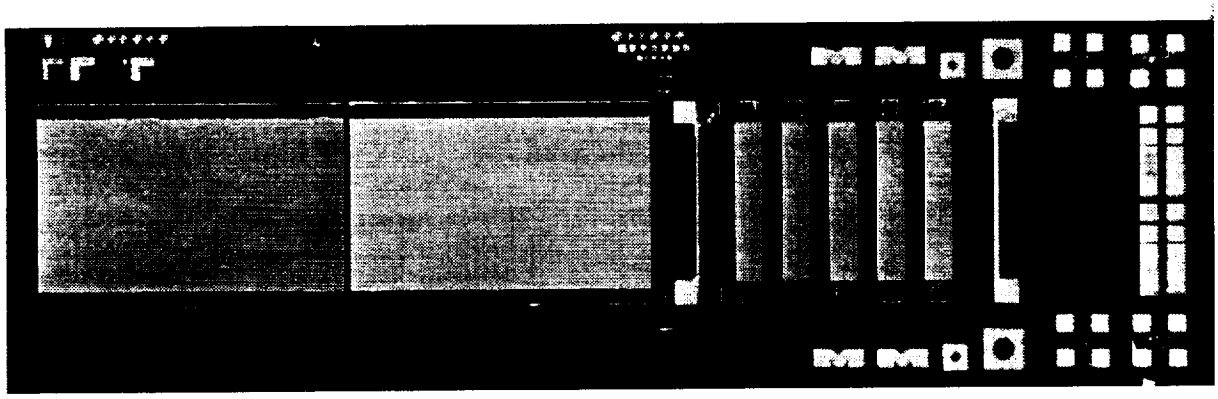


Figure 2.20 Micrograph of fabricated HACT.

The current layer structure is designed to be completely depleted, as a result, work was done on high resistivity materials. Techniques to measure critical parameters were evaluated.

Wafer Processing

The fabrication process for the new mask set is basically the same as before. With the new layer structure, the ohmic contacts and Schottky diodes need optimization, possibly with a different process; however, for now they work fine.

Work is currently in progress to investigate these issues: Posistrip 830, made by EKC Technologies is a positive photoresist stripper commonly used during lift-offs. There was evidence to suggest that the Posistrip was altering the properties of the material. Slight etching of the GaAs was observed during long soaks in the stripper. After consulting with EKC Technologies, it was learned that the 802 product is specifically formulated for III-V semiconductors and safe for GaAs. EKC Technologies verified that 830 was known to etch GaAs. Posistrip 802 was ordered as a replacement.

The new layer structure is designed to be fully depleted by carefully balancing the charge in each layer. This makes pre-processing validation of material difficult. In the previous structure, a sheet resistance was measurable using a Tencor M-Gauge. However, the M-Gauge at Georgia Tech did not have the range to measure the wafers. An extended range was ordered that would read up to 20,000 ohms. This equipment was installed, calibrated and verified sheet resistance measurements taken with another system. Now that the layer structure has changed, the effective sheet resistance of the material is in the megohm range. The Tencor M-Gauge is not capable of measuring the new material.

Due to the depth of the AlGaAs layer in the material, the mesa etch was increased from 0.17 mm to 0.5 mm. To achieve this height the photoresist must be hardbaked at 110°C for 10 minutes after developing. This makes the photoresist pattern more resilient to the etchant and allows it to withstand the long etch times.

Since the bonding pads are no longer included in the Schottky layer, the gold used for the Schottky layer has been reduced from 3000Å to 1000Å. The bonding pad layer consists of 500Å titanium and 4000Å gold.

Electrical Characterizations

As stated previously, the new HACT design relies on a completely depleted material structure which makes material and process characterizations difficult. Two wafers were completely fabricated and packaged for testing. The first sample, B93-105, produced working HACT devices and the second, B94-15, did not.

The first process control measurement taken is the ohmic contact resistance. This verifies that the mesa, metallization and alloy steps were all successful. With the current material, accurate contact resistances were difficult to measure. As seen in Figure 2.21, the ohmic IV plot for B-93-105 is not linear and resembles two back-to-back diodes. Modeling data confirmed that this type of curve was expected. The ohmic contacts, as expected, were photosensitive, as shown in Figure 2.22. The resistance decreased when illuminated due to photogenerated carriers. B94-15 however, showed a very high resistance with and without illumination.

After the Schottky metal is deposited, diode curves can be measured. Typical characteristics for the two completed devices include V_T from 2.5-3.5 volts and reverse leakage current of -1.0 mA at -4 V. Forward resistances of the diodes were between 1 and 2 kilohms (see Figure 2.23). Future work will include attempting to decrease to V_T to less than 1 volt and improve the ideality factors which are above 10 due to the high V_T.

Statistical Process Control

In order to maintain consistent process results, excluding variations related to material properties, a method of statistical process control (SPC) was implemented. A spreadsheet was developed, and is being revised for the new layer structure, to track the wafer from growth, through fabrication, to final RF

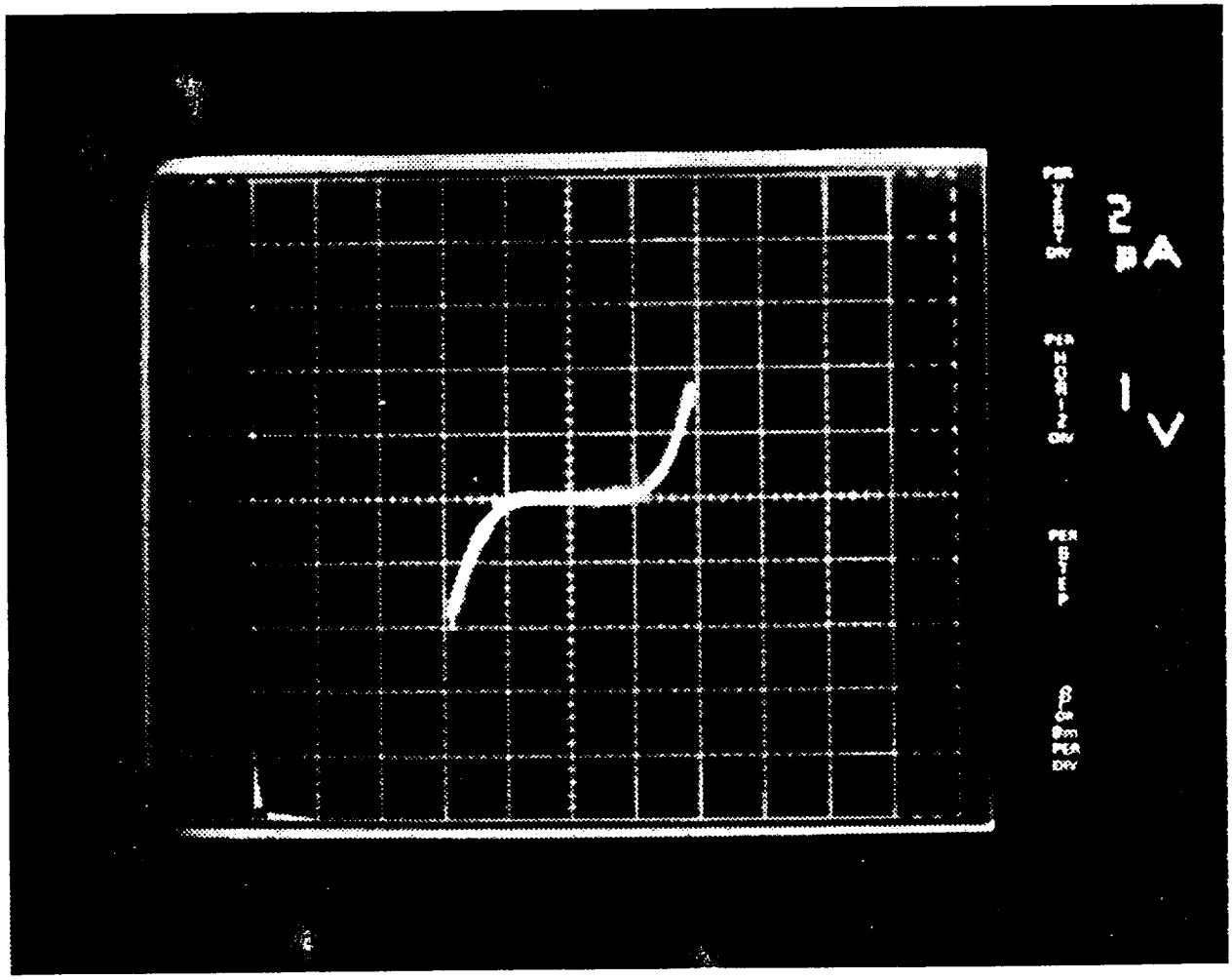


Figure 2.21 Ohmic contact IV plot for B93-105.

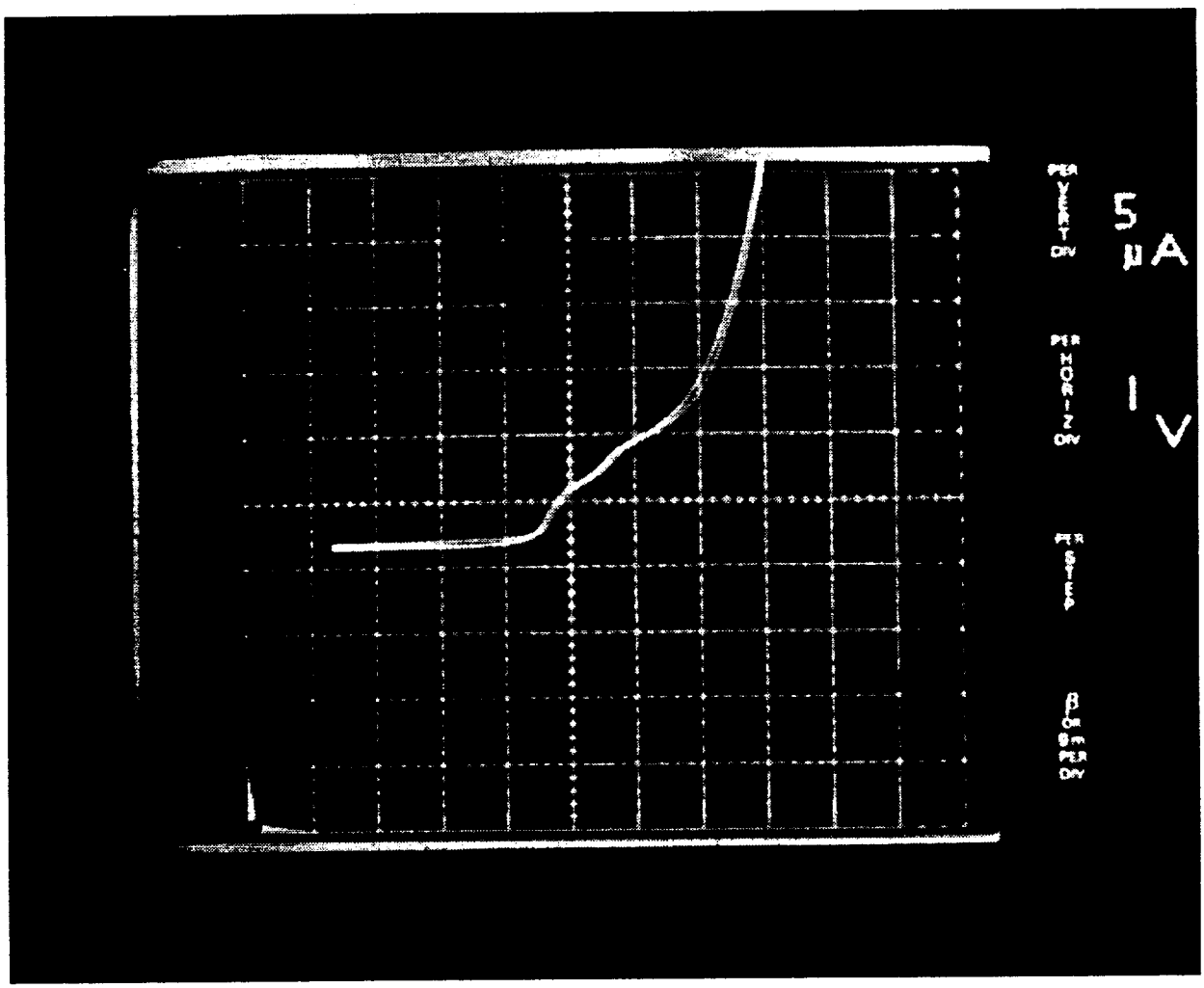


Figure 2.22 Light sensitivity of ohmic contact IV on B93-105.

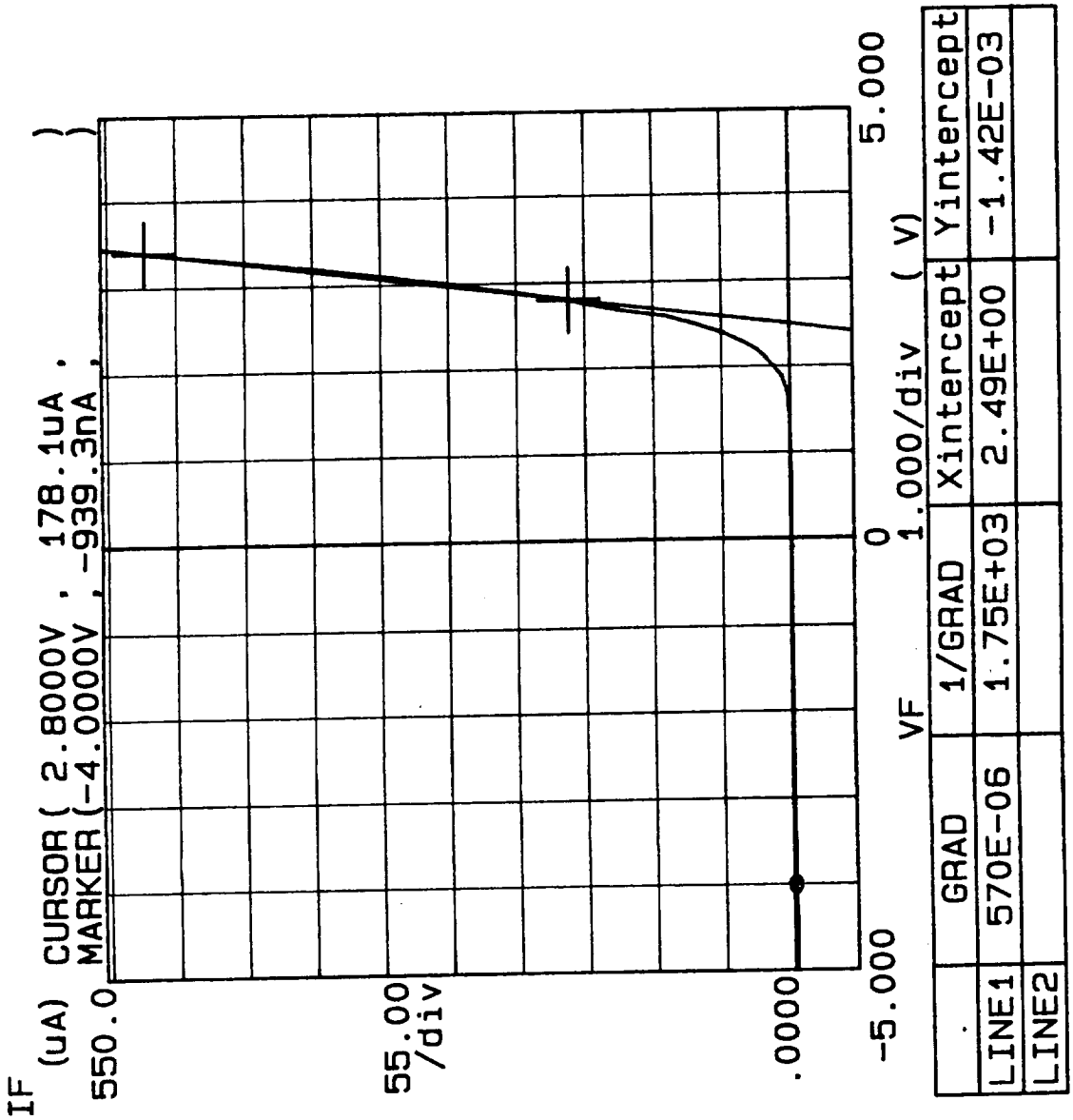


Figure 2.23 IV characteristics of Schottky contacts showing forward bias resistance characteristics.

testing. Critical parameters such as contact resistance, diode characteristics and drain-source resistance are measured and documented. This system allows for a better comparison of devices an later for developing models.

To aid in the SPC, an automated testing station and computerized test equipment is proposed to automate the process of probing an entire wafer. Wafer variations can be found and the devices can be analyzed faster and with more accuracy with this system.

Future Plans

The next six months will be dedicated to refining the process for the new layer structure. Ohmic contact and Schottky diode optimization will continue to be a top priority. The acquisition of an automated probe station will improve Georgia Tech's ability to completely characterize the devices prior to packaging and help provide a connection between device performance and dc characteristics. This will allow out of spec. material to be identified as soon as possible and prevent futile processing. Additional work is planned to validate the quality of the material before processing, as well as each at processing step, to find problems and acquire characterizations before the devices are completed.

2.3 Acoustoelectric Amplifiers

Acoustic charge transport is caused by the non-linear acoustoelectric (AE) effect. In a piezoelectric semiconductor with mobile carriers, amplification of an RF signal is possible due to the linear AE effect. AE amplification in piezoelectric semiconductors was first published in the 1961 [1], and many subsequent devices were developed over the next twenty years. The basic AE amplifier is shown in Figure 2.24. The device has launching and receiving transducers on either end and ohmic contacts in between which apply the drift voltage. As illustrated, the electrical signal launches an acoustic wave via the transducer. The wave is amplified by the interaction with the drifting carriers and converted back to an electrical signal at the output transducer.

2.3.1 Device Physics

In order to briefly describe the physical origin of AE amplification, a very clear understanding may be had from an analytical model by Adler [2]. The analytical explanation goes as follows. Consider the acoustic wave as a comb of charge brought in contact with the surface of a conducting material as in Figure 2.25. If the material is a metal, a pattern of induced image charges is set up almost immediately. If the material is a poor conductor, then this happens after some time interval. In any case after some time constant,τ, which is dependent on the distributed resistance and capacitance, the image charge pattern mirrors the induced charge pattern. Now, if the comb pattern is moved parallel to the surface, as would be the case for an acoustic wave, the image charge pattern will follow the comb forming a wave of induced charge. In order to create such a wave charges must move back and forth along the surface. If the material is a poor conductor the charges will always lag the comb and the charge wave will never reach its normal amplitude. This is analogous to a simple R-C circuit where the quadrature component is:

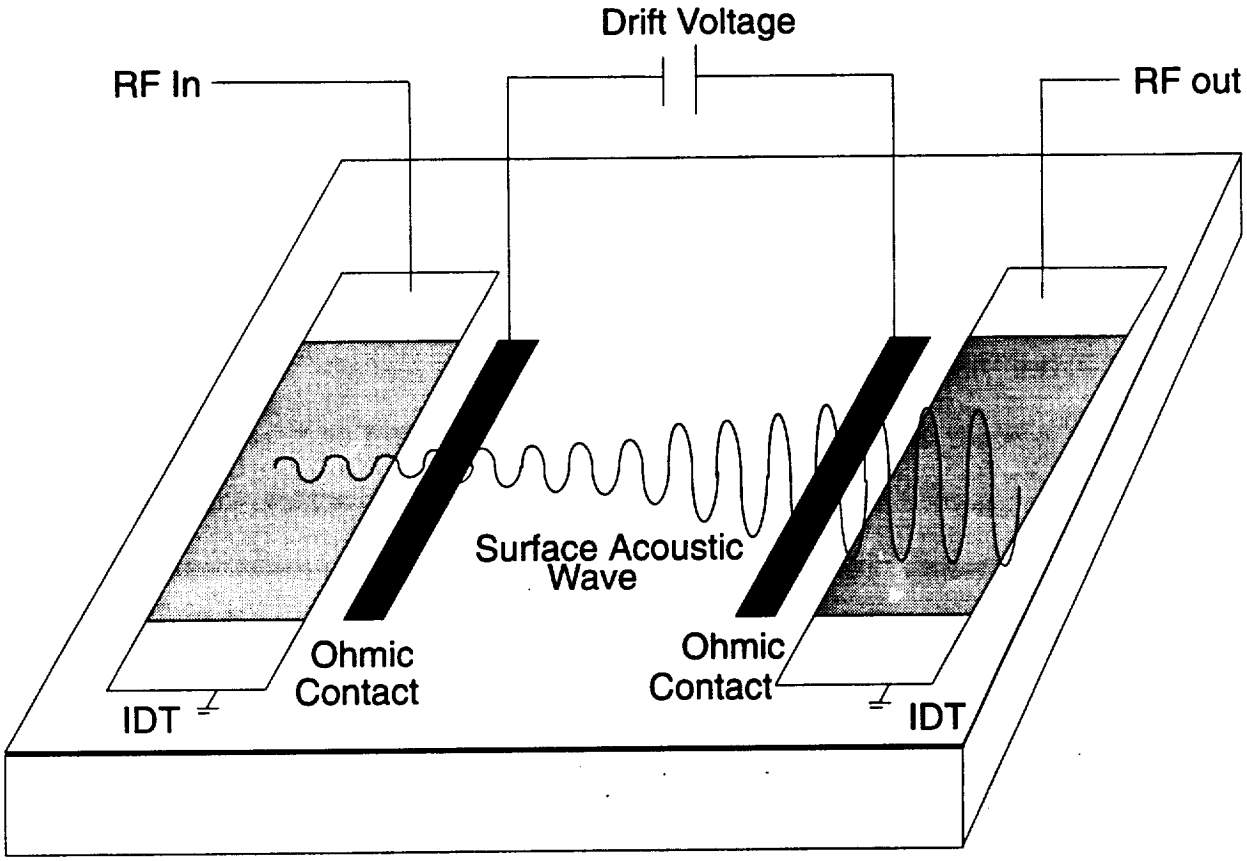


Figure 2.24 AE amplifier.

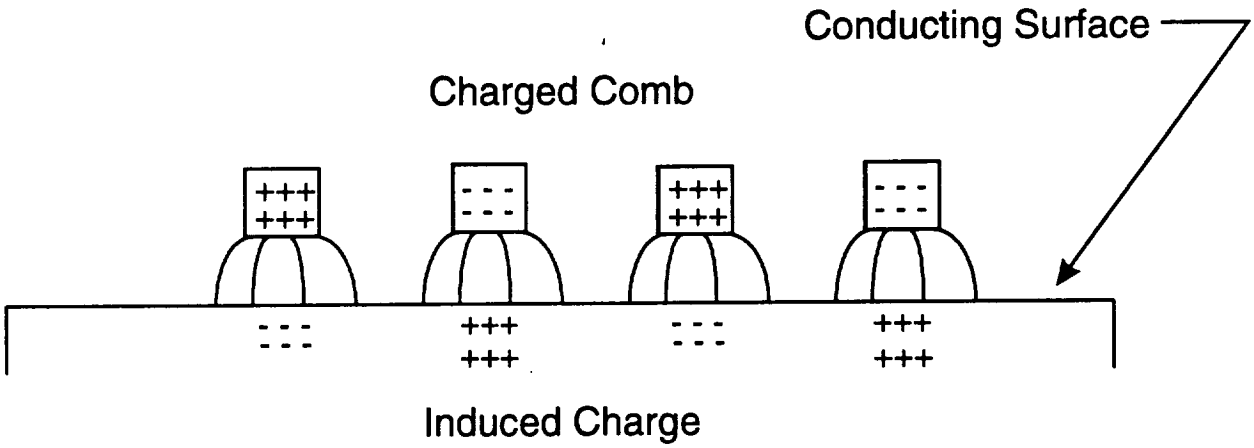


Figure 2.25 Acoustic wave as comb of charge

$$V_0 = V_i \frac{\omega\tau}{1+(\omega\tau)^2}$$

Now assume that this poor conductor is in fact a semiconductor with mobility μ. If we apply a dc field E, the mobile carriers will drift forward with a velocity v = μE. If the comb is moved along with the same velocity, then relative to the comb, the carriers appear to stand still. Thus, a semiconductor with an applied drift field looks like a shorted surface to an acoustic wave traveling at the same speed as the drifting carriers.

If the comb is moved forward faster than the drifting carriers at velocity V, then the charges lag the comb and the material appears as a lossy medium. If the comb moves slower than the drift velocity and the carriers lead, then the quadrature component can again be found from our simple R-C equation by substituting for ϖ_a, the Doppler shifted frequency of the traveling wave. If we define ϖ as the frequency measured in the laboratory frame and V_0 as the phase velocity of the comb, then we have $\varpi_a=\gamma\varpi$ where $\gamma=(V_0-V)/V_0$.

When this model is applied to an acoustic wave propagating on a semiconductor a quadrature component α/k_0 arises that constitutes a loss per radian equal to

$$\frac{\alpha}{k_0} = \frac{1}{2}K^2 \frac{\gamma\varpi\tau}{1+(\gamma\varpi\tau)^2}$$

where K^2 is the acoustic coupling coefficient.

From this simple equation, the origin of the gain is evident. If the drift velocity is less than the acoustic velocity then γ is positive and so the loss α is positive. However, if the drift velocity exceeds the acoustic velocity, then the loss α is negative indicating that we have gain. There are many more complicated models which describe AE amplification, but this simple analytic technique illustrates the

effect most clearly. For any piezoelectric semiconductor, the gain may be determined to first order simply from the mobility, the dielectric constant and the applied drift voltage.

2.3.2 Motivation

AE amplifiers, although initially thought to be a miniaturized traveling wave tube had limited commercial applications due to high noise figures and limited linearity compared to transistor amplifiers. However, they were used in special applications in very long acoustic delay lines and convolvers. The ACT imager is acoustically just a very long delay line and our primary motivation is to incorporate AE amplifiers in between pixels to maintain the acoustic wave amplitude along the length of the imager. With regards to power consumption it is more efficient to distribute the gain along the length of the device than to have all the gain in the RF source driving the input transducer. Gains of 50 dB per cm (0.7 dB/ wavelength) have been reported on O_2-doped GaAs [3] utilizing a surface shear mode. Given the ability to achieve high mobility and high resistance in heterostructure GaAs and the improved acoustic coupling coefficient from using a ZnO overlay, we are confident that very high gains will be achieved in a short distance with low drift voltage.

There are spin-off applications in other HACT devices, oscillators and tunable filters for commercial telecommunications products.

2.3.3 Research Plan

Research is presently underway to characterize the acoustic loss on thick epilayers of doped GaAs. Subsequently, ohmic structures will be fabricated on these same substrates to characterize the gain. From these studies a model will be developed which can then be extended to multilayer structures such as those which are used in the HACT device.

Ultimately we intend to develop AE amplifiers on multilayer structures which will occupying small areas of real estate and provide high gain for low drift voltage leading to a more efficient HACT device.

