NASA CR-145091

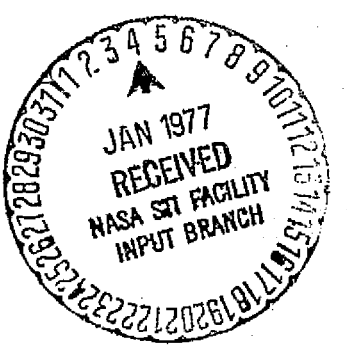
INVESTIGATION OF FIELD INDUCED TRAPPING ON FLOATING GATES

By

W. Milton Gosney, Principal Investigator
Walter T. Matzen and Gary E. Nelson

Prepared under Contract No. NAS1-13610
by
TEXAS INSTRUMENTS INCORPORATED
P. O. Box 5012
Dallas, Texas 75222

for



National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23665

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (continued)

LIST OF TABLES

SUMMARY

In a series of five experimental design iterations, an electrically-alterable read-only memory (EAROM) or reprogrammable read-only memory (RPROM) technology has been developed which can be fabricated using a single-level, metal-gate p-channel MOS process with all conventional manufacturing steps. Given the acronym DIFMOS for Dual Injector, Floating-gate MOS, this technology utilizes the floating-gate technique for nonvolatile storage of data. Avalanche injection of hot electrons through gate oxide from a special injector diode in each bit is used to charge the floating gates. From 30% to 40% of the electron injector avalanche voltage can be attained by the floating gate in about 50 ms for 500 μA avalanche currents. Once charged, the floating gates will stay charged almost indefinitely. Measurements of charge decay from the floating gates indicate a log time dependence of the loss of approximately 0.06% per decade of time. Thus, over 90% of the original charge will remain after 100 years at room temperature. A second injector structure included in each bit permits discharge of the floating gate by avalanche injection of holes through gate oxide. A special bootstrap capacitor must be operated with the hole injector to provide electric fields favorable for hole injection. Like the charging operation, the discharge operation is accomplished within 50 ms using 500 μA avalanche currents. The size of the bootstrap capacitor and its associated voltage must be sufficient to provide at least a 15-volt change on the floating gate to assure complete discharge. Thus, the overall design of the DIFMOS bit is dictated by the physical considerations required for each of the avalanche injector types.

The presence of electron and hole traps in the injector oxides causes a slow build-up of charge in the injector oxides as the injectors are operated, resulting in a decreasing efficiency of injection. This means that DIFMOS bits can only be charged and discharged a finite number of times—at present, the limit is about 10^4 cycles.

But this limitation is sufficiently large to permit application of the DIFMOS technology to electrically alterable, read-mostly memories. Since it is basically a PMOS plus three diffusion technology, and since its cell size is in the 0 to 16 nm^2 to 16 nm^2 size range, circuits of 1K to 2K bits should be economically feasible. The ease with which DIFMOS can be used in system designs makes it attractive for many applications such as calculator and microprocessor program storage. Standard processing, electrically erasable by row or by bit, fully decodable without substrate isolation, unlimited reading capability, and a variety of possible organizations tend to offset the programming limitations. Additional reductions in cell size and improvements in device performance and circuit density will be realized as the floating gate, avalanche-injected, nonvolatile semiconductor memories continue to evolve.

INVESTIGATION OF FIELD INDUCED TRAPPING ON FLOATING GATES

W. Milton Gosney, Walter T. Matzen, and Gary E. Nelson

SECTION I

INTRODUCTION

This report outlines the development of a technology for building electrically-alterable read-only memories (EAROMs) or reprogrammable read-only memories (RPROMs) using a single-level metal-gate p-channel MOS process with all conventional processing steps. Nonvolatile storage of data is achieved by the use of charged floating gate electrodes. The floating gates are charged by avalanche injection of hot electrodes through gate oxide, and discharged by avalanche injection of hot holes through gate oxide. Three extra diffusion and patterning steps are all that is required to convert a standard p-channel MOS process into a nonvolatile memory process. For identification, this nonvolatile memory technology has been given the descriptive acronym DIFMOS which stands for Dual Injector, Floating gate MOS.

DIFMOS devices are similar to FAMOS[1] devices in that both device types feature writing by avalanche injection of hot electrons through gate oxide; and both feature data storage using floating MOS gates. Thus, all the storage and reliability characteristics of FAMOS are applicable to DIFMOS. But FAMOS devices cannot be electrically erased, whereas DIFMOS devices can be erased. The ability to erase electrically makes DIFMOS an attractive choice for applications such as program storage in microprocessors where nonvolatility is important, and where ease of program modification is desired.

DIFMOS devices are also similar to MNOS[2] devices in that both device types can be electrically written and erased. But there are major fundamental differences in physical structure and operation. The MNOS devices require a tunneling oxide with special characteristics to control the nonvolatile storage and programming specifications. The present few MNOS products on the market require an extremely complicated process consisting of a basic PMOS process on epitaxial substrate, two extra diffusion and patterning steps, growth of the tunneling oxide, and nitride dielectric deposition. On the other hand, DIFMOS requires only three extra diffusion and patterning steps to be added to the basic PMOS process. The excellent nonvolatile behavior of the DIFMOS devices, together with their more conventional processing and their operational simplicity, gives DIFMOS a significant edge over MNOS for most circuit applications.

The original concept for the DIFMOS device, as discussed in the First Experimental Series, called for the injection of electrons from a p^+/n^+ junction into oxide for writing, and the injection of holes from a p^+/n^+ junction into nitride for erasing. Discrete devices to test this concept were fabricated, and these devices were functional. However, the erase mechanism was found to be thermal in nature, rather than by hole injection. The final device configuration was different from the original concept. The electron injector consists of a p^{++}/n^+ junction, and the hole injector is p^+/n^{++}. Both the hole and electron injector avalanche their carriers into gate oxide. In addition to the injectors, a special bootstrap capacitor is required to provide the electrical fields favorable for hole injection. The injector designs and the bootstrap capacitor differ in detail, but not in concept, from the original DIFMOS idea which called for: (1) a floating gate for storage, (2) an MOS transistor for sensing, (3) an electron injector for writing, and (4) a hole injector for erasing.

The original contract for DIFMOS development specified three series of experiments. First, an initial device type was to be studied. Then, information from this prototype device was to be incorporated into a series of devices on a special test bar. After these devices were tested, a third design was to be built, based upon information from the second test bar. The original contract has been modified twice by the addition of add-on work packages. The first add-on package called for special series of device tests on the properties of silicon nitride, to be completed prior to the final device design series of the original contract. The second add-on called for another design iteration to give further improvements in the device characteristics. The experimental series can be summarized as:

First experimental series	Preliminary study, prototype results	(original contract)
Second experimental series	Test bar design (10 structures)	(original contract)
Third experimental series	Special nitride evaluation	(first add-on)
Fourth experimental series	Structure redesign (2 structures)	(original contract)
Fifth experimental series	Structure redesign (5 structures)	(second add-on)

Design iterations of this nature are a time-consuming but necessary portion of the evolution of electronic technology. The rate of progress is limited by the design cycle time of circuit layout, photomask generation, slice fabrication, and testing. The five experimental series of devices developed with this contract support have provided vehicles for the achievement of steady progress in the development of a successful product technology from the original device concept.

SECTION II

FIRST EXPERIMENTAL SERIES

A. DIFMOS DEVELOPMENT

The Intel FAMOS device was the first commercially available MOS integrated circuit to utilize charge stored on a floating gate electrode as a nonvolatile data storage technique for making programmable read-only memories (PROMs). However, the FAMOS device could not be electrically erased, and this disadvantage prevented its use in electrically alterable ROMs (EROMs) which can be electrically programmed and reprogrammed at will.

The concept of building a FAMOS type device with electrical erase intrigued us. Since the FAMOS device was written by injecting hot electrons onto the gate electrode through the gate oxide from an avalanche plasma in a p-n junction biased into breakdown, it seemed obvious that the charge could be removed by injecting hot holes in a similar manner from a special hole injector structure included with each bit. All that would be required would be to find a structure to inject the hot holes which could be included in the memory array. The boundary conditions for such a structure were: (1) that its processing steps be compatible with the overall array process, and (2) that its electrical characteristics be similar to the electron injector such that the hole injector could be addressed and switched by the array decode circuitry.

It was suggested by Tarui et al.[3] that hole injection could be obtained by avalanching an n^+/p junction, similar to Frohman-Bentchkowsky's electron injection from p^+/n junctions. In 1971, a series of experiments conducted at Texas Instruments studying avalanche injection into oxide from p^+/n and n^+/p junctions on a CMOS test chip indicated that both junction types injected electrons into the oxide. This result seemed to contradict Tarui's supposition that there was asymmetry in the injection characteristics of n^+/p and p^+/n junctions. Furthermore, Nicollian et al.[4] showed that electron currents injected into thermal SiO_2 were more than three orders of magnitude greater than hole currents under equivalent field but complementary semiconductor conditions. Snow et al.[5] have found that electrons travel rather freely through thermal oxide while holes are quickly trapped. Their measurements indicate a hole trap density of 10^{18} cm^{-3}. Williams[6] found the density of electron traps in thermal oxide to be 3×10^{14} cm^{-3}, which is more than three orders of magnitude less than the density of hole traps. Curiously, this ratio is nearly the same as the ratio between avalanche-injected hole and electron currents measured by Nicollian et al.[4] It

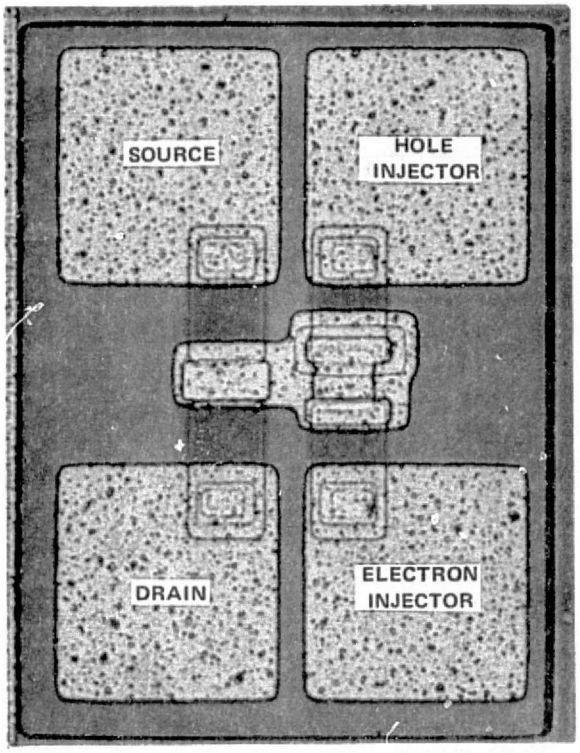
would appear the nature of the traps and energy bands in thermal oxide is such to favor the avalanche injection of electrons having sufficient energy to charge a floating gate electrode. Operation of the FAMOS device is direct experimental proof of this statement.

On the other hand, Wallmark and Scott[2] report electron trap densities of 2×10^{18} cm^{-3} in silicon nitride film. (No value was reported for the hole trap density.) Thus, there is some evidence that in thermal oxide, electrons are mobile and holes are trapped, and that the opposite may prevail in nitride.[7] It was believed that the key to achieving selective avalanche injection of electrons and holes to a floating gate was to provide for injection of holes through nitride, and electrons through oxide with only a single junction type (n^+/p, p^+/n, or p^+/n^+).

Figure 1 shows an avalanche-injection nonvolatile memory structure that features p^+/n^+ low avalanche voltage junctions for charge injectors. The avalanche voltage can theoretically be made as low as 6 to 7 volts before tunneling dominates the reverse-bias characteristics. A more typical value would be in the 12-15-volt range. Voltages in this range can be switched by MOS address transistors. Notice that the hole injector features silicon nitride with no oxide over its avalanche region, while the electron injector features gate oxide between the floating gate structure and the avalanche region. Sensing the charge on the floating gate is accomplished by a conventional MOS transistor whose gate is part of the floating gate which overlaps the hole and electron injector regions. If the floating metal gate is not covered by a deposited silicon oxide or nitride overcoat, the gate can be probed directly to measure the avalanche currents at fixed bias conditions. Figure 2 shows detail of the hole and electron injector regions. The process for fabricating the dual avalanche injector structure shown in Figure 1 is basically a PMOS process having two extra masking steps, an extra n^+ diffusion, and a nitride deposition steps over the conventional low-V_{tx} [100] PMOS process.

In addition to featuring the same p^+/n^+ low avalanche breakdown junctions, both the hole and electron injectors operate from the same polarity. It is expected that the hole injection erasure will work, even if the trap distribution in the nitride is such that both holes and electrons are approximately equally passed by the nitride, because the drift field from electrons already captured will favor the injection of holes and cause the discharge of the floating gate.

A number of lots of the discrete nonvolatile memory device were processed. Nonvolatile memory operation was demonstrated; however, avalanche currents larger than expected were required for writing and erasing. Most experimental devices required erase pulses of 25 V at 80 mA to operate the hole injectors, even though the junction BVs were in the 12 to 14 range. Because of the power required to erase, it appeared that the erasable discrete devices were being erased by a thermal effect, rather than an injection of holes effect. Also, such large programming current pulses could not be switched by minimum-geometry on-chip MOS address transistors; hence, such memory devices could not be addressed within an array organization.



SCALE: BOND PADS ARE
50 μm X 50 μm

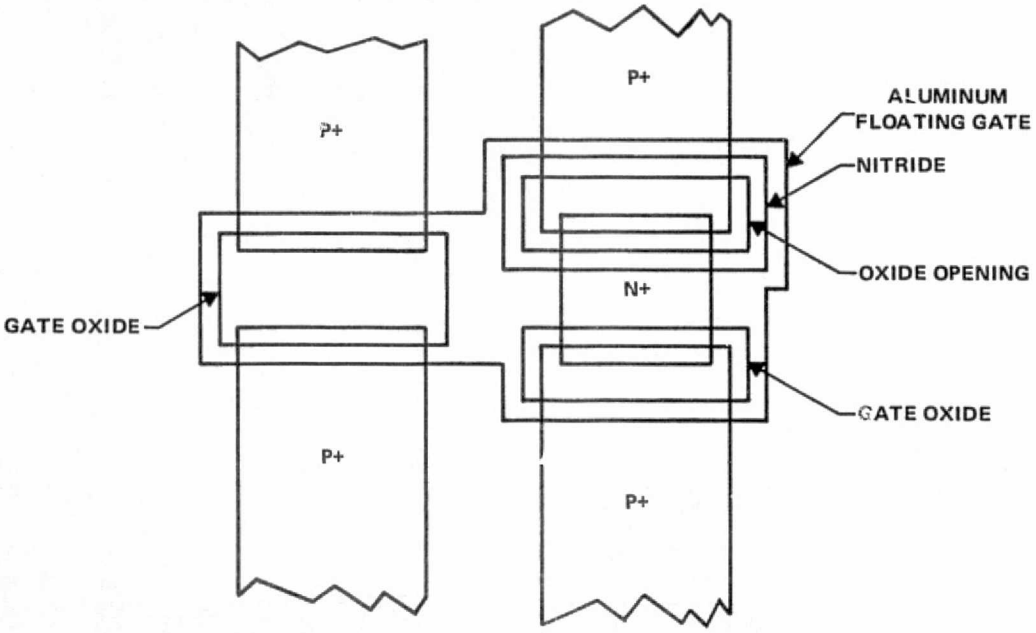


Figure 1. Dual Avalanche Injection Nonvolatile Semiconductor Memory Device Featuring Low-Voltage Hole and Electron Avalanche Injection to a Floating Gate

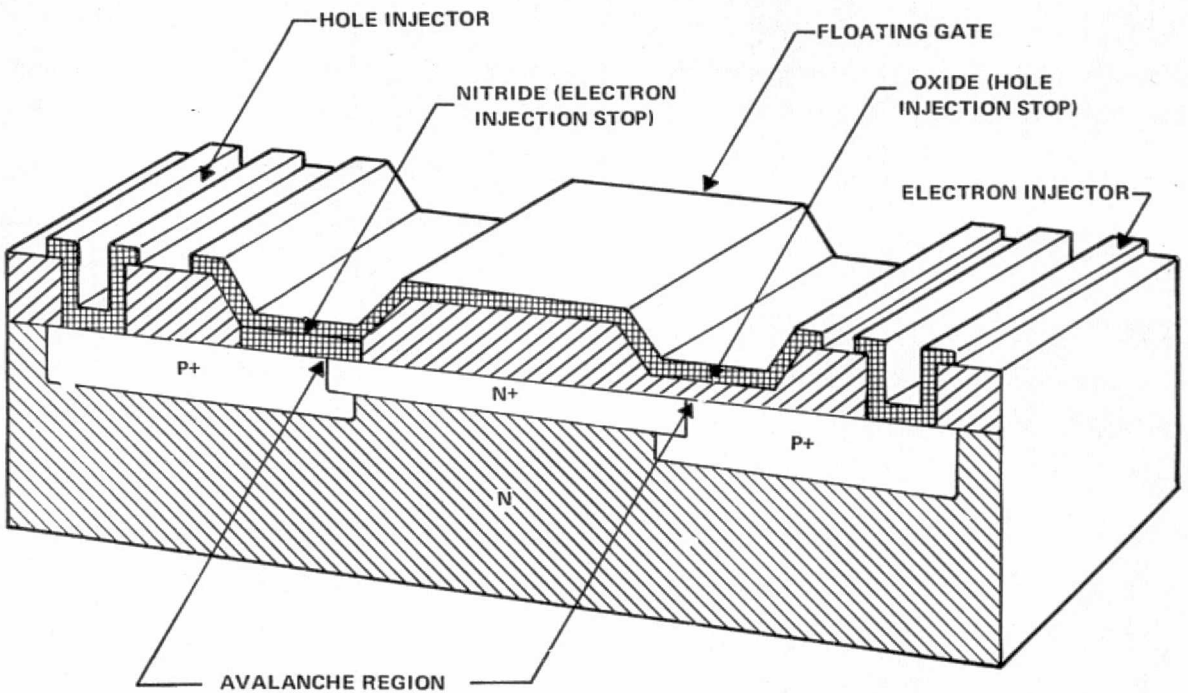


Figure 2. Isometric Drawing of a Cross Section
Showing Details of the Hole and Electron Injection Regions

SECTION III

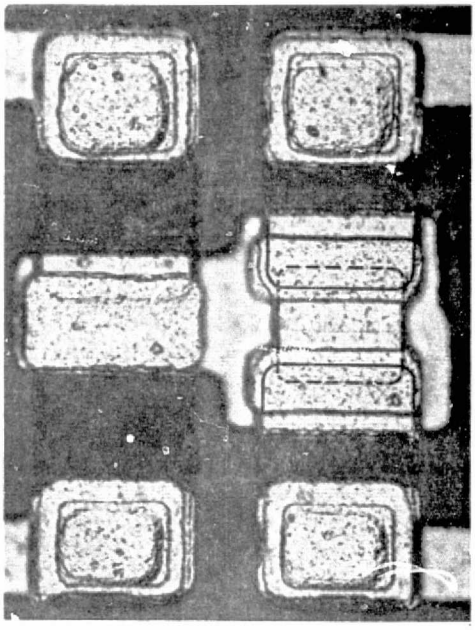
SECOND EXPERIMENTAL SERIES

A. DIFMOS DEVELOPMENT

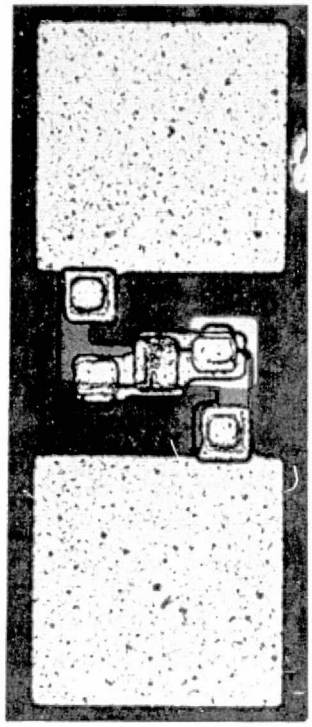
The results obtained with the first discrete DIFMOS nonvolatile memory cell were at least partially encouraging. The ability to write, store, and erase charge on the floating gate was demonstrated. However, the erase avalanche requirements of over 20 mA current precluded implementation into integrated arrays because of the physical size of the address transistors required to switch this amount of current. It was believed, and later more conclusively shown, that the erase mechanism in this device was thermal in nature.

But there was also the chance that hole injection through nitride might still work. Hence, we designed a series of test structures for hole and electron injection into nitride and oxide. We also included in these structure/designs some junction types which could be used to verify the injection asymmetry associated with junction type as reported by Tarui et al.[3] As shown in Figure 3, many of these test structures contained multiple variations of the injector designs, so that the ten design variations required by the contract actually consisted of many more possible design variations. Nine of the ten designs for the nitride hole injector studies were integrated on a single special chip. These slices were to be processed according to Table I. The remaining design was representative of typical p- and n-channel MOS transistors found on a special CMOS test bar wherein source/drain spacing and diffusion/channel spacing were test parameters on both n- and p-channel devices as listed in Table II.

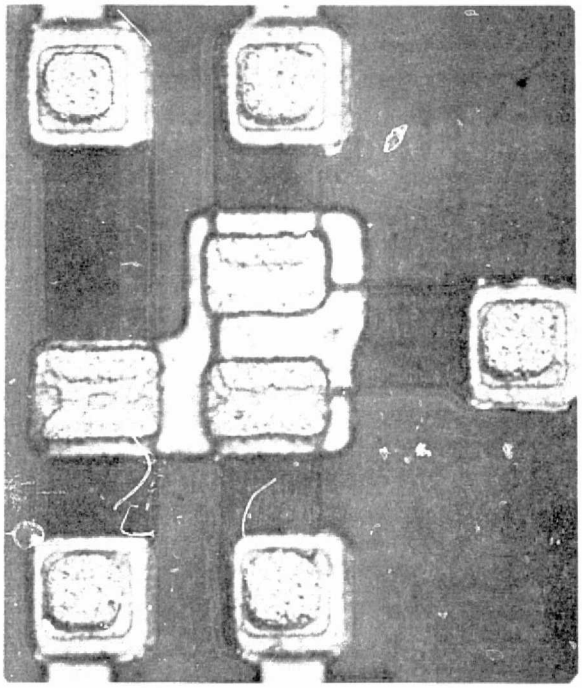
Like the first series of experiments with the discrete DIFMOS device, only limited success was achieved with the first nine of the device designs in Figure 3. A wide variety of writing performance was obtained using these various electron injectors, with the $p^+/n/n^+$ junction, hereafter called the gap injector because of the n-type gap between p^+ and n^+ regions, being the best injector. These so-called gap injectors were included in structures 7 and 8, where the gap is readily visible as shown in Figure 3. Gap widths of 2.5, 5.1, and 7.6 μm were designed into the mask set, with normal misalignment providing grading between these values. The performance of the gap injector can be summarized as follows: avalanching the electron injector tends to charge the floating gate voltage to a value no more than approximately one-half the value of the avalanche voltage. Figure 4 shows floating gate voltages measured on a number of gap injectors as a function of the avalanche voltage. The dotted line is the locus of $V_g = 1/2\ V_A$, and this line forms the least upper bound of the data points.