2.4 Analysis of Reflection of Metal Gratings on ZnO/GaAs

In ACT devices, the charge packet can be detected by non-destructive sense (NDS) electrodes on the surface for a signal processing application. Consequently, it is important to analyze the reflection properties of the NDS electrodes, i.e. metal gratings.

An exact theoretical analysis of the reflection properties is extremely difficult. The reflectivity and velocity shift of metal gratings on a single material can be analyzed by using Datta and Hunsinger's perturbation theory [4] when the metal thickness is small comparing to the SAW wavelength, λ.

Obtaining effective material constants for a multilayered substrate enables to directly apply the perturbation theory developed for a single material to a multilayered substrate (ZnO/GaAs). The effective constants, c_{eff}, may be obtained by integrating multilayered material constants weighted with SAW energy distribution as follows,

$$c_{eff} = \frac{\int dy\, c(y) P(y)}{\int dy\, P(y)}$$

where $c(y)$ is a constant of the multilayered material tensor and $P(y)$ is the energy distribution as a function of the depth, y. Since the substrate is piezoelectric, the mechanical, mechanical-electrical, and electrical energy distribution are utilized for stiffness, piezoelectric, and permittivity tensor, respectively. A computer program was developed so that it could be cooperated with the Laguerre polynomial technique [5]. The validity of the obtained effective constants may be checked by comparing the SAW velocity by using the effective constants with that by using the multilayered material constants. Fig. 2.26 shows a comparison between the SAW velocity of ZnO/GaAs and that of its effective single material. The maximum difference between both velocities is 0.009 %.

The reflectivity of gratings can be divided into two components: the piezoelectric and the mechanical. The former depends on the piezoelectric coupling constant K^2, and the latter depends on the

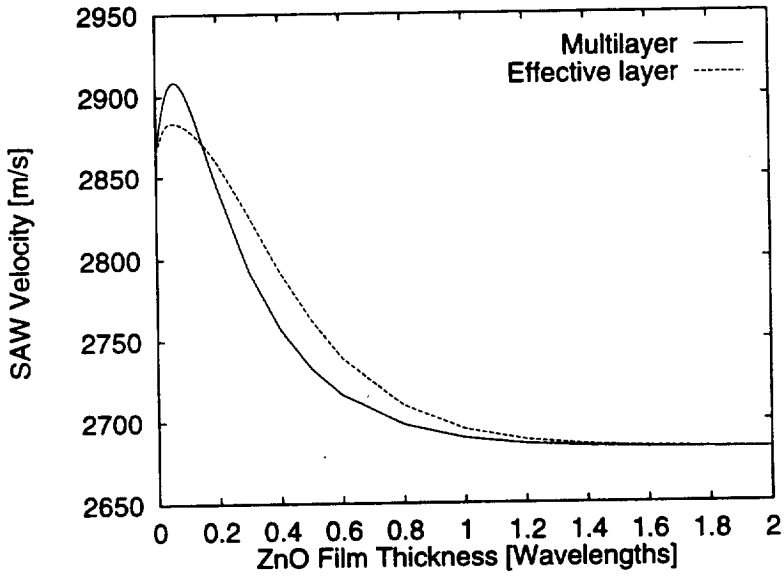


Figure 2.23 Comparison of SAW velocity of multilayer ZnO/GaAs and effective layer models.

thickness h of the metal in wavelength. The reflectivity, r, of a metal strip with a metallization ratio of η at a frequency f can be given by

$$r = j[\ P_z(\eta)\frac{K^2}{2} + F_z\frac{h}{\lambda}\ ]\ \sin\frac{\eta\pi f}{f_o}$$

where f_0 is the center frequency, P_z is the piezoelectric scattering coefficient, and F_z is the first order mechanical scattering coefficient. P_z is the same for all substrate-electrode combinations and depends only on the metallization ratio. On the other hand, F_z depends only on the substrate and electrode materials and are independent of the metallization ratio. More detailed explanation about these parameters can be found in Datta and Hunsinger's report [4].

The value of P_z at f_0 is given as a function of η and shown in Fig. 2.27. It is to be -0.75 when the value of η is 0.5. Figs. 2.28 and 2.29 show the values of K^2 for ZnO/GaAs and F_z for aluminum gratings on ZnO/GaAs as a function of ZnO film thickness, respectively. The value of K^2 increases from 0.07 % (GaAs) to 1.0 % (ZnO), and that of F_z decreases from 0.9 (GaAs) to 0.2 (ZnO) as the film thickness increases. Thus, the piezoelectric scattering and the mechanical scattering compensate each other due to their opposite signs if the magnitudes of each other are comparable. This zero reflectivity can be a great advantage for ACT devices with many numbers of NDS electrodes.

Fig. 2.30 shows a theoretical calculation for the reflectivity of an aluminum grating at f_0 with 1/4 λ width and 1/4 λ spacing on a ZnO/GaAs substrate as a function of the metal thickness. For the calculation, the material constants for the film were assumed to be the same as those of the bulk material. Note that the sign of the reflectivity changes and zero reflectivity occurs at some thicknesses of the ZnO film when the metal thickness is thin.

In order to verify the theoretical expectation, a set of characterization devices shown in Fig. 2.31 is designed to measure the effect of the gratings on ZnO/GaAs substrates. Several thicknesses of ZnO film

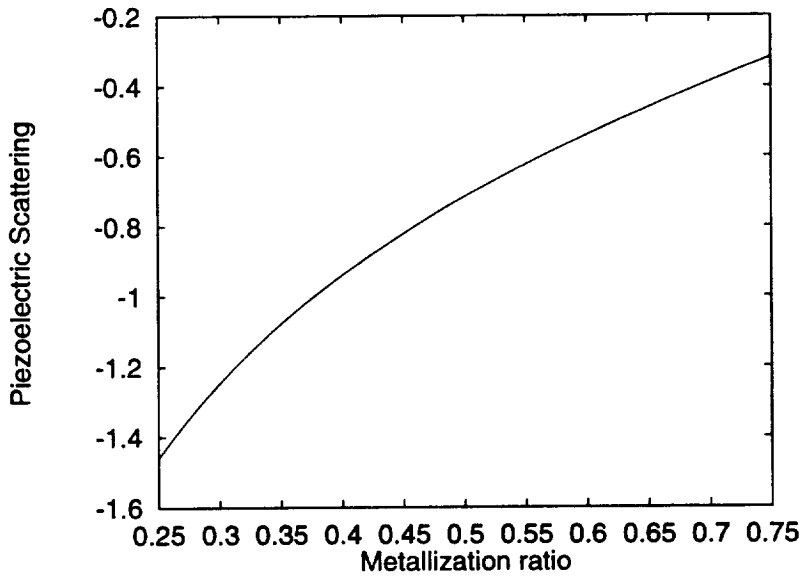


Figure 2.24 Piezoelectric scattering coefficient as function of metallization ratio.

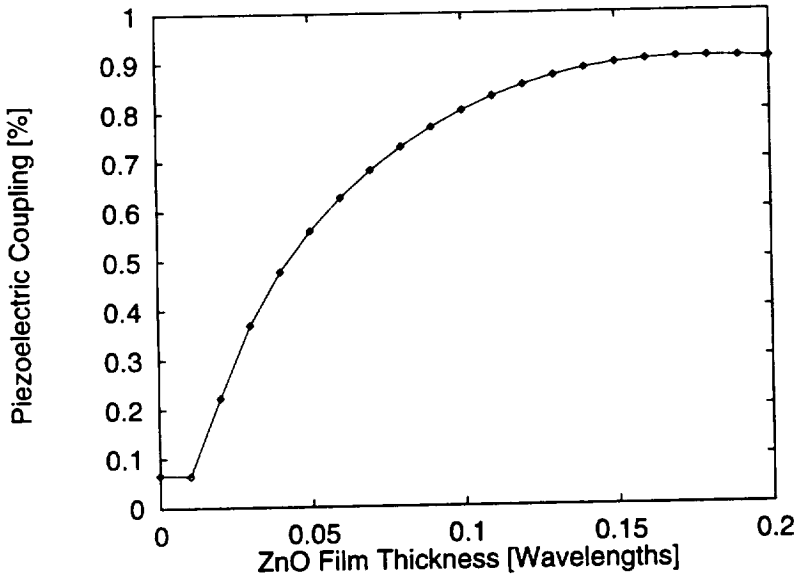


Figure 2.25 Piezoelectric coupling as function of ZnO film thickness for aluminum gratings.

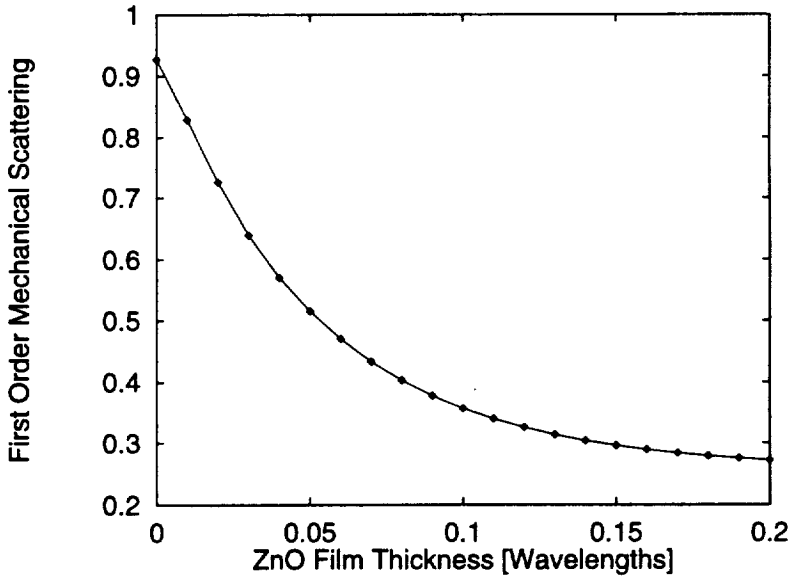


Figure 2.26 Mechanical scattering as function of ZnO thickness for aluminum gratings.

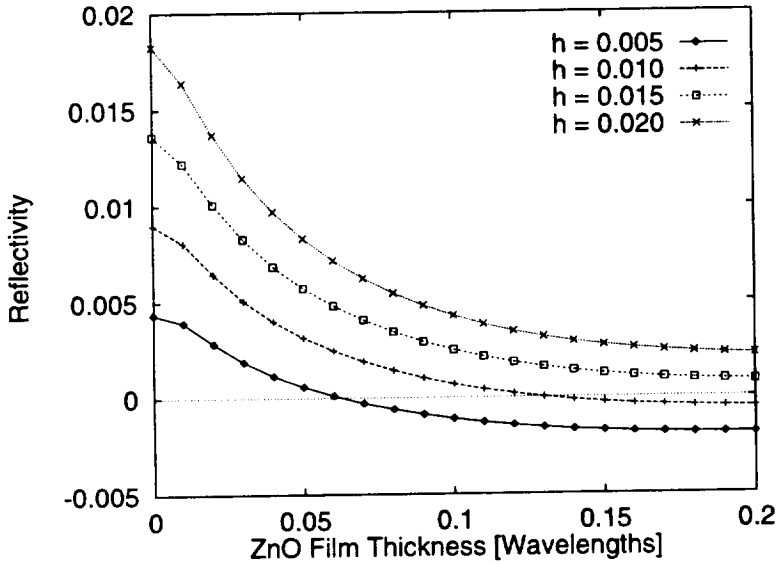


Figure 2.27 Theoretical reflectivity of aluminum grating at f_o with 1/4λ width and 1/4l spacing and thicknesses of hλ.

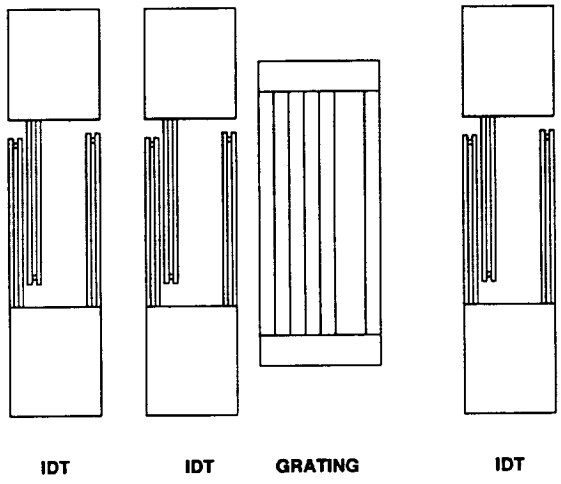


Figure 2.28 Experimental device layout for measuring reflectivity of metallic gratings.

(0.8, 1.6, and 2.4 μm) were provided by Dr. Hickernell at Motorola, Inc. The probe IDTs are 100 λ wide and centered in frequency around 180 Mhz. Each of the probe IDTs is 20 λ long with split finger electrodes in order to reduce the internal reflection. Both shorted and open gratings with the width of 108 λ are investigated, and each of the gratings consists of 200 strips with 1/4 λ width and 1/4 λ spacing. The distance from the center of the front two IDT electrodes with same polarity to the edge of the shorted grating is designed to be $(1/2n+1/4)$ λ and $1/2n$ λ, where n is an integer, to measure the sign of the reflectivity since the phase construction of a standing wave occurs at the distance of $(1/2n+1/4)$ λ for a negative sign of reflectivity and $1/2n$ λ for a positive one.

The properties of the open grating is compared with those of the shorted one. Unlike the short gratings, the SAW regeneration on the open grating should be considered in the analysis. At the center frequency, the reflectivity, r_o, of an open metal strip with the metallization ratio of 0.5 is given as [6]

$$r_o = r + j0.64K^2$$

The total reflectivity, R, of N metal strips is given by [6]

$$R = \tanh(N|r|)$$

The value of R when N is 200 is shown in Fig. 2.32. Again, the zero reflectivity should be noted.

The identical IDT probes with free and metallized surface instead of the gratings are also fabricated to calibrate the measured data of the gratings. Moreover, from these calibration sets, the SAW properties of velocity, K^2, and attenuation for the ZnO/GaAs substrates will be able to be characterized. The fabrication of devices with several thicknesses of the metal grating is currently on going process. The measurements will be performed with a HP network analyzer using a time-gating technique with the same manner described in Wright's paper [7].

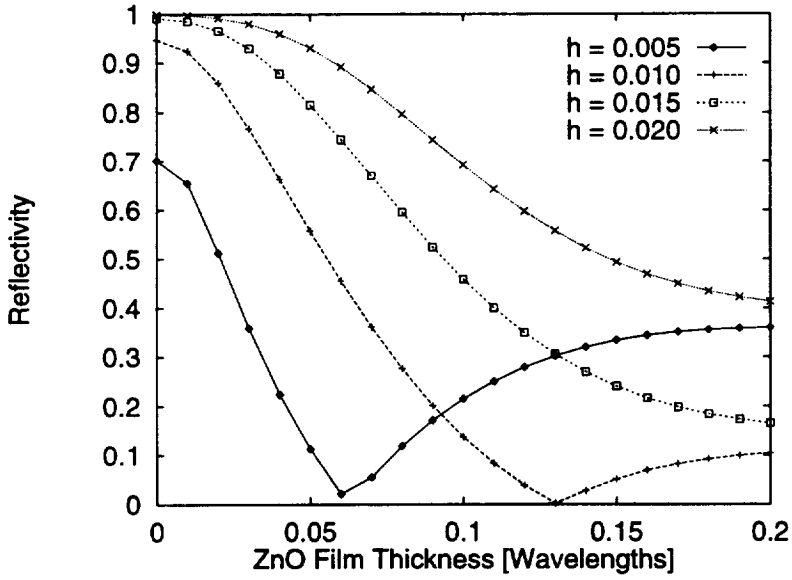


Figure 2.29 Total reflectivity of grating (hλ thick) as function of ZnO thickness for 200 metal strips.

3.0 Charge Transfer Device

The HDTV imager chip consists of three components: the avalanche photodiode detector (APD), the charge transfer device (CTD), and the acoustic charge transport device (HACT). Among these three devices, the CTD plays a major role in the operation of the acoustic charge transport imager. In the imager, the CTD acts to collect charge from the APD, then transfer the accumulated charge to the HACT for readout. The CTD also has a mechanism for removal of excess charge to prevent pixel blooming. The electron storage capacity, storage time and discharge time of the CTD device will directly affect image quality and the speed of image readout.

In this section the progress made on the CTD device will be presented. Complete electrical characterization on the vertical CTD (V–CTD) devices described at the previous review will be presented, along with improvements made in the fabrication technology to address issues with ohmic contacts. The efforts to define an architecture for the integration of the APD, CTD and HACT into the ACT imager are then described, which leads to a new CTD structure, the lateral CTD (L-CTD). The implementation of the L-CTD and the development of a mask set for testing various features of the L-CTD are then described. This section will conclude with future goals for CTD research and imager development.

3.1 Vertical CTD

3.1.1 Introduction and Review

In the previous program review, we had developed the fabrication process and had completed fabricated of two-terminal devices for electrical testing, with some initial results being presented. Since that time, these devices have been extensively characterized and the results interpreted using the STEBS-2d hydrodynamic simulator. Fabrication process development has continued with the demonstration of Pd/Ge ohmic contacts necessary for the V-CTD. We are currently testing a new V-CTD structure which was developed using our results from the previous devices and the STEBS-2d simulator and which uses the Pd/Ge ohmic contacts.

3.1.2 Electrical Characterization

In this section the results of electrical characterization on the current V-CTD devices is described along with a presentation of the measurement system developed for the CTD. The characteristics of the CTD device are obtained using a capacitance test system, consisting of a Hewlett-Packard 4280A 1 Mhz Capacitance Meter, a Tektronix DSA602 Digitizing Signal Analyzer, and a Tektronix AWG2020 Arbitrary Waveform Generator. This test system is controlled by an 486DX-33 personal computer through an IEEE-488 GPIB bus. The accuracy of the HP4280A is 0.1%, and the resolution is 1 fF. All the tests were done at room temperature and under dark conditions. The measurements for individual devices are repeatable. The Hewlett-Packard 4280A 1 MHz Capacitance Meter uses an internal dc bias source to perform the C-V measurement.

Figure 3.1 illustrates the C-V characteristics of a CTD device. This data was taken at 0.2 V intervals. Changing the voltage sweep rate (dV/dt) from 0.02 V/s to $2x10^4$ V/s yields only little change in the C-V characteristics which may be caused by measurement error. The capacitance readings obtained at V=0 from the measurement are close to the steady state V=0 readings. The current V-CTD device has

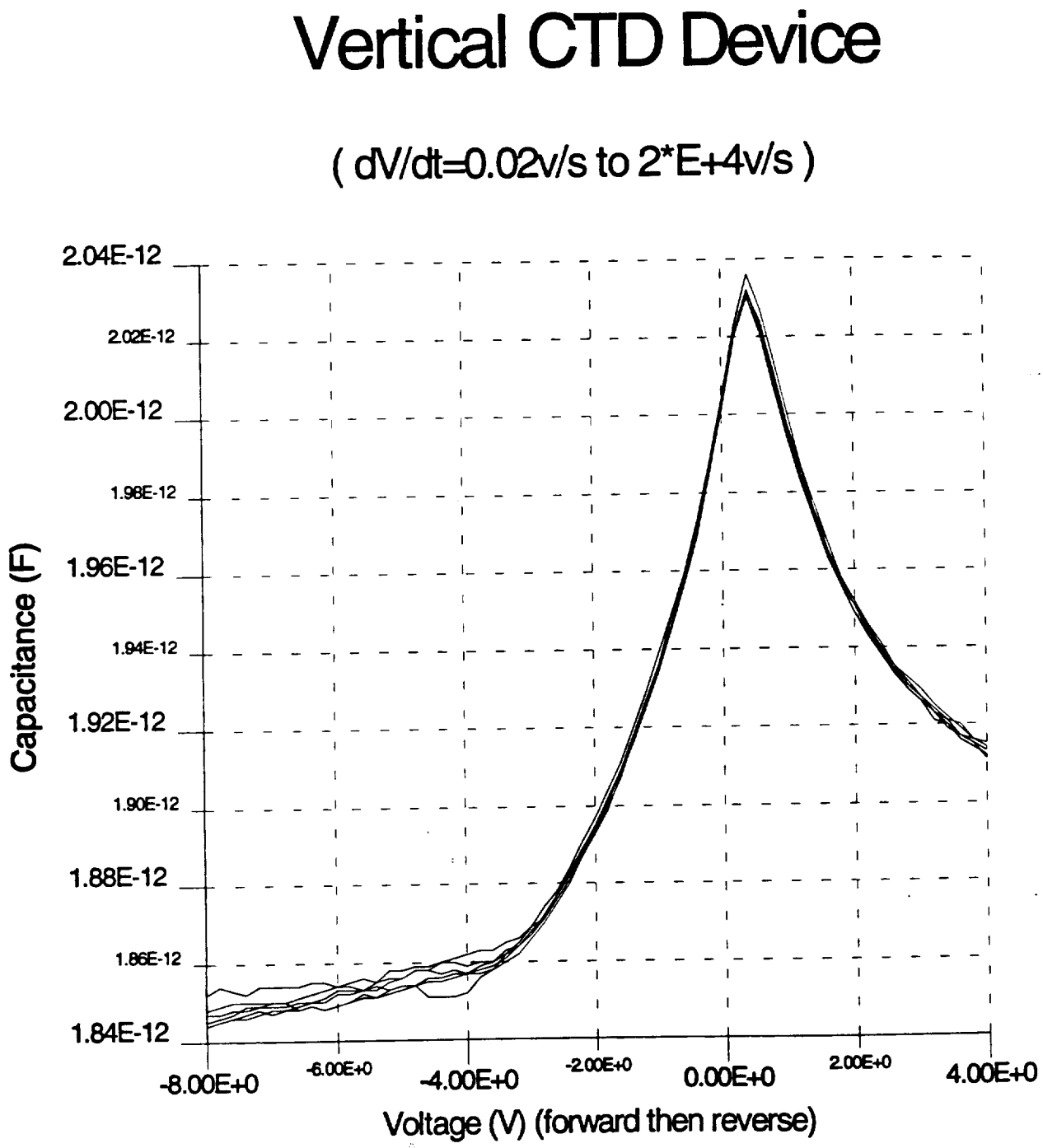


Figure 3.1 C–V characteristics of the V-CTD device.

an n-type readout layer, a p-type readout barrier, and a n-type storage well structure, essentially two pn junctions. The asymmetry in the C-V curve is due to different doping concentrations in the readout and well layers and small variations of doping concentration in the readout barrier region.

The C–V data can be understood as follows. As a positive voltage is applied to the readout layer with the n^+ substrate grounded, the readout–readout barrier junction becomes reverse biased, and the readout barrier–well junction becomes forward biased. The free electrons stored in the charge storage well diffuse into the readout barrier, and then drift out of the readout contact. So the readout of the charge in the storage well is accomplished by a positive voltage being applied to the readout contact. Since the total capacitance consists of the two junction capacitances in series at V=0, the readout barrier–well junction, which has the smaller capacitance, dominates the total capacitance.

Initially, when applying a positive bias to the readout contact, the forward biased readout barrier–well junction dominates the electron flow. Increasing the positive bias causes the total capacitance to increase with the decreasing depletion layer width of the readout barrier–well junction. When the applied positive voltage increases further, the capacitance of the reverse biased readout–readout barrier junction becomes smaller than the forward biased readout barrier–well junction capacitance, and thus starts to dominate the capacitance of the device. A further increase in positive voltage causes the total capacitance to decrease with the increasing depletion layer width of the readout–readout barrier junction.

Because of the approximately symmetric n-p-n structure, by applying a negative voltage to the readout contact with the substrate grounded, the readout–readout barrier junction becomes forward biased as the readout barrier–well junction becomes reverse biased. The conduction band energy of the readout layer is lifted, and electrons diffuse through the readout barrier layer into the charge storage well. With a negative voltage applied to the readout contact, the readout barrier–well junction controls the electron flow. Increasing negative voltage causes the total capacitance to decrease with the increasing depletion width of the reverse biased junction.

Another basic characterization of CTDs is by C-t measurements. Because of the high sampling rate needed for fast capacitance transient measurement, an external pulse generator was used and a 10 μs capacitance transient can be measured. The external pulse generator supplies a square pulse of a preset magnitude to bias the device. The HP4280A capacitance meter is programmed to provide a synchronization signal to the external pulse generator, and control the bias pulse width and time interval during C-t testing. The C-t measurement provides information on the charge and discharge capabilities of the CTD.

Figure 3.2 shows a representative C-t curve for positive and negative input pulses of 4 V. When a positive 4 V pulse is applied to the readout contact, at t<0 the capacitance has the value which corresponds to the 4 V steady state capacitance value. After the input pulse, the capacitance returns to its equilibrium value within 0.01 s. At t=0, electrons in the well are driven out of the readout contact, the readout–readout barrier junction is reverse biased, and the capacitance is smaller than the steady state value. At t>0, the energy band of the readout layer returns to the zero bias level. The electron distribution in the two pn junctions eventually moves back to its equilibrium configuration and the final capacitance reading reaches its equilibrium value. When a negative 4 V pulse is applied to the readout contact with the substrate grounded, at t=0 the capacitance value is the same as the -4 V steady state capacitance value, and electrons flow into the storage well. At t>0, the applied voltage becomes zero, and the device moves back to its equilibrium condition. Since some electrons which flow into the well by applying the negative voltage do not have sufficient energy to leave the well, extra electrons are trapped for at least 5 s (the limit of our ability to measure in this configuration). An internal reverse bias remains so the equilibrium capacitance is less than the steady state value. When the input pulse voltage is 0.5 V, at t=0 the readout barrier–well junction is forward biased, and the capacitance is higher than its steady state value. At t>0, the device becomes zero biased and the capacitance starts to decrease back to its equilibrium value. When a negative pulse of the same magnitude is applied to the device, the capacitance again returns to a level

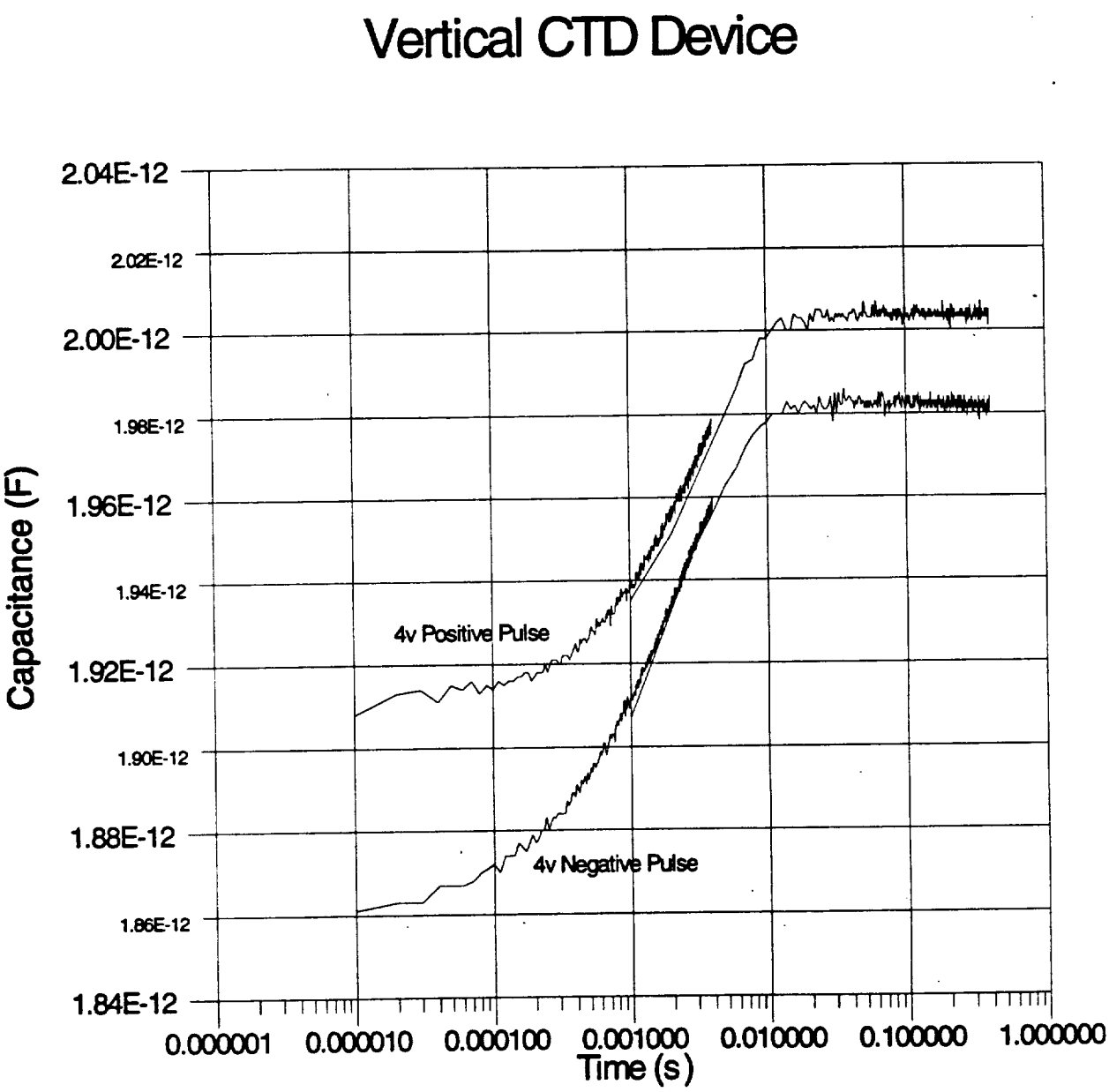


Figure 3.2 C–t characteristics of the V-CTD device.

less than the equilibrium value. Therefore, charge storage has been observed in the V-CTD device with characteristic lifetimes of at least 1 s.