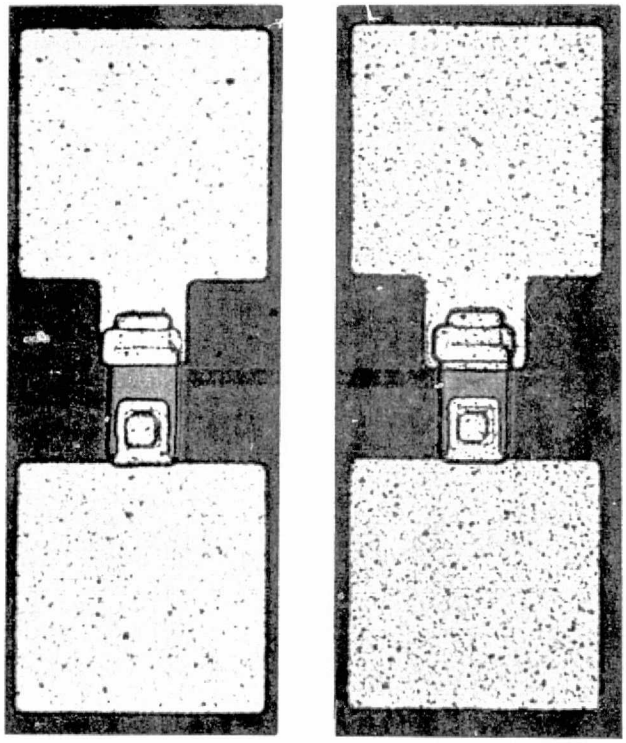
STRUCTURE #1
DIF BASIC FOUR TERMINAL DUAL INJECTOR.
FLOATING GATE MOS STRUCTURE. SCALE: 720X.



STRUCTURE #2
DIF ϕ2 TWO-TERMINAL ADAPTATION OF BASIC
INJECTORS AND SENSE DEVICE. SCALE: 320X.

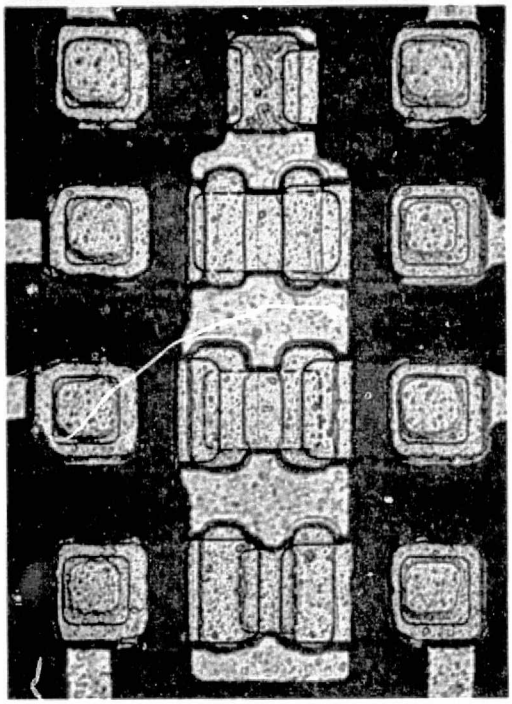


STRUCTURE #3
DIF ϕ4, FIVE-TERMINAL DIF MOS DEVICE WITH
PUNCH-THROUGH INJECTORS. SCALE: 720X.

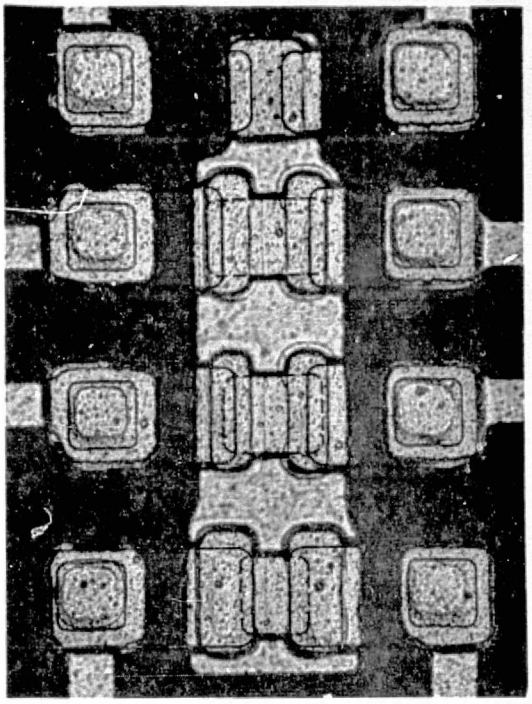


STRUCTURE #4A, 4B
DIF 15, DIF 16 ISOLATED ELECTRON (LEFT) AND
HOLE (RIGHT) INJECTORS. SCALE: 320X.

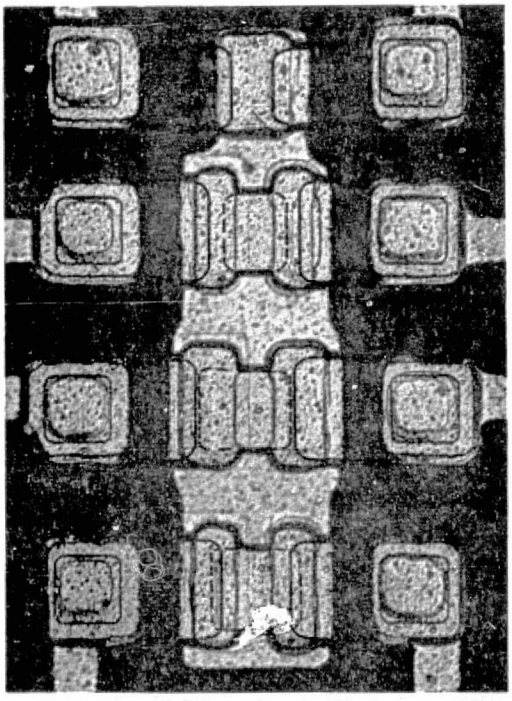
Figure 3. Test Structures 1-4 (Sheet 1 of 3)



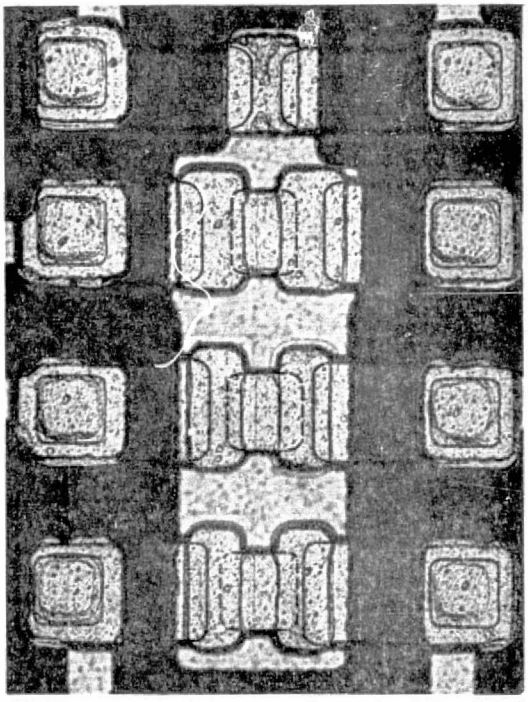
STRUCTURE #5
DIF 18 STANDARD SENSE DEVICE WITH MULTIPLE INJECTORS HAVING 2.5 μm n+/p+ OVERLAP. INJECTOR POSITION WRT WINDOW IS VARIED. SCALE: 575X.



STRUCTURE #6
DIF 20 STANDARD SENSE DEVICE WITH MULTIPLE INJECTORS HAVING 2.5 μm OVERLAP. INJECTOR POSITION WRT WINDOW IS VARIED. SCALE: 575X.

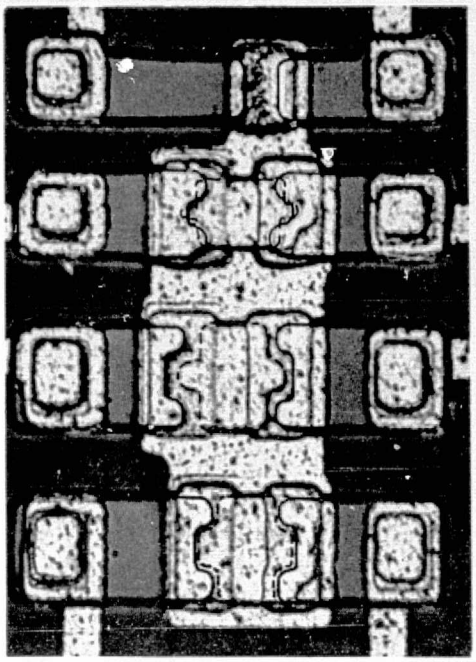


STRUCTURE #7
DIF 19 STANDARD SENSE DEVICE WITH MULTIPLE INJECTORS HAVING 2.5, 0.0, AND −2.5 μm OVERLAP. SCALE: 575X.

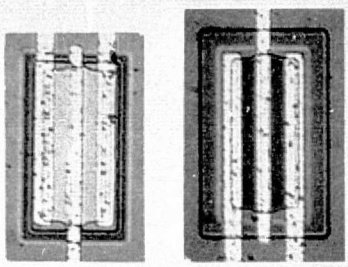


STRUCTURE #8
DIF 21 STANDARD SENSE DEVICE WITH MULTIPLE INJECTORS HAVING −2.5, −5.1 AND −7.6 μm OVERLAP. SCALE: 575X.

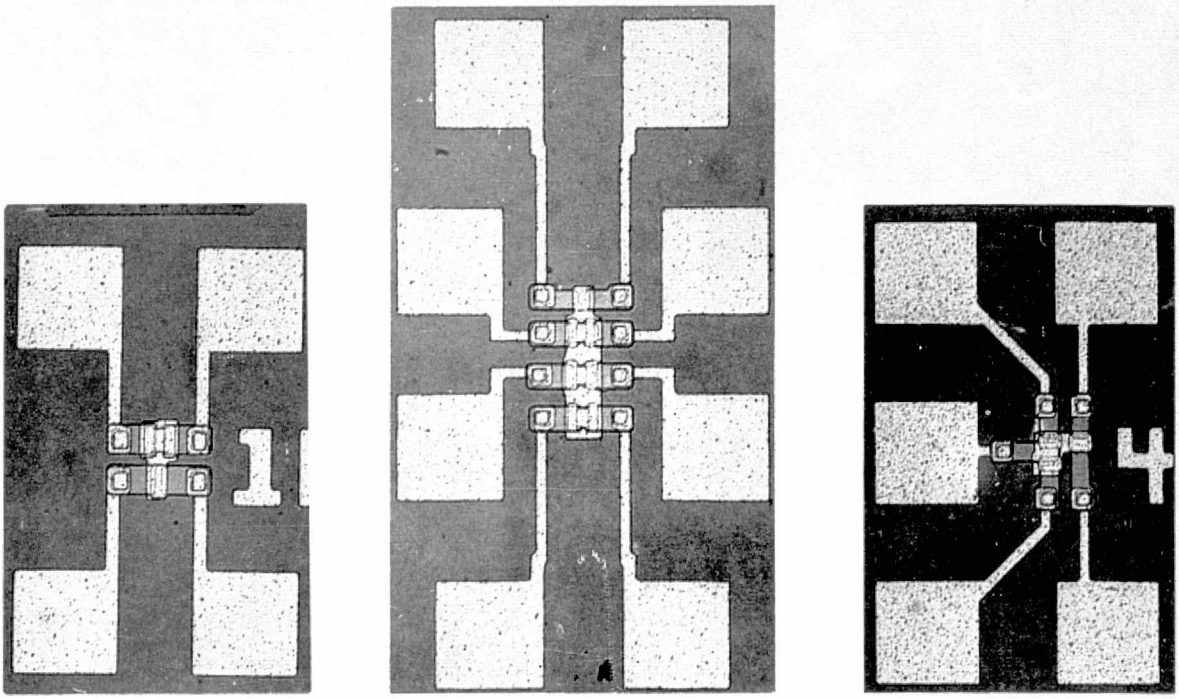
Figure 3. Test Structures 5-8 (Sheet 2 of 3)



STRUCTURE #9
DIF 22 STANDARD SENSE DEVICE WITH MULTIPLE INJECTORS HAVING EXTRA CORNERS. SCALE: 512X.



STRUCTURE #10
P-CHANNEL (LEFT) AND N-CHANNEL (RIGHT) CMOS TEST TRANSISTORS FOR p+/n AND n+/p INJECTOR COMPARISONS. SCALE 205X.



BOND PADS AND CONTACT METALLIZATION FOR:
DIF10 (UPPER LEFT). DIF ϕ4 (UPPER RIGHT). DIF 18, 19, 20, 21, AND 22 (LOWER) SCALE: BOND PADS ARE 100 μm X 100 μm.

Figure 3. Test Structures 9-10 (Sheet 3 of 3)

TABLE I. PROCESSING SEQUENCE: TEN STRUCTURE DESIGNS

Structures 1 - 9 DIFMOS Nitride Process	Structure 10 DIFMOS CMOS Process
Initial slice cleanup: 4-6 Ω-cm 100 n-Si	Initial slice cleanup: 4-6 Ω-cm 100 n-Si
Initial oxidation	Initial oxidation
p^+ pattern and etch	p-tank pattern and etch
p^+ deposition and drive	p-tank deposition and drive
n^+ pattern and etch	p^+ pattern and etch
n^+ deposition and drive	p^+ deposition and drive
Gate oxide pattern and etch	n^+ pattern and etch
Nitride deposition	n^+ deposition and drive
Nitride pattern and etch	Gate oxide pattern and etch
Gate oxidation	Gate oxidation
Contact pattern and etch	Contact pattern and etch
Metal deposition	Metal deposition
Metal pattern and etch	Metal pattern and etch
Sinter and anneal	Sinter and anneal
Protective overcoat deposition	Protective overcoat deposition
Bond pad pattern and etch	Bond pad pattern and etch
p-channel threshold voltage to be in the −1.5 to −2.0 volt range. p^+/substrate breakdown voltage to be over −60 volts.	
Adjust n^+ diffusion for p^+/n^+ BV of −12 to −20 volts	Adjust p-tank to set n^+/p BV to −12 to −20 volts

TABLE II. STRUCTURE 10 JUNCTION SPACINGS (μm)
(p- and n-Channel Devices))

Source-to-drain	25.4	20.3	15.2	10.2	7.6	6.4	5.1	3.8
Diffusion to guard ring	7.6	10.6	12.7	15.2				

However, no erase action was noted with these devices. None of the nitride passivated hole injectors was found capable of either injecting holes or electrons, or could effect any level of usable erase action. Because none of these structures could be erased, a series of experiments requiring successive write/erase operations to be performed could not be done.

Having achieved only limited success, we reexamined our basic concepts. We still believed that hole injection could be utilized to accomplish the erase action, so we set out to find any evidence of hole injection using either the oxide passivated electron injector structures or the nitride passivated hole injector structures with suitable bias as suggested by Verwey.[8] In all fixed-gate bias experiments on nitride passivated junctions where electric fields were externally superimposed with the gate upon the avalanche junction, no evidence of any carrier injection of either holes or electrons was observed. The only currents measured were due to leakage of the nitride films. The failure to observe any avalanche injection was probably due to the large densities of traps in the

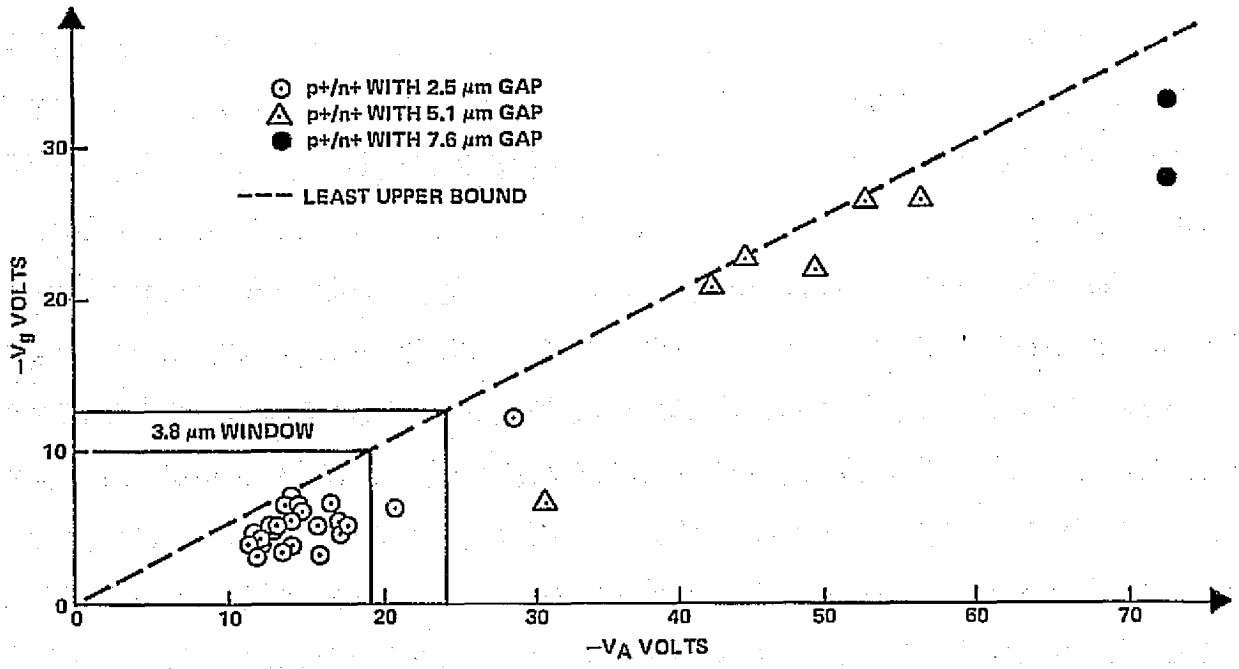


Figure 4. Final Floating Gate Voltage as a Function of Avalanche Voltage for p^+/n^+ Electron Injectors. Avalanche pulse was 500 μA at 100 ms

nitride which quickly captured the first few hot carriers, with their resulting field retarding any further injection.

But when Verwey's experiments were repeated with the oxide passivated electron injector junctions, both electron and hole injection into thermal SiO_2 were observed. Electron injection occurred with zero gate bias, while hole injection took place when gate bias was –60 volts. An example of the avalanche injection of holes and electrons into SiO_2 from a single junction type is shown in Figure 5. These experiments indicated that we could indeed inject holes through the oxide, and that the hole current could be measured. Thus, we had a quantative method of evaluating hole injector performance.

As shown in Figure 5, excessively high (–60 volt) gate voltages were required to obtain any usable hole injection from the p^+/n junctions. Such voltages would be too great to permit addressing individual capacitors in an integrated array. Therefore, it was decided to reexamine the results of Tarui et al.[3] on junction injection asymmetry in light of Verwey's methods. In other words, we were going to compare electron and hole injection into SiO_2 from an n^+/p junction (using NASA structure 10) against electron and hole injection from p^+/n junctions on the other NASA structures. The experimental setup is shown in Figure 6. The n-channel isolation tank was

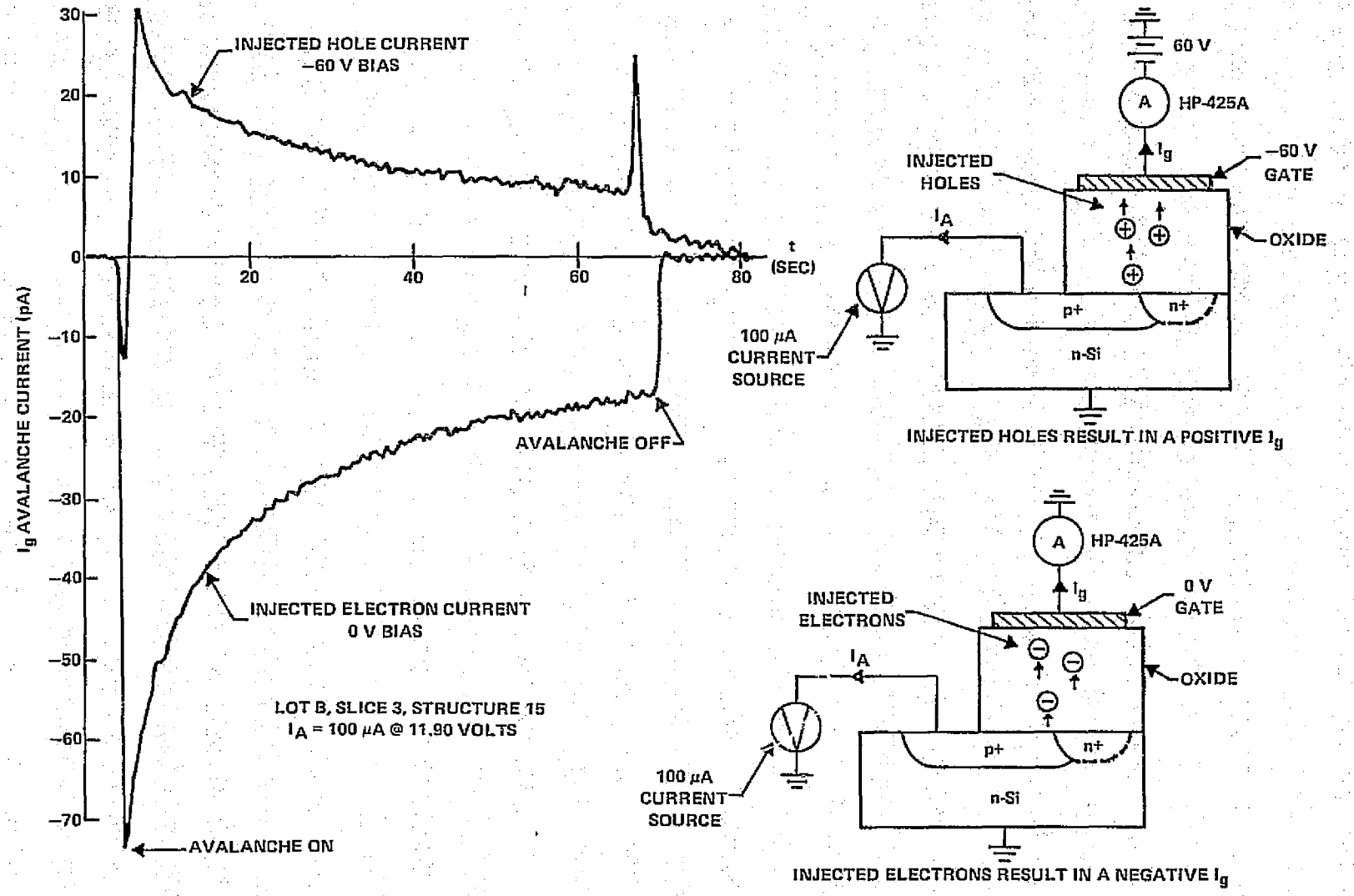


Figure 5. Avalanche Injection Experiments on Structure DIF15 (NASA Structure 4a)

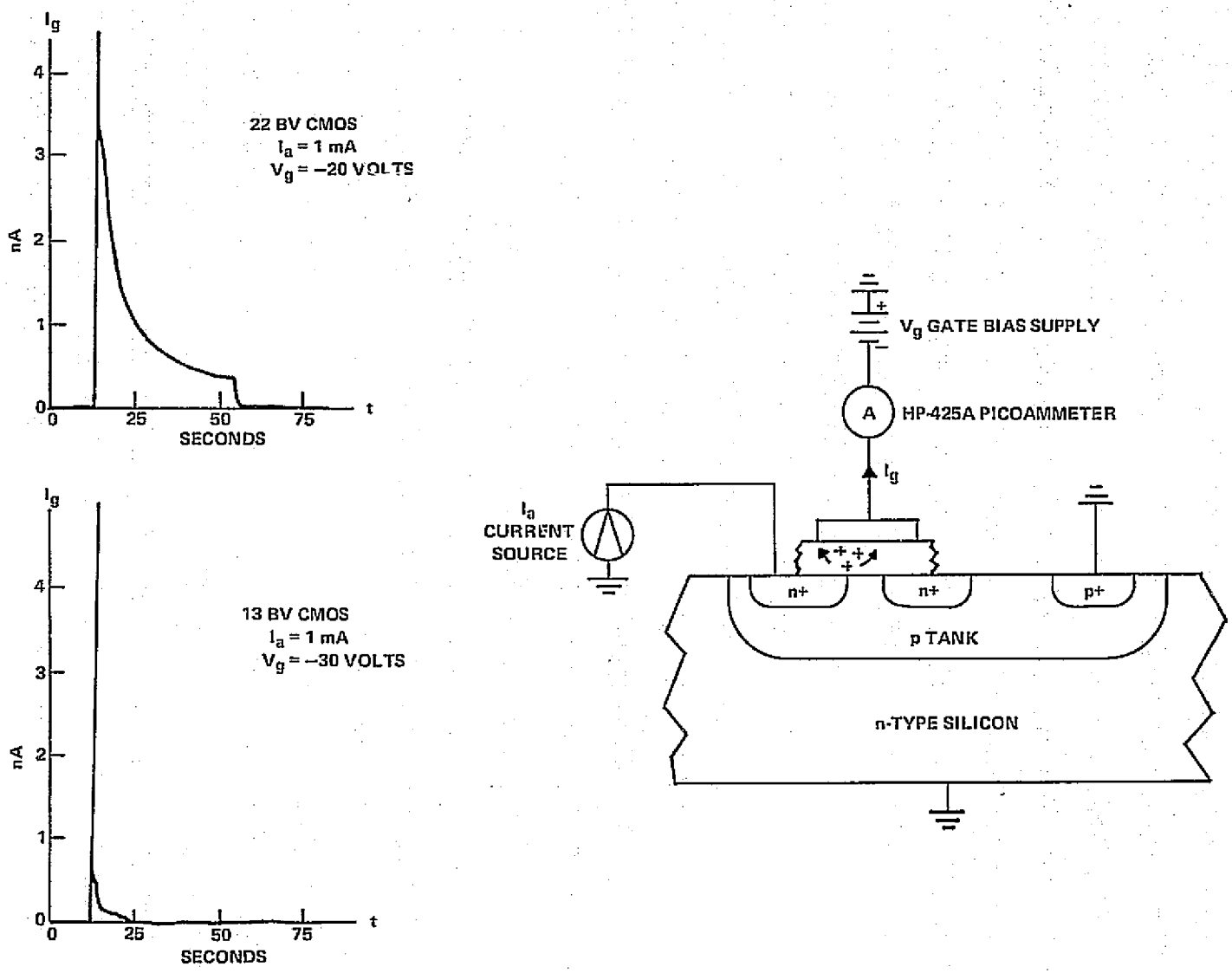


Figure 6. Comparison of 22 and 13 BV CMOS Slices –
Fixed Gate Bias Hole Injection Experiments

grounded, and avalanche was initiated using a positive current source. Injected current was measured using a gate bias supply in series with a picoammeter.

The hole injection results were dramatic! In comparison, the peak hole currents from the n^+/p junctions in Figure 6 were more than 100 times greater than hole currents from the p^+/n^+ junctions in Figure 5. Figure 6 showed injected current characteristics from two slices; one had an n^+/p breakdown of 22 volts, the other was 13 volts. In either case, approximately equal peak currents were obtained with n^+ diffusion to gate potential differences of −42 volts. Also, the 13 BV slices current exhibited a faster rate of decay than the 22 BV slice. This was due to more traps in the SiO_2 of the 13 BV slice due to heavier dopant concentration in the gate oxide from out-diffusion as the 13 BV slice had a heavier p-type tank predeposit than did the 22 BV slice.

Finally, no electron injection current could be measured using these n^+/p avalanche junctions. Unfortunately, we could not directly compare the n^+/p injectors with p^+/n injectors on the same slice because the p^+/n junctions required over 80 volts for avalanche. Thus, the best pair of injectors appeared to be the $p^+/n/n^+$ gap electron injector on the DIFMOS process slices, and the n^+/p hole injector from the CMOS slices. Since substantial hole injection into SiO_2 was observed using the n^+/p junctions, we wanted to determine if this injection could be used for erasing charge from the floating gates.

The large hole currents observed in Figure 6 were obtained using a positive avalanche current. Integration of DIFMOS into arrays would require that negative current sources be used to excite the avalanche plasma. Figure 7 shows the experimental setup used for measuring the relative discharge ability of the CMOS n^+/p hole injector junctions. In this case, the n^+ diffusion was grounded, and the avalanche was excited by biasing the p-type tank negatively, consistent with PMOS addressing logic. The bonding diagrams for negative and positive avalanche currents are shown in Figure 8. Negative avalanche currents require a bond wire to connect an n-channel drain or source to ground. This connection is shown as the long bond wire in Figure 8(b). The other bond wire in Figure 8(b) is used to connect the n-channel gate covering the n^+/p avalanche region to a p-channel transistor on the same chip for sensing purposes. The measurement of the ability of the n^+/p junction for erasing requires the use of a floating gate type of experiment in which the gate is initially charged negatively. The p-channel transistor is then used as an electrometer to measure the discharge, if any, when the n^+/p junction is avalanched.

The results of such discharge experiments are shown in Figure 7. The true floating gate voltage has been determined by calibrating the p-channel sense transistor by plotting the sense transfer characteristic sense (output voltage versus floating gate voltage) when a known voltage is imposed on the floating gate by an external probe. Figure 7 shows that indeed the n^+/p avalanche does inject holes to the negatively charged floating gate, but only a partial discharge is affected. Using a 13 BV CMOS slice avalanched for 100 ms at 1 mA, it was found that the floating gate can be

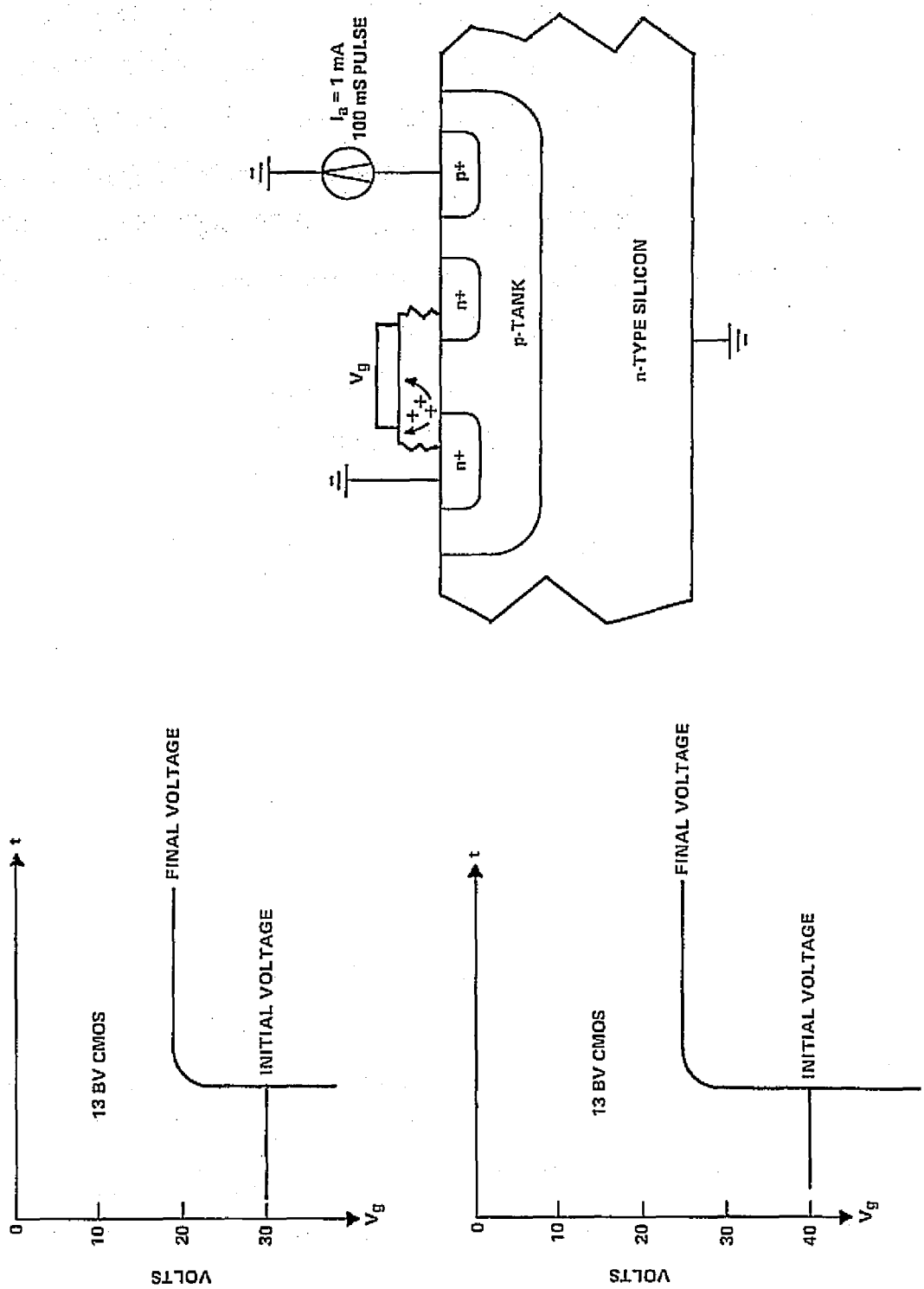
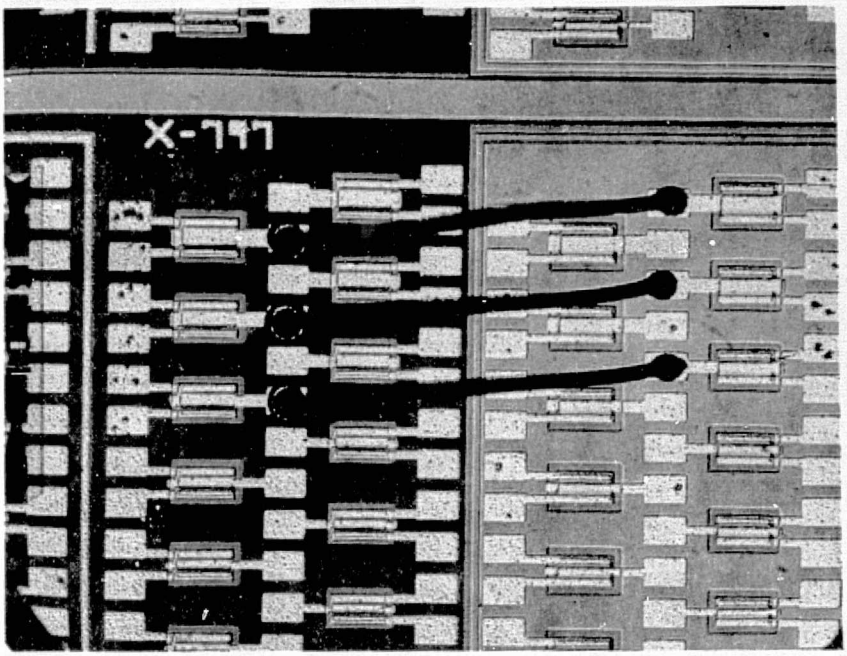
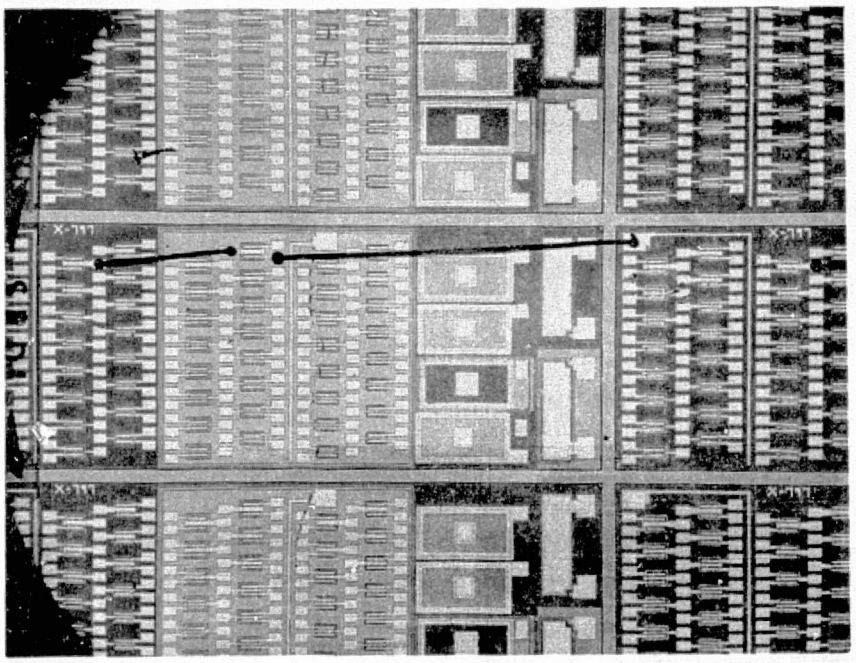


Figure 7. Floating Gate Discharge Experiments, 13 BV CMOS, Pulsed Negative Avalanche



(a) BONDED FOR POSITIVE AVALANCHE EXPERIMENTS



(b) BONDED FOR NEGATIVE AVALANCHE EXPERIMENTS

Figure 8. CMOS Test Bar Bonded for Avalanche Experiments

discharged from −30 to −18 volts in the 100-ms period, or from −40 to −24 volts in like time. The 100-ms avalanche is seen as a spike in the output waveforms in Figure 7. For longer avalanche pulses, it was found that the amount of gate discharge saturated with about −20 volts remaining.

B. BOOTSTRAPPING

We found that we could negatively charge a floating gate using avalanche injection of electrons from a p^+/n gap/n^+ junction. The gate can be charged up to about one-half the avalanche voltage. For the optimum avalanche voltage of −20 volts for an array which is capable of being switched by an address transistor, the floating gate can be charged to about −10 volts. In discrete experiments, high-voltage punch-through of CMOS p-channel transistors was capable of charging the floating gate to −30 volts or more. Because of the high punch-through voltages required for such charging, the high-voltage storage cannot be implemented into an array because the voltages are too high to switch with address transistors. It was also found that the floating gate voltages could be partially discharged using avalanche injection of holes from an avalanched n^+/p junction. The gates could be discharged only from voltages more negative than −20 volts up to about −20 volts. Even though we have both electron and hole injectors, those which can be integrated together have incompatible charge/discharge voltage ranges. That is, the electron injector could not charge the floating gate up to the minimum level that the hole injector can discharge the gate. Or can it?

It was postulated that a negative voltage could be capacitively coupled to the floating gate to aid the erasure process. The circuit element needed to accomplish this coupling is called a bootstrap capacitor. The floating gate would be extended over an extra diffusion area, and the diffusion coupled to the outside world through a pin connection. Thus, any voltage connected to the diffusion would be fractionally coupled to the floating gate; the fractional coupling coefficient depending upon the ratio of bootstrap capacitance to total gate capacitance and bootstrap capacitance. The bootstrap coupling coefficient should be maximized to minimize the external bootstrap voltage requirements.

Such a capacitive element was included on the CMOS test bar for gate to diffusion measurements. The bonding diagram for the inclusion of this capacitor is shown in Figure 9. Figure 10 shows the structural and schematic representation of the bootstrap device. Figure 11 shows experimental results obtained in bootstrap voltage experiments. In this experiment, the CMOS cell is written by quickly pulsing the sense transistor into punch-through using a −80-volt pulse for 100 ms. Then, the hole injector is avalanched at 500 μA for 100 ms while the bootstrap voltage is applied. After the hole injector avalanche is complete, the bootstrap voltage is returned to ground. Figure 11 shows that the bootstrap begins to aid the erase when V_B is −30 volts. By the

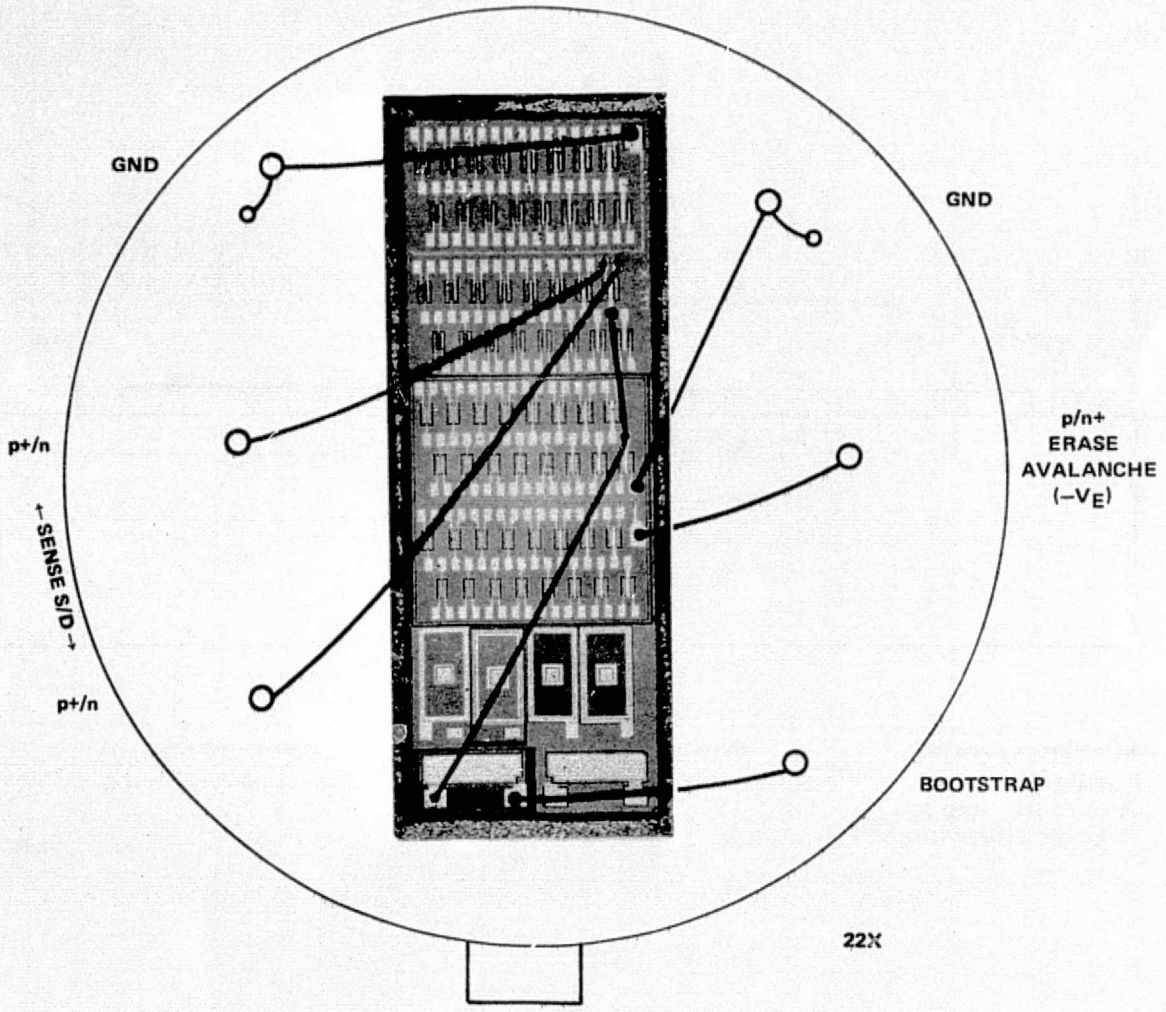
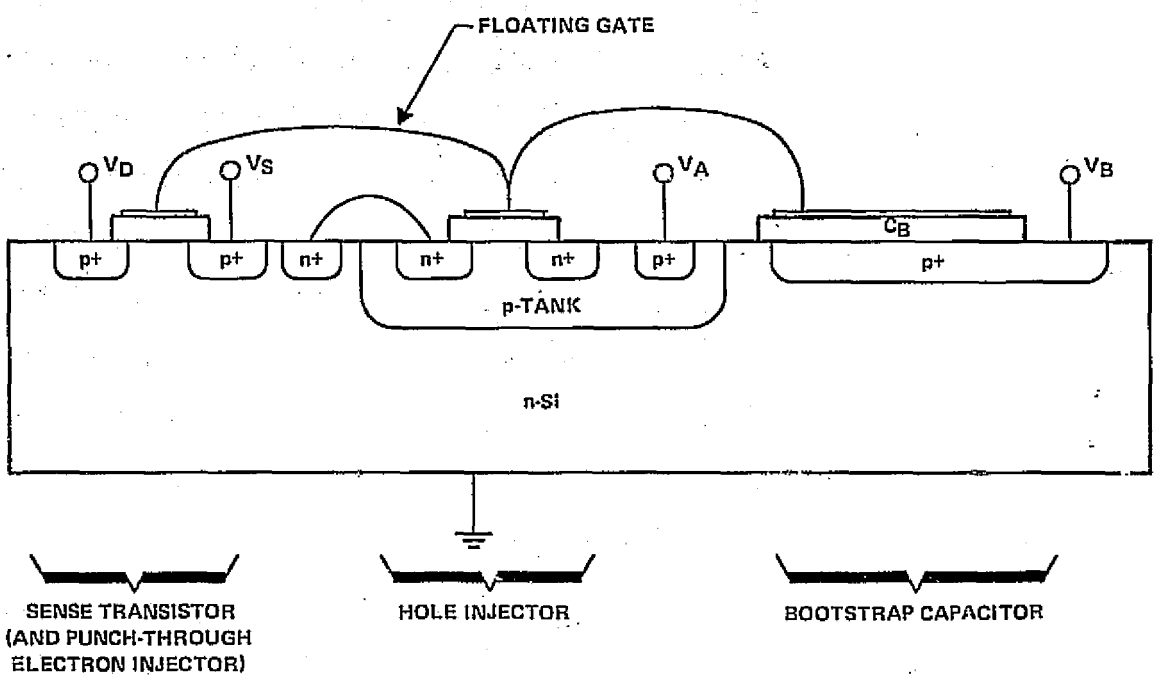


Figure 9. Bonding Diagram for the CMOS Bootstrap Capacitor Experiments

time V_B is −35 volts, the gate can be completely discharged. The CMOS structure bootstrap capacitor had a coupling coefficient of approximately 90%, so that −31.5 volts was coupled to the floating gate during the 100-ms discharge.