3.1.3 Device Fabrication

As presented previously, about 50% of the CTD–5 devices were lost due to shorting of the readout–readout barrier junction, while the other devices show nonuniform breakdown voltages. We believe this behavior resulted from our ohmic contact technology, which was based on AuGe/Ni/Au. In this contact a heavily doped region is formed through liquid phase alloying of a Au-Ga compound. However, this liquid phase reaction causes significant, nonuniform penetration of the Au-Ga compound into the device structure to depths greater than 2000 Å. If any extended crystalline defects (dislocations, etc.) are present, the contact penetration depth could be much greater. Although contact penetration is not a concern for lateral devices such as FETs and the HACT, it can not be tolerated for vertical devices such as the V-CTD.

Therefore, we have implemented an ohmic contact technology that will not penetrate through shallow p-n junctions and heterojunctions, which is based on Pd/Ge [8]. The heavily doped contact region is formed through the solid phase epitaxy of a Pd/Ge compound on the GaAs. Because it is a solid rather than liquid phase reaction, contact penetration depths are less than 50 Å. However, there are other consideration when using Pd/Ge contacts, specifically a higher contact resistance compared to alloyed AuGe contacts.

Test were performed to evaluate Pd/Ge contacts for use in the V-CTD. Contact definition was by lift-off, with e-beam evaporation of the Pd and Ge, followed by annealing using a standard hotwall furnace at 325°C. There are two critical parameters that must be controlled to obtain low resistance Pd/Ge contacts.

- Order of Metallization: For the contact to be ohmic, the Pd must be in direct contact with the GaAs, followed by deposition of the Ge. The Pd reduces oxides on the GaAs surface and the Ge will diffuse through the Pd to react at the GaAs surface
- Ratio of Metal Thickness: For a Pd thickness of 500 Å, the Ge thickness must be greater than 1265 Å. This criterion results from the need to form the PdGe alloy and have sufficient Ge remaining to dope the GaAs surface.

Contact resistance was evaluated using transmission line measurements. We obtained a minimum specific contact resistance of $7.5x10^{-5}$ ohms-cm^2, compared to a best reported value for Pd/Ge contacts of $2.0x10^{-6}$ ohms–cm^2 and our HACT AuGe based ohmics of ~$5x10^{-7}$ ohms–cm^2. Unfortunately data interpretation is very dependent on the test material structure, which makes comparison difficult. However, from our test results it appear that the Pd/Ge contact is acceptable for V-CTDs, and all subsequent fabrication runs will utilize this contact. More V–CTD device testing will determine if further optimization of the contacts is necessary.

3.1.4 Current Status of V–CTD

From our testing of the current V–CTD structure, we have concluded that the large amount of equilibrium free charge in the charge storage well leads to relatively high bias requirements and small well capacity. Also, the complex structure of these devices makes it difficult to precisely characterize. Therefore, new vertical device structures, that include a substantial barrier to current flow through the substrate and lower doping levels in the well region, have been designed, grown and fabricated. These devices should isolate the charge storage region of the V–CTD from the substrate side of the structure, thereby simplifying analysis of the test results. The new V–CTD structure consists of 2000 Å GaAs n=$2x10^{18}$ cm^{-3}, a 2000 Å wide undoped $Al_{0.45}Ga_{0.55}As$ barrier, a 2000 Å GaAs n=$2x10^{15}$ cm^{-3} storage well, a delta-doped readout barrier with equivalent doping of p=$5x10^{18}$ cm^{-3}, a 2000 Å GaAs n=$2x10^{15}$ cm^{-3}

readout layer, and a 500 Å $n=10^{18}$ cm^{-3} readout ohmic contact region. The conduction band diagram obtained using the STEBS-2d simulator indicates that the charge storage well for this device is more depleted of free charge than the current structure. By depleting the well of free charge, the dynamic range of the structure is increased by essentially lowering the background signal level in the V–CTD. These devices are currently being tested and the results will be presented in the next report.

3.2 Device Integration for Imager

Current research in this project has focused on the optimization of the three individual device components. However, with the progress being made on the HACT, APD, and CTD, concerns will turn to the integration of these devices into the HDTV imager. Because the CTD is the link between the APD and HACT, its final structure is constrained by the layout of these devices in the imager. Therefore, a conceptual study was made on the integration of the three component devices into the HDTV imager, with the goal of determine a realistic structure for the CTD. Two general architectures were considered, vertical and lateral integration. The advantages and disadvantages of these architectures are described below.

The three components can be arranged in a variety of different formats, as will be shown. The first option for total integration is the vertical structure in Figure 3.3, where the APD and CTD are on top of the HACT channel. The advantages and disadvantages of the structure are listed in Table 3.1. Due to the disadvantages of mass loading and the inability of using the ZnO coating to aid in supporting the surface acoustic wave this design is not considered viable.

A second vertical design considered is where the HACT readout channel is on top of the CTD and APD, as shown in Figure 3.4. The advantages and disadvantages of this structure are listed in Table 3.2. None of the disadvantages listed in this table preclude the use of this structure.

A third option for total integration is shown in Figure 3.5. In this design the CTD and HACT channels are horizontally adjacent to each other with the APD on top of the CTD region of the device. The advantages and disadvantages of this structure for total integration are listed in Table 3.3. This design also allows flexibility during the testing of the individual devices because the CTD and HACT structures can be fabricated and tested without the APD. In the same vein, the APD and CTD can be fabricated without the HACT portion of the device and still be tested. The principle advantages of this design are that the current p-HACT design is compatible and that the Schottky barrier isolating the CTD from the HACT is a binary operator (ie. either on or off), thereby reducing the likelihood of tailoring overflow or

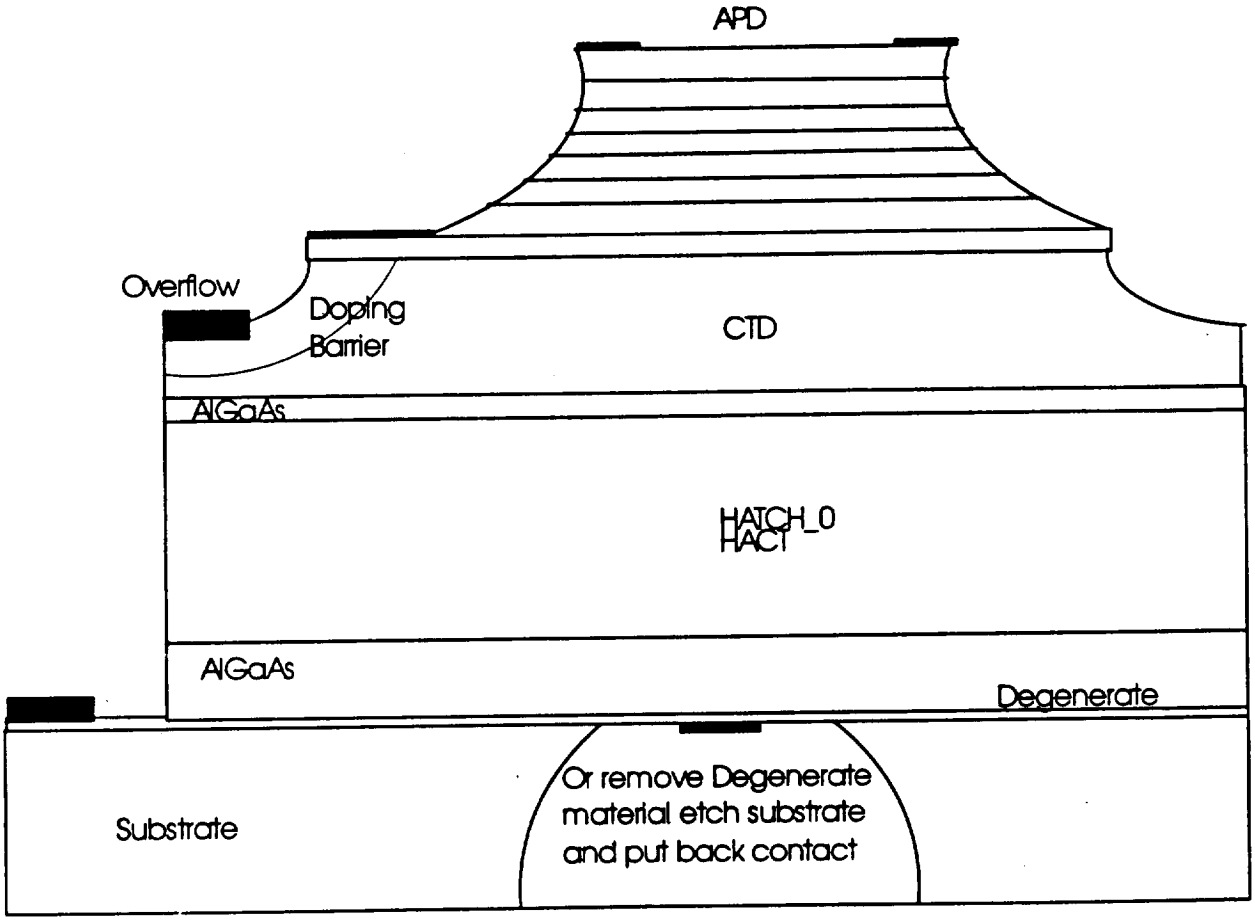


Figure 3.3 Total integration of CTD, HACT, and APD into HDTV imager. In this design the APD and CTD are on top of the HACT channel.

Table 3.1 Advantages and disadvantages of using the vertical configuration with the APD on top of the HACT channel

Advantages	Disadvantages
Good isolation of pixels due to mesas	Degenerate material is required to provide a means of injection into HACT
HACT wave potential will not penetrate the CTD or APD	Tailoring of overflow is required by ion implantation
No regrowth is required	Mass loading of the HACT channel
No back side etching required, if using degenerate material	Possible premature dampening of the SAW, when using degenerate material
Good mechanical strength	Can not use ZnO coating
1" by 0.5" chip for 1k by 2k array of pixels	Current HACT structure is not appropriate
	Tunnelling is required for injection
	Image charge located in degenerate material
	When using a back side contact there is the possibility of SAW reflections at etched portion

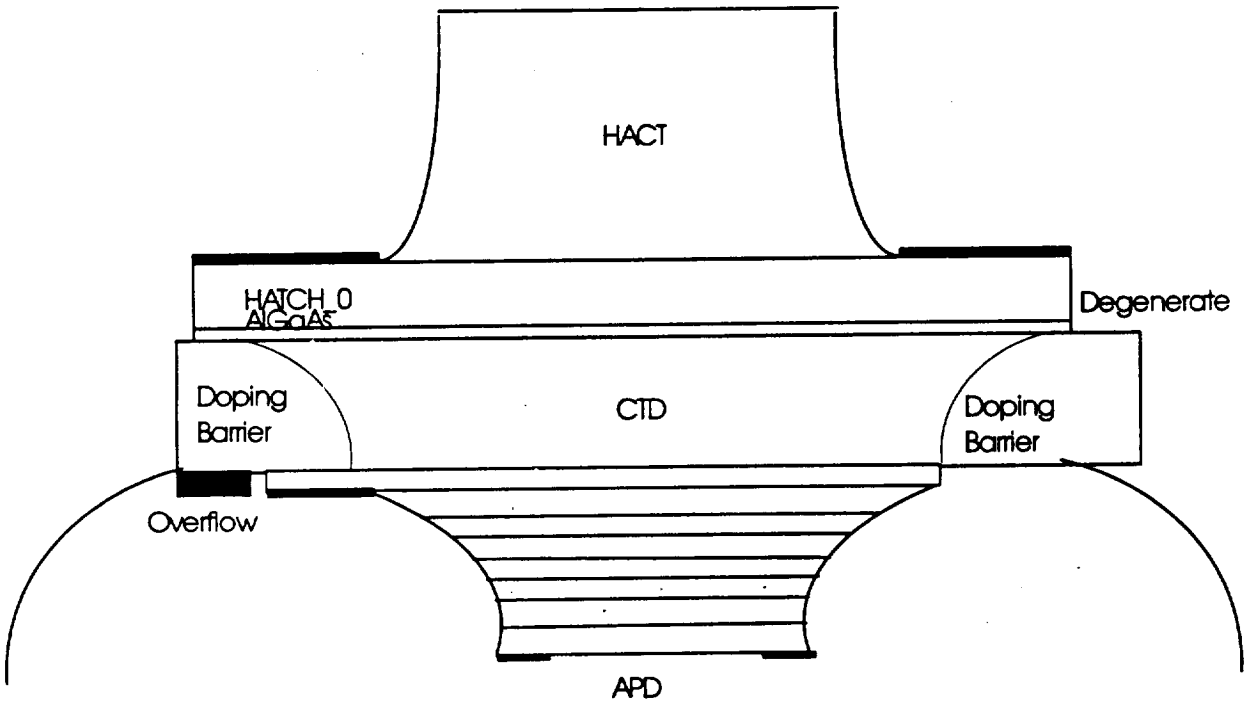


Figure 3.4 HDTV imager layout for the case where the HACT channel is located on the top of the CTD and APD structures.

Table 3.2 Advantages and disadvantages of using the vertical configuration with the HACT channel on top of the CTD and APD structures

Advantages	Disadvantages
Good isolation of pixels due to mesas	Degenerate material is required to provide a means of dampening the SAW
HACT wave potential will not penetrate the CTD or APD with degenerate material	Tailoring of overflow is required by ion implantation
No regrowth is required	Poor mechanical strength
Readout is on top of the structure	Possible premature dampening of the SAW when using degenerate material, without the degenerate material multiple reflections occur
ZnO coating can be used in the HACT design	Extensive etching is required to expose material layers
1" by 0.5" chip for 1k by 2k array of pixels	Current HACT structure is not appropriate
	Tunnelling is required for injection
	Image charge located in degenerate material

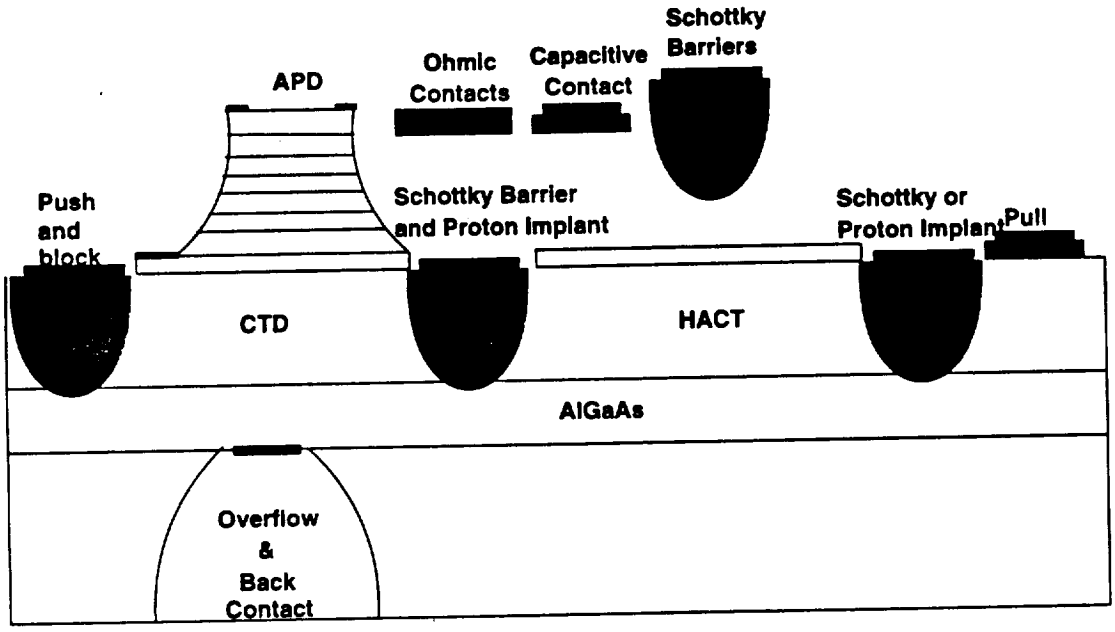


Figure 3.5 HDTV layout configuration where the CTD and HACT channel are horizontally adjacent to each other. The APD is on top of the CTD.

Table 3.3 Advantages and disadvantages of using the lateral layout for the HDTV imager.

Advantages	Disadvantages
CTD and HACT are the same dimensions	Area of pixel is pushing limits
Binary switching (contacts are either on or off), no tailoring of barriers	Regrowth is required for the APD
No tunnelling required for injection	Capacitive contacts (MOS) are required
No degenerate material required	Possible injection spreading
P type HACT structure is compatible with CTD design	Back side etching is required
Optical isolation of HACT channel is possible	1" by 1" chip required for a 1k by 2k array of pixels
ZnO coatings can be utilized	
Good mechanical strength	

injection barriers. This greatly simplifies the device design. A top view of one pixel is displayed in Figure 3.6, in this view the acoustic wave is propagating in the plane of the paper.

From this discussion, it is our belief that the lateral imager architecture has the greatest possibility for rapid and successful implementation. Therefore, CTD research has branched into the design and testing of a lateral CTD (L–CTD).

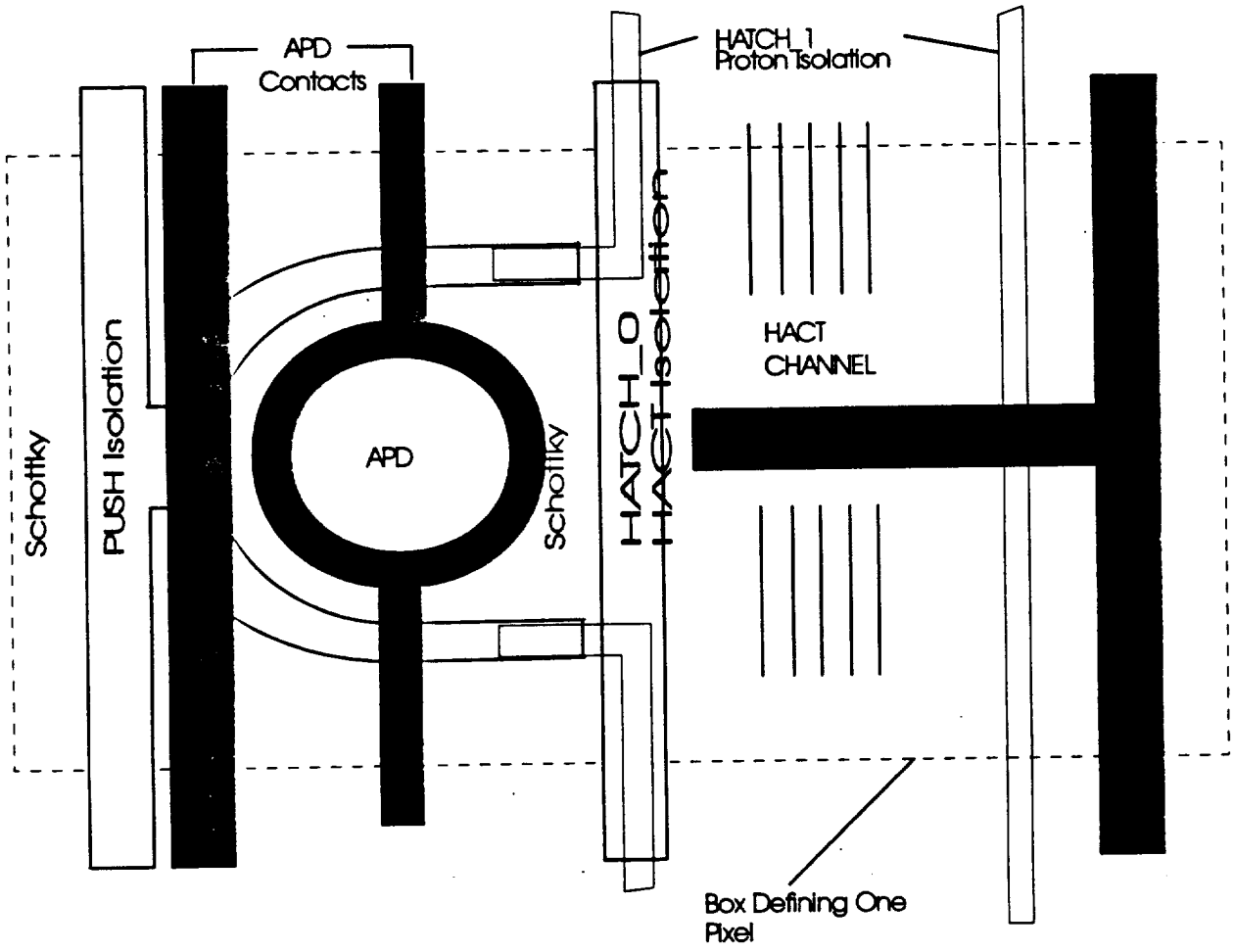


Figure 3.6 Layout of the contacts for one pixel in the lateral design of the HDTV imager.

3.3 Lateral CTD

Our work on the L–CTD has focused on two areas, the modelled of the device for charge storage capacity and lifetime and the development of a mask set for determining the feasibility and operating parameters of the L–CTD. These areas are described in detail in the sections to follow.

3.3.1 Modelling

In the design of the CTD the number of photogenerated electrons and the time duration for which the electrons must be held are important design constraints. For normal HDTV operation a frame of two million pixels is read every 33 msec. To achieve slow motion of 6 to 1, the frame of 2M pixels is read in 5.5 msec. The HACT portion of the imager is designed to operate with a 250 MHz surface acoustic wave. Using 1000 pixels in a line the time to readout the entire line would be 4 μsec. To maintain 2M pixels with 1000 pixels per line requires 2000 lines to be read. Multiplexing the 2000 lines to 50, 10 and 2 readout lines leads to frame readout times of 0.16 msec, 0.8 msec and 4 msec respectively. These times give an upper limit on the duration which the electrons must be held in the CTD. All of these readout rates are sufficient to obtain the 5.5 msec for slow motion effects.

Using the readout times an estimate of the number of signal electrons is calculated based on the spectrum of light and the size of one pixel. The spectrum is AM1.5 direct solar spectrum and corresponds to 1000 Watts/m^2, with no amplification or concentration assumed. The pixel size is 10 μm by 15 μm, which is a 75% fill factor. Table 3.4 is the number of electrons which must be held in the CTD if the pixel is collecting carriers between readout cycles. Based on the numbers of electrons in Table 3.4 the CTD must be capable of holding from 3000 to 1.385×10^9 electrons. These numbers of electrons are representative and will scale with amplification, shuttering, reflective losses, and light intensity.

Figure 3.7 is the simulated layer structure of the CTD based on the optimized p-HACT structure of Section 2.1. This structure is analyzed to find the background electron concentration in the CTD and

Table 3.4 Millions of electrons which must be stored in the CTD at three different wavelengths and different colors.

Readout time → Color ↓	4 msec	0.8 msec	0.16 msec
Blue	17.01	3.402	0.680
Green	21.63	4.324	0.865
Red	23.80	4.761	0.952
Black + White	1385	277	55.4

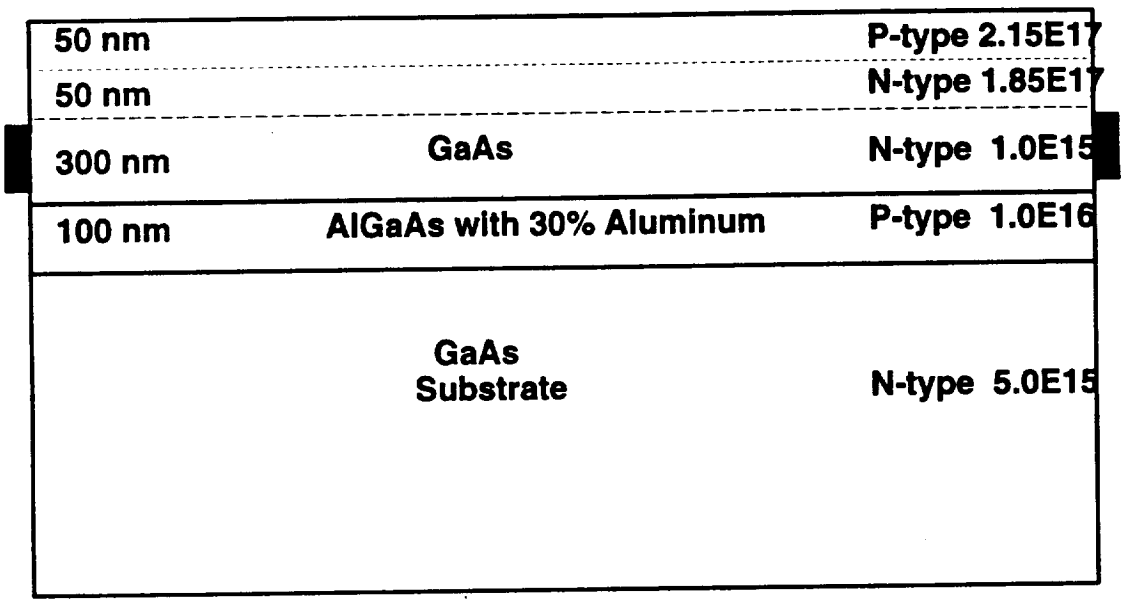


Figure 3.7 Device structure used to model the channel region of the lateral CTD design.

to model the channel conductance. The background electron concentration is the first step in determining the signal to noise ratio of the imager. If the background density is as high or higher than the number of signal electrons in Table 3.4 then the signal is lost in the thermal noise of the CTD. The background concentration is determined by integrating the equilibrium electron density obtained from hydrodynamic simulations of the structure. For a 10 μm wide pixel the background density is 20 electrons/cm, which is several orders of magnitude below the signal electrons listed in Table 3.4. Non-equilibrium simulations of the structure in Figure 3.7 reveal that at low biases the channel will conduct. The conductive channel is the first step to determining if the Push/Pull contacts in Figures 3.6 and 3.7 are capable of injecting charge from the CTD into the HACT structure.