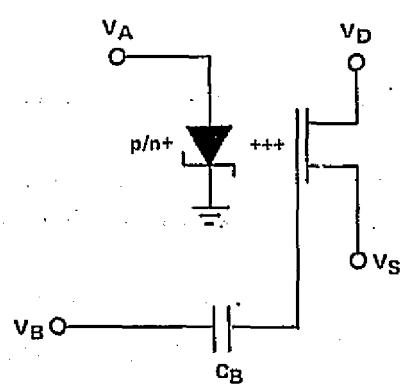


Figure 10. Schematic Diagram and Physical Structure of the Bootstrap CMOS Test Device

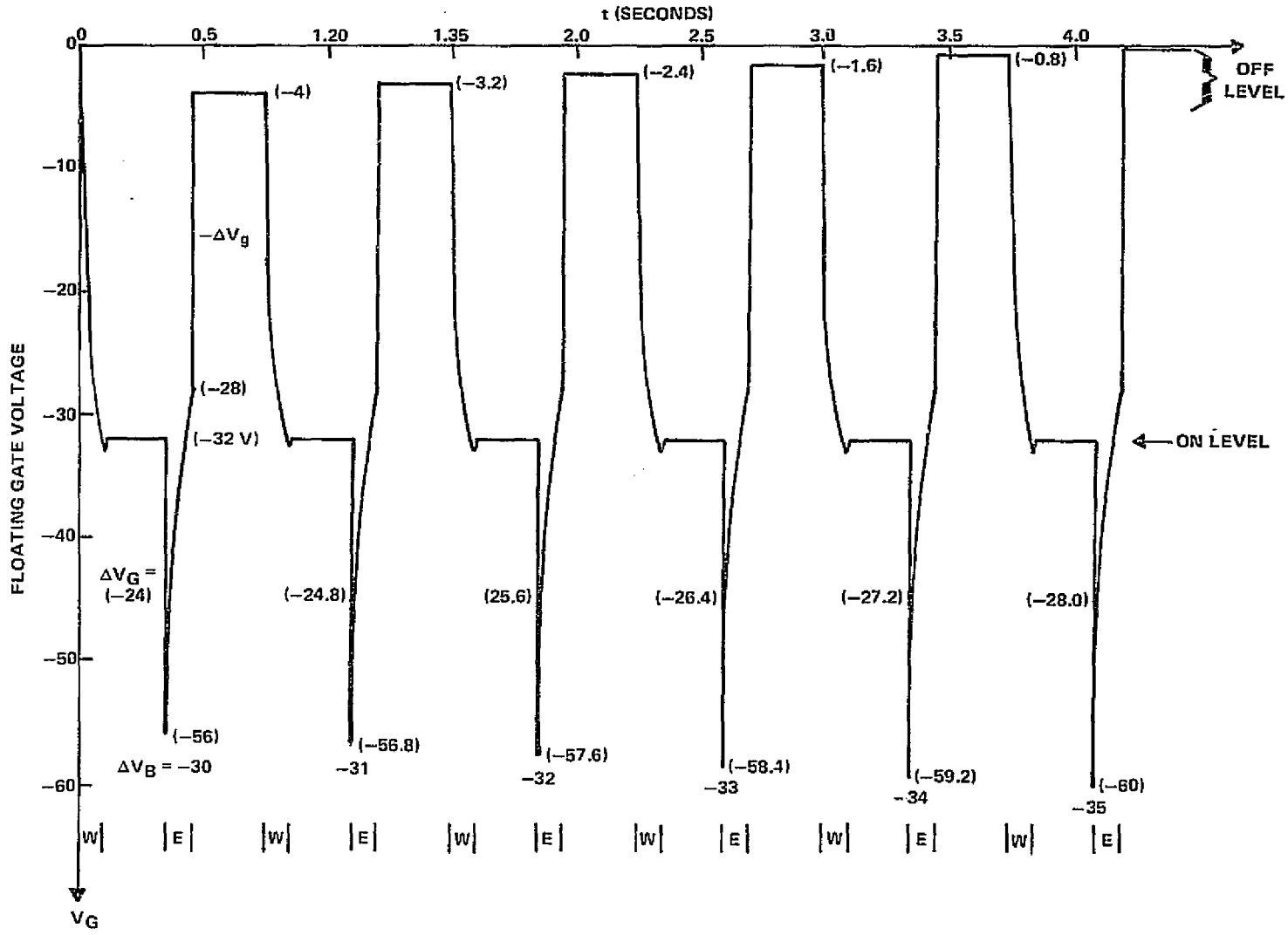


Figure 11. Measurement of the Write/Erase Window as a Function of the Bootstrap Bias Voltage Using Discrete CMOS Test Structures

The use of a bootstrap capacitor in conjunction with the n^+/p hole injector gave us the ability to achieve hole injection capable of discharging the floating gate. Together with the gap-type electron injector and the sense transistor, we now have all the elements required to design a nonvolatile element.

However, before we describe the latest nonvolatile cell design, we want to reexamine the possiblity of using nitride passivated p^+/n^+ junctions for hole injection. A series of experiments will be described where differing oxide and nitride films will have their hole and electron injection characteristics measured. This will be a last attempt to see if the original DIFMOS concept can be made to work.

SECTION IV

THIRD EXPERIMENTAL SERIES

A. DESCRIPTION OF TASK

This task was funded as an amendment to the initial contract. The objective of the task was to perform an experimental study of the gate insulator which facilitates the passage of positive charge between the semiconductor and the floating gate. It was specified that deposition be by chemical vapor deposition and by sputtering, and that injector efficiency should be measured as a function of thickness, refractive index, etch rate, dielectric constant, and resistivity of the insulator.

In addition to the hole injector dielectrics specified for the task, two oxide types were fabricated for evaluation as electron injectors. The experimental processing which was carried out to meet these objectives is outlined in Table III.

In a preliminary run, process variables were adjusted for fabrication of DIFMOS devices. The most critical process step was optimization of an n^+ profile for the desired avalanche voltage (16-20 volts). The basic process was used to fabricate devices with various insulator types for the electron and hole injectors. The DIFMOS test bar was used for each of these runs.

The electron-injector insulator variation run included steam and dry oxides, with three thicknesses of each. For this run, the electron injector oxide was used in both the electron injector window and the hole injector window.

Insulators for the hole injector window included several types of silicon nitride and nitrous oxide as described below:

Furnace Tube Nitride: Silane, ammonia and nitrogen passed over slices heated in a furnace tube to 850°C.

Plasma Nitride: Nitride deposited in an rf plasma reactor at a temperature of 300°C. Two refractive-index ranges were used, with two anneal temperatures for each range of refractive index.

Reactor Nitride: Nitride formed in a Nitrox reactor at 850°C.

Nitrous Oxide: SiO_2 deposited in an rf plasma reactor at 300°C using silane and nitrous oxide as source gases.

TABLE III. EXPERIMENTAL PROCESS VARIATIONS

Preliminary Run

Adjust process flow for fabrication of slices on CDD pilot line
- Optimize n^+ profiles for desired avalanche voltage
- Process slices and verify performance of devices

Electron Injector Variations

Fabricate slices varying deposition process and thickness for electron injector oxide. (Nitride will not be used in this run.) The oxide deposition processes and thicknesses to be investigated are shown below:

Oxide Deposition Process	Oxide Thickness
950° Steam oxide	50 nm, 100 nm, 200 nm,
1100° Dry oxide	50 nm, 100 nm, 200 nm,

Hole Injector Variations

Fabricate slices varying deposition process and thickness for hole injector insulator. (950°C steam will be used for the electron injectors on all slices.) The insulator deposition processes and thicknesses to be investigated are shown below:

Hole Injector Deposition Process	Insulator Thickness
850° Tube nitride	50 nm, 100 nm, 200 nm,
850° Reactor nitride	50 nm, 100 nm, 200 nm,
300° Plasma nitride	
Low index	
450° anneal	50 nm, 100 nm, 200 nm,
850° anneal	50 nm, 100 nm, 200 nm,
Standard index	
450° anneal	50 nm, 100 nm, 200 nm,
850° anneal	50 nm, 100 nm, 200 nm,
Sputtered nitrous oxide	50 nm, 100 nm, 200 nm,

Slices were processed as described in Table III and devices were evaluated for each slice. Pilot slices were processed with each injector insulator variation for evaluation by NASA and Texas Instruments.

B. ION IMPLANTED INJECTOR JUNCTION

It was observed that injection of carriers for DIFMOS devices increased with the avalanche voltage of the injector junction. In order to compare the various electron and hole injectors of this task on a common basis, it was desired that avalanche voltage be independent of process variables. For compatibility with MOS circuitry, the avalanche voltage should be in the range of 16-20 volts.

Breakdown voltage for p^+/n junctions formed by diffusing p-regions into the n-type substrate was in excess of 50 volts. For DIFMOS circuits fabricated previously, avalanche voltage was adjusted by fabricating $p^+/n/n^+$ junctions. Breakdown voltage was adjusted to the desired value by varying the gap; i.e., the width of the n region between the p^+ and n^+ regions. This structure is shown in Figure 12(a). Avalanche voltage in the range of 20 volts was obtained; however, spacing and avalanche voltage depended critically upon alignment of the p^+ and n^+ oxide removals.

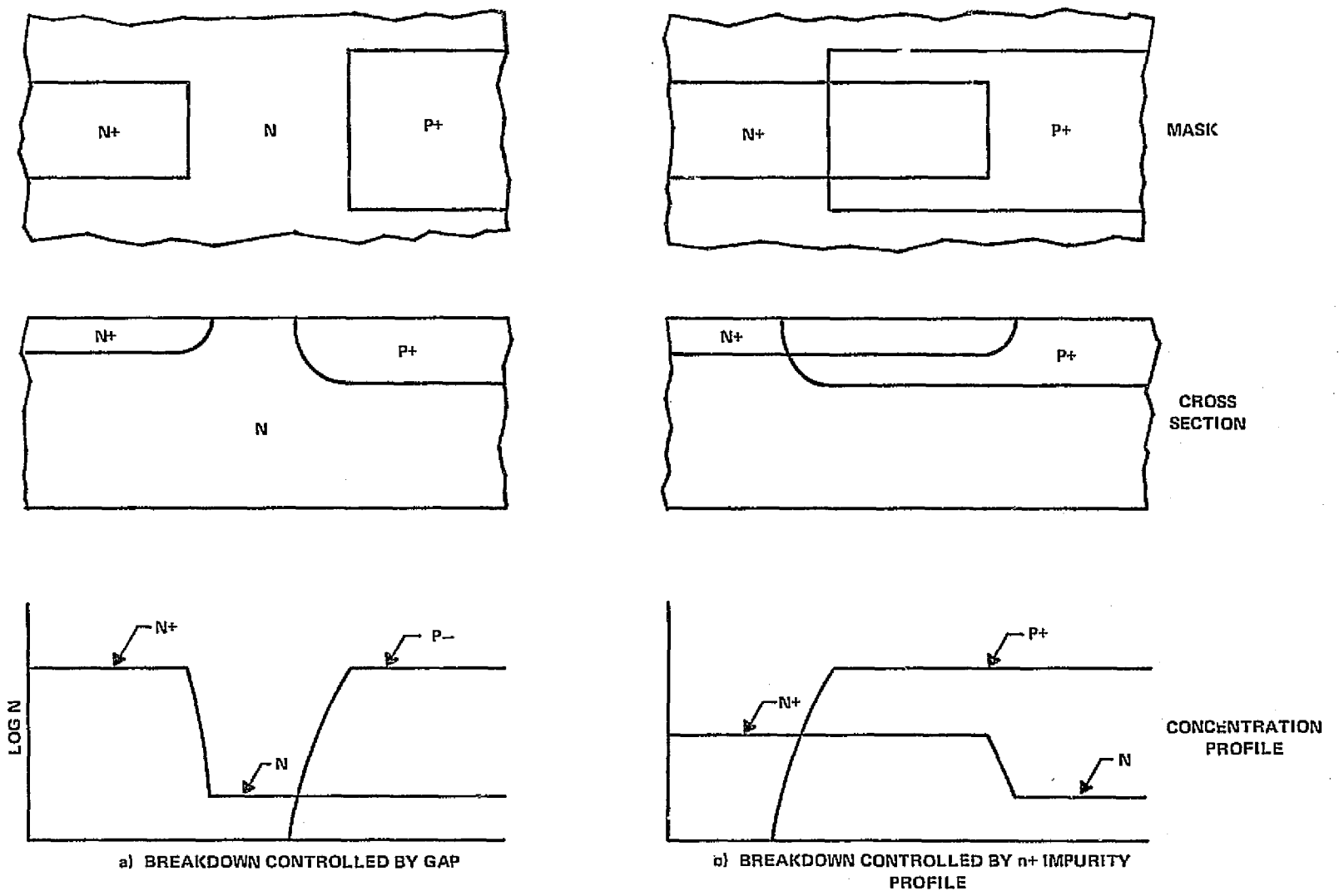


Figure 12. Avalanche Breakdown in DIFMOS Structures

Devices for this task were fabricated with overlap junctions to obtain the desired avalanche voltage. For the overlap junction, shown in Figure 12(b), breakdown voltage is determined by the concentration of the n-region and does not depend upon alignment.

An ion-implanted n-region was used to provide the required n-region profile with process reproducibility. Implant dose and anneal cycle were varied in the preliminary run to adjust avalanche voltage.

C. ELECTRON INJECTOR INSULATOR VARIATIONS

Average characteristics of two slices for each oxide condition were measured in slice form using a 3-point probe. The characteristics, shown in Table IV were quite uniform across each slice and for different slices processed with the same oxide conditions.

TABLE IV. CHARACTERISTICS OF ELECTRON INJECTORS FOR VARIOUS OXIDES

Slice No.	Oxide Type	Oxide Thickness nm	V_b Volts	g_m - μmhos @ V_D = 10V I_D = 1mA	V_{tx} Volts	Device No.	I(t=o) nA	I(t=50s) nA
2	Steam	50	19.3	220	3.1	20b	2.2	0.09
3	Steam	50	18.6	210	2.9	18b	1.8	0.08
5	Steam	100	19.6	130	4.2	20b	1.0	0.05
6	Steam	100	19.5	130	4.3	18b	0.8	0.04
9	Steam	200	20.4	70	6.7	18b	0.3	0.02
10	Steam	200	19.1	70	6.7	20b	0.4	0.03
12	Dry	50	19.2	200	3.3	20b	1.3	0.05
13	Dry	50	20.0	200	3.2	18b	1.6	0.08
15	Dry	100	21.4	120	3.9	20b	1.1	0.05
16	Dry	100	22.?	140	3.9	18b	0.9	0.04
18	Dry	200	28.0	80	5.7	18b	0.7	0.04
20	Dry	200	28.1	80	6.2	20b	0.6	0.05

Electron injection currents were measured with fixed-gate bias using the circuit and measurement conditions shown in Figure 13. Avalanche current was held constant at 0.1 mA and avalanche voltage was monitored with a digital voltmeter. Injected gate current was amplified with an HP-425A microammeter, and applied to an XY recorder. Avalanche voltage increased by a few tenths of a volt during the avalanche pulse; final values of avalanche voltage were recorded Plots of electron-injection current versus time as measured on the X-Y recorder are shown in Figure 14. Amplification of the HP-425A was increased by a factor of 10 after 50 seconds to give a more accurate reading at this point.

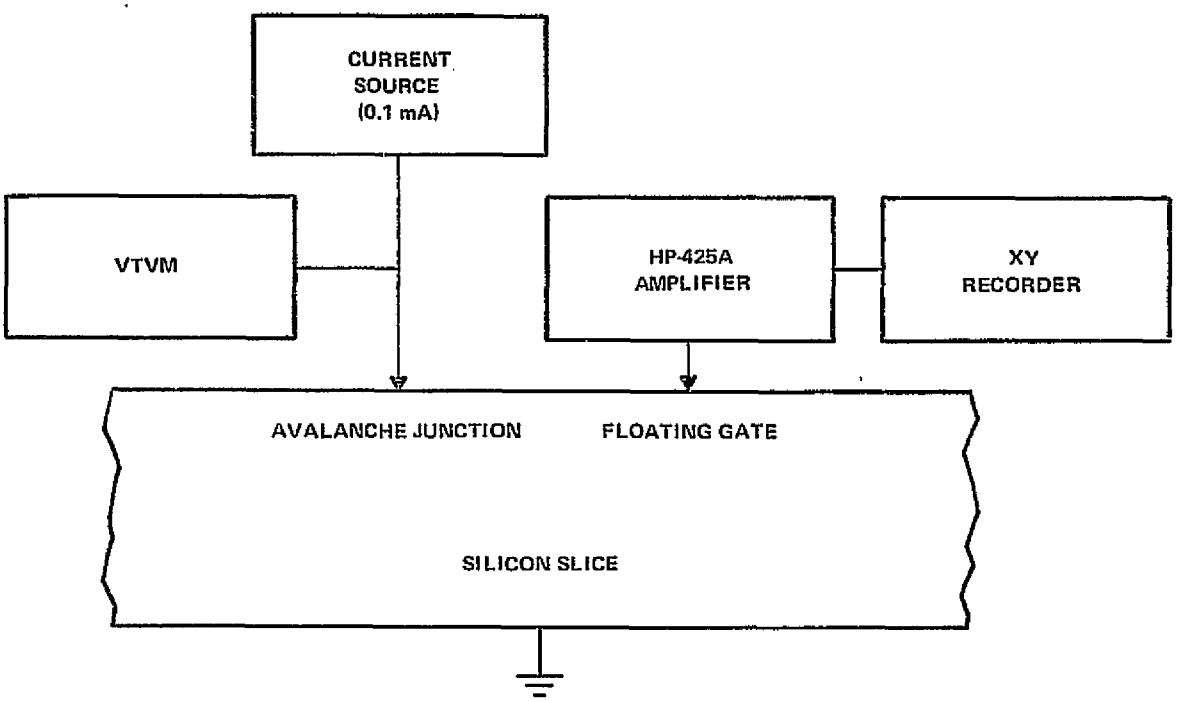


Figure 13. Circuit for Measuring Electron Injection Current

Average values of peak (initial) injected current, current after 50 seconds, and avalanche voltage for each slice are recorded in Table IV. Injected current decreased for thicker oxides as expected. For the 50 nm oxides, steam oxide devices exhibit slightly higher injected current than dry oxide devices. However, injected current for dry oxide decreases less rapidly as oxide thickness is increased. The latter effect is probably due to the fact that avalanche voltage is higher for the 100 nm and 200 nm dry oxide slices. Peak concentration of the implanted phosphorus region decreases because of the long temperature cycle required to grow thick oxides in dry oxygen, causing avalanche voltage to increase.

D. HOLE INJECTOR INSULATOR VARIATIONS

Evaluation of devices for each film type was carried out on devices 1 and 1A of the test bar by probing in slice form. Several devices of each type were measured. The characteristics measured are described below:

Threshold Voltage V_{tx} was measured for a drain current of 10 μA.

Charge Storage capability was measured qualitatively by charging the floating gate with a probe. Source-drain VI characteristics were observed on a curve tracer to indicate whether gate voltage remained charged at a fixed value after the gate probe was removed.