The hydrodynamic model simulations are utilized to determine upon which layer of the structure the isolation and Push Schottky barriers should be located. If the Schottky barriers are located on the p-type layer of the structure, simulations show that there is no charge confinement unless a large voltage (> 1 volt) is applied. If the top p-type layer is etched off and no surface states are present, simulations indicate that the Schottky barriers provide confinement. But, the background concentration is very high due to the residual n-type doping. If the effect of surface states are included in the calculation, there is no lateral confinement in the equilibrium case, however, the bias voltage required to provide the lateral confinement is less than when the Schottky barriers are located on the p-type layer, (<1 volt). Based on these calculations and assuming that the APD is left off of the structure as in Figure 3.7, the isolation Schottkys must be on the n-type layer of the structure but the p-type layer must remain over the rest of the CTD area to provide channel depletion. This information was then used for the development of the L-CTD test mask set design.

3.3.2 Test Mask Set Design

The test structure developed for the L-CTD will yield information on storage capacity, charge lifetime, the ability of the Schottky barriers to laterally confine charge, the lateral transit time of stored charge, and the switching time of the device. The requirements set out for the design of the test structure were flexibility in terms of being able to evaluate many different L-CTD scenarios with same mask set, having the set be applicable to different material structures, and having the ability to easily add other masks to the set. The L–CTD test mask set was fabricated by BNR, in keeping with the integrated approach for device fabrication in this project.

The L–CTD mask set consists of 7 layers: Alignment, P-Mesa, N-Mesa, Ohmic, Schottky, Passivation and Cap. Device structures for both optical and electrical charge injection are present, as are Schottky or ohmic push contacts and ohmic or capacitive pull contacts. Some devices have Schottkys present over the storage well to remove photogenerated holes from the well during optical charge injection. Each step and repeat pattern includes 30 devices, 2 transmission line type ohmic test structures, 2 C–V test structures and alignment marks. The thirty devices are divided into two types: those with capacitive outputs and those without. The overall dimension of the array is 1500μm x 2900 μm. The minimum feature size in the mask set is 2 μm. The bottom half of the die contains devices with the capacitive contacts substituted for the output ohmic. The use of a separate alignment layer allows complete interchanging of the mask order, in case of new mask insertion or process changes, as well as extra alignment marks for future mask additions. A schematic of a L-CTD test device is show in Figure 3.8. Details of the test mask set and its processing are described below.

The alignment layer will create, via titanium deposition and liftoff, alignment marks for all subsequent layers plus two extra layers to allow for additional masks. Each array of devices is serialized with a row and column number and 250μm wide dicing streets are also patterned on this layer. The P-Mesa layer will etch away the p^+ GaAs down to the n^+ layer. The areas masked will create a 80μm x

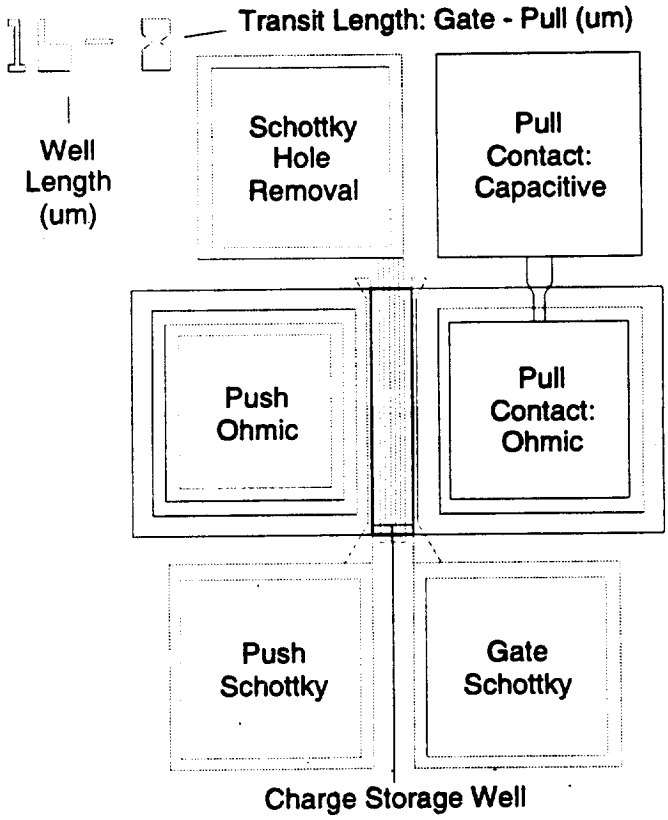


Figure 3.8 Schematic of a lateral CTD test device.

80μm, p^+ doped mesa which will later become part of an output capacitive contact, which will replace the output ohmic contact and provide another method of measuring the charge transfer on 15 of the devices. The N-Mesa layer will electrically isolate all thirty devices on a 216μm x 96μm mesa by masking the mesa and etching the n^+ GaAs.

The Ohmic layer creates 80μm^2 n-type ohmic contacts on the exposed n^+ layer (only one contact is deposited on the capacitive devices), for the removal of charge. The distance between the input and output pads are a function of the well and drift widths. The Schottky layer defines the charge confinement region and, on some devices, an additional gate to electronically inject charge. The gates are 2μm long and stretch 112μm across the mesa. They increase in width to 6μm as they run off the mesa edge to increase yield. Bonding pads are also included in this layer to allow bonding to all Schottky gates and ohmic pads. The well width, defined as the distance between the Schottkys, and drift width, defined as the distance between the last Schottky and the ohmic, are indicated by the numbers in the upper left hand corner of each device. The dimensions are in microns. The well widths are 20, 16 and 10μm and the drift widths are 12, 8 and 4 μm.

A layer of passivation will be deposited for the capacitive contacts and the Passivation mask will allow for the removal of the passivation in the bonding pad areas. A 1000Å layer of SiO_2 will produce a capacitance between 0.245 pF and 0.980 pF. The dicing streets are also etched to prevent cracking of the dielectric during the dicing operation. The Cap layer will create, via liftoff, the top contact and a bonding pad for the capacitive contacts on these devices. The bonding pad is located off the mesa to prevent cracking the passivation during the bonding. The upper portion of the array contains alignment marks and test structures. The three transmission line test structure consist of 5 x 100μm^2 ohmic pads, with spacing between the pads of 5, 10, 20 and 40 μm. Two different sizes of C–V structures are also included. The larger is 250μm^2 with a 100μm diameter Schottky diode in the center of a 150μm diameter opening. The smaller is 150μm^2 with a 50μm diameter Schottky diode within a 75μm diameter opening.

3.4 Future Work

We are currently testing the latest V-CTD design which should yield greater understanding of the charge storage capabilities of this device. After obtaining the necessary results for the V-CTD, work will concentrate on the fabrication and testing of the L-CTD. This change in focus in justified by our analysis of APD, CTD, and HACT device integration for the HDTV imager, where the lateral imager architecture appeared most feasible for rapid implementation. The test mask set for the L-CTD has been received and with process development underway and p-HACT material available, initial results on devices should be presented in the next report.

4.0 Avalanche Photodiode Development

Of the three major components of the ACT imager, the avalanche photodiode (APD) provides for the conversion of the optical signal into charge. Because of the high frame rate and high pixel density requirements of the HDTV imager, it is necessary to provide some gain in the imager to allow operation under reasonable light conditions. The APD must, therefore, supply low power, low noise, front end gain. The APD structures under investigation to meet these demands are volume-doped barrier multi-quantum well (MQW), volume-doped well MQW, delta-doped MQW and doped i-region pin APD's. This research has obtained low voltage (10-15V) devices with low noise (excess noise factor~2) at moderate gains (G=5).

It was previously reported that these APD's were not operating at their full potential due to incomplete depletion of the gain stages. However, fully depleted structures have now been obtained and are reported below.

To optimize these devices by increasing their gain while decreasing their noise, it is necessary to obtain an insight into the basic physical mechanisms controlling the gain processes as well as the macroscopic parameters controlling the performance of the APD's. We discuss below in Section 4.1 (and Appendix A) recent results obtained on the modeling of the gain processes. Also, in Section 4.2, we report on the performance characteristics of some of these devices.

4.1 Modeling and Theory of Impact Ionization

We have made significant strides towards refining the theory of interband impact ionization with the expressed purpose of improving the predictive capability of the device simulators. Specifically, we have investigated three different approaches for evaluating the interband impact ionization transition rate. These approaches are all based on Fermi's golden rule. Traditionally, the impact ionization transition rate is most commonly expressed using an analytical expression known as the Keldysh formula [9]. The Keldysh formula, though universally accepted, is not totally satisfactory since it is derived assuming parabolic energy bands, contains two parameters which must be determined empirically, and fails to account for any possible wave-vector dependence of the rate. Alternatively, we have developed a completely numerical approach [10,11] with which the impact ionization transition rate can be determined without recourse to parameterization. In this method, the ionization transition rate is evaluated by directly integrating the transition rate over the full Brillouin zone. The overlap integrals in the transition rate are evaluated directly from a k•p calculation. The numerical method has the obvious advantages that it does not require any parameterization and can be thus applied to any material system for which the band structure is known. Unfortunately, the numerical method requires extensive computation. This is due mostly to the complex evaluation of the overlap integrals needed to determine the transition rate for those states satisfying the energy and momentum conservation requirements.

Recently, another analytical treatment of the impact ionization transition rate, which also includes the **k**-vector dependence of the impact ionization transition rate, was presented by Quade et al [12]. They calculated the impact ionization transition rate for the general case of an arbitrarily shaped energy band for the impact-ionizing conduction electron, and three anisotropic parabolic bands for the final states with their extrema located at different points in the **k**-vector space. The assumption of parabolic energy bands for the final states is acceptable since the final states are generally at low energy near the band minimum where the parabolic approximation is typically good. Other approximations used in the derivation include

nondegeneracy and a constant value for the overlap integrals. Though the formulation developed by Quade et al. [12] contains some parameterization due to the fact that the overlap integrals are not evaluated but are assumed to be constant parameters, it offers a far more computationally efficient means of evaluating the ionization transition rate with the inclusion of the **k**-dependence than the direct, numerical method of Wang et al. [10,11]. We have investigated how the method of Quade et al. [12] compares to the direct numerical approach. The results of these investigations are discussed at length in the enclosed manuscript entitled, "Theoretical investigation of wavevector dependent analytical and numerical formulations of the interband impact ionization transition rate for electrons in bulk silicon and GaAs".

Based on the results discussed in the enclosed paper, we are proceeding with incorporating the Quade formulation into our existing macroscopic, hydrodynamic simulator for advanced analysis of the APD devices. Using the hydrodynamic simulator, the terminal characteristics of the APD devices, i.e., current-voltage, capacitance-voltage, gain, bandwidth etc. can be evaluated to enable comparison to experiment. This work is currently underway and will be reported in the next project summary.

4.2 Characterization Results

During the last six months, a significant characterization capability (transient response) has been added which enables both the determination of a relevant device operating parameter and a more direct comparison of experimental and theoretical noise results. This characterization tool enhances the several other computer automated experiments developed to study and optimize the different parameters of the APDs in order to determine their gain, noise level, breakdown voltage, spectral response, and speed. A brief description of some of those experiments and the information they provide about the photodiode is presented below.

Experiment	Description	Information Provided
Spectral Response	Responsivity as a function of wavelength and gain	Sensitivity of photodiode throughout the optical spectrum
I-V response	Diode current output as a function of applied bias (10 - 373 K)	Photodiode gain and variation as function of incident light intensity
C-V response	Diode capacitance as a function of applied bias (10 - 373 K)	Profile of the carrier concentration vs. depletion width in doped APD's
Noise Measurement	Noise level as a function of laser power density (10 - 373 K)	Excess noise factor at various diode gains and temperatures
Transient Response	APD response to a high power 50 ps laser pulse	APD bandwidth limitations and heterojunction interface quality

The recently added transient response experiment has provided us with valuable information on the speed performance and the various factors that affect the bandwidth of our APDs. In order to better understand the results that were obtained, we will briefly discuss the various physical effects that limit the frequency response of a photodiode.

Bandwidth Considerations

Considerable work has been done in the literature on the study of photodetector bandwidth limitations [13-17]. To summarize the results, the bandwidth of a SAM (Separate Absorption and Multiplication Regions) APD is determined by 5 physical effects:

1. Transit time: This is the time it takes the generated carriers to travel through the depleted region under the effect of the electric field. This is usually calculated using the ratio of the distance travelled to the "saturation" velocity of the appropriate carrier.

2. Carrier diffusion time: in the undepleted regions of the device, carrier transport must take place by diffusion rather than drift, with a resultant slowing of device response.

3. RC time constant: there is a fundamental limit on bandwidth due to the capacitive transient charging effects which arise from the depletion region capacitance of the device and the combined resistance R of the load and the device.

4. Carrier trapping: in heterojunction APDs, there is a possibility of carrier delay caused by traps present at the heterojunction interface. This effect is related to the abruptness of the heterojunction, the barrier height, the temperature, and the effective mass of the carrier. Because the effective mass of holes is larger by an order of magnitude than that of electrons, trapping is more likely to occur for holes than electrons. This phenomena is known as "hole trapping".

5. Avalanche buildup time: For single carrier ionization, one only need to consider the transit time through the multiplication layer. For dual carrier ionization, however, there is a feedback process that introduces a time delay through the multiplication region. This is called the avalanche buildup time.

The experimentally observed frequency response of an APD is actually a complicated function of all of the above processes. We will present some of the experimental data below and attempt to provide the proper interpretations as they relate to the above processes.

4.2.1 Transient Response Results

Figure 4.1a shows the pulse response for an unbiased doped barrier MQW device. The APD was mounted on a 50 GHz Tektronix sampling scope and was excited with an 810 nm 50 ps laser pulse. As can be seen from the figure, the unbiased APD output pulse has a rise time of about 317 ps, a fall time of 2.5 ns and a full width at half max (FWHM) of about 1.4 ns. The oscilloscope trace is actually a convolution of the 50 ps Gaussian laser pulse with the output response of the APD. The fall time of the pulse shows a "fast" and a "slow" component. The slow component is due to trapping by interface states at the heterojunction, while the fast component is due to "hot" carriers that are not affected by the traps [17].

Figure 4.1b shows the response of the above APD under bias (low gain). The fall time and the FWHM have now dropped to 819 and 952 ps respectively which resulted in about 32% increase in the speed of the device. If the bias is further increased, as shown in Figure 4.1c, those values drop to 570 and 593 ps, respectively, with a speed increase of about 58%. This is to be expected since hole trapping and diffusion effects are only dominant at low gains. Diffusion will limit the speed of the device as long as there are undepleted regions in the structure and a separate absorption layer is being used. At high bias, trapping is no longer an issue, and the device response is limited by the transit time and the RC time

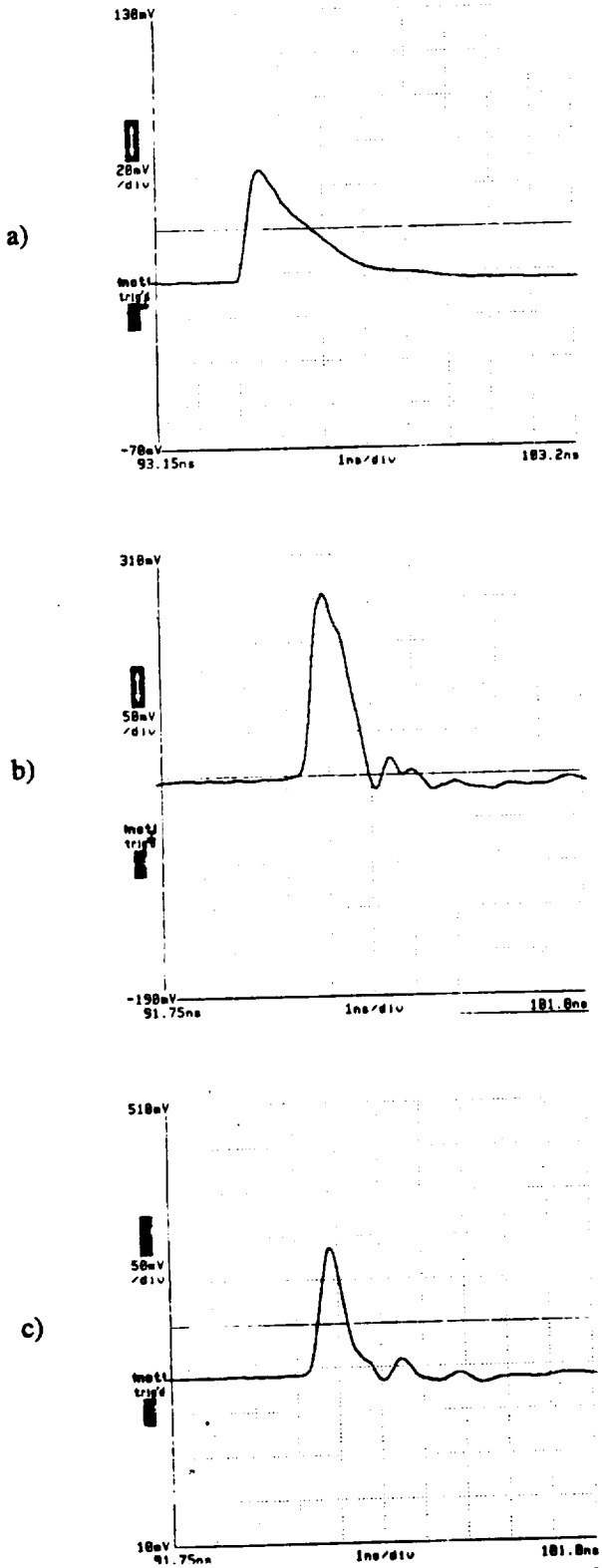


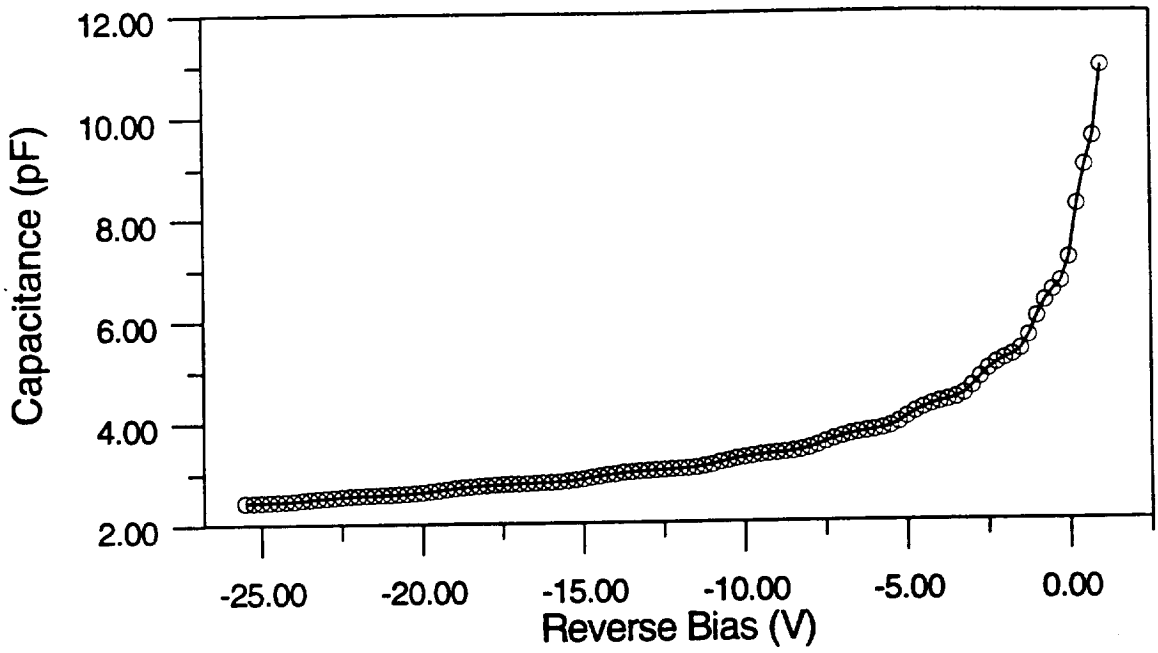
Figure 4.1 Transient pulse response for an a) unbiased doped barrier MQW APD, b) the same device under low bias and c) under higher bias.

constant. The "ringing effect" seen in the oscilloscope trace following the output pulse is due to impedance mismatch between the APD circuit and that of the oscilloscope sampling head. Note that the relative magnitude of the pulses is not representative of the gain of the device since a variable resistor was used in an attempt to match circuit resistances and thus limited the voltage applied at the oscilloscope.

4.2.2 CV Results

As discussed earlier, C-V measurements allow us to obtain the carrier concentration profile as a function of the depletion width throughout the MQW region. Our first APDs showed only partial depletion which greatly limited the low-noise gain of the initial devices. Recently, however, we were able to fabricate doped-well APDs in which all stages were fully depleted. This is demonstrated in Figure 4.2 for a 10 period doped-well MQW device. Figure 4.2a shows the capacitance-voltage (CV) data, and the dashed line in Figure 4.2b is the resulting carrier concentration profile as a function of the depletion width. Note that for the original data (dashed), the distance between the peaks decreases with increasing depletion width. This is due to the fact that the standard CV analysis used in the literature uses a one-sided junction approximation which does not apply for this particular device [18]. In order to take into account the two-sided depletion, one has to consider the dependence of the depletion width on the doping concentrations in the individual wells. This was done numerically using the measured peak values to adjust the depletion width throughout the structure. The results of this analysis are shown with the solid line in Figure 4.2b where the peak distance are now a more realistic representation of the actual wells spacing which is known to be about 1000 Å.

a)



b)

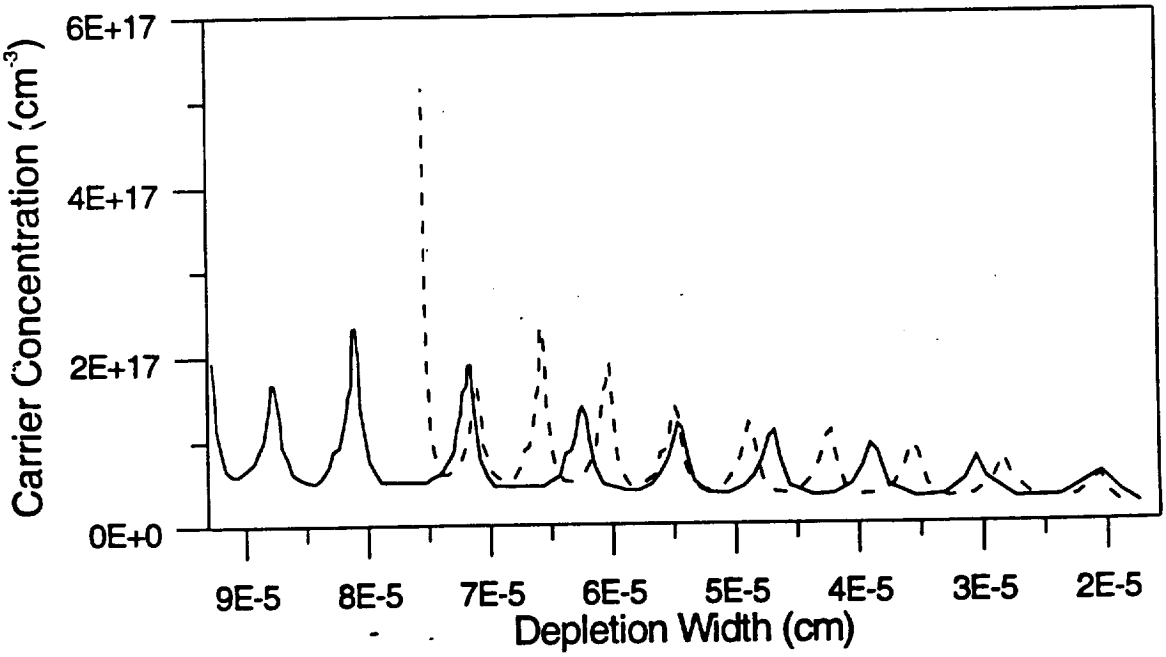


Figure 4.2 Doped well MQW APD a) C-V plot and b) resulting calculated carrier concentration profile. The dashed line represents the one sided depletion approximation while the solid line accounts for doping in the wells.

5.0 References

[1] A.R Hutson, J.H. McFee, and D.L. White, "Ultrasonic Amplification in CdS", Phys. Rev. Lett, Vol. 7, no.7, 1961, pp. 237-239.

[2] R. Adler, "Simple Theory of Acoustic Amplification", IEEE Trans. Sonics and Ultrason.,Vol. Su-18, no.3,July 1971, pp115-118.

[3] S. Ludvik and C.F. Quate, "Amplification of Surface Shear-Wave Mode in GaAS", J.Appl.Phys, Vol 43, no.9, Sept 1972, pp 3619-3622.

[4] B.J.Hunsinger, *Research to provide a theoretical determination of surface acoustic wave velocity and impedance differences between metal strips and free surface regions of metallic gratings, Technical Report*, ROME Air Development Center, 1981.

[5] Y.Kim and W.D.Hunt, "Acoustic fields and velocities for surface acoustic wave propagation in multilayered structures: An extension of the Laguerre polynomial approach," *J. Appl. Phys.*, vol.68, pp.4993-4997, 1990.

[6] S.Datta, *Surface Acoustic Wave Devices*, Prentice-Hall, 1986.

[7] P.V.Wright, "Modeling and experimental measurements of the reflection properties of SAW metallic gratings," *IEEE Proc. Ultrason. Symp.*, pp.54-63, 1984.

[8] E.D. Marshall, B. Zhang, L.C. Wang, P.F. Jiao, W.X. Chen, T. Sawada, K.L. Kavanagh, and T.F. Keuch, J. Appl. Phys. **62**, 942 (1987).