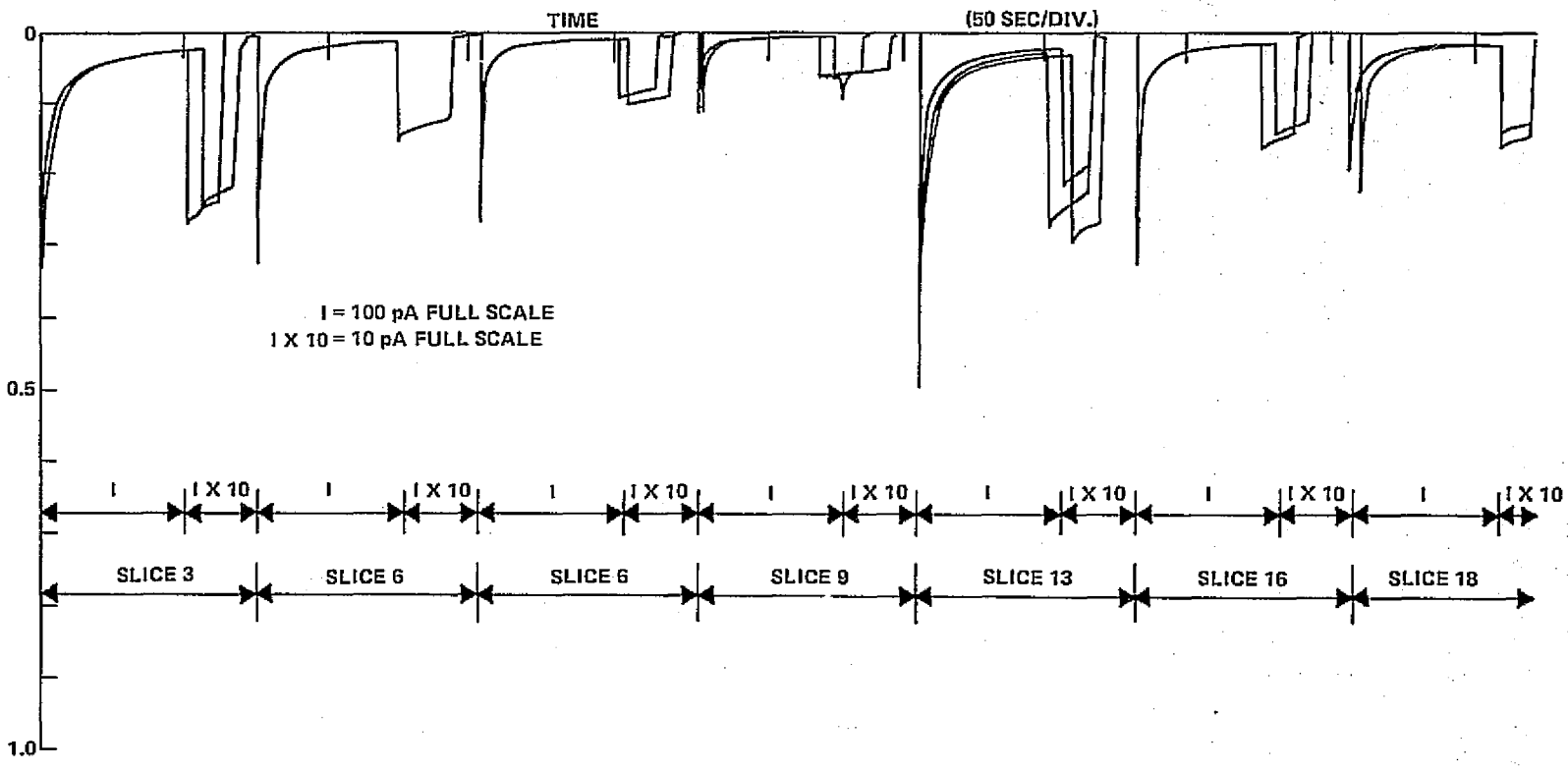


Figure 14. Electron Injection Gate Current versus Time
for Various Oxide Conditions

"Write" capability for the electron injectors was tested using the circuit of Figure 15(a). The current source was adjusted for a current of 0.5 mA and a compliance voltage of 18 volts, which was slightly greater than the avalanche voltage (16 volts) of the injector junction. The source-drain VI characteristic was observed on a curve tracer as the injector junction was avalanched for a few seconds. Drain current for a 3-volt sweep was recorded after avalanche.

"Erase" capability for the hole injectors was evaluated using the circuit of Figure 15(b). The hole injector was avalanched to determine whether the floating gate was discharged.

The maximum gate voltage which could be applied was limited by rupture of the hole injector dielectric film. The value for which rupture occurred was measured by applying a negative voltage by probe to the gate.

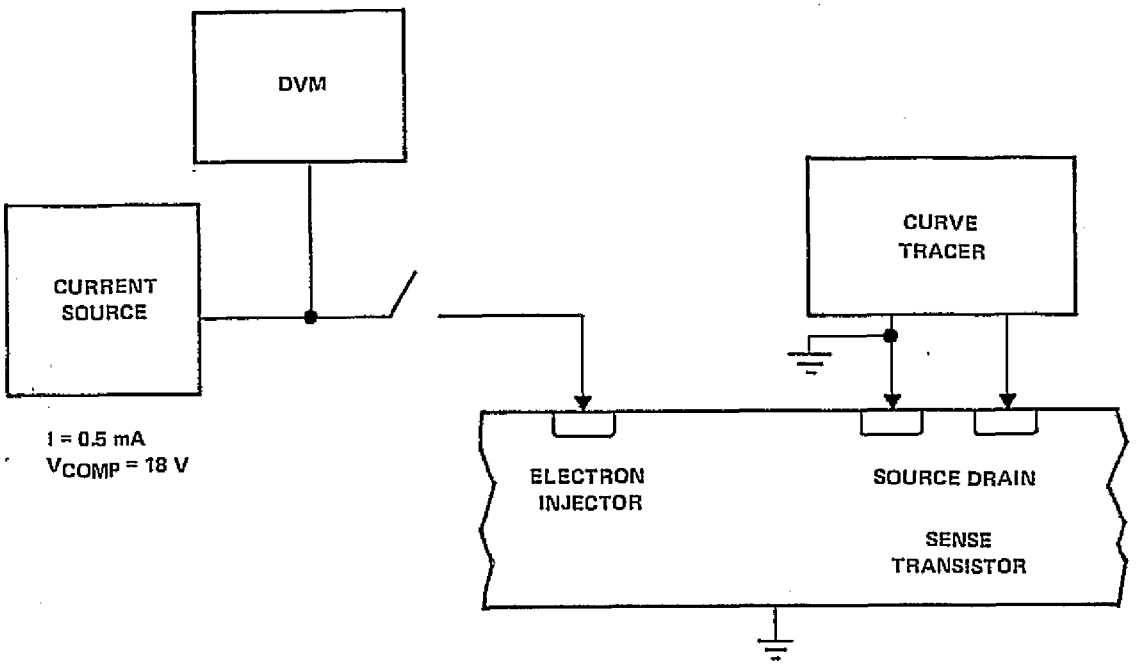
The results obtained are listed in Table V for the various hole injector films. "Writing" by avalanching the electron injector was observed for some devices on nearly all the slices. Those devices which did not write exhibited high gate-to-substrate leakage.

Erasure was obtained for some devices from the tube nitride and the reactor nitride slices. For both nitride types, the devices with thinner films erased more rapidly than those with thicker films. However, the devices with thin films would not sustain as high a voltage on the floating gate. These results are as expected, since conductivity of nitride films increases with electric field. A given value of conductivity will correspond to a higher value of gate voltage for thicker films.

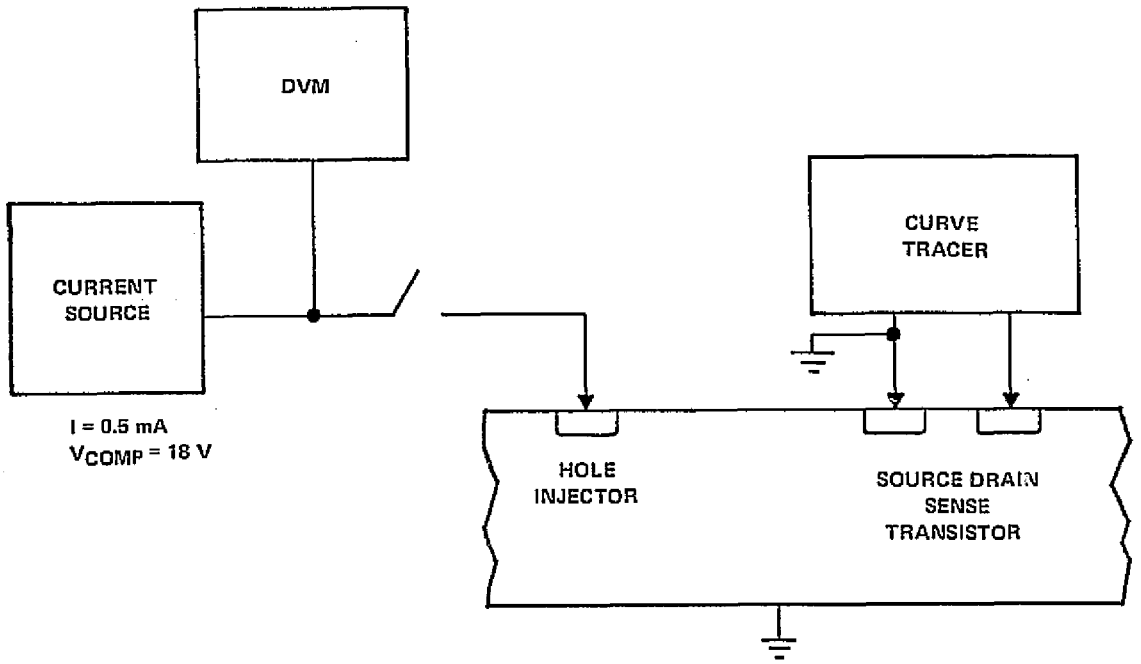
Erasure with nitride films had been observed on earlier DIFMOS devices. As discussed previously, considerable power was required for erasure in those devices and the effect was attributed to heating of the nitride film.

The results described here appear to be a different effect. Power dissipated during the erase pulse was about 8 mW, which should not produce sufficient heating for conduction in the nitride film. Furthermore, no erase was observed for injector junction voltages slightly less than the avalanche value. Erase occurred when injector voltage was increased to the value where avalanche current was initiated.

In order to determine whether erasure was due to hole injection, a microammeter was connected by probe to the gate during avalanche of the injector junction. No hole current was detected, indicating that the erase was probably due to field emission from the gates rather than hole injection from the junction.



(a) "WRITE" CIRCUIT



(b) "ERASE" CIRCUIT

Figure 15. Circuits for Evaluating "Write" and "Erase" Injectors

TABLE V. DEVICE CHARACTERISTICS FOR VARIOUS HOLE-INJECTOR DIELECTRIC FILMS

Run No.	Slice No.	Insulator Type	Film Thickness nm	V_{tx} @ 10μA	Charge Storage	Write I_D (μA)	Erase	Max Gate Volts
4A		Tube Nitride						
	1		50	3.8	Yes	50	Complete	18
	2		50	3.6	Yes	50	Complete	17
	3		100	3.5	Yes	110	Incomplete	19
	6		100	3.5	Yes	100	Slight	18
	7		200	3.5	Yes	250	Slight	23
	8		200	3.6	Yes	250	No	27
4B		Plasma Nitride						
	9	Low Index	50	2.4	No	No	No	6
	11	Anneal 450°	100	2.3	Yes	500	No	38
	13	Anneal 450°	200	2.5	Yes	400	No	52
	10	Low Index	50	4.2	Yes	400	No	45
	12	Anneal 850°	100	4.2	Yes	400	No	54
	14	Anneal 850°	200	4.2	Yes	400	No	48
	15	Standard Index	50	3.2	Leaky	Leaky	No	34
	18	Anneal 450°	100	3.4	Leaky	350	No	>55
	20	Anneal 450°	200	3.4	Yes	500	No	>55
	17	Standard Index	50	5.2	Leaky	No	No	27
	19	Anneal 850°	100	5.0	Leaky	No	No	36
	21	Anneal 850°	200	5.3	Leaky	No	No	54
4C		Reactor Nitride						
	22		50	3.9	Leaky	Leaky	Leaky	14
	23		100	3.9	Slight Leak	80	Incomplete	16
	24		200	4.0	Slight Leak	120	No	26
		Nitrous Oxide						
	25		50					
	26		100					
	27		200					

SECTION V

FOURTH EXPERIMENTAL SERIES

In this series of experiments, we have designed, built, and tested nonvolatile memory cells incorporating the gap-type electron injector, the n^+/p hole injector, and the bootstrap capacitor. In two separate designs incorporating the same basic designs of injectors, address and sense transistors, there are three variations of the bootstrap capacitor size. The full-sized capacitor cell has been designed to have the same 90% bootstrap coupling coefficient as the CMOS test structures. In addition, capacitors with one-half and one-fourth that capacity are also included.

One boundary condition of the DIFMOS cell design spelled out by the contract SOW is that the cell design be expandable to integrated arrays. To verify that this can be done, each cell design has included address transistors for switching the avalanche voltages. In addition, there is a 4-column by 8-row array of the full-sized bootstrap capacitor cells, as the basic cell design was designed to permit implementation into a row-and-column framework.

As seen in Figures 16 and 17, there are three column lines, with a memory cell located on each side. Two row address lines are designed to select either the right or left-side cell. In operation, a cell (or row) is selected by connecting the appropriate row address line to −45 to −50 volts. All other rows are held at ground.

The three column lines operate the sensing, writing, and erasing functions. Application of −45 to −50 to the middle column line (marked W/S) will result in the formation of an avalanche plasma in the electron injector. As the injector is designed to break down in the −20 volt range, approximately −25 to −30 volts are dropped across the address transistor. The finite conductivity of the address transistor will limit the avalanche current to several hundred microamperes. The electron injector avalanche will charge the floating gate electrode to about −10 volts.

The state of the floating gate charge can be determined by sensing the conductivity between the W/S and S column lines. Alternatively, the S column line can be eliminated entirely, provided the sense transistor source is internally grounded. Sensing would then be accomplished by measuring the low voltage (voltages less than the injector BV) conductivity of the W/S line to ground.

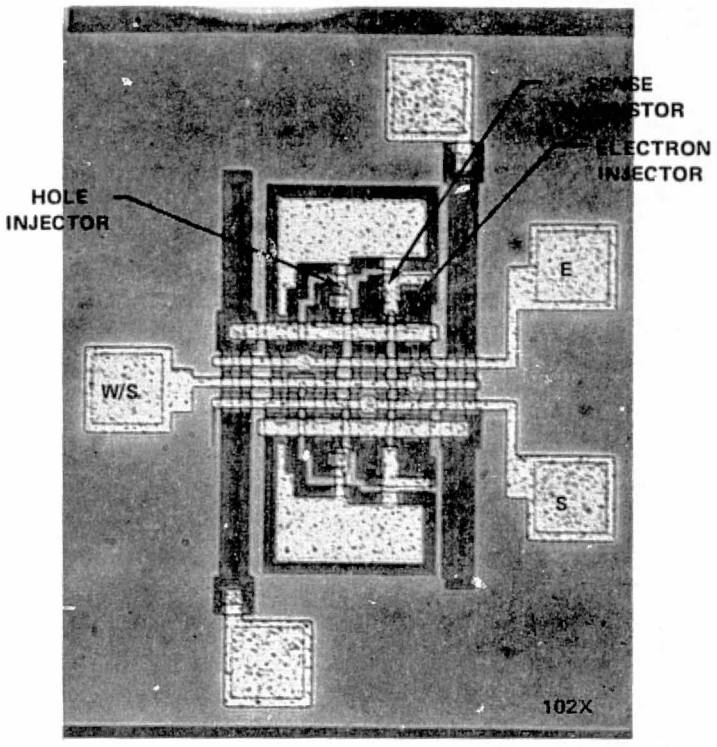


Figure 16. DIFMOS Memory Cell Design –
Dual Bits with Left/Right Select C_b Sized for 90% Coupling

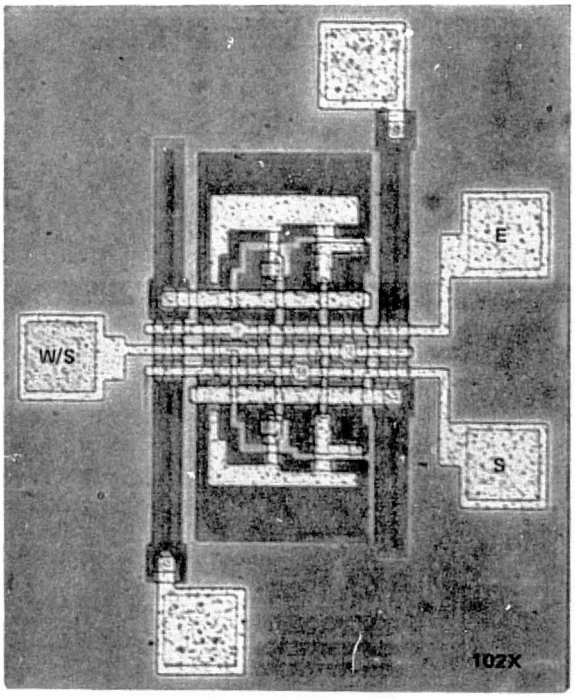


Figure 17. DIFMOS Memory Cell Designs – 0.50 C_b (left), 0.25 C_b (right)

The floating gate charge can be erased or removed by applying −45 to −50 volts to the erase column line. Voltage from this line is routed to both the hole injector junction and the bootstrap capacitor through two address transistors. The row address and column voltage lines must be sufficiently large to apply −35 volts to the bootstrap capacitor diffusion to effect the proper erase; this is why −45 to −50 volts is required on both lines. (Even though the row address line does not actually participate in events within the memory cell, the bootstrap capacitor diffusion and the hole injector diodes are actually loads connected to the address transistors as source followers. It is a well known fact of circuit design that the maximum output voltage of a source follower is the gate voltage or row address voltage less the device threshold voltage with all back gate bias considerations included.)

It is very important that two separate address transistors be included for the erase function. To minimize the cell size, the bootstrap voltage should be as large as possible. Thus, the bootstrap voltage is switched through one address transistor. Since the capacitor draws only a few pA leakage current after the capacitor is charged, the source follower output voltage will be maximized. On the other hand, the hole injector diode is supplied through a separate address transistor. The source follower output voltage is clamped to a maximum value by the n^+/p avalanche voltage. Thus 20 to 25 volts are dropped across the injector address transistor, and like the electron injector, the finite conductivity of the address transistor limits the hole injector avalanche current to several hundred microamperes. So from a single erase column line, we supply maximum voltage at zero current to the bootstrap capacitor node, and avalanche voltage and current to the hole injector junction.

Several lots of the DIFMOS cells were fabricated and tested. In general, excellent performance of individual cells was attained, but there were several problems. First, a small design error in the 32-bit array causes all the even-numbered row address lines to exhibit a 7-volt breakdown. This was caused by a violation of the diffusion to guard ring spacing requirement. Normally, this is 12.7 μm, and is dropped to 3.8 μm in the electron injector. The design error actually resulted in the n^+ and p^+ diffusions touching, resulting in the 7-volt BV. So all arrays are limited to only the 16 bits, found on the 4 odd rows. This problem did not affect the discrete cells on the chip.

Second, it was found that we had very poor control of the electron injector avalanche voltage. Computations show that a 254 nm error in alignment can result in a 6-volt differential in electron injector breakdown voltage between the left and right halves of the cell designs. The layout of the injectors is such that the misalignment is added to the spacing of half the injectors, and subtracted from the others. While individual cells and some arrays of 16 bits could be found, it is doubtful that many 32-bit arrays would have worked, even had the design error previously mentioned not occurred. And probably those 32-bit units that might have worked would have exhibited a different BV between the odd and even rows (or left and right-handed cells).

In spite of these problems, we have characterized a number of arrays and cells. In general, storage time is quite long with charge obeying a log time dependence. For every decade of time, the same percentage of original charge is lost. Figure 18 shows an example of storage characteristics. The rate of decay has been determined to be less than 1%/decade at 85°C; this extrapolates to more than 90% of the original charge remaining after 100 years. Of course, we cannot actually predict that the charge will remain that long unless we actually wait the full time; but to date, we have not measured any examples of deviation from this law, unless a particular device was exhibiting poor storage characteristics to start with.

Figure 19 shows an example of degradation of the write/erase window with repeated write/sense/erase/sense cycling. In each case, the write/erase pulses are −50 volts at 100 ms. Figure 20 shows the cycling continued to 10^5 cycles. In spite of significant decay of the W/E window, there is still a measurable and usable window. For applications requiring a larger number of W/E cycles, differential sensing can be used as opposed to the more desirable direct sensing; but the devices would at least be usable.

Figure 21 shows a comparison of the W/E degradation between the three bootstrap cell designs. As expected, the smaller sized bootstrap device shows less erase with the same voltage conditions because the bootstrap coupling coefficient is smaller. But it also shows less W/E degradation because less current is required to be injected through the oxide to achieve the W/E operations. Hence, there is less trapping because the trapping depends only on the amount of charge passed through the oxide, and not on the avalanche time.

Figure 22 shows that the W/E window is increased by lengthening the width of the W/E pulses. In all these experiments, devices which show partial bootstrap erase were selected so that the initial effects of decay at the off-level can be seen. A device which completely erased would pass through the threshold level, and off-level decay would not be seen.

Figure 23 shows an effect called recovery wherein a device is initially cycled and the W/E window is recorded. After storage, some of the traps are thermally emptied, and the W/E window may partially or completely recover its original characteristic. Again, the smaller capacitor devices appear to show more complete recovery than the larger capacitor devices because fewer of their traps are filled under equivalent cycling conditions.

Significant progress has been achieved toward development of a semiconductor nonvolatile reprogrammable memory technology. From the original concept, three design iterations have been required to achieve workable cells. The development of the low-voltage electron injector, the search for usable hole current, the development of the n^+/p hole injector, the concept of the bootstrap capacitor, and the implementation of all these ideas into a single, addressable integratable cell design are highlights which clearly stand out as milestones in our progress. The major remaining problems,

including (1) alignment sensitivity of the electron injector, (2) excessive cell size, and (3) degradation of the W/E window will be studied in the next phase of this contract, where improvement of the DIFMOS device characteristics will be the overall goal.