[9] L. V. Keldysh, Zh. Eksp. Teor. Fiz., **48**, 1962, (Sov. Phys. JETP, **21**, 1135 (1965)).

[10] Y. Wang and K. F. Brennan, J. Appl. Phys., **75**, xxxx (1994).

[11] Y. Wang and K. F. Brennan, J. Appl. Phys., **75**, xxxx (1994).

[12] W. Quade, E. Schoell, and M. Rudan, Solid-State Electron., **36**, 1493 (1993).

[13] R.B. Emmons, "APD frequency response", J. Appl. Phys. 38, 3705 (1967)

[14] J.C. Campbell, W. T. Tsang, G. J. Qua and J. E. Bowers, "InP/InGaAsP/InGaAs APDs with 70 Ghz gain-bandwidth product", Appl. Phys. Lett. 51, 1454 (1987).

[15] J. N. Hollenhorst, "Frequency response theory for multilayer photodiodes", IEEE J. Lightwave Technology LT-8, 531 (1990)

[16] B.C. Roy and N.B. Chakrabarti, "Pulse response of APDs", IEEE J. Lightwave Technology LT-10, 169 (1992).

[17] R.D. Dupuis, J.C. Campbell and J.R. Velebir, "InGaAs/InP photodiodes grown by metalorganic chemical vapor deposition", J. of Crystal Growth 77, 598-605 (1986)

[18] N. Yamamoto, K. Yokoyama and M. Yamamoto, "Carrier profile evaluation for a Zn-doped InGaAsP/InGaAsP MQW using a low-temperature C-V method", Appl. Phys. Lett. 62, 253-254 (1993)

6.0 Appendix

Theoretical investigation of wavevector dependent analytical and numerical formulations of the interband impact ionization transition rate for electrons in bulk silicon and GaAs

Ján Kolník, Yang Wang, Ismail H. Oğuzman, and Kevin F. Brennan

School of Electrical and Computer Engineering

Georgia Institute of Technology

Atlanta, Georgia

30332-0250

Abstract

In this paper calculations of the interband impact ionization rate in bulk silicon and GaAs are presented using three different formulations of the ionization transition rate within an ensemble Monte Carlo model. The transition rate is expressed as a wavevector-dependent (**k** - dependent), analytical formulation, as a completely numerical formulation and as the traditional Keldysh formula. Though the completely numerical formulation contains no adjustable parameters and as such provides a very reliable result, it is computationally too intensive for many applications. Alternatively, the Keldysh formula, though inherently simple and computationally efficient, fails to include the **k**-dependence as well as the details of the energy band structure. The **k**-dependent analytical formulation overcomes the limitations of both of these models but at the expense of some new parameterization. It is found that the **k**-dependent analytical method produces very similar results to that obtained with the completely numerical model. Specifically, both models predict that the effective threshold for impact ionization in GaAs and silicon is quite soft and that the majority of ionization events originate from the second conduction band in both materials. Therefore, it is concluded that the **k**-dependent analytical model can qualitatively

reproduce similar results to those obtained with the numerical model yet with far greater computational efficiency. The only major limitation of the analytical approach is that it requires determination of a new parameter, related physically to the overlap integrals of the Bloch states. This parameter can only be adjusted by comparison to experiment.

I. Introduction

The theoretical study of interband impact ionization has been greatly aided by the advent of numerical methods. Owing to the complexity of the energy band structure at high energies, at which impact ionization is initiated, as well as the complicated carrier-phonon scattering mechanisms, simplified analytical formulations, such as the early theories of Shockley [1] , Wolff [2] and Baraff [3] have limited validity. Numerical models for calculating the ionization rate were then advanced based on the Monte Carlo method [4-5]. However, these approaches utilized parabolic or nonparabolic analytical energy bands which are of questionable validity at high carrier energies. The full details of the energy band structure were first accounted for in the Monte Carlo model of Shichijo and Hess [6]. In their model, the dynamics of the electrons in bulk GaAs were simulated within the first conduction band, calculated based on an empirical pseudopotential model. Later, Fischetti and Laux [7] developed a more advanced Monte Carlo simulator for studying impact ionization which improved the numerical accuracy as well as incorporated transport within higher conduction bands.

In order to calculate the overall impact ionization rate using the Monte Carlo technique, it is necessary to formulate an expression for the impact ionization transition rate. This formulation is then incorporated into the Monte Carlo simulator and is treated as an additional scattering mechanism. An impact ionization event is chosen stochastically in the usual way [8] through use of a random number based on the relative magnitude of the ionization transition rate compared to competing phonon scattering events. Though there has been recent work on formulating the ionization transition rate using a full order quantum mechanical approach [9-11], the impact ionization transition rate is more typically determined from use of Fermi's golden rule. Fermi's

golden rule can be expressed as [12],

$$W_{ii} = \int \frac{2\pi}{\hbar} |M|^2 \delta(E_f - E_i)\, \delta(\Delta k)\, dS_f \quad , \tag{1}$$

where dS_f represents integration over all the final states, E_i and E_f are the energies of the initial and final states, respectively, Δk is the momentum change during the interaction and M is the matrix element of the interaction given as

$$M = \frac{e^2}{\epsilon V} \frac{I(k_1, k_1')\, I(k_2, k_1')}{|k_1' - k_1|^2 + \lambda^2} . \tag{2}$$

In Eq.2, k_1, k_1' are the wave vectors of the incident electron before and after the interaction. k_2 represents the target electron before the collision. λ is the static screening factor. e is the electronic charge; V is the crystal volume; and ϵ is the dielectric constant. $I(k_1, k_1')$ and $I(k_2, k_1')$ are overlap integrals [12].

In most previous Monte Carlo simulations [e.g 6,13], the impact ionization transition rate is calculated using the Keldysh formula [14], which can be obtained from (1) by adopting several simplifying assumptions. These are that the transition is calculated assuming a direct semiconductor with parabolic bands for all of the carriers, the overlap integrals are considered to be constant and the denominator in the matrix element in Eq. (2) is also taken as a constant with threshold values of k_1, k_1'. This yields the very well known quadratic dependence of the transition

rate, W_{ii}, on the energy E of the initiating particle, usually expressed as

$$W_{ii} = p\, W(E_{th}) \left(\frac{E-E_{th}}{E_{th}}\right)^2 , \qquad (3)$$

where E_{th} is the threshold energy and $W(E_{th})$ is the total phonon scattering rate at threshold. E_{th} and the prefactor p cannot generally be determined from first principles. Instead they are determined through comparison of the calculations to experimental data. Aside from this obvious limitation of the Keldysh formula, the neglect of an accurate accounting of the high energy region of the band structure also results in the failure to properly assess the **k**-dependence of the ionization rate itself. Recent theoretical studies have indicated that there is a significant deviation in the ionization rate with respect to the initiating carrier's **k**-vector in many different materials [15,16].

An alternative approach to the Keldysh formula for calculating the interband impact ionization rate via Monte Carlo simulations has been recently presented by Wang and Brennan [15,16]. They determined the transition rate directly from Fermi's golden rule by numerically integrating Eq. 1 at different **k**-points within the first Brillouin zone. Their analysis includes the direct calculation of the overlap integrals from the numerically generated wavefunctions using a **k.p** method. Using this approach, the transition rate was found to be dependent on the initiating electron wave vector. The impact ionization rate can be obtained without relying on any adjustable parameters, thus making possible the investigation of the ionization rate in materials for

which no reliable experimental data of the impact ionization rates are available. The disadvantage of this technique however, is that extensive computations are needed. This is due mostly to the complex evaluation of the overlap integrals needed to determine the transition rate for those states satisfying the energy and momentum conservation requirements.

Recently, another analytical treatment of expression (2), which also includes the **k**-vector dependence of the impact ionization transition rate, was presented by Quade et al [17]. They calculated the impact ionization transition rate for the general case of an arbitrarily shaped energy band for the impact-ionizing conduction electron, and three anisotropic parabolic bands for the final states with their extrema located at different points in the **k**-vector space. The assumption of parabolic energy bands for the final states is acceptable since the final states are generally at low energy near the band minimum where the parabolic approximation is typically good. Other approximations used in the derivation include nondegeneracy and a constant value for the overlap integrals. Though the formulation developed by Quade et al. [17] contains some parameterization due to the fact that the overlap integrals are not evaluated but are assumed to be constant parameters, it offers a far more computationally efficient means of evaluating the ionization transition rate with the inclusion of the **k**-dependence than the direct, numerical method of Wang et al. [15,16]. Subsequently, it is of interest to examine how the **k**-dependent, analytical model of Quade et al. [17] compares to the direct, numerical model of Wang et al. [15,16].

In this paper, we incorporate the formulation of the impact ionization transition rate of Quade et al. [17] within an ensemble Monte Carlo simulator to determine the electron impact ionization rate in bulk silicon and GaAs. The impact ionization rate calculated in this way is then compared to experimental data as well as calculations made in an otherwise identical Monte Carlo

simulator using the Keldysh formula and the direct, k-dependent transition rate. The nature of the ionization rate is further probed using the Quade model to determine if it yields similar physical results to that of the direct, numerical model. Specifically, the transition rate within each band is calculated and the percentage of ionization events originating from each conduction band is determined. The number density distributions as a function of energy are determined and compared between the different theoretical models as well. In Section II, the details of the theoretical models are reviewed. The calculated results are presented in Section III and conclusions are drawn in Section IV.

II. Model Description

The total impact ionization rate as a function of inverse electric field is calculated using an ensemble Monte Carlo simulation, based on the original work of Shichijo and Hess [6]. The details of the simulator have been extensively described elsewhere [15,16]. Therefore only the main features are outlined here. The simulation includes the full details of the first and second conduction bands for silicon or gallium arsenide. The scattering rate in either case is determined using the improved phonon scattering technique of Chang et al. [18] and the improved band structure interpolation technique of Fischetti and Laux [7]. The total phonon scattering rate is determined in the following way. Within the low energy region, the scattering rate is calculated from Fermi's golden rule for all the relevant mechanisms present [19]. Within the high energy range, deformation potential scattering is assumed to be the dominant scattering mechanism. The scattering rate in this region is obtained by integrating over the final density of states, calculated from the numerically generated band structures, including collision broadening of the final state. The deformation potential is assumed constant and is selected to match the scattering rate

calculated from Fermi's golden rule at a low specific energy. Interband electronic transitions are enabled via the action of the deformation potential scattering as described in Ref. [15].

The impact ionization transition rate for an initiating electron of wavevector $\mathbf{k}_1$, colliding with another electron in the valence band is determined by Quade et al. [17] to be

$$W_{ii}(k_1)_{\alpha,\beta,\gamma} = \frac{1}{\tau}\int_{-\infty}^{K_1^2} d\varepsilon\, W_{ii}(K_1,\varepsilon)_{\alpha,\beta,\gamma}\,\delta(g(k_1)-\varepsilon), \tag{4}$$

where

$$W_{ii}(K_1,\varepsilon)_{\alpha,\beta,\gamma} = \frac{1}{\tau}\left[\frac{1}{2}\left(\frac{K_1}{\sqrt{\varepsilon_\lambda}}+\frac{\sqrt{\varepsilon_\lambda}}{K_1}\right)-1\right]\Theta(K_1^2-\varepsilon), \tag{5}$$

$$\frac{1}{\tau} = \frac{1}{4\pi^2}\,\frac{\sqrt{\mu}}{\prod_{i=1}^{3}(\alpha_i+\beta_i)^{1/2}}\left(\frac{e^2}{\varepsilon}\right)^2\frac{m_0^* F}{\hbar^3}, \tag{7}$$

$$\mu_i = \frac{\alpha_i\beta_i+\alpha_i\gamma_i+\beta_i\gamma_i}{\alpha_i+\beta_i}, \tag{8}$$

$$g(k_1) = \frac{1}{\mu}\left[\varepsilon_0+\sum_{i=1}^{3}\frac{\alpha_i\beta_i}{\alpha_i+\beta_i}(k_{\alpha,i}-k_{\beta,i})^2+\sum_{i=1}^{3}\gamma_i(k_1^2-k_{\gamma,i})^2-\varepsilon(k_1^2)\right], \tag{8}$$

$$K_{1,i}=\frac{1}{\sqrt{\mu\mu_i}}\left[\gamma_i\left(k_{1,i}^{L}-Z_{\gamma,i}\right)\frac{\alpha_i\beta_i}{\alpha_i+\beta_i}\left(Z_{\alpha,i}-Z_{\beta,i}\right)\right], \tag{9}$$

and

$$e_\lambda=\frac{1}{2}\left[\sqrt{(\lambda^2-e)^2+4K_1^2\lambda^2}-(\lambda^2-e)\right] \quad . \tag{10}$$

In the above set of equations, α, β and γ represent a parabolic valence band and two possibly different parabolic conduction bands for the final states of the hole and electrons after the impact ionization event, respectively. The anisotropic bands are parametrized by ratios of effective masses $\alpha_i = m_0/m_{\alpha,i}$, $\beta_i = m_0/m_{\beta,i}$, $\gamma_i = m_0/m_{\gamma,i}$, where i labels the Cartesian coordinate, and the extremal points of each band are positioned at $\mathbf{Z}_\alpha$, $\mathbf{Z}_\beta$, $\mathbf{Z}_\gamma$; m_0 is an isotropic effective mass of the conduction band for the initiating electron. F represents the product of the squared overlap integrals in Eq.2 and is taken as constant for each initiating band but may vary between bands. μ stands for the geometrical average of the quantities μ_1, μ_2, μ_3 defined by Eq. (7). The arbitrarily shaped function $\epsilon(\mathbf{k}_1)$ represents the energy of the initiating electron, ϵ_0 represents the gap plus offsets of the extremal points of the bands α, β, γ. $\mathbf{k}_1^L=\mathbf{k}_1-\mathbf{L}$, where $\mathbf{L}$ is a reciprocal lattice vector and Θ is the Heaviside function.

The total impact ionization scattering rate for the initiating particle with wavevector $\mathbf{k}_1$ is calculated taking into account all of the possible final states within the full Brillouin zone. The total impact ionization transition rate can then be expressed as

$$W_{11}(k_1) = \sum_{\alpha}\sum_{\beta}\sum_{\gamma} W_{11}(k_1)_{\alpha,\beta,\gamma} \quad , \tag{11}$$

where the summation goes over all the bands α and all possible valleys β, γ in the first and second conduction bands. The summations in Eq.11 are taken so as to include both direct and exchange terms.

The material parameters used in the calculations have been taken from Refs. [20-22] for silicon and [23-27] for gallium arsenide. For the case of the second conduction band in both materials, the longitudinal, m_l, and transverse, m_t, effective masses for the X-valleys have been obtained from the numerically calculated band structures determined by a **k.p** calculation. The longitudinal and transverse masses used for silicon and gallium arsenide in the second conduction band are reported in Table 1. In the case of L-valleys, Γ-valleys and all the valence bands, the isotropic effective masses have been used and their values are also listed in Table 1, as well as the energies of symmetry points relative to the top of the valence band.

The comparison between the numerically calculated interband impact ionization transition rate and the rate calculated using the approach based on Quade's results assuming that the overlap integrals are simply unity is presented in Figs. 1 - 4 for the first and second conduction bands in silicon and gallium arsenide. Aside from the difference in magnitude, which is caused by the fact that the effect of the overlap integrals was not included in the analytical model, very good qualitative agreement between the data calculated using the analytical formula and the data obtained by direct numerical integration (which includes the overlap integrals evaluation) can be seen for the first (Fig. 1a,b) and second (Fig. 2a,b) conduction bands in silicon. Notice that the number of points for which a nonzero transition rate has been calculated is higher in the case of

the numerical approach. A similar effect can be seen in the case of the first (Fig. 3a,b) and second (Fig. 4a,b) conduction bands in gallium arsenide. Again notice that there is a significant difference in the number of nonzero transition rate points between the two models for the first conduction band. This is probably due to the fact that these points are not close to the band extrema and the corresponding energy cannot be found using the parabolic relation between the energy and the distance from the band minimum in the **k**-vector space. However, the rate from these points is relatively small compared to the competing scattering rates.

Further examination of Figures 1-4 reveals that the transition rate is greater for electrons originating from the second conduction band than from the first conduction band. This is particularly true for GaAs. As will be discussed below, due to the fact that the transition rate is greatest for the second conduction band, most ionization events occur from electrons which originate from within the second conduction band.

III. Results

The impact ionization rate as a function of inverse electric field, calculated using the ensemble Monte Carlo simulation, is presented in Figs. 5 and 6 for silicon and gallium arsenide, respectively. In each case, the rate is calculated using the Keldysh formula, the **k**-dependent numerical formulation and the **k**-dependent analytical formulation for the impact ionization transition rate. The experimentally obtained data are also shown in both figures. The experimental results of Overstraeten [28], Grant [29] and Woods [30] for electron initiated impact ionization in silicon are plotted in Figure 5. These data comprise a representative set of data which cover the full range of experimental measurements for electron impact ionization in

silicon. The experimental measurements of Bulman et al. [31] are plotted for gallium arsenide in Figure 6.

In the case of silicon, good agreement with the data of Woods is obtained for the **k**-dependent numerical formulation, as was also reported previously [15]. No adjustable parameters for the impact ionization transition rate are used in this case. Very similar results can be obtained using the analytical **k**-dependent formulation by choosing the values for the parameter F in Eq. (6) equal to 0.24 for both the first and second conduction bands. F corresponds to an "average" value of the squared overlap integrals that appear in Eq. 2. The squared overlap integrals always have value between 0 and 1. In the analytical model of Quade et al. [17] and that presented here, F is treated as a parameter which is chosen so as to yield good agreement with experiment. Best agreement with the low range of silicon experimental data occurs for F=0.24 given the phonon scattering rate adopted here. The calculated silicon impact ionization rate obtained using the Keldysh formula with p=0.02 and E_a=1.13 eV is also shown in Figure 6. A value of p of about 0.02 is typically considered to correspond to a soft threshold. In contrast to both **k**-dependent models, better agreement with the higher range of experimental measurements can be seen for this particular choice of threshold and p. However, a different choice of the parameter F within the analytical **k**-dependent formulation would also provide a better fit to the higher range of experimental data as well.

In gallium arsenide, the dependence of the impact ionization rate on the inverse electric field obtained using the numerical, the analytical **k**-dependent models with F=0.1 and 0.02 for the first and second conduction band, respectively, and the Keldysh formula with p=0.02, E_a=1.85eV for the impact ionization transition rate, shows in all three cases fairly good agreement with the

experimentally measured data (Fig.6). It should be noted that in the case of the k-dependent analytical approach, the selection of the fitting factor F for the first conduction band essentially does not play any role due to the negligible impact ionization transition rate of the electrons in the first conduction band with respect to the second conduction band.

The percentage of ionization events originating from the first and the second conduction band in bulk silicon as a function of applied electric field is plotted in Fig. 7a for the numerical, and Fig. 7b for the k-dependent analytical model. At low applied electric fields, the impact ionization events originate predominantly in the second band in both cases. Therefore, on the basis of these results, it appears that most electron initiated ionization events originate from electrons within the second conduction band. The number of events originating in the first conduction band is found to be higher for the numerical model than for the k-dependent analytical model at the field of 250 kV/cm, while for higher fields this difference vanishes. This slightly different behavior of the two models at low fields is caused by the fact that the threshold energy for impact ionization is about 1.32 eV in the analytical k-dependent model for the particle originating in the first conduction band and 2.11 eV for the second conduction band, while for the numerical model the threshold energy value is not well defined. Instead, in the numerical model ionization is possible at all states with energy greater than the band gap which are found to satisfy the momentum and energy conservation conditions. Therefore, there is no sharp threshold energy, other than the band gap, below which ionization cannot occur and above which ionization does occur. Thus the contribution of the first conduction band at low fields can be expected to be higher when the numerical model is used.

The effect of the second conduction band on the impact ionization rate in gallium arsenide

is even more important than in silicon. The impact ionization transition rate for electrons in the first conduction band is much lower than the rate within the second conduction band in both models as can be seen from Figs 3 and 4. As a result, virtually all the electrons which impact ionize originate within the second conduction band. The effect is even stronger using the **k**-dependent analytical formulation than the numerical formulation. Though the value of F used in the analytical formulation is unknown a priori, i.e. without first comparing the calculations to experiment, it should be noted that F can never be greater than 1. Therefore, given that the transition rate of electrons within the first conduction band is several orders of magnitude lower than from the second conduction with F=1, as shown by Figs. 3 and 4, it is clear that the first conduction band plays no significant role in impact ionization in bulk GaAs regardless of the value of F. Hence, in either the numerical or analytical **k**-dependent models the predominant source of ionizing electrons in bulk GaAs are carriers within the second conduction band.

The electron number density function defined as the density of states function multiplied by the electron distribution function, is shown in Figs. 8a and 8b for silicon at an applied electric field of 500 kV/cm for the analytical **k**-dependent model and the numerical **k**-dependent model, respectively. In both cases, the number of electrons which survive to high energies is relatively high. The second peak in each of the figures corresponds to the electron population in the second conduction band. Both models essentially confirm previous results [32-34] that the threshold in silicon is exceedingly soft and that the electrons can reach energies substantially higher than the band gap before they impact ionize.

The electron number density functions in gallium arsenide are plotted in Figs. 9a,b calculated using the **k**-dependent numerical and analytical models, respectively, for the applied

electric field of 500 kV/cm. In both cases the distribution functions are similar and the effect of the second conduction band (the second peak at higher energies) is again slightly larger in the analytical model. Notice that the number distribution remains substantially large at high carrier energies indicating an exceedingly soft effective threshold for impact ionization in GaAs. Again these results are consistent with previously reported observations [34].

The energy dependent impact ionization transition rate, derived from the k-dependent rate in a manner described in [15], is plotted in Figs 10a, b for the numerical and analytical k-dependent formulations, respectively, in silicon. Analogous curves are shown in Figs. 11a,b for gallium arsenide. As can be seen from Figures 10a and 10b the ionization rate increases monotonically in silicon for both cases. Inspection of Figs. 11a and 11b shows that the transition rate reaches a maximum near 3 eV in GaAs, and then decreases. The maximum region is more pronounced and the rate at the maximum is greater for the analytical formulation than the corresponding rate in the numerical case. The origin of the maximum can be understood within the analytical formulation since the transition rate from each valley can be individually evaluated. The maximum in the transition rate is due to the collision of an electron within the second conduction band with an electron in the heavy hole band, after which the electrons occupy states either both in the Γ-valley (the lower peak at the energies 2.8-3.0 eV) or one is in the Γ-valley and the other in the L-valley (the higher peak at 3.4 eV). In both cases, a transition of the initiating electron from the second conduction band to the first one with only a slight change of the k-vector (a "near vertical" transition, as explained in [17]) contributes to the final values of the transition rate. Near vertical transitions have a significantly greater rate owing to the fact that the denominator in Eq. 2 becomes very small under these conditions [35]. This region may be

somewhat unrealistically pronounced due to the fact the value of the overlap integrals were taken as an averaged constant. The actual values of the overlap integrals in this region may be quite small thereby reducing the rate if calculated. It is much more difficult to isolate the physical origin of the maximum in the numerical model. In this case, a similar classification would be questionable, mainly due to the uncertainty in which the "valleys" are defined in the real band structure.

The value of the overlap integrals used in the evaluation of the transition rate given by Eq. 2 are evaluated directly in the numerical approach. As mentioned above, this requires extensive numerical computation. By approximating these overlap integrals as a constant, the factor F in Eq. 6, a new single parameter is introduced greatly reducing the numerical complexity required to evaluate the transition rate but at the expense of a new unknown parameter. In the calculations presented above, F is adjusted such that the analytical model yields satisfactory agreement with the experimental data. To check to see if the determined values of F are reasonable, it is useful to compare these values to those determined using the sum rule quoted by Ridley [12]. Using the sum rule, the squared overlap integral between the first conduction band and the heavy hole band in gallium arsenide is determined to be 0.38. More elaborate calculations made by Burt et al. [36,37] using wavefunctions obtained from a k.p calculation show that this value is $\sim 10^{-2}$, which is in good agreement with the results presented here for GaAs, F=0.02 for the second conduction band. In the case of silicon, the value F=0.24 obtained for the impact ionization originating in both the first and second conduction bands, is higher than the value of F~0.05 [38] or F=0.01, quoted in [39,40] obtained by theoretical evaluations. It should be noted though that a discrepancy between the experimentally determined recombination coefficients for

Auger recombination, and the theoretically obtained ones, which also contain the same overlap integrals, was observed in bulk silicon. This discrepancy is thought to be possibly due to an underestimation of the overlap integrals by roughly an order of magnitude [41].

IV. Conclusions

In this article, the electron interband impact ionization rate calculated by an ensemble Monte Carlo simulation has been presented for both bulk silicon and gallium arsenide. The impact ionization transition rate is determined from a semiclassical description of the two-body screened interaction with energy and momentum conservation. The transition rate has been evaluated from Fermi's golden rule in three different ways. A **k**-dependent, analytical formulation based on Quade's [17] technique, a **k**-dependent completely numerical evaluation including the direct determination of the overlap integrals, and the traditional Keldysh formula approach have been used. Of the three techniques, the completely numerical calculations do not contain any adjustable parameters, and as such yield the most reliable result. Unfortunately, the completely numerical technique, though fundamentally based, requires extensive numerical computation making it somewhat unattractive. As is well known, the Keldysh formula, though commonly used in Monte Carlo calculations of the impact ionization rate owing to its relative simplicity and computational efficiency, suffers from several limitations as well. Chief among these is the fact that it is derived assuming parabolic energy bands for all of the participating carriers, and includes no **k**-dependence in the rate. Additionally, due to its reliance on two adjustable parameters, more than one set of parameters can be used within the Keldysh formula to produce satisfactory agreement with experiment. As a result, the Keldysh formula cannot by itself predict whether a material has a soft or hard ionization threshold. Therefore, there is a need to develop a more

computationally efficient transition rate formulation than the direct, numerical technique, yet retain the **k**-dependence and details of the band structure in its formulation. In this way, the physical nature of the ionization process can be preserved (soft or hard threshold, etc.) while maintaining computational efficiency. The Quade formula, though it contains one adjustable parameter corresponding physically to the squared overlap integrals of the periodic parts of the Bloch functions, includes the full details of the band structure for the initiating carrier. In this paper, we have provided the first comparison of this formula to the direct numerical technique to determine if this approach yields qualitatively similar results while providing far better computational efficiency.

The calculations presented here show that the **k**-dependent analytical formulation produces very similar results to the completely numerical calculations. In contrast to the Keldysh formula, where no **k**-vector dependence of the impact ionization transition rate is included, the transition rate in the **k**-dependent numerical and analytical cases is shown to be strongly dependent on the initiating electron **k**-vector. It is also shown that using all three models a relatively good fit to the experimentally determined impact ionization rate can be obtained. In the case of silicon, the value of the fitting parameter for the **k**-dependent analytical formulation for both the first and second conduction bands is taken as F=0.24. Good agreement with the low range of the experimental data represented by the measurements of Woods et al [30], is obtained, which is consistent with the results obtained using the numerical formulation. For comparison, the ionization rate in silicon obtained with the Keldysh formula (p=0.02, E_a=1.13 eV) is also determined.