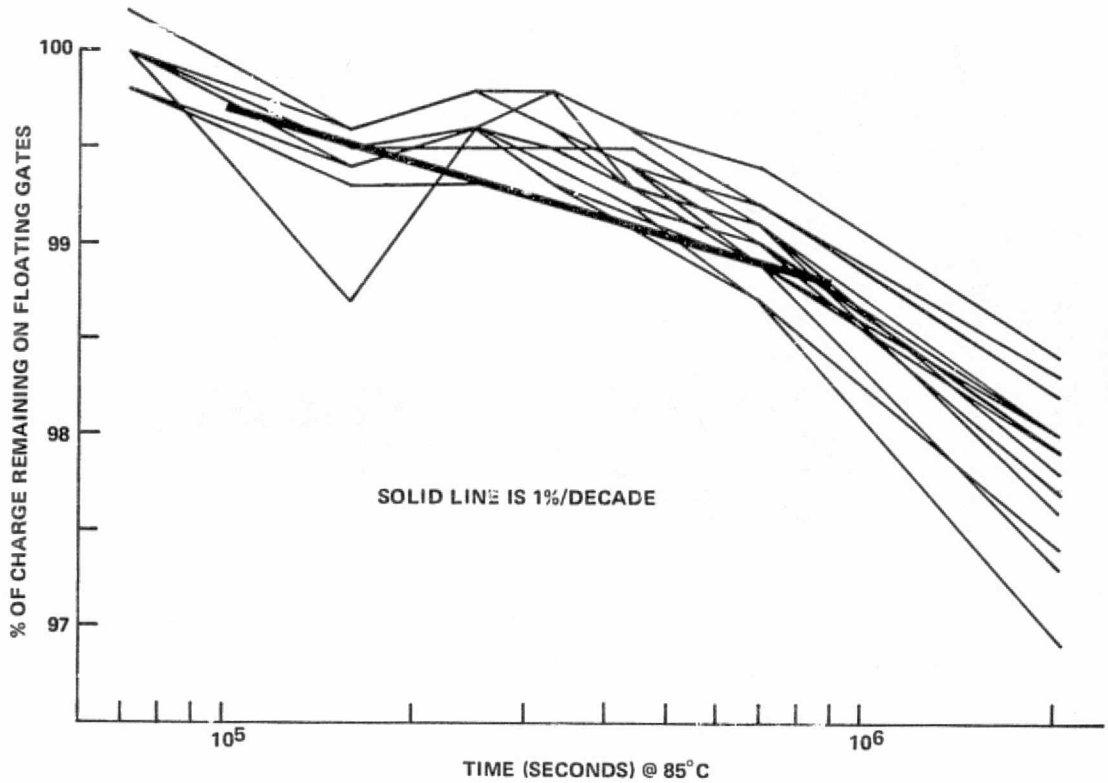


Figure 18. Example of Charge Retention on Floating Gate Array

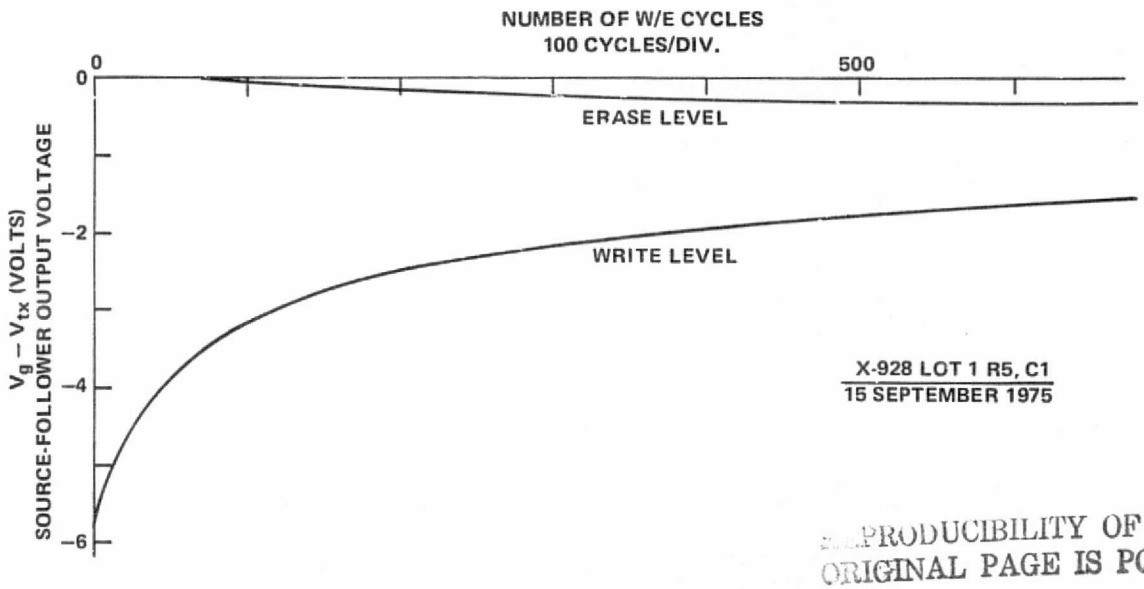


Figure 19. Degradation of W/E Envelope versus Number of W/E Cycles

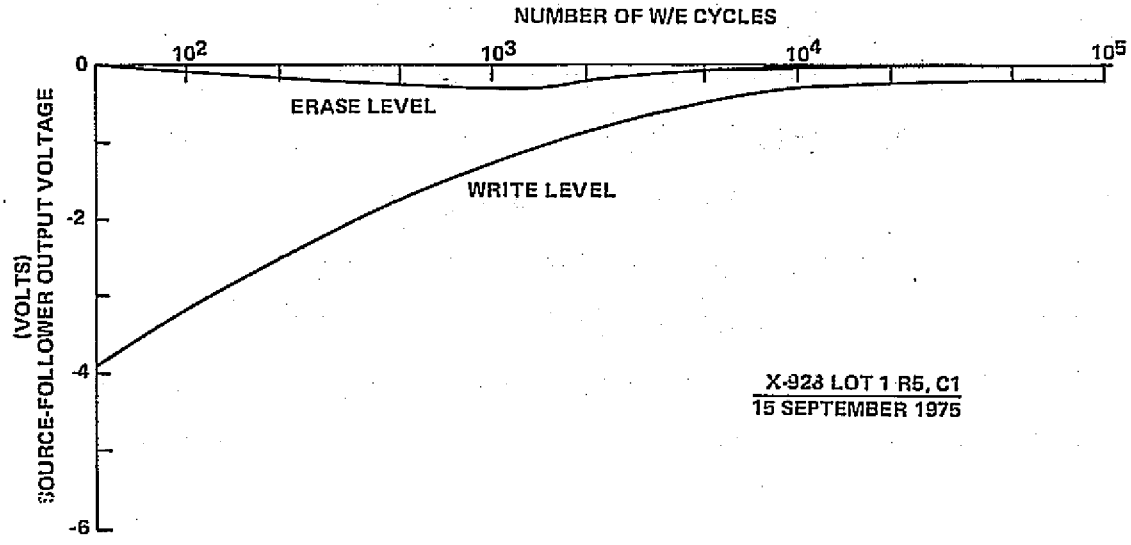


Figure 20. Log Plot of Degradation of W/E Envelope versus Number of W/E Cycles

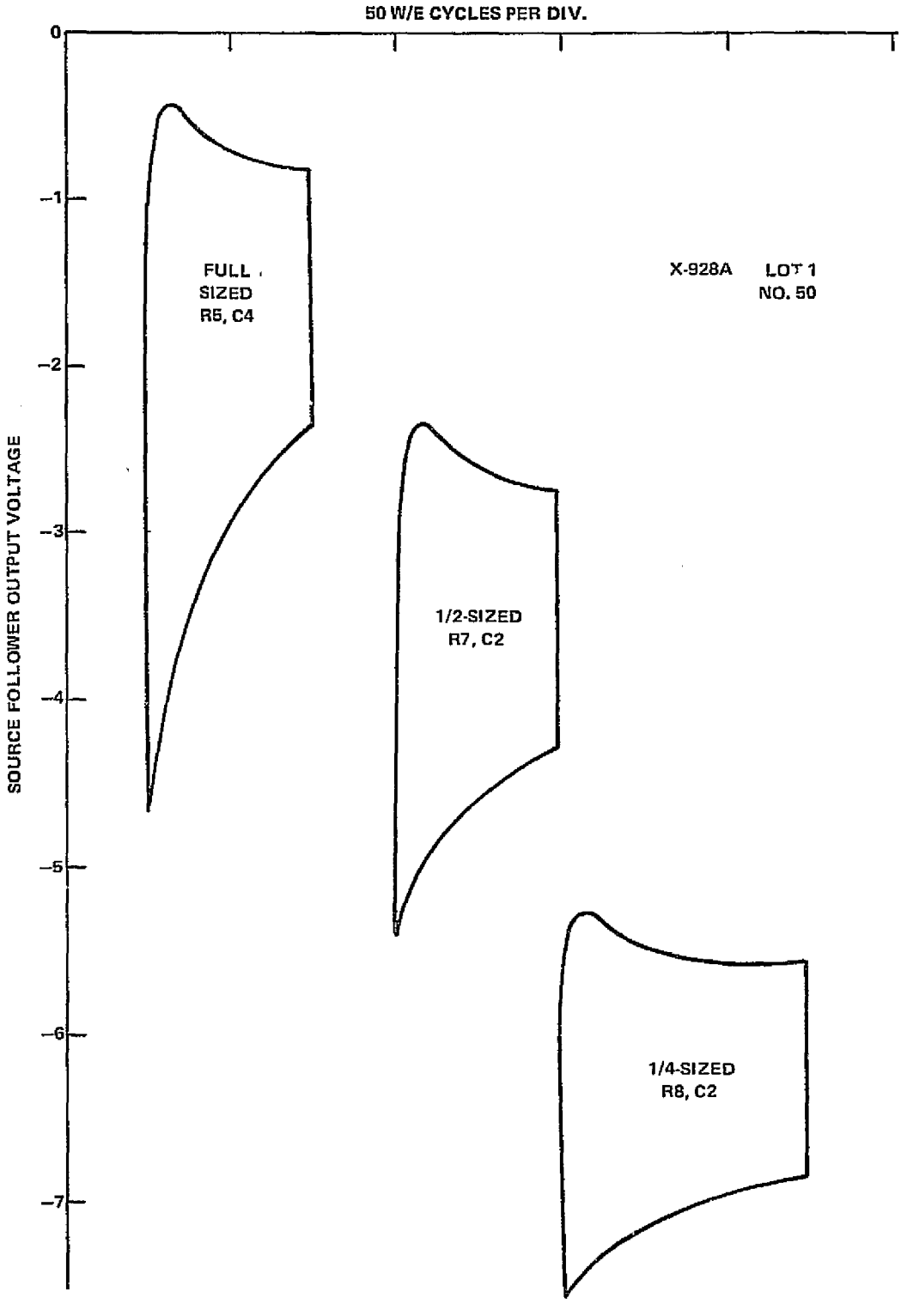


Figure 21. Comparison of W/E Characteristics of Full, One-Half, and One-Quarter Sized DIFMOS Cells

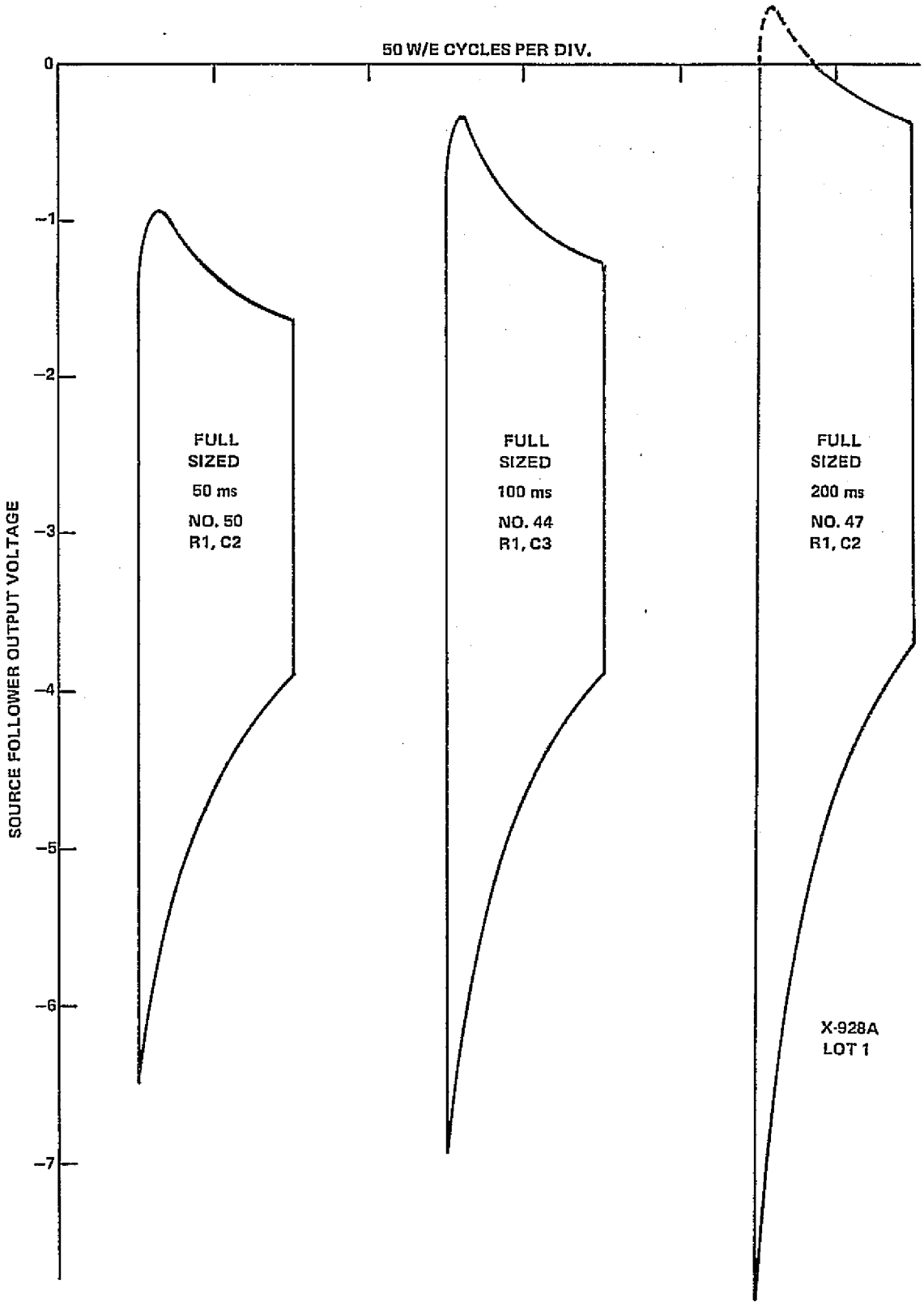


Figure 22. Comparison of W/E Characteristics of Full-Sized DIFMOS Cells versus Pulse Widths

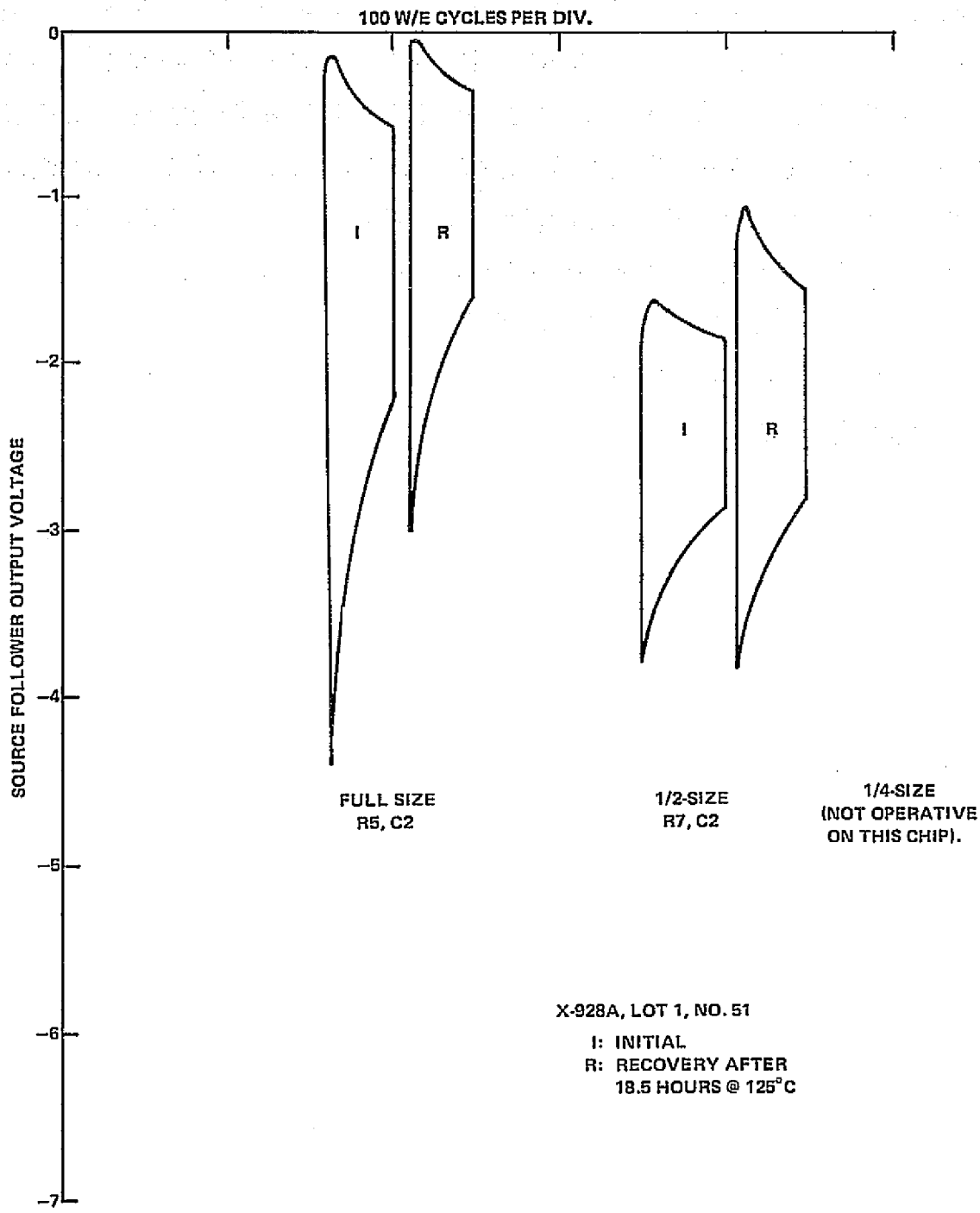


Figure 23. Recovery of W/E Window Following Storage at 125°C

SECTION VI

FIFTH EXPERIMENTAL SERIES

In this last experimental series, significant improvement in the performance of the DIFMOS devices was attained through a combination of design and process improvements. Cell size was reduced from 77.4 nm^2 per bit down to approximately 10.3 nm^2 per bit. The alignment sensitivity of the gap electron injector was eliminated by the use of an implanted electron injector. Also, the deep hole injector diffusion has been replaced with a shallow implanted diffusion which results in a shortened cycle time. Smaller cell size, simpler array designs, and process improvements result in faster programming with lower voltages and reduced degradation with write/erase cycling over the previously designed devices.

A. THEORETICAL STUDY OF PREVIOUS DEVICES

The electric fields required for programming the DIFMOS devices are considerably greater than the normal operating voltage encountered in PMOS circuits. The worst-case voltages are encountered during the bootstrap operation of the floating gate voltage which is required to initiate hole injection. There is some concern that peak voltages encountered during the operation of the DIFMOS device may be dangerously approaching the maximum dielectric field strength for the gate oxide. Therefore, it is desirable to determine what the electric fields are in the DIFMOS device under all operating conditions.

A concise closed-form solution for the electric fields in the insulators of the DIFMOS devices does not exist, because the regions of interest in the insulator involve a two-dimensional boundary value solution of Poisson's equation. However, a number of approximations to the fields may be used. One of the most detailed solutions of the fields in an oxide-passivated, avalanching junction has been done by Bulucea et al.[9] Bulucea's solution was achieved by relaxation techniques in arrays of 80 by 60 and 80 by 20 points. He concluded that under the influence of a control gate field, the peak electric field in the oxide occurs at the Si/SiO_2 interface at the metallurgical junction of the avalanching diode where strong field crowding occurs. Bulucea et al.[9] found that the maximum field strength in the oxide near the edge of the avalanching junction is in excess of 1.4×10^6 V/cm for 100 nm oxides. Although this value of maximum field strength is much greater than the 5×10^5 V/cm value used by Grove,[10] such fields are consistent with experimental findings of avalanche injection. This reported peak field is still well below the breakdown strength of silicon dioxide which is approximately 10^7 V/cm.

Referring to Figure 24, the average peak fields present in the gate oxides over the sense transistor, the hole and electron injectors, and the bootstrap capacitor can be calculated easily under conditions of reading, writing, and erasing. These average peak field strength values are listed in Table VI. The term "average peak field strength" is defined as the average electric field in the gate dielectric under peak voltage conditions. The average field is obtained simply by dividing the voltage difference by the dielectric thickness. Average field strength does not account for higher localized field strengths which may arise as a result of field crowding. Typical numerical values have been inserted in Table VI to permit an estimation of the oxide fields during sense, write, and erase cycles. Note that the gate voltage V_G is affected both by injected negative charge (0 to −10 volts) at $V_B = 0$ and by the bootstrap voltage ($\Delta V_G = K \times \Delta V_B$ where K is the bootstrap coupling coefficient). Since Bulucea et al.[9] showed that the peak field is greater than the average field in the avalanching injector, the peak field coming closest to the oxide breakdown occurs in the hole injector during the erase cycle. The average field in the oxides during the erase cycle can be as high as -4×10^6 volts/cm, which means that the oxide integrity of the DIFMOS devices must be good. Note also that the DIFMOS devices do not erase by bootstrap action alone, since both bootstrap field *and* avalanche are required for hole injection. We thus conclude that the erase action is not due to a field-dependent bulk oxide leakage mechanism.

B. AVALANCHE INJECTOR ANALYSIS

Injection of electrons into SiO_2 from an avalanching p-n junction is analogous to thermionic emission from a heated cathode into vacuum in that production of a population of energetic carriers results in some carriers having sufficient energy to overcome the potential barrier at the interface. Thus, avalanche injection can be analyzed in a manner similar to thermionic emission, except there are many physical differences. The avalanche plasma produces the energetic or hot carriers by the action of high electric fields on the free carriers. Thus, the carriers can be "hotter" than the lattice. In thermionic emission, both the carriers and the lattice are heated. Also, the barrier heights for injection into SiO_2 are lower than the vacuum barrier. Finally, thermionic emission into vacuum has no direct analog for hole injection which can occur under suitable conditions into an insulator.

Perhaps the most comprehensive analysis of avalanche injection into SiO_2 has been done by Bulucea.[11] This paper considers the effects of surface field, barrier height, perturbation of barrier height by surface field, and energy spectrum of the carriers to obtain a number for the injection efficiency, which is defined as the ratio of injected current to avalanche current. Correlation of the injection efficiency between theory and experiment can be used to indicate the validity of the analysis.