All three models can be used to fit the experimentally obtained data in gallium arsenide as

well. Again good qualitative agreement between the **k**-dependent analytical and numerical models is obtained. In GaAs, F=0.02 is used for the second conduction band for the **k**-dependent analytical formulation (the choice of F for the first conduction band is irrelevant), while in the Keldysh formula p=0.02 and E_g=1.85eV are used.

Further comparison of the **k**-dependent analytical and numerical models reveals that they predict similar physical characteristics for the impact ionization rate. Specifically, both models indicate that the majority of ionization events originate from the second conduction band in both silicon and GaAs. Additionally, both models predict that the effective threshold for impact ionization is relatively soft, implying that the carriers drift to substantially high energies before suffering an ionization event. It is shown that the average energy of the electrons is high when compared to that expected for a hard threshold model. Thus the important qualitative results obtained by adopting the computationally intensive numerical approach, i.e. the dominance of the second conduction band and the importance of the **k**-dependence, can be recovered if the analytical **k**-dependent formulation is used. Although the **k**-dependent analytical formulation cannot fully replace the numerical approach, which contains no parameterization and hence is more of a first principles theory, it can be used for many applications where the experimentally determined impact ionization rate is available and the parameter, F, which appears in this model can be ascertained. It should be emphasized that the **k**-dependent, analytical model can reproduce many of the important results of the numerical model with far greater computational efficiency.

Acknowledgements

The authors would like to thank Drs. C. J. Summers and B. K. Wagner for many helpful

discussions. This work was supported in part by the National Science Foundation through a Presidential Young Investigator Award made to K. Brennan, through Bell Northern Research grant I25A3J and by ARPA through contract to NASA, NAGW-2753. Computing support was provided by the Digital Equipment Corp. through contract E21-H36.

References

[1] W. Shockley, Solid-State Electron., **2**, 35 (1961).

[2] P. A. Wolff, Phys. Rev., **95**, 1415 (1954).

[3] G. Baraff, Phys. Rev. **128**, 2507 (1962).

[4] R. C. Curby and D. K. Ferry, Phys. Status Solidi A, **15**, 319 (1973).

[5] P. A. Lebwohl and P. J. Price, Solid State Commun., **9**, 1221 (1971).

[6] H. Shichijo and K. Hess. Phys. Rev. **B 23**, 4197 (1981).

[7] M.V. Fischetti and S.E. Laux, Phys. Rev. **B 38**, 9721 (1988).

[8] C. Jacoboni and P. Lugli, The Monte Carlo Method for Semiconductor Device Simulation, Springer-Verlag, Wien-New York, 1989.

[9] J. Bude, K. Hess, and G. J. Iafrate, Phys. Rev. B, **45**, 10958 (1992).

[10] W. Quade, F. Rossi, and C. Jacoboni, Semicond. Sci. Technol., 7, B502 (1992).

[11] P. D. Yoder, J. M. Higman, J. Bude and K. Hess, Semicond. Sci. Technol., 7, B357 (1992).

[12] B.K. Ridley, Quantum Processes in Semiconductors, 2nd ed., Oxford University Press, Oxford, 1988.

[13] J.Y. Tang and K. Hess, J. Appl. Phys. **54**, 5139 (1983).

[14] L.V. Keldysh, Zh. Eksp. Teor. Fiz. **48**, 1962 (1965) [Sov. Phys. JETP **21**,1135 (1965)].

[15] Yang Wang and K. F. Brennan, J. Appl. Phys, **75**, xxxx-xxxx (1994).

[16] Yang Wang and K. F. Brennan, J. Appl. Phys., **75**, xxxx-xxxx (1994).

[17] W. Quade, E. Scholl and M. Rudan, Solid St. Electron. **36**, 1493 (1993).

[18] Y.C. Chang, D.Z.-Y. Ting, J.Y. Tang, and K. Hess, Appl. Phys. Lett., **42**, 26, 1983.

[19] W. Fawcett, A. D. Boardman and S. Swain, J. Phys. Chem. Solids, **31**, 1963 (1970).

[20] J.Y. Tang, PhD. thesis, University of Illinois, Urbana, IL, 1982.

[21] Y. Ohno, J. Appl. Phys. **64**, 4549 (1988).

[22] C. Jacoboni, C. Canali, G. Ottaviani and A.A. Quaranta, Solid St. Electron., **20**,77 (1977).

[23] K.F. Brennan, D.H. Park, K. Hess, and M.A. Littlejohn, J. Appl. Phys., **63**, 5004 (1988).

[24] K.F. Brennan, N. Mansour and Yang Wang, Computer Phys. Comm., **67**, 73 (1991).

[25] D.H. Park, PhD. thesis, Georgia Inst. of Technology, Atlanta, GA, 1986.

[26] Landbolt-Börnstein, Numerical Data and Functional Relationships in Science and Technology, K.-H. Hellwege and O. Madelung, eds., Springer Verlag, New York, 1982.

[27] M.L. Cohen and T.K. Bergstresser, Phys. Rev., **141**, 789 (1966).

[28] R. van Overstraeten and H. DeMan, Solid State Electron. **13**, 583 (1970).

[29] W.N. Grant, Solid State Electron. **16**,1189 (1973).

[30] M.H. Woods, W.C. Johnson and M.A. Lambert, Solid State Electron., **16**, 381 (1973).

[31] G.E. Bulman, V.M. Robbins, K.F. Brennan, K. Hess and G.E. Stillman, IEEE Electron Dev. Lett., **EDL-4**, 181 (1983).

[32] N.Sano, T.Aoki, and A.Yoshii, Appl. Phys. Lett. **55**, 1418 (1989).

[33] N.Sano, M.Tomizawa, and A. Yoshii, Appl. Phys. Lett., **56**, 653 (1990).

[34] J. Bude and K. Hess, J. Appl. Phys. **72**, 3554 (1992).

[35] Y. Wang and K. F. Brennan, J. Appl. Phys., **71**, 2736 (1992).

[36] M.G.Burt and C.Smith, J.Phys.C, **17**, L47 (1984).

[37] M.G. Burt, S.Brand, C. Smith, and R.A. Abram, P.Phys. C, **17**,6385 (1985).

[38] J. Huldt, Phys. Stat. Sol.(a) **8**, 173 (1971).

[39] D. Hill and P.T. Landsberg, Proc. R. Soc. A, 347, 547 (1976).

[40] D. Hill, Proc. R. Soc. A, **347**, 565 (1976).

[41] J. Huldt, **1972 Int. Conf. Semicon. Phys.**, Warsaw 1097 (1972).

Figure Captions

Fig. 1 Calculated interband impact ionization transition rate for the first conduction band in bulk silicon as a function of k_x and k_y (k_z is fixed at 0.0) within the reduced zone of the first Brillouin zone. a) results obtained by direct numerical calculations; b) k-dependent analytical calculated results.

Fig. 2 Calculated interband impact ionization transition rate for the second conduction band in bulk silicon as a function of k_x and k_y (k_z is fixed at 0) within the reduced zone of the first Brillouin zone. a) results obtained by numerical calculations; b) analytically calculated results.

Fig. 3 Calculated interband impact ionization transition rate for the first conduction band in bulk gallium arsenide as a function of k_x and k_y (k_z is fixed at 0) within the reduced zone of the first Brillouin zone. a) results obtained by numerical calculations; b) analytically calculated results.

Fig. 4 Calculated interband impact ionization transition rate for the second conduction band in bulk gallium arsenide as a function of k_x and k_y (k_z is fixed at 0) within the reduced zone of the first Brillouin zone. a) results obtained by numerical calculations; b) analytically calculated results.

Fig. 5 Calculated and experimental results for the electron impact ionization coefficients in bulk silicon plotted as a function of inverse electric field. Results of the calculations of the ionization rate including the Keldysh formula, k-dependent numerical and k-dependent analytical model are presented. Experimental results of Overstraeten [28], Grant [29] and Woods

[30] are also shown. F1 and F2 represent the values of the averaged squared overlap integrals for bands 1 and 2, respectively.

Fig. 6 Calculated and experimental results for the electron impact ionization coefficients in bulk gallium arsenide plotted as a function of inverse electric field, using the Keldysh formula, **k**-dependent numerical and **k**-dependent analytical model. Experimental results shown are those of Bulman et al. [31]. F1 and F2 represent the values of the averaged squared overlap integrals for bands 1 and 2, respectively.

Fig. 7 Histogram showing the percentage of ionization events originating from the first and second conduction bands in silicon for (a) the **k**-dependent numerical formulation, (b) the **k**-dependent analytical formulation.

Fig. 8 The electron number density function in silicon, defined as the product of the density of states function and the electron distribution function, at an applied electric field of 500 kV/cm, calculated using (a) the **k**-dependent numerical model, (b) the **k**-dependent analytical model.

Fig. 9 The electron number density function in gallium arsenide, defined as the product of the density of states function and the electron distribution function, at an applied electric field of 500 kV/cm, calculated using (a) the **k**-dependent numerical model, (b) the **k**-dependent analytical model.

Fig. 10 Calculated impact ionization transition rate in silicon as a function of electron energy measured from the conduction band minimum using (a) the **k**-dependent numerical model, (b) the **k**-dependent analytical model. The energy dependent rate is calculated by performing a further integration of the **k**-dependent rate following the approach explained in [15].

Fig. 11 Calculated impact ionization transition rate in gallium arsenide as a function of electron energy measured from the conduction band minimum using (a) the **k**-dependent numerical model, (b) the k-dependent analytical model. The energy dependent rate is calculated by performing a further integration of the **k**-dependent rate following the approach explained in [15].

TABLE I. Energies of symmetry points (relative to the top of the valence band, in eV) and values of relative effective masses used in the calculations of the impact ionization transition rate in silicon and gallium arsenide

	Silicon	Gallium Arsenide
$E(\Gamma_{c1})$	3.42[a]	1.423[d]
$E(X_{c1})$	1.12[a]	1.898[d]
$E(L_{c1})$	1.92[a]	1.706[d]
$E(X_{c2})$	1.14[f]	1.973[f]
$E(\Gamma_{so})$	-0.044[b]	-0.340[d]
$m_t(X_{c1})$	0.19[c]	0.27[e]
$m_l(X_{c1})$	0.92[c]	1.98[e]
$m_t(X_{c2})$	0.21[f]	0.23[f]
$m_l(X_{c2})$	0.29[f]	0.40[f]
$m(L_{c1})$	0.284[c]	0.18[e]
$m(\Gamma_{c1})$	---	0.067[e]
$m(\Gamma_{HH})$	0.537[b]	0.45[e]
$m(\Gamma_{LH})$	0.153[b]	0.082[e]
$m(\Gamma_{so})$	0.234[b]	0.170[e]

[a] Ref. [27] [d] Ref. [25]

[b] Ref. [21] [e] Ref. [26]

[c] Ref. [20] [f] values obtained from the numerically calculated band structure

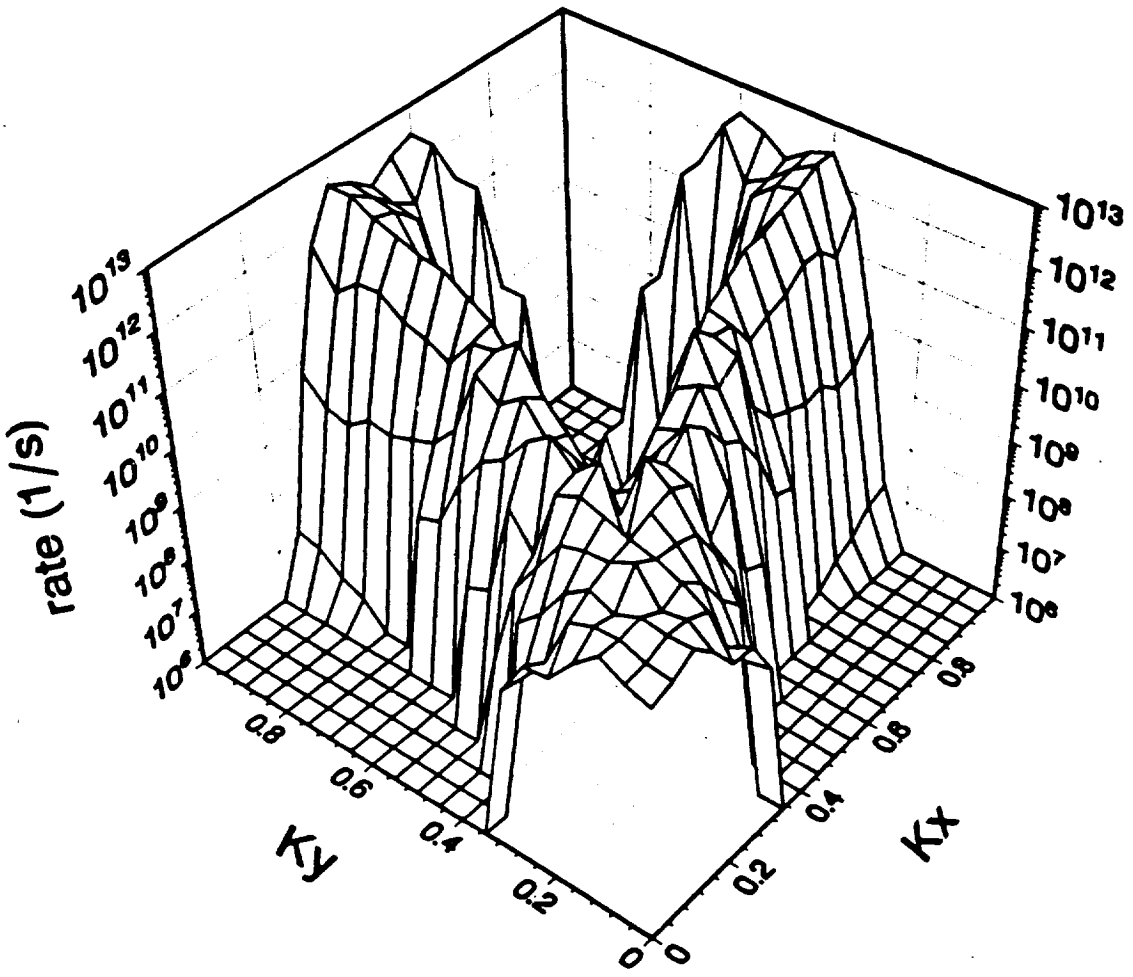


FIG 1A

KOLNIK u al.

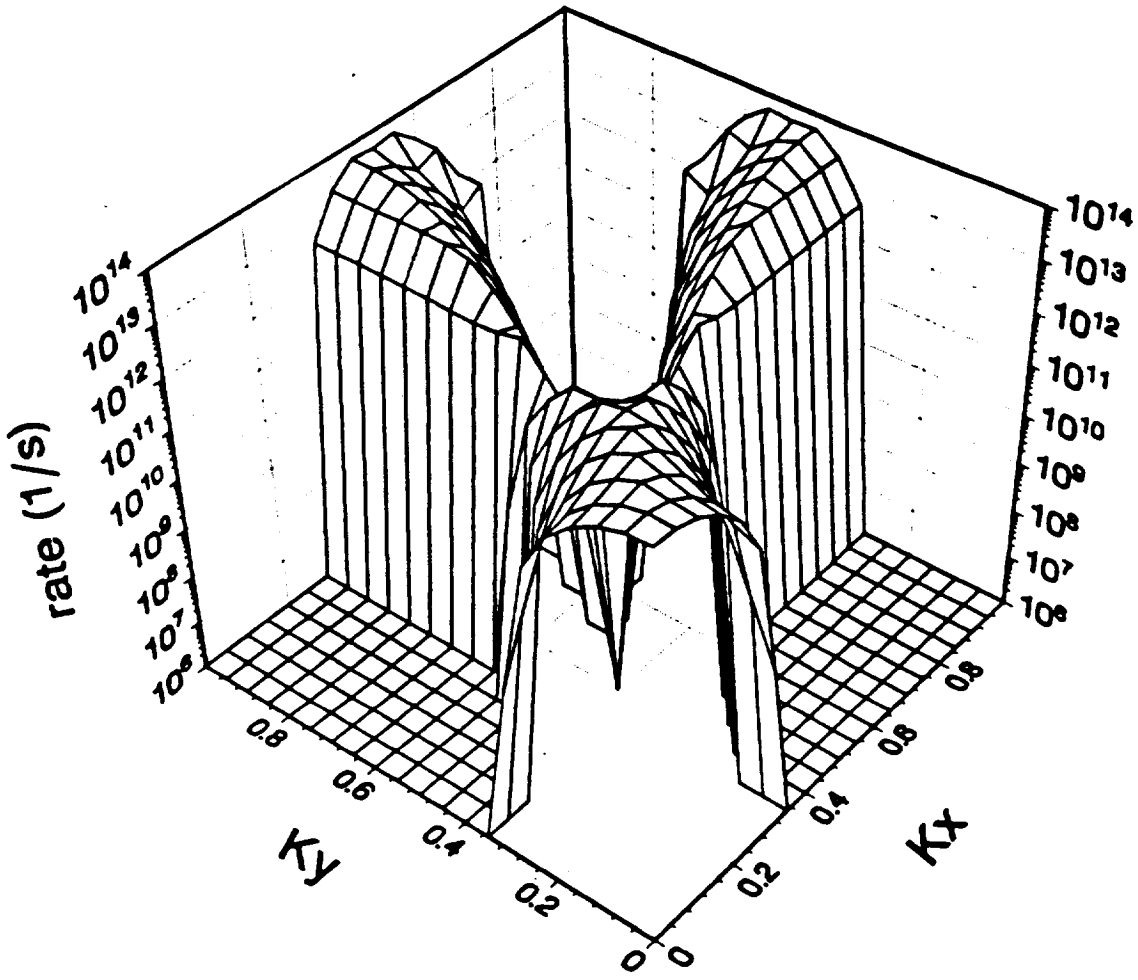


FIG 1B

KOLNIK et al

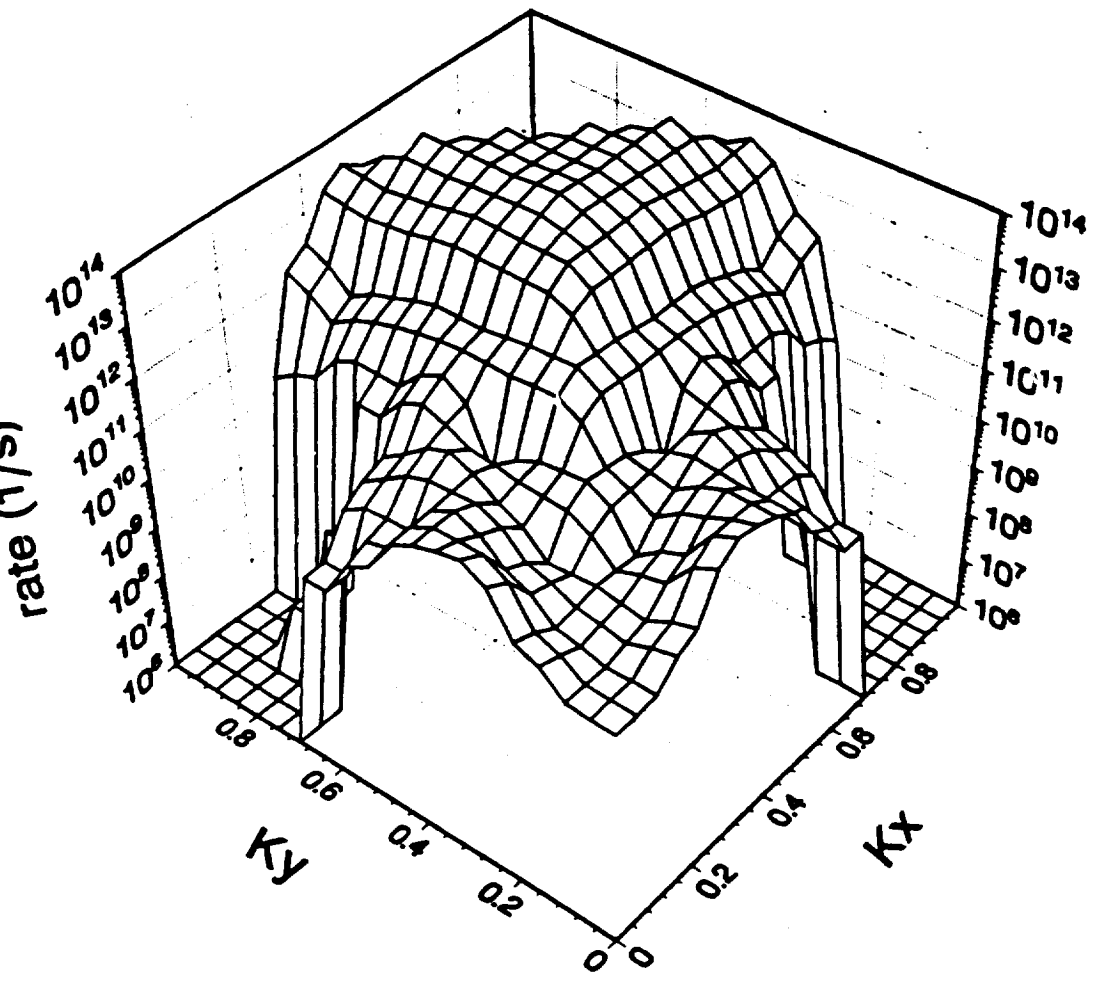


FIG 2A

KOLNIK et al

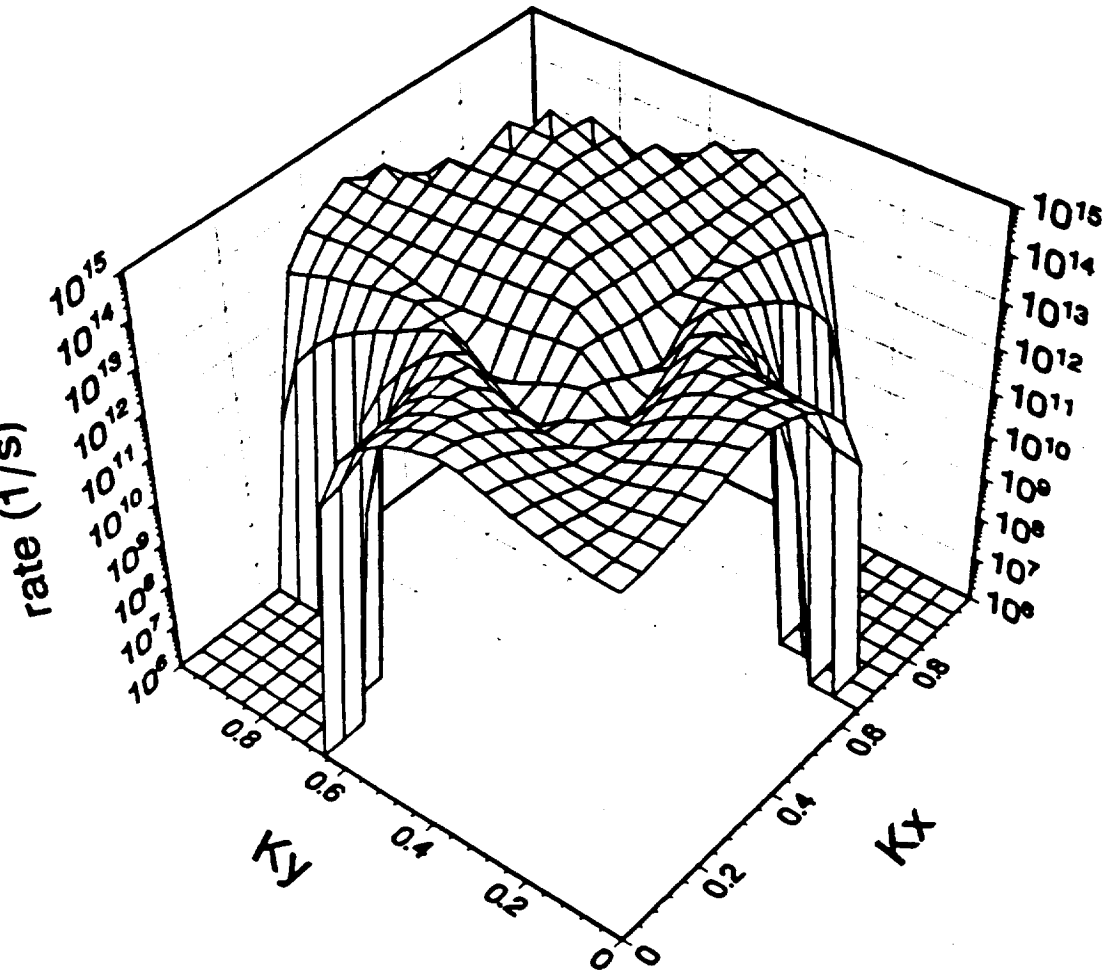


FIG 2B

KOLNIK et al

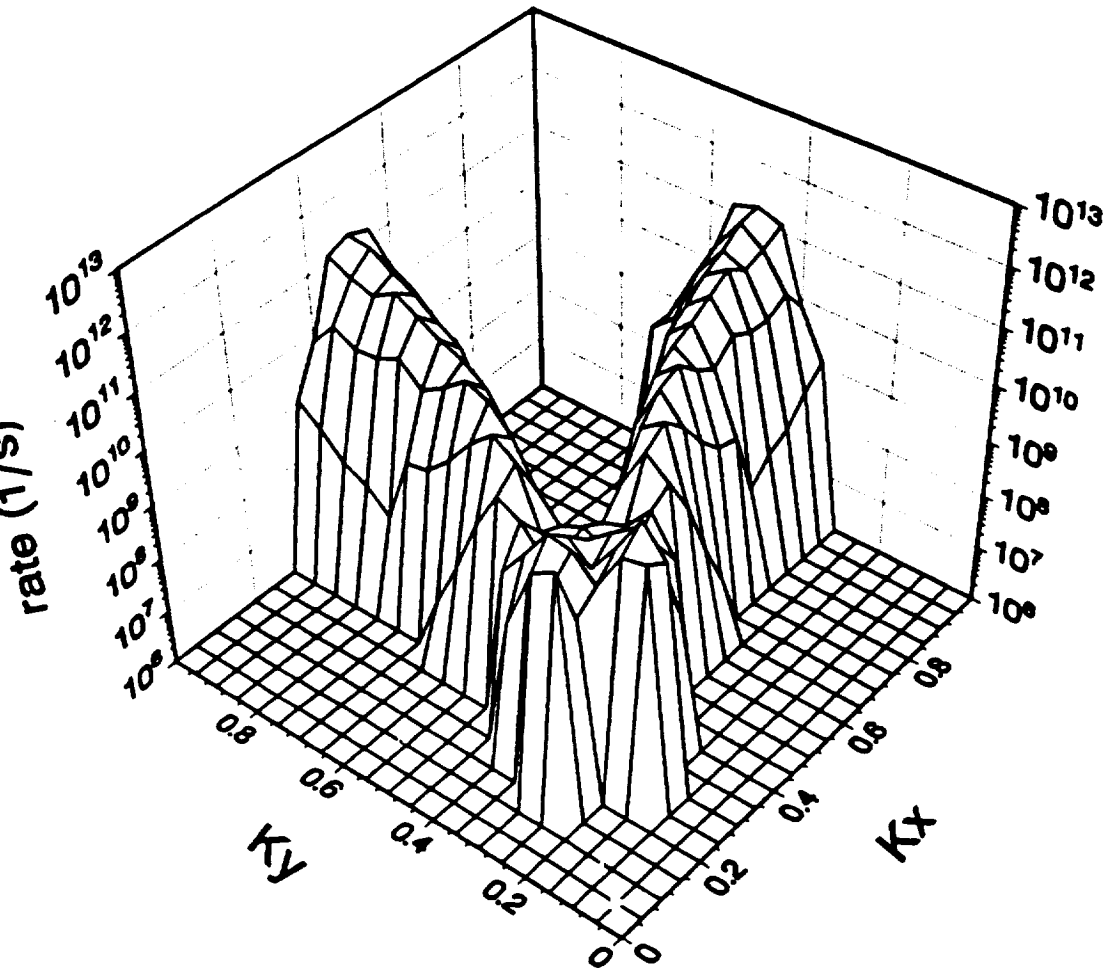


FIG 3A

KOLNIK et al

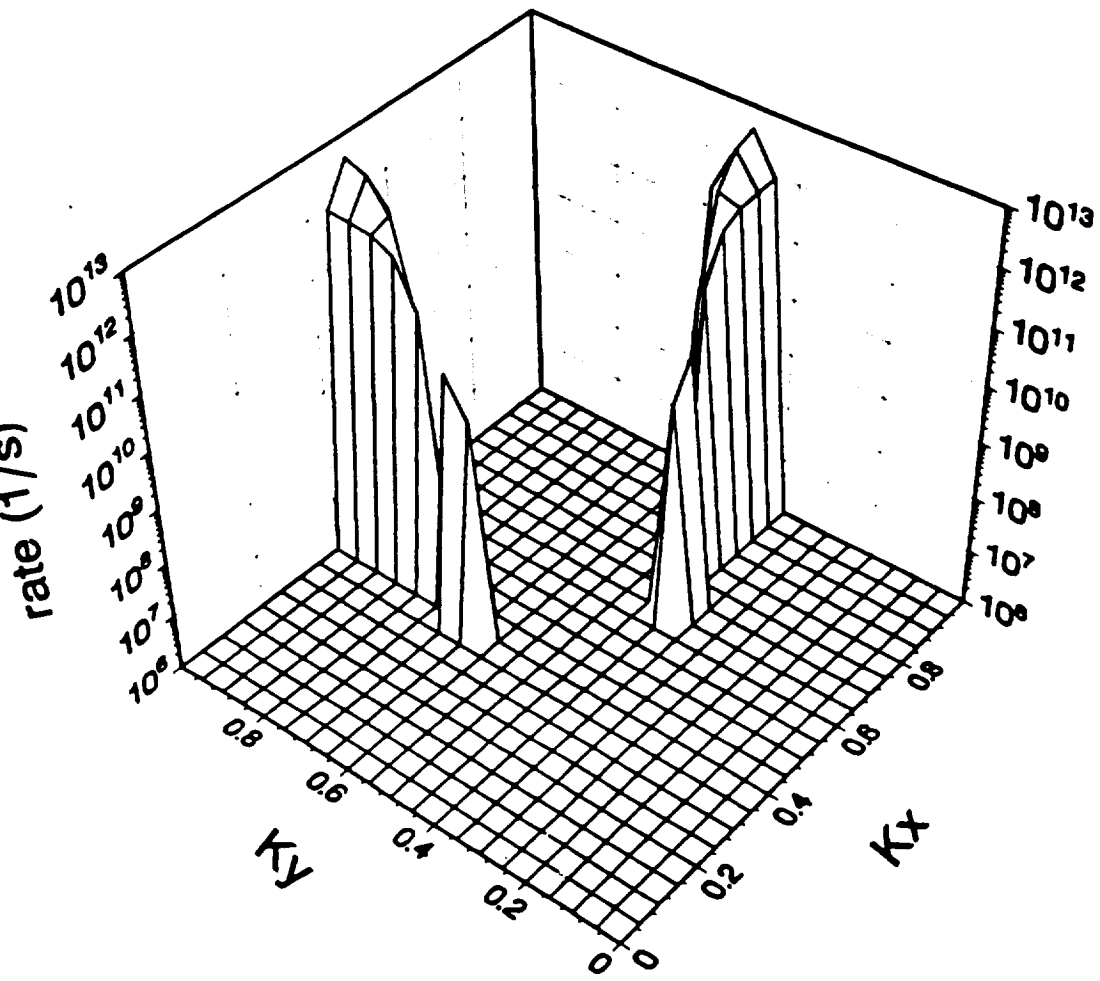


FIG 3B
KOLNIK et al

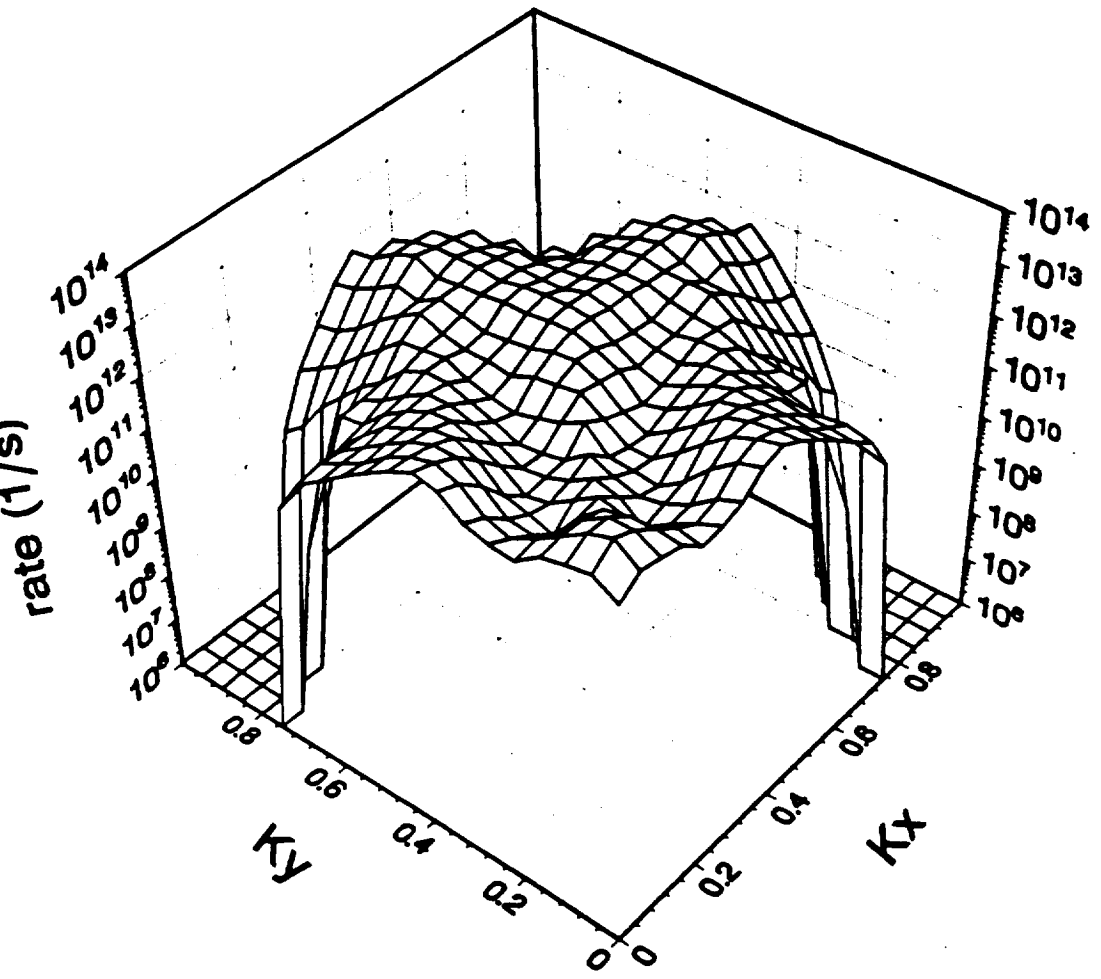


FIG. 4A

KOLNIK et al

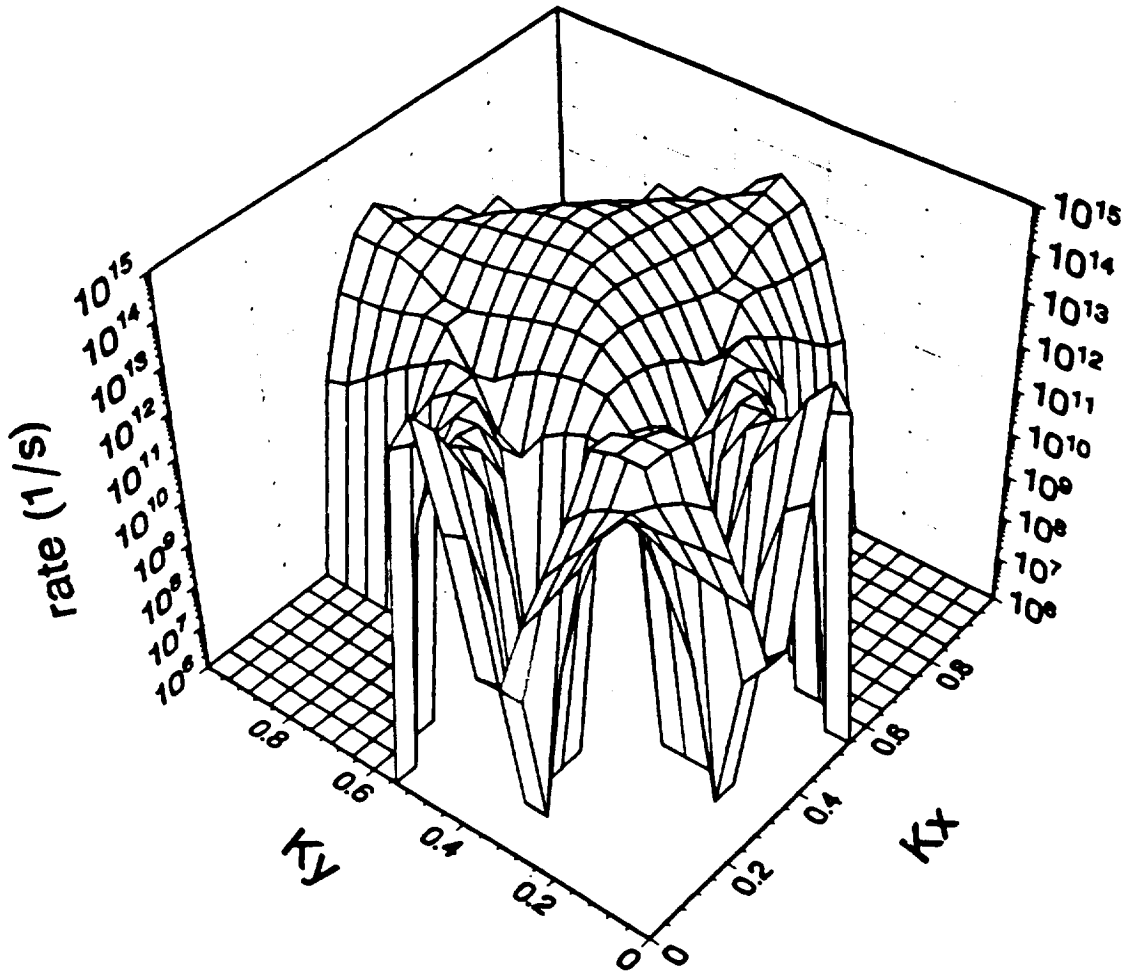


FIG 4B

KOLNIK 21 G.

FIG. 5

KOLNIK et al

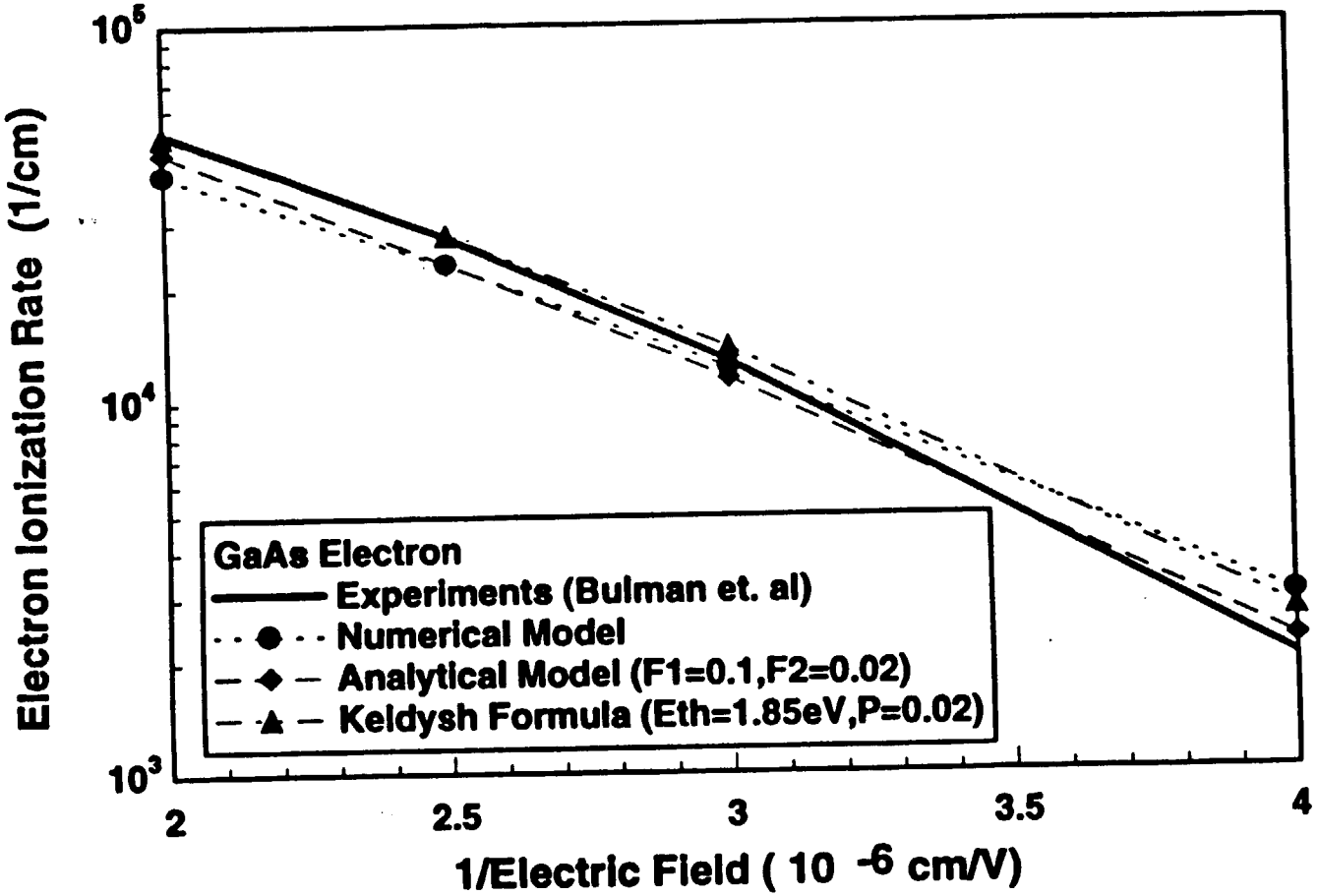
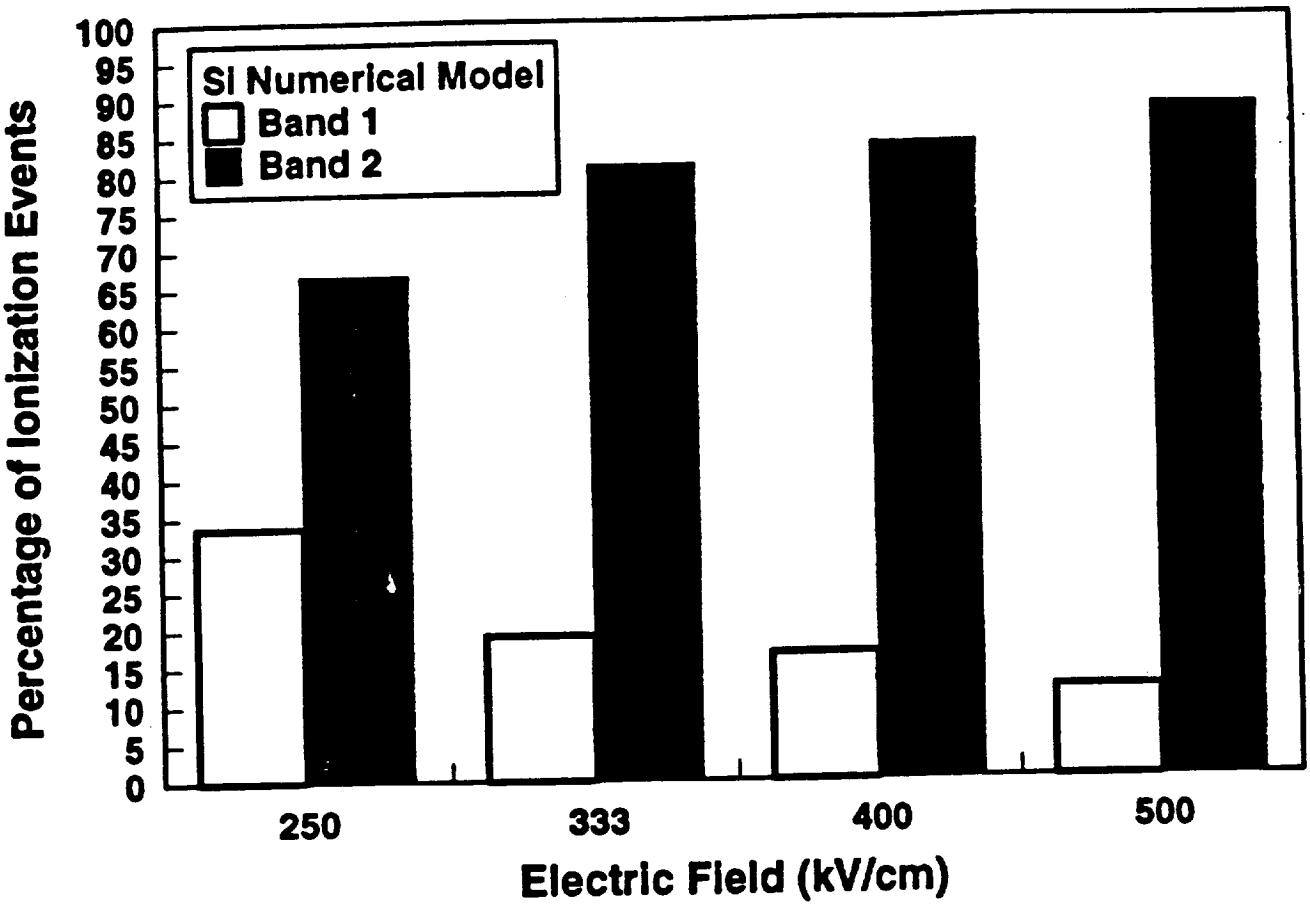
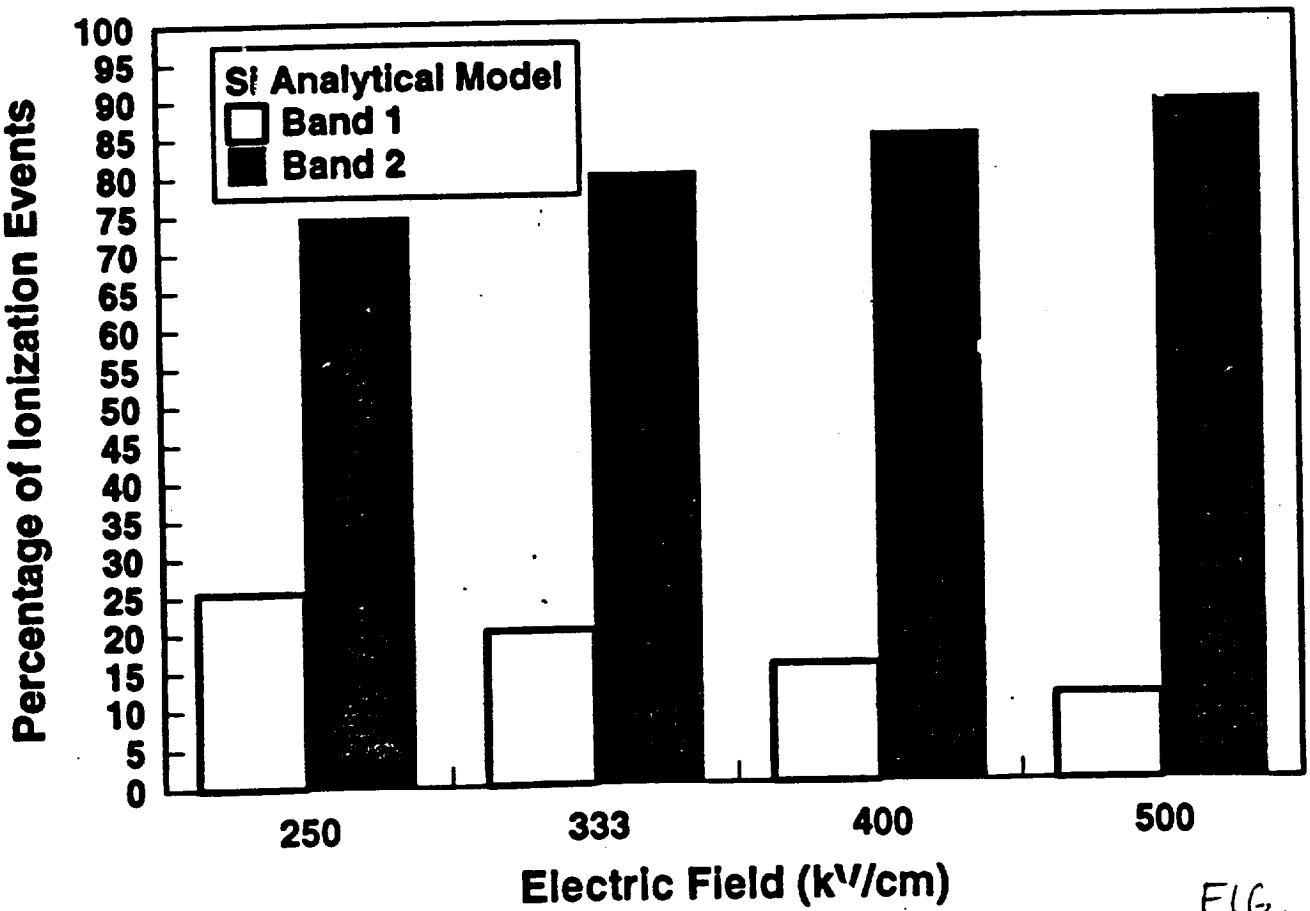


FIG. 6 : KOZNIK et al

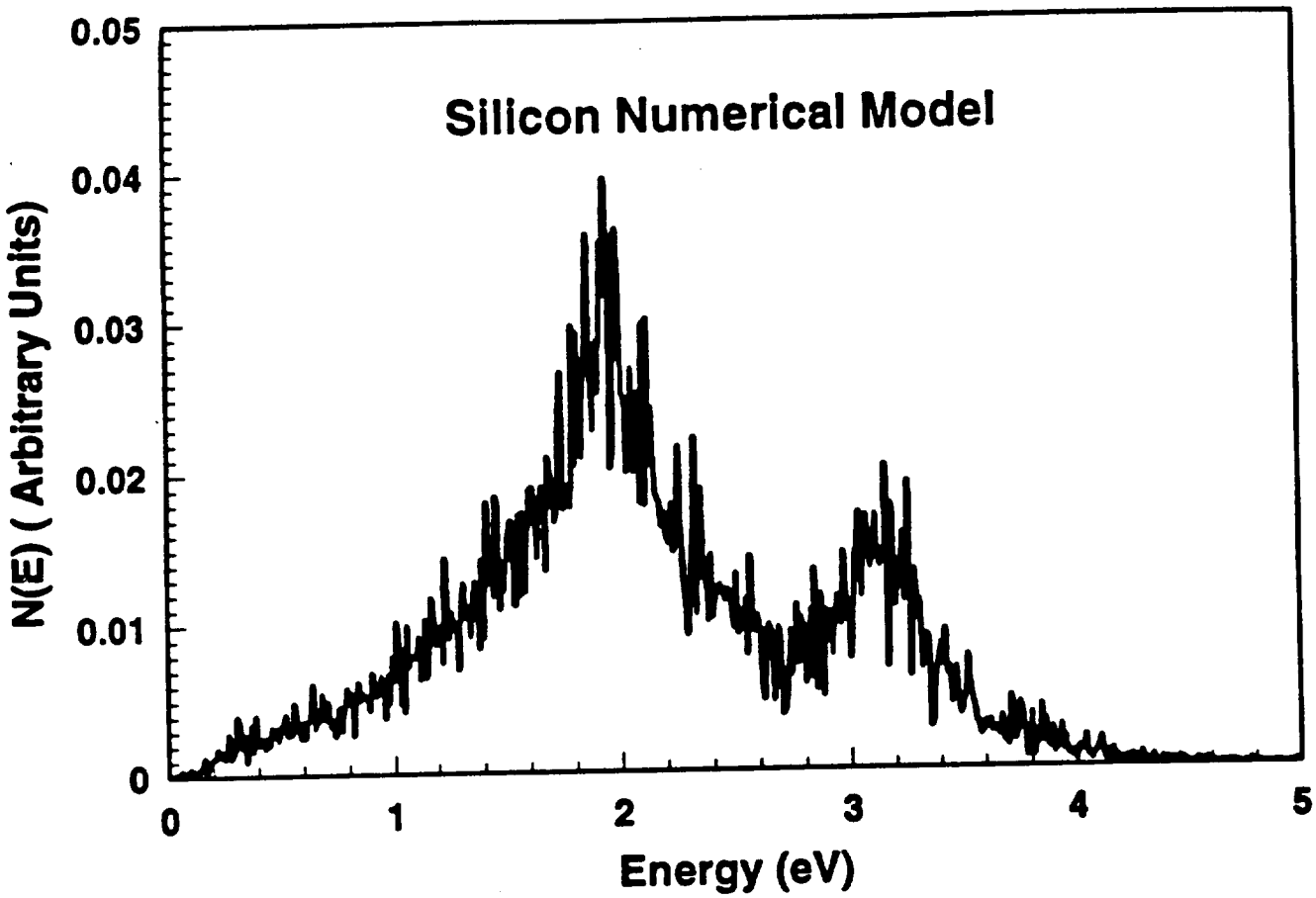

Si Numerical Model
Band 1
Band 2
Percentage of Ionization Events
100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
250
333
400
500
Electric Field (kV/cm)

(a)