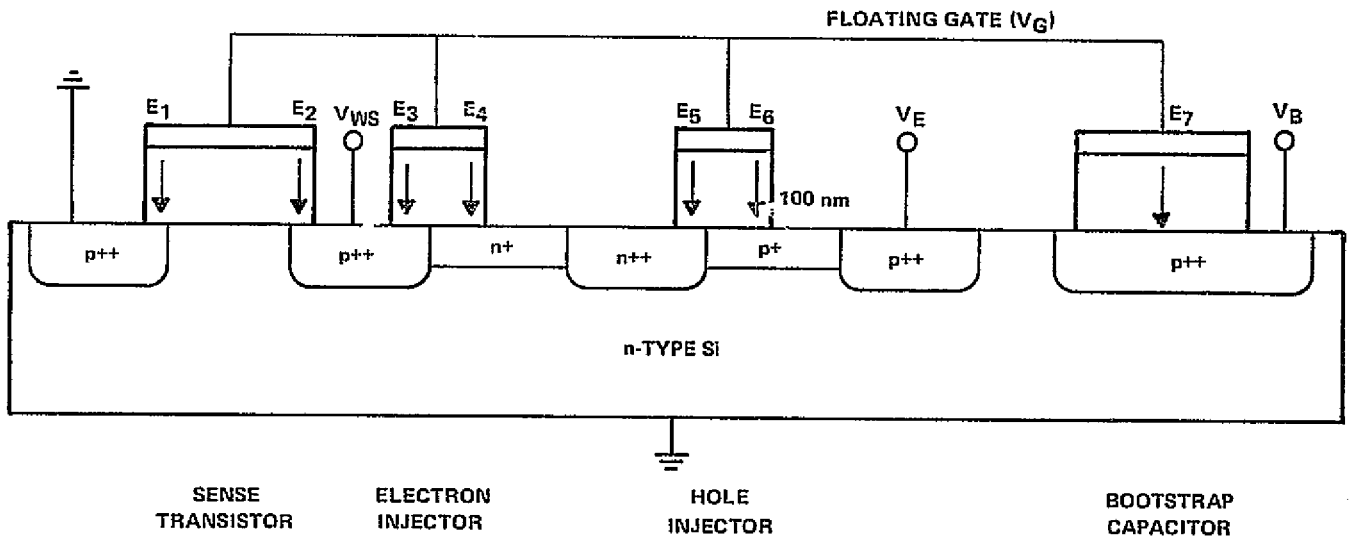


Figure 24. Schematic Representation of DIFMOS Nonvolatile Memory Cell

TABLE VI. AVERAGE PEAK ELECTRIC FIELDS IN DEVICE INSULATORS DURING READ (SENSE), WRITE AND ERASE CYCLES

		Sense		Write		Erase	
	V_G**	V_{Data}	0 to −10 V		0→ −10		−40→ −30*
	V_{WS}	V_{Data}	0 to −10 V	A_{Ve}	−20	0	
	V_E		0	0		A_{Vh}	−20
	V_B		0	0		BV	−35*
Field			V/cm	V/cm		V/cm	
E_1 (gate to source)		$\frac{V_G-V_S}{X_0}$	0 to -10^6	0 to -10^{+6}		-4×10^6 to -3×10^6	
E_2 (gate to drain)		$\frac{V_G-V_{WS}}{X_0}$	$+10^6$ to -10^6	$+2 \times 10^{+6}$ to $+1 \times 10^{+6}$		-4×10^6 to -3×10^6	
E_3 (gate to electron injector p^{++})		$\frac{V_G-V_{WS}}{X_0}$	$+10^6$ to -10^6	$+2 \times 10^6$ to $+1 \times 10^6$ (causes elec. inj)		-4×10^6 to -3×10^6	
E_4 (gate to electron injector n^+)		$\frac{V_G}{X}$	0 to -10^6	0 to -10^6		-4×10^6 to -3×10^6	
E_5 (Gate to hole injector n^{++})		$\frac{V_G}{X_0}$	0 to -10^6	0 to -10^6		-4×10^6 to -3×10^6 (causes hole inj.)	
E_6 (gate to hole injector p^+)		$\frac{V_G-V_E}{X_0}$	0 to -10^6	0 to -10^6		-2×10^6 to -1.5×10^6	
E_7 (gate to bootstrap capacitor)		$\frac{V_G-V_B}{X_0}$	0 to -10^6	0 to -10^6		$-.5 \times 10^6$ to $+.5 \times 10^6$	

* It is assumed that BV = 35 places $\Delta V_G = -30$ volts

**V_G is the floating gate voltage

The most important aspect of the injector analysis is the energy spectrum of the hot carriers. Holes and electrons in the depletion region of a reverse-biased junction gain energy from the electric fields present, and lose energy to the lattice in the form of optical phonons and in ionizing lattice collisions. The latter produces additional hole-electron pairs which then become free carriers under influence of the electric fields. The energy imparted by the electric field is superimposed upon the normal distributions of energy of the carriers in the semiconductor. Solutions of the Boltzmann transportation indicate that the overall effect of the electric field on the energy distribution of the carriers is equivalent to their having a characteristic temperature hotter than that of the lattice, and the distribution of these carriers in energy is Maxwellian at the characteristic temperature. Thus, there is a relationship between electric field and electron temperature given by:

$$kT_e = \frac{\mathcal{E}_o}{\frac{1}{2} + (\frac{1}{4} + \mathcal{E}_o/\mathcal{E}_r)^{\frac{1}{2}}} \tag{1}$$

$$\mathcal{E}_o = (qEl_r)^2/r\mathcal{E}_r \tag{2}$$

where E = electric field

l_r = mean free path for phonon-emission interactions

$\mathcal{E}_r$ = energy of the zero-wave-vector optical phonons involved in these interactions

r = ratio between the mean free path for ionization interactions l_i and l_r.[11]

The most likely values for $\mathcal{E}_r$ and l_r as used by Bulucea[11] are

$$\mathcal{E}_r = 0/063 \text{ eV}, \; l_r = 60 \text{ A}$$

Bulucea's value for r is 3.2, which means that the optical phonon generation is more probable than ionization interactions.

For the abrupt one-sided junctions used for hole and electron injectors in this study, the doping concentration was set in the 4 to 5 X 10^{16} cm^{-3} range to provide avalanche breakdown of approximately 20 volts. The maximum electric field strength in such diodes is approximately 5 X 10^5 volts/cm, occurring at the metallurgical junction. The characteristic temperature calculated for these junctions using equation (1) is 2500 K. Thus, the energy distribution of the carriers in the avalanching junctions is given by

$$n(\mathcal{E}) = \text{const } \mathcal{E}^{\frac{1}{2}} \exp(-\mathcal{E}/kT_e) \tag{3}$$

where T_e = 2500 K and $\mathcal{E}$ is the energy of the hot electrons. This is the characteristic temperature for a one-dimensional diode. Field crowding that exists near the surface of three-dimensional injector diodes can result in higher characteristic temperatures near the injection region.

Figure 25 shows the energy band diagram of an MOS structure with conditions favorable for electron injection. The barrier height of 3.25 eV from the silicon conduction band to the oxide conduction band is perturbed by the oxide electric field. The perturbation is due to Schottky or image force barrier lowering and is approximately 0.28 eV for junctions of the type used in DIFMOS device charge injectors.

Bulucea's[11] analysis of the avalanche injection phenomena indicates that the avalanche injection ratio should be between 4.3 X 10^{-4} and 4.7 X 10^{-3} for typical surface potential values.

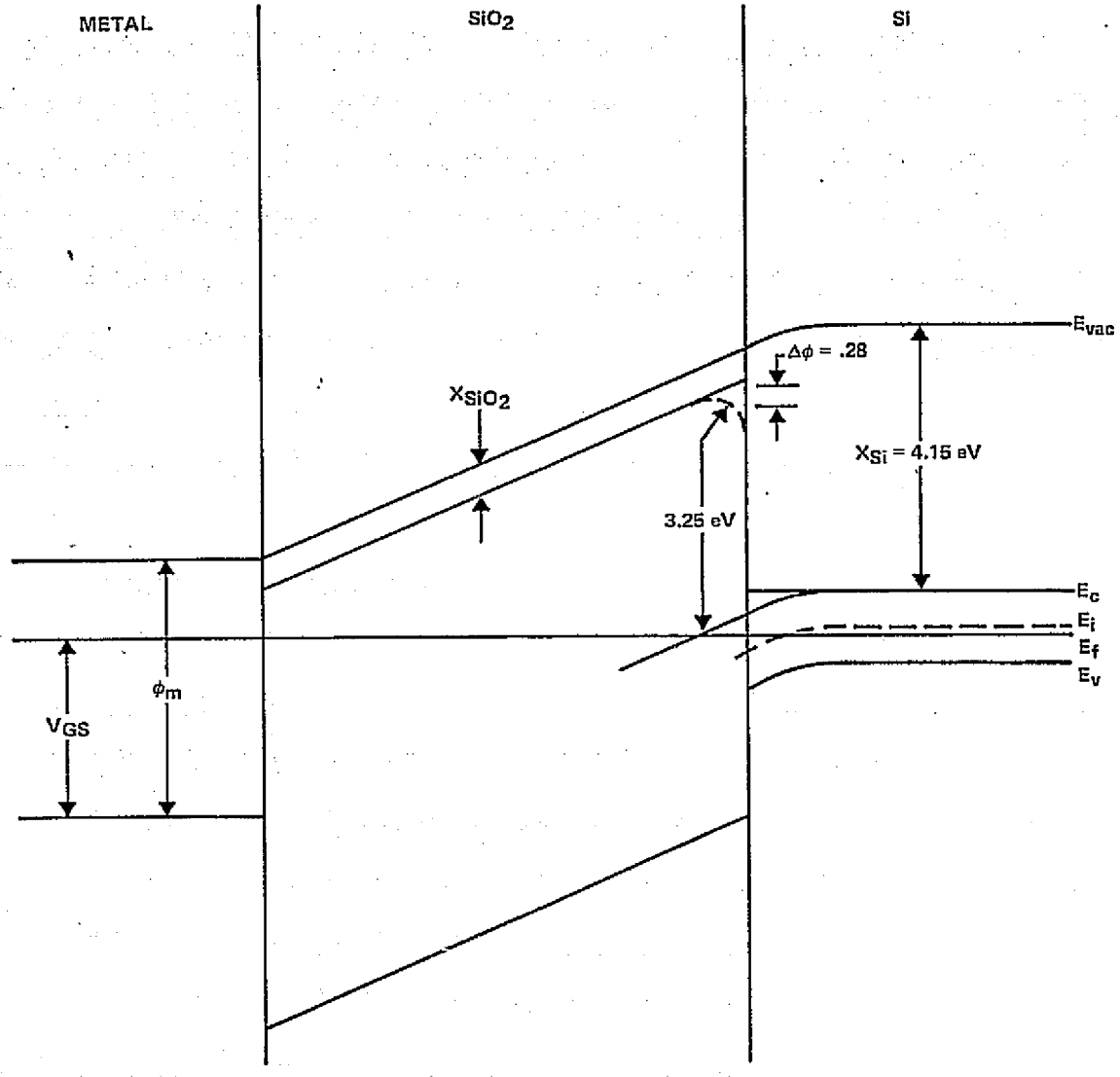


Figure 25. Energy Band Diagram for Electron Injection
Showing Perturbation of the Barrier Height by the Insulator Field (Schottky lowering)

Figure 26 shows the results of fixed-bias measurements of avalanche injected gate current from both hole and electron injector junctions. In both cases, the injection ratio is within an order of magnitude of the anticipated values calculated by Bulucea. Carrier trapping in the oxide does not permit the injection of constant current unless a variable gate voltage bias is used. The peak value of the injection ratio was calculated by fitting dominant time constants and amplitudes to the experimental data and then calculating the injected current at $t = 0$. Because the trapping rate is maximum at $t = 0$, estimates of the initial injected current are apt to be low, which agrees with the measured results being lower than the calculated values. However, the relatively close agreement between experiment and theory indicates that Bulucea's analysis is an excellent model of the injection phenomena.

1. Consideration of the Physical Reality of Avalanching Holes and Injecting Positive Charges

The physical reality of avalanche injecting hot holes from a reverse-biased junction into silicon dioxide is often questioned. This reality can be accepted if one considers the true nature of the hole which is the motion of an unoccupied state below the Fermi level.

Thermally oxidized silicon surfaces consist of an insulator having a great deal of short range order that is sufficiently crystalline in nature for the usual quantum-mechanical considerations for holes and electrons apply. But the question still remains, how does one inject an empty state into the oxide?

Referring to Figure 27, observe the energy band structure of the gate, insulator, and semiconductor under avalanche and field conditions favoring hole injection. Bulucea[11] has shown that the fields of the avalanche produce Maxwellian distributions of holes and electrons near the surface of the silicon. Since hole energy in the band diagram increases in the downward direction, one can expect some small number of very energetic holes to be formed deep in the valence band near the SiO_2 interface. These deep states can be quickly occupied by neighboring electrons. But for an electron to fill the state, it, in turn, must also create a hole. Therefore, the holes have velocity in accordance with Maxwell-Boltzmann statistics. We can thus postulate that the creation of very energetic holes near the oxide interface results in the finite (but nontrivial) possibility that the hole state is filled by an electron from the oxide. Thus, we inject holes into the oxide by creating a sufficiently deep unoccupied state in the silicon near the oxide that an electron in the valence band of the oxide can move into the unoccupied state deep in the silicon band. Hole injection takes place by the motion of unoccupied electron states.

An obvious objection to this idea is the slight probability of filling the deep state from the oxide rather than the silicon. This is exactly the case that was observed experimentally. The measured injection probability, defined as injected gate current/avalanche current, is typically of

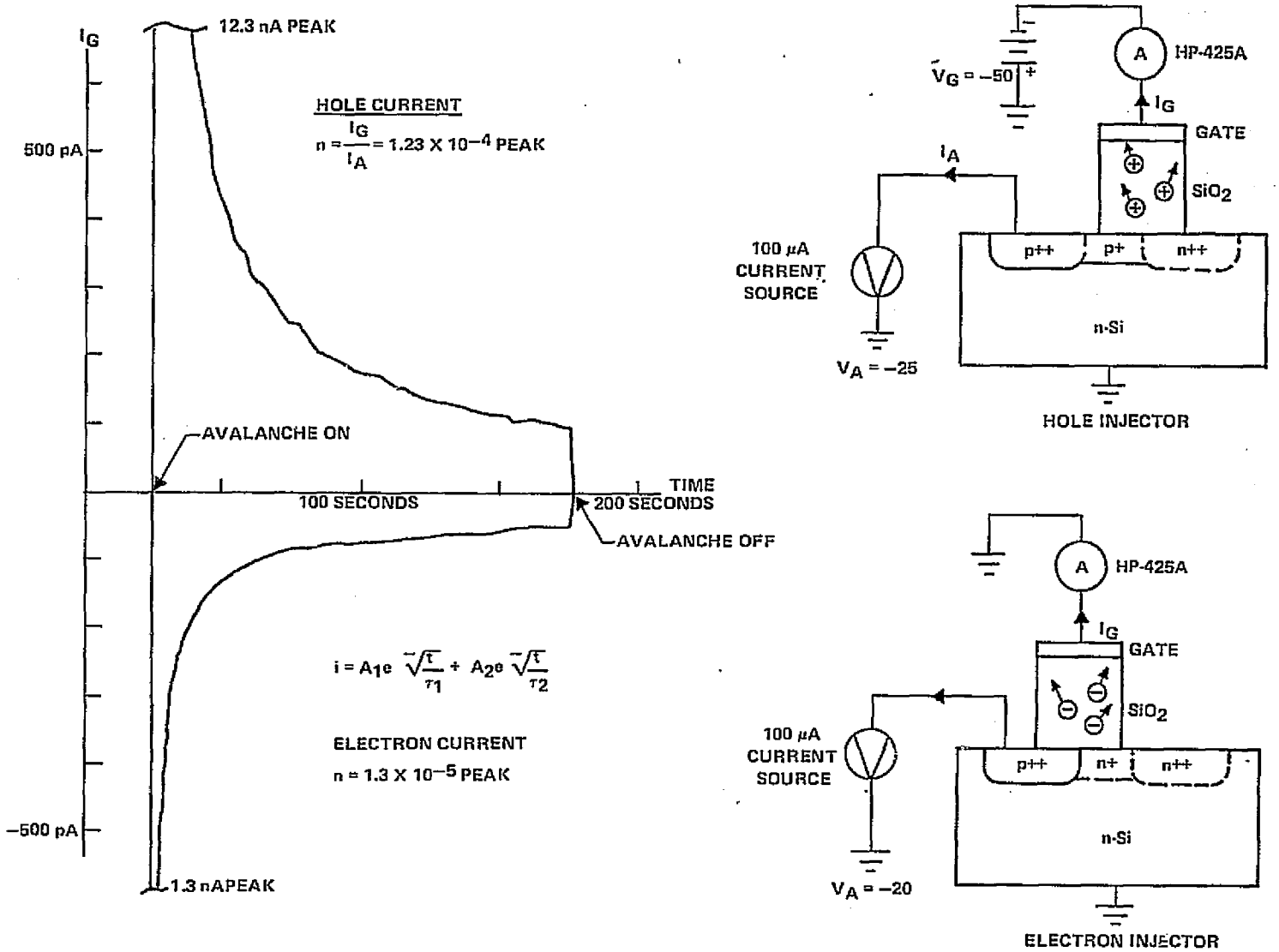


Figure 26. Avalanche Injection of Holes and Electrons into SiO_2 Fixed Gate Bias

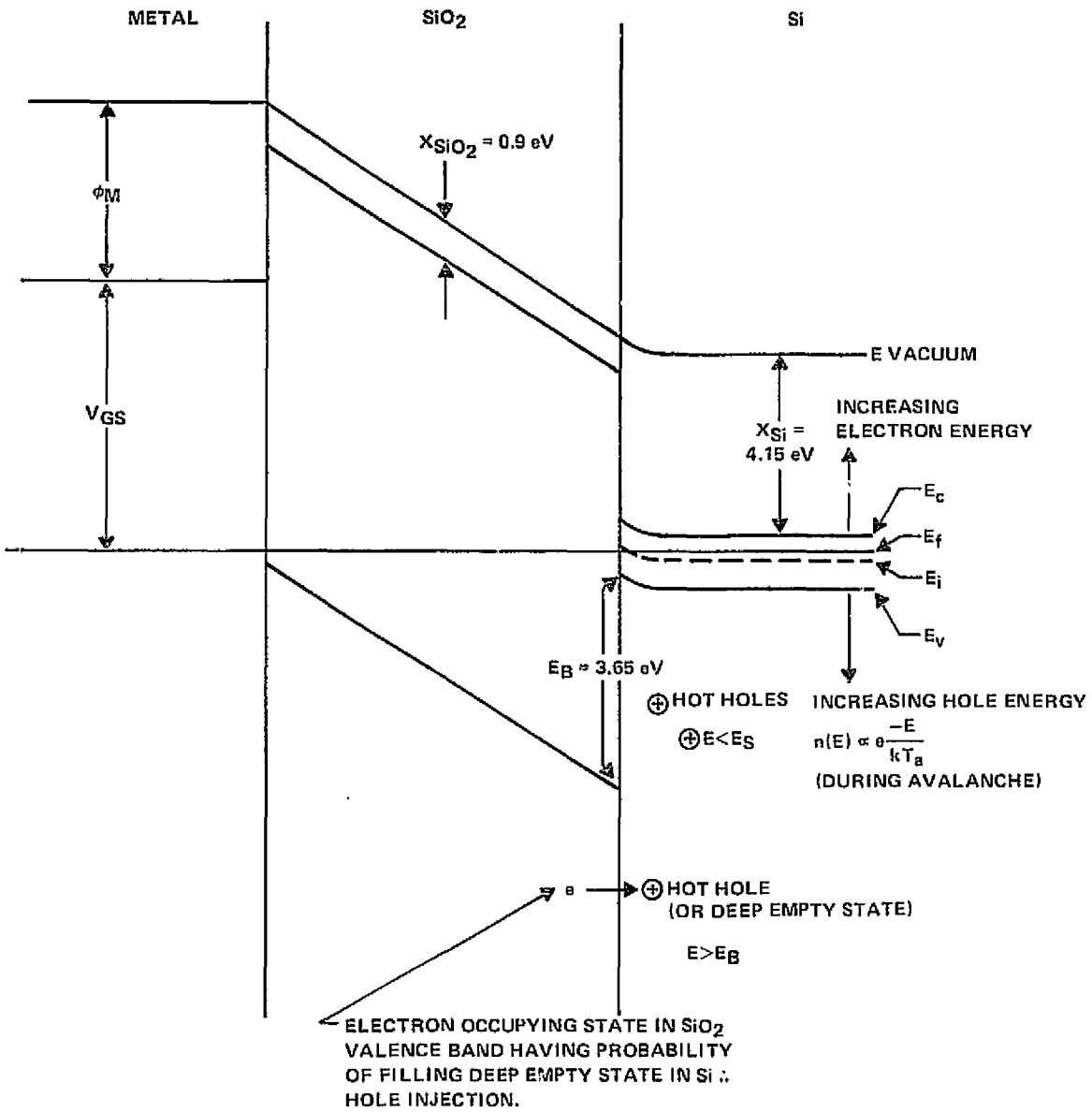


Figure 27. Energy Band Diagram Showing Mechanism of Hot Hole Injection Whereby Empty States are Created Deep in the Valence Band During the Avalanche. Filling these states by electrons from the SiO_2 constitutes hole injection.

the order of 10^{-4} for both holes and electrons. Thus, only a small fraction of the energetic carriers participating in the avalanche actually end up being injected into the oxide with sufficient energy to overcome the surface barriers.

C. FIGURE OF MERIT

Relative performance of nonvolatile memory devices can be compared by the use of a figure of merit. A figure of merit for the DIFMOS devices can be arbitrarily defined by the following expression:

$$f_M = \frac{\Delta V_g R}{N_y N_m N_p A(t_w + t_e) t_r V_B (V_w + V_e)(\Delta V_w + \Delta V_e)(i_w + i_e) D L_r L_c V_{pp}(1 + C)}$$

Where

f_M = figure of merit

ΔV_g = change in gate voltage from on to off states

R = number of write/erase cycles to failure

N_y = age of the technology in years

N_m = number of mask levels

N_p = number of major process steps (including mask levels)

A = size per bit in an array (nm^2)

t_w = write avalanche time (seconds)

t_e = erase avalanche time (seconds)

t_r = read time (seconds))

V_B = erase bootstrap voltage (volts)

V_w = electron injector avalanche voltage (volts)

V_E = hole injector avalanche voltage (volts)

ΔV_w = degree of process control of the electron injector avalanche voltage (volts)

ΔV_e = degree of process control of the hole injector avalanche voltage (volts)

i_w = electron injector avalanche current (amps)

i_e = hole injector avalanche current (amps)

D = decay rate of the data in %/decade

L_r = number of row lines/bit in an array

L_c = number of column lines/bit in an array

V_{pp} = program voltage which must be applied to an array for writing and erasing (volts)

C = number of metal-to-silicon contacts/bit in an array

In general, every factor which improves DIFMOS by getting bigger was placed in the numerator, and every factor which improves DIFMOS by getting smaller was placed in the denominator.

Let us also define a relative figure of merit f_R for comparing two differing devices as the ratio of their respective figures of merit. Thus, a 2X improvement in almost any one factor will result in a 2X improvement in the relative figure of merit.

Using this figure of merit, the performance of previous designs has been rated. The best device designs shown in the Fourth Experimental Series exhibited a figure-of-merit rating of 0.145, based upon the gap-injector process. Substitution of an implanted injector for the electron injector to eliminate the alignment sensitivity raises the rating 0.183, a 1.25X relative improvement. Other devices prior to these were not functional or exhibited a poor relative rating compared to the 0.145 value listed above. The 1.25X improvement was predominantly due to reducing the electron injector avalanche voltage variation caused by process-induced gap variations in the gap-type electron injector.

The Fourth Experimental Series device designs were operating devices but they had a number of problems. The major problem encountered was the extreme sensitivity of the avalanche voltage of the $p^+/n/n^+$ gap-type electron injector. In addition, the cell size was extremely large (77.4 nm^2/bit), imposed by the 1 row/3 column array organization and the 12.7 μm p^+-to-n^+ guard ring spacing. Finally, operating the bootstrap capacitor as part of the source-follower load through an address transistor required extremely high voltages to operate the array (over 45 volts for the undecoded arrays). While the basic injectors performed well, it was felt that improvements could be made in process and design to raise the device figure of merit.