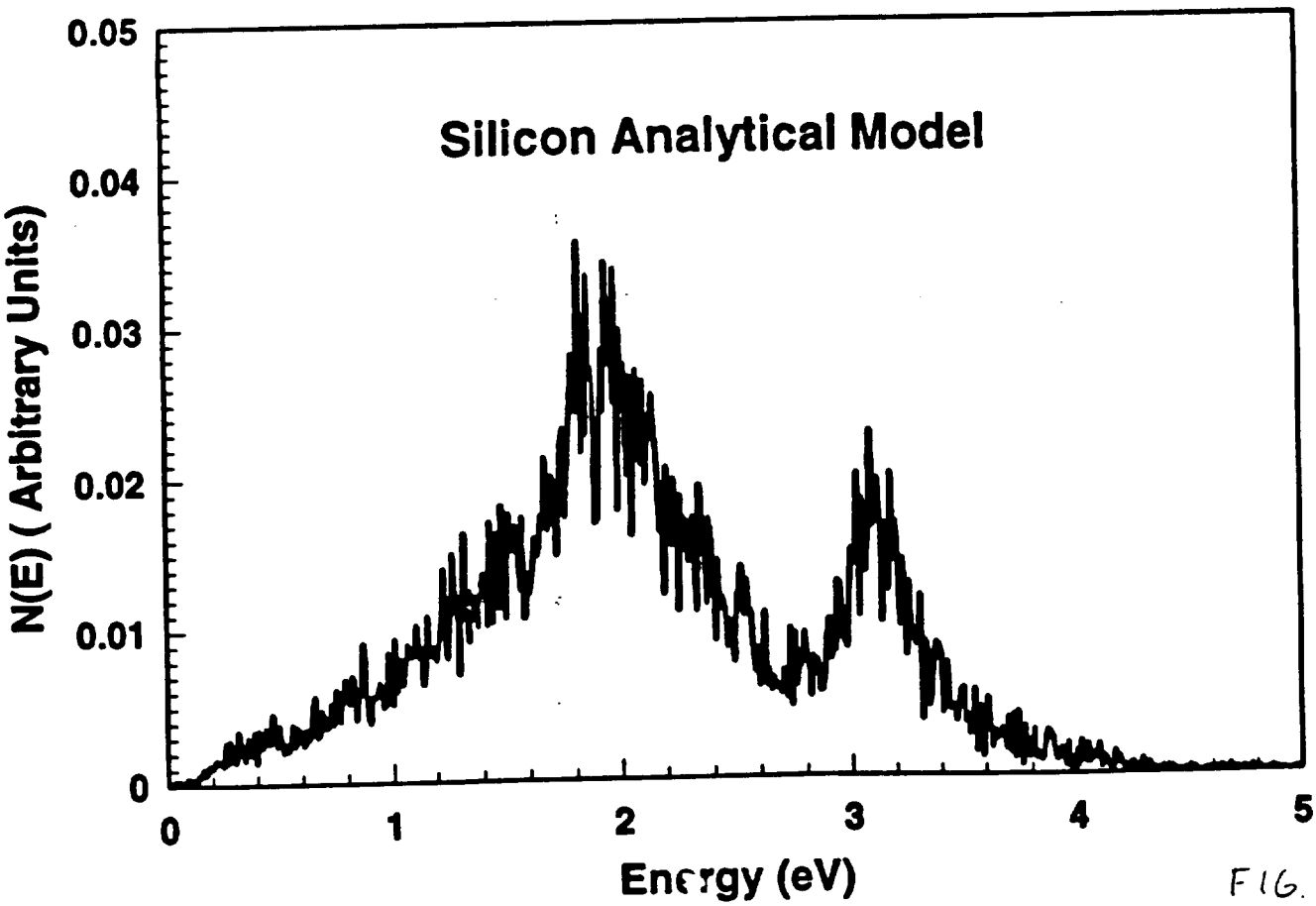

Si Analytical Model
Band 1
Band 2
Percentage of Ionization Events
100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
250
333
400
500
Electric Field (kV/cm)

(b)

FIG. 7

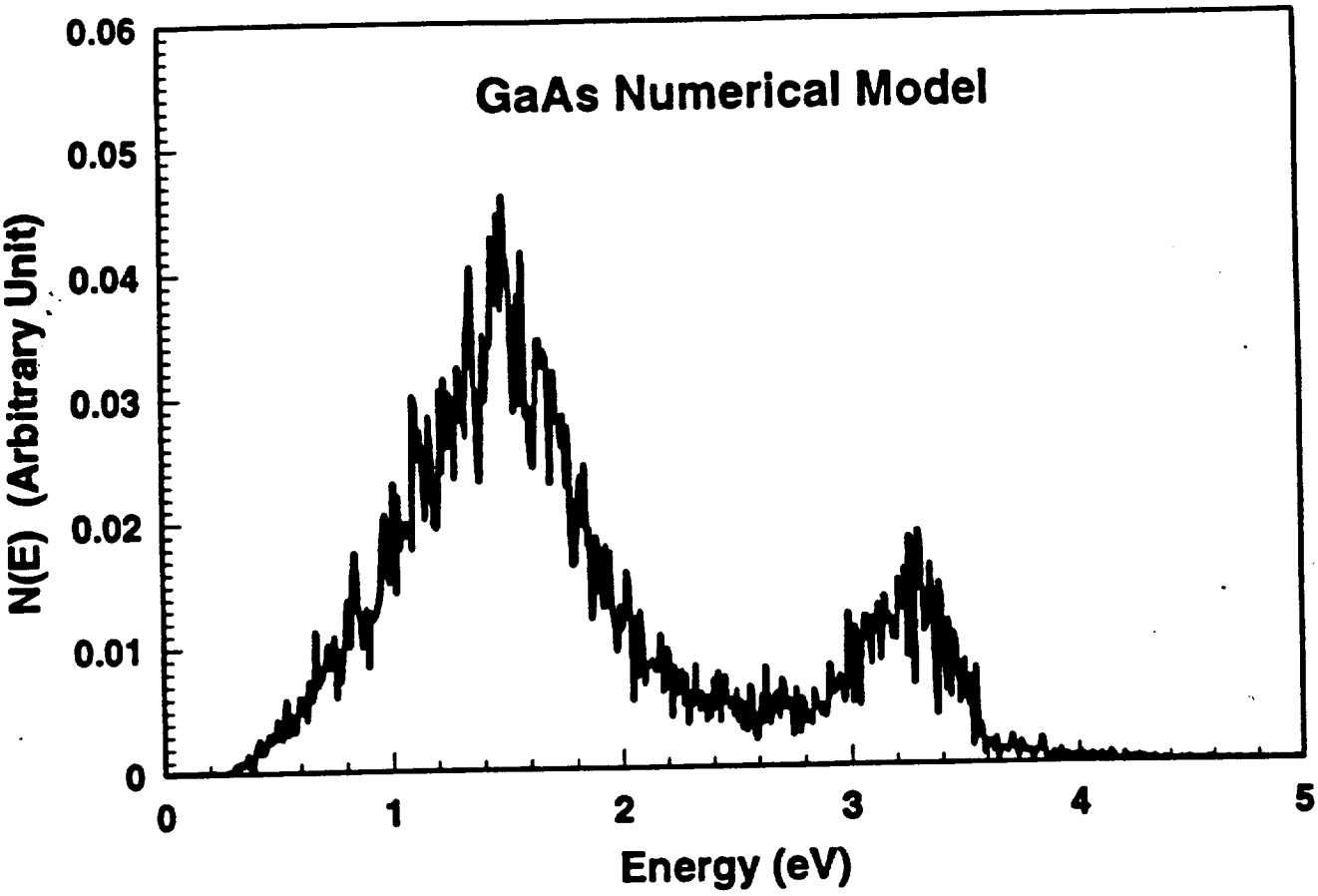

Silicon Numerical Model
N(E) (Arbitrary Units)
0.05
0.04
0.03
0.02
0.01
0
0
1
2
3
4
5
Energy (eV)

(a)

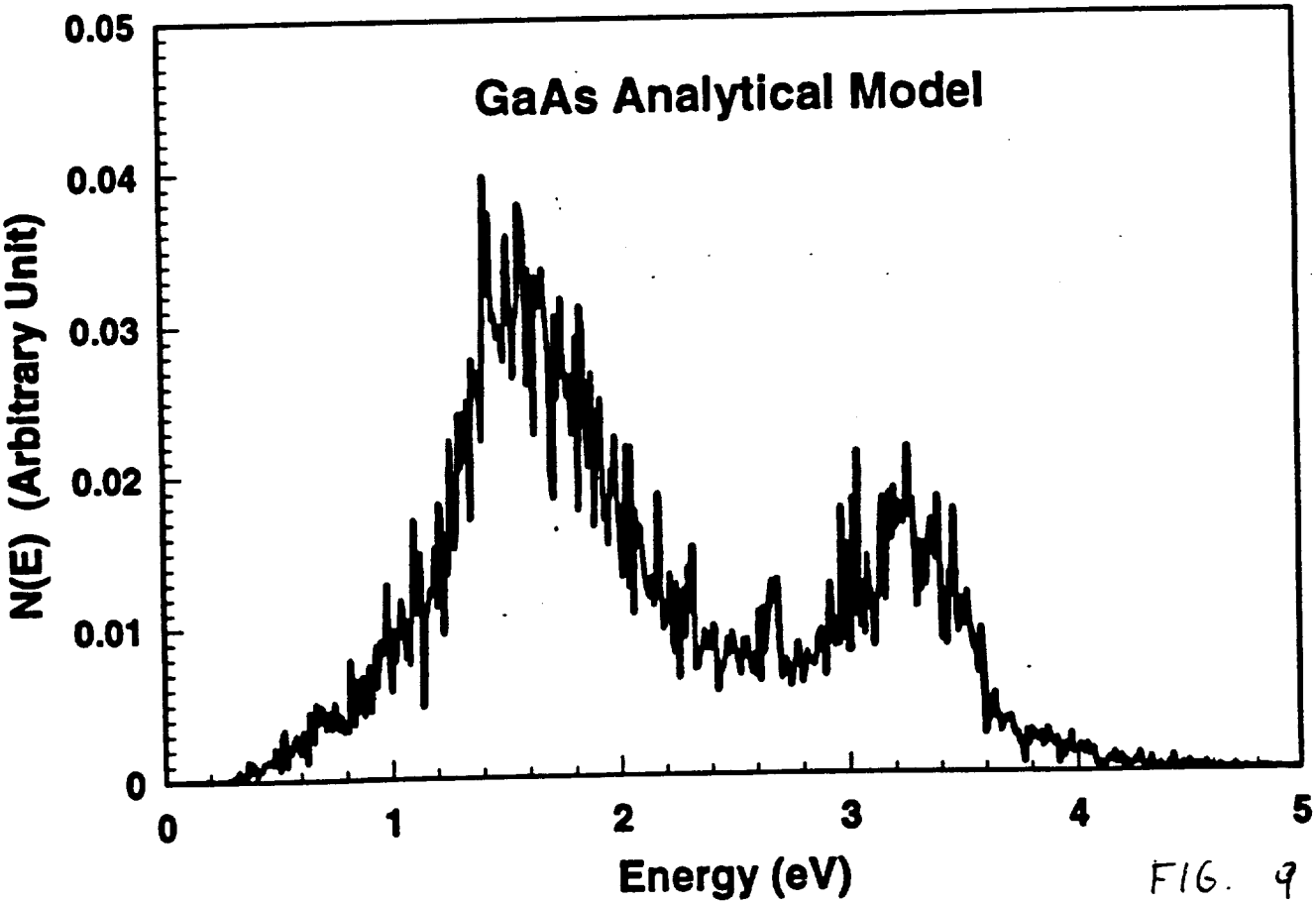

Silicon Analytical Model
N(E) (Arbitrary Units)
0.05
0.04
0.03
0.02
0.01
0
0
1
2
3
4
5
Energy (eV)

(b)

FIG. 8



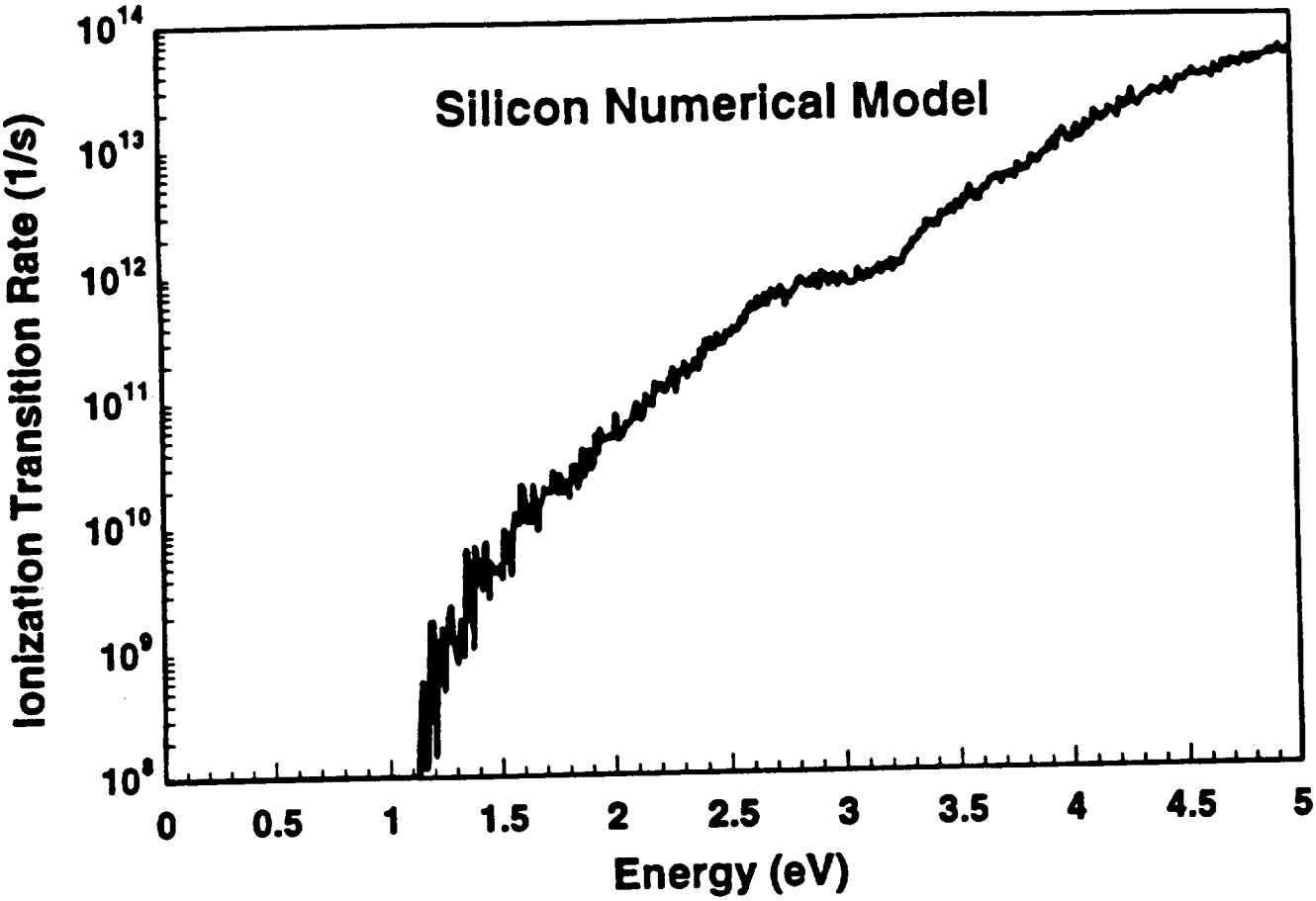
(a)



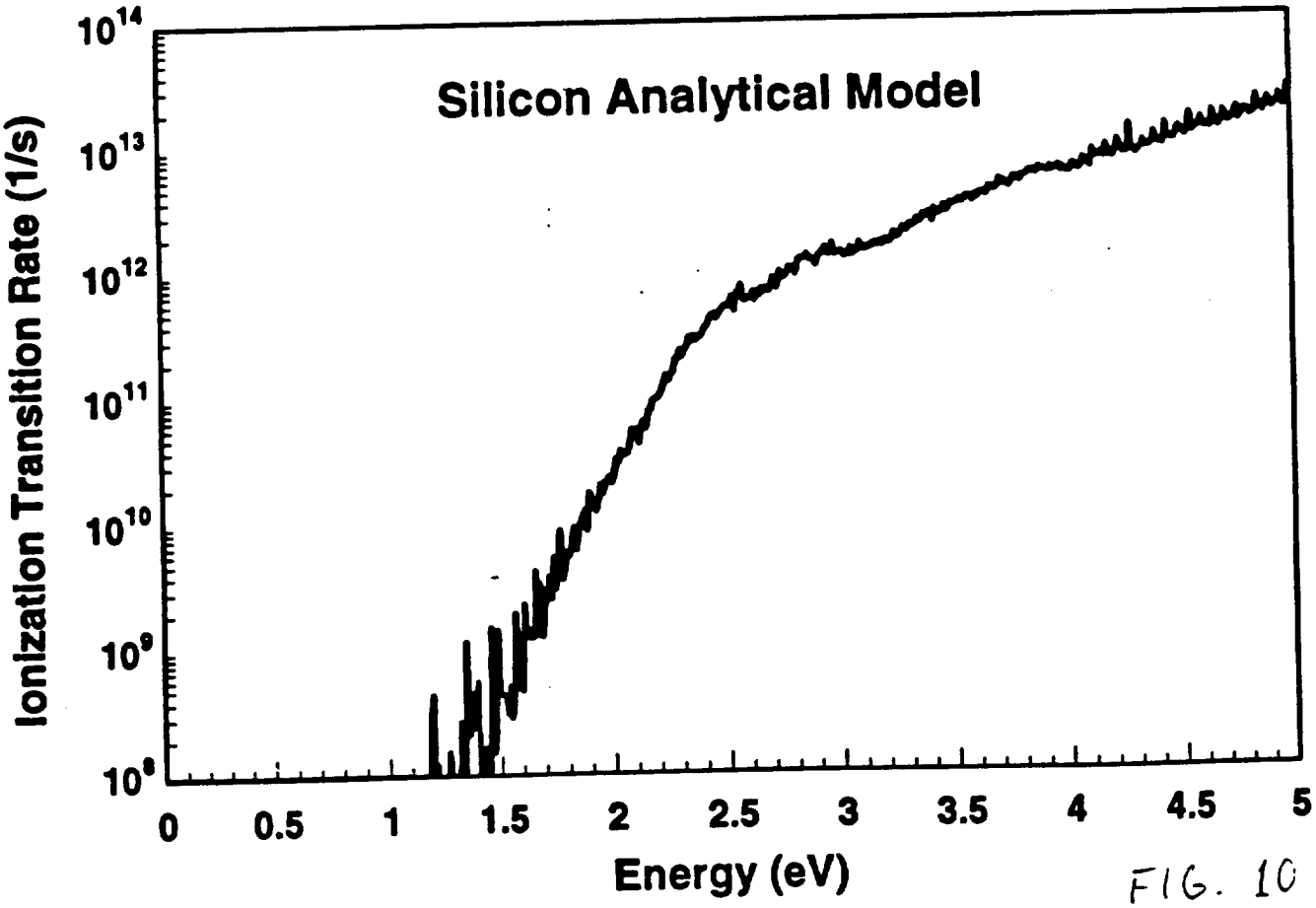
(b)

FIG. 9

KOLNIK 2122



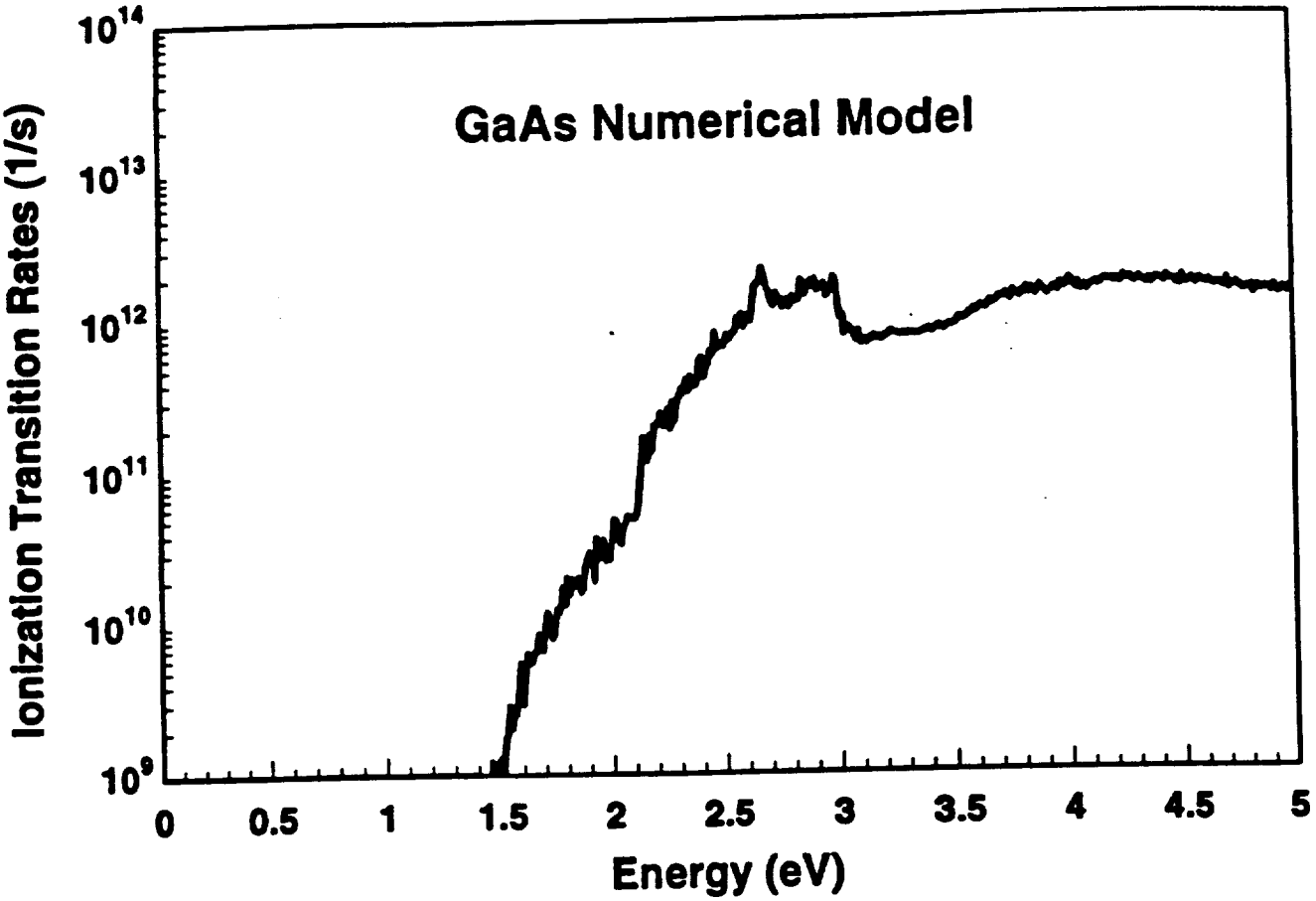
(a)



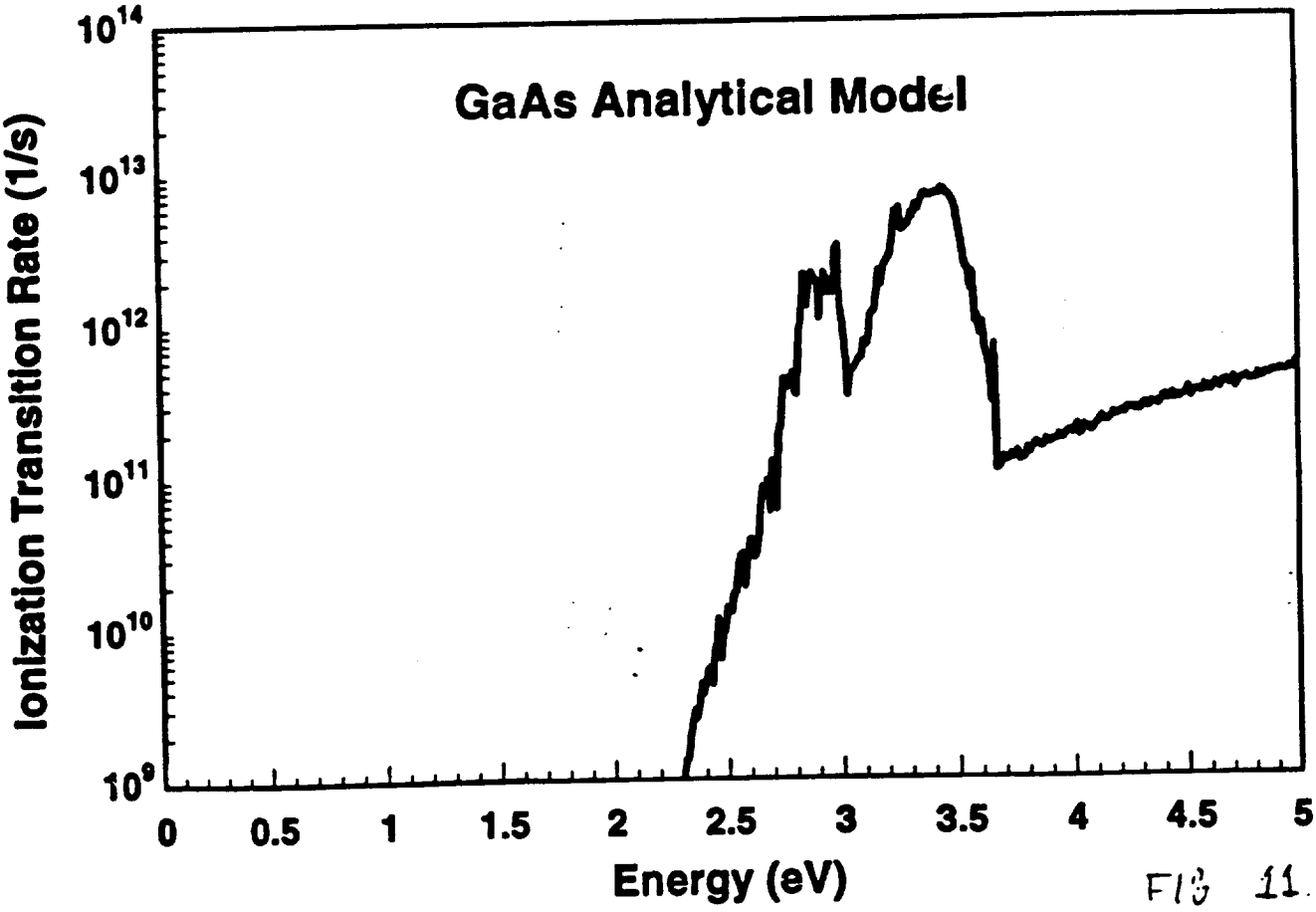
(b)

FIG. 10

KOLNIK 21 ai



(a)



(b)

FIG 11.

KOLNIK et al.