D. DESIGN RULE CHANGES

One idea developed in the Third Experimental Series and later incorporated into the Fourth Experimental Series devices was the elimination of the gap-type electron injector by implanting into the gap region. The implant dose could be adjusted to set the avalanche breakdown voltage at nearly any desired level completely independent of alignment problems. Thus, voltage variations in the injectors were eliminated.

In the move toward higher densities, the 7.0 μm deep p-type diffusion which sets the hole injector avalanche voltage can be eliminated in favor of a shallow 1.25 μm deep implanted p-type region. Both of these injector diffusions can be implanted just before gate oxidation, and activated during the gate oxide growth. This gives good independent control over both injectors, shortens the process, and offers the potential for higher circuit density by reducing the p-type region lateral diffusion.

The p^+ to n^+ guard ring spacing can be reduced from 12.7 μm to 10.2 μm without any jeopardy, even though the substrate resistivity is in the 6 to 8 ohm-cm range. The $p^+/n/n^+$ diffusion to guard ring breakdown mechanism is similar to that of pin-type diodes, and overlap of their space-charge regions is not harmful. (This differs from space-charge overlap from p^+ to p^+ which amounts to punch-through in an MOS transistor.)

Finally, reorganization of the cell into arrays in either a one row/two column or two row/one column addressing scheme, rather than the previous one row/three column design, permits more compact array designs. Bootstrap voltage requirements can be lowered by operating the bootstrap diffusion directly from the row or column line without passing through an address transistor. Finally, combining as many functions as possible on the row and column lines, and attempting as dense a layout as possible will result in smaller cell sizes.

Ion implantation is a technique that will be used extensively for setting the injector avalanche voltages. It also has the potential use for adjusting device V_{tx}, forming depletion load transistors, and partially eliminating guard rings by implanting field shields to give maximum field threshold voltage and junction breakdown voltage. For this study, only injector avalanche voltages will be adjusted by ion implantation.

E-beam lithography offers the possibility for smaller pattern definition, but pattern definition is not presently the limiting size factor in DIFMOS. Instead, space-charge widths required to support the programming voltage impose pattern sizes which can adequately be defined with conventional photomasks; therefore, e-beam lithography is not required to improve cell size.

Alloy insulators have been considered in the third experimental series without any encouraging results. It appears that the oxidized silicon surface provides the optimum insulator, with minor variations in the growth process being the key to improving the insulator properties.

All previous designs have utilized aluminum metal for the floating gates. However, the storage time of aluminum metal gate devices is limited by surface leakage of charge which is affected by moisture. Polysilicon gates which can then be thermally oxidized may provide a superior gate from a storage standpoint. However, no other significant factor is expected to be affected. Designs incorporating both silicon and aluminum gates will be submitted, but we see no advantage in other gate materials at this time. Unless an obvious advantage for some other material arises, the semiconductor industry will probably consider aluminum or silicon as the accepted gate material in the near future.

E. DEVICE DESIGNS

The following device designs are submitted to fulfill the requirements of Task XIV. Although only five designs were required, seven designs were submitted. These designs offer variation in design layout rules, bootstrap capacitance ratio, and metal gate versus silicon gate. One approach offers a different circuit approach which excludes contacts from the cell while utilizing aluminum gate technology. In each case the design of the electron and hole injectors is the same proven, implanted design as that used in the Fourth Experimental Series.

The electron injector is designed with 7.6 μm spacing between p^{++} and n^{++} regions with an n^{+} implant between them to reduce the effect of misalignment on V_a. The e– avalanche injection actually occurs at the p^{++}/n^{+} junction. The hole injector is designed with a p^{+}/n^{++} junction. In all test devices except one, the hole injectors maintain large spacings between the p^{++} and n^{++}. In laying out the contactless cell (CNLES), the p^{++} to n^{++} spacing was reduced to the minimum allowed 7.6 μm (aggressive layout rules). This design is still the same as the others in that it relies on p^{+}/n^{++} avalanche, but it also makes maximum use of the layout rules to reduce the size of the memory cell.

The remaining paragraphs will give specifics concerning each test device. The test devices include STAGG, AGGNP (which is populated by memory cells AGGM1 and AGGM2), NOCON (which is populated by memory cell CNLES), COSGT (which is populated by memory cell CONSG), and AGSGT (which is populated by AGGSG). NOCON is the metal gate, contactless design. COSGT and AGSGT are silicon gate designs.

STAGG (Figure 28) was designed to allow comparison of the conservative layout rules used in the layout of the cell in the 256-bit array and the more aggressive layout rules which allow:

1) 5.1 μm X 7.6 μm contacts
2) 5.1 μm minimum diffusion widths
3) 7.6 μm minimum spacing between diffusions
4) 7.6 μm/7.6 μm minimum metal line widths and spacings

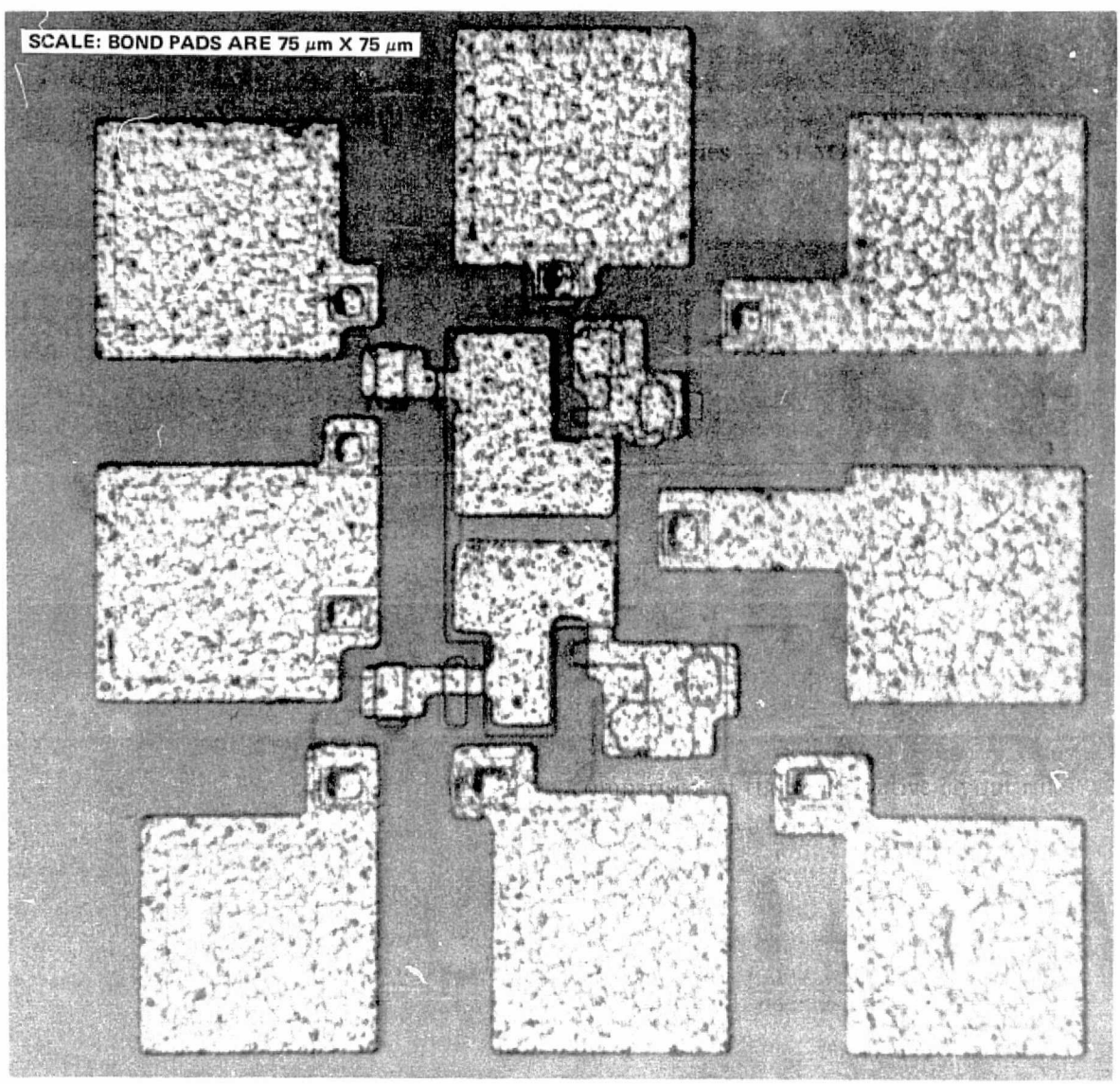


Figure 28. Fifth Experimental Series – STAGG

The test device is drawn in discrete form to allow direct access to each injector, the bootstrap capacitor, and the sense transistor. The test device also makes minimum use of n^+ guard rings (i.e., placed only in areas where metal would cause V_{tf} turn-on between unrelated p^{++} lines). This approach to guard rings was used in all test cells.

AGGNP (Figure 29) was designed to house memory cells AGGM1 and AGGM2. It simulates address transistors leading to each row and column, accesses both injectors in each cell type, accesses both plates of the bootstrap capacitor in each cell type, and accesses the drain of the sense transistor in each cell type. AGGNP, AGGM1, and AGGM2 utilize the aggressive layout rules.

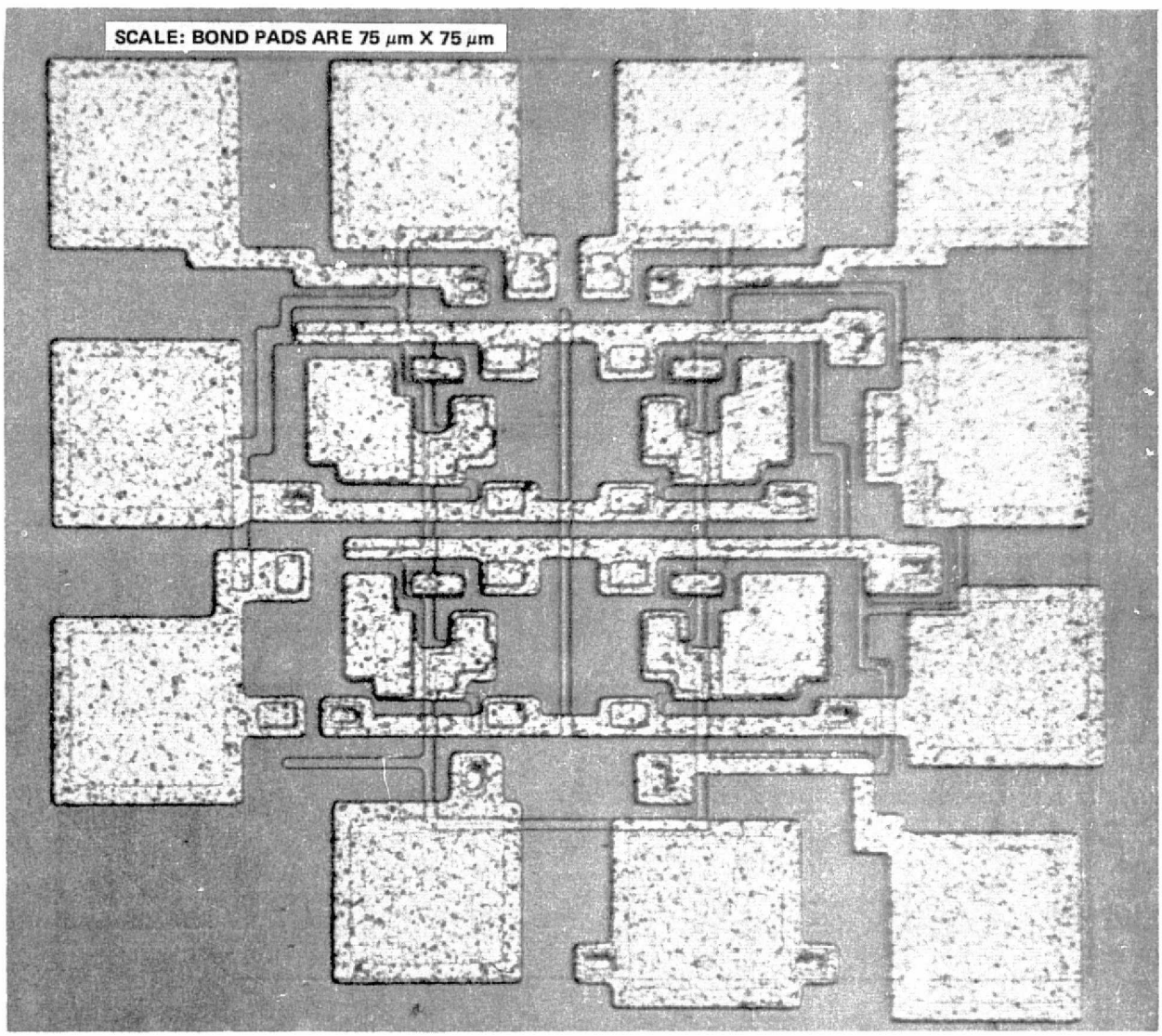


Figure 29. Fifth Experimental Series – AGGNP

AGGM1 and AGGM2 are identical except for the size of the bootstrap capacitor:

$$\text{AGGM1 has } \frac{C_B}{C_B + C_L} = 80\% \tag{4}$$

$$\text{AGGM2 has } \frac{C_B}{C_B + C_L} = 60\% \tag{5}$$

The layer cell sizes of AGGM1 are apparent in Figure 29. The purpose of altering the C ratios is to determine what minimum ratio is required for satisfactory erasure. Other than the use of aggressive layout rules, the differences between AGGM1 and AGGM2 and standard cell used in the 256-bit array are:

1) Column line is all diffusion rather than alternating metal and diffusion. The cell size reduction allowed by the omission of two contacts results in a shorter cell with consequent reduction of p^{++} resistance. Additionally, the $\approx$100 ohms of contact resistance is also removed. The net effort is to make the cell diffusion columns quite competitive with the hybrid metal/diffusion column line.

2) Zero spacing between sense transistor channel and electron injector. This has not been done in the other cells because of uncertainty about the effect of the n^- implant on the sense transistor turn-on characteristics.

In both memory cell types the p^+ to p^+ spacing between the column and the sense transistor drain were maintained at 10.2 μm to assure punch-through did not occur when both are brought to a large negative voltage.

NOCON (Figure 30) was designed to house a 3 X 4-bit array of 12 CNLESS cells (Figure 31). No effort was made to obtain direct access to the injectors, bootstrap capacitor, or sense transistor on the assumption that injector operation and charge storage could be assured as functional on the other test devices. The unusual arrangement of the CNLESS cell made us much more desirous of knowing its behavior in an actual array. The array is sufficiently large to get a view of (1) possible cross coupling between bits on all sides of the bit, and (2) voltage drop in the two column lines.

No column or row address transistors were used because of the unknowns about voltage drop in the column lines and voltage required in the row line to sufficiently turn on the two select transistors to allow sensing between the two column lines. The unusual structures under the four row control pads are input static charge devices consisting of a V_{TF} device in series with a large resistor.

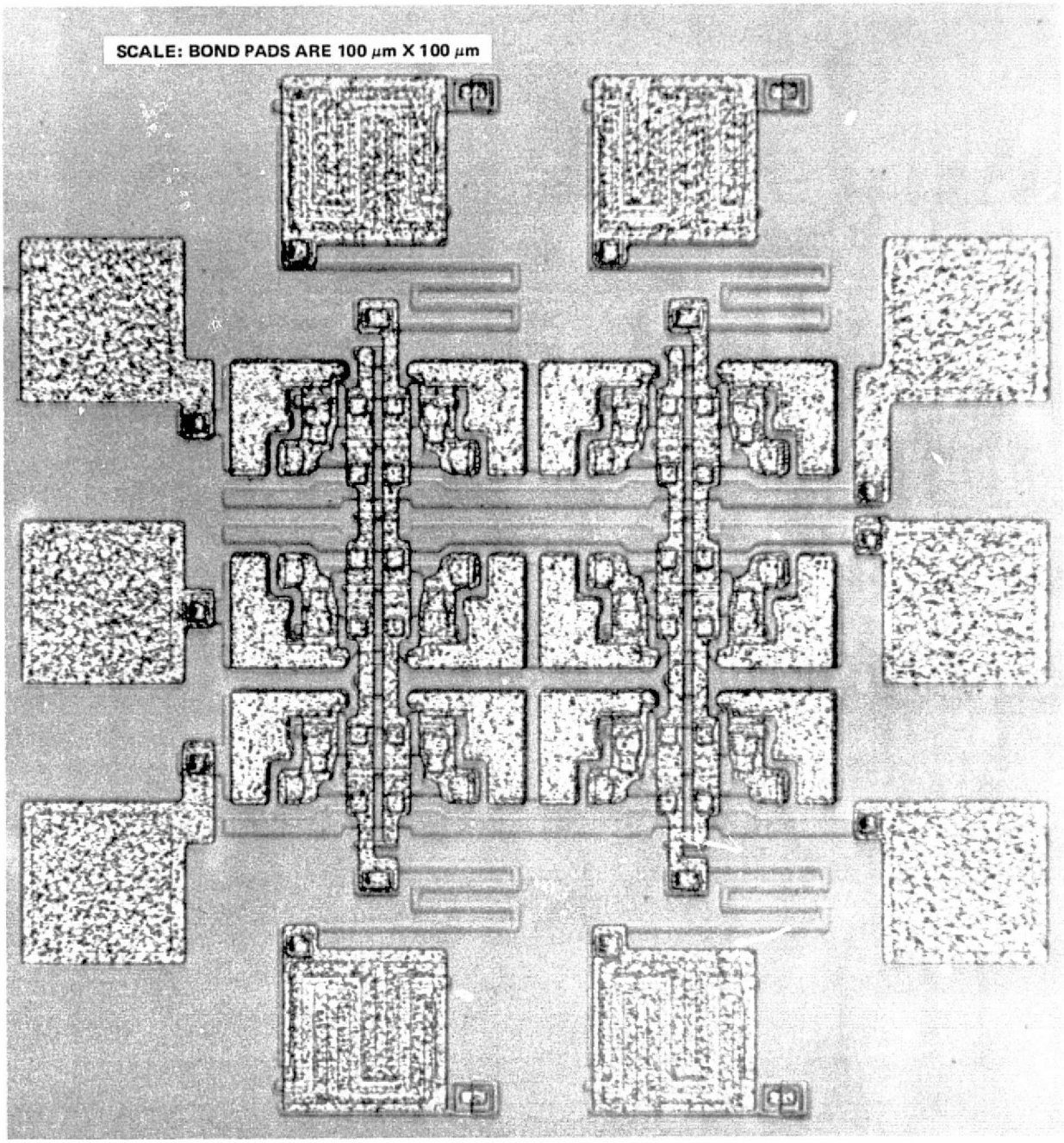


Figure 30. Fifth Experimental Series – NOCON

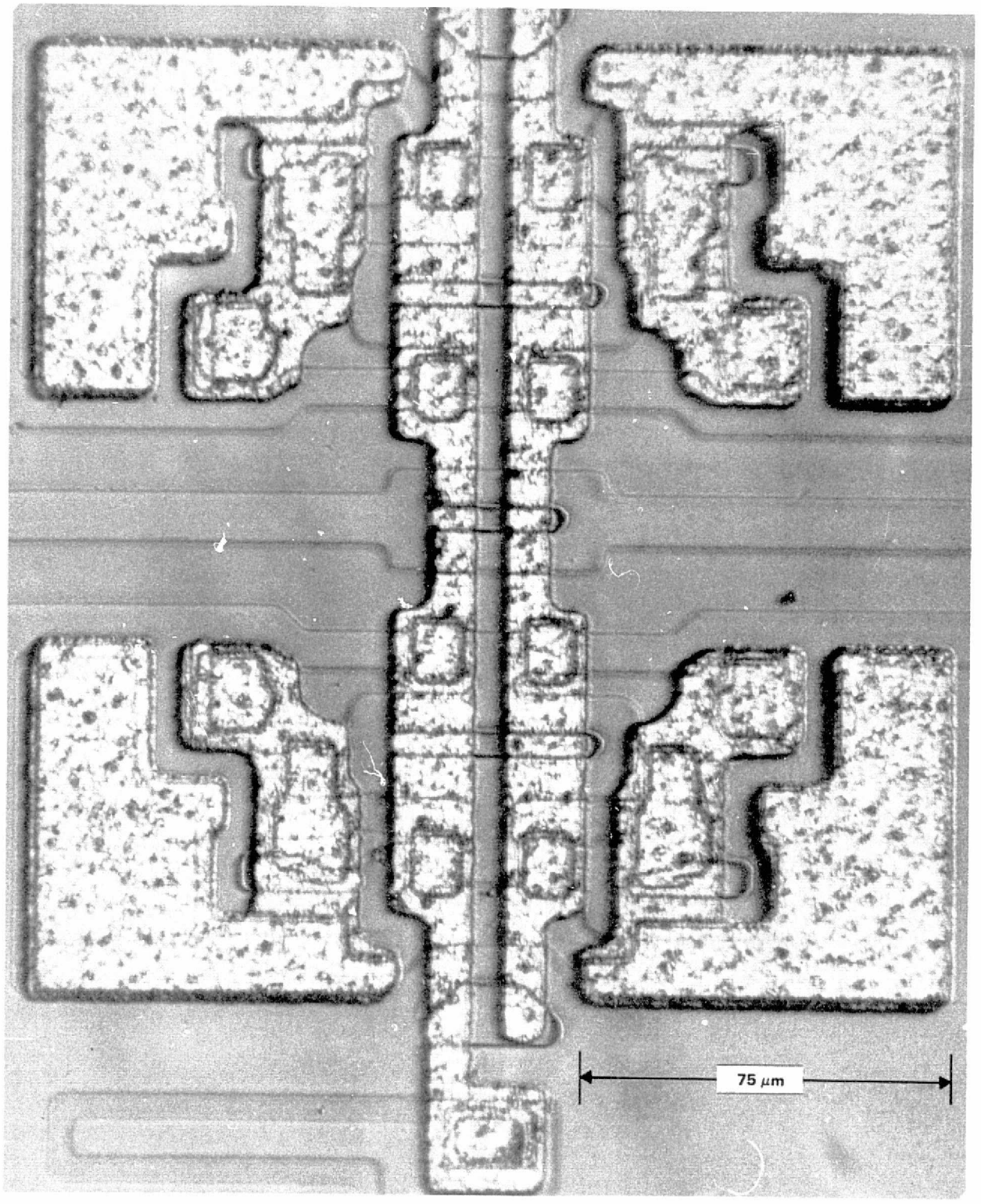


Figure 31. Fifth Experimental Series – NOCON

COSGT (Figure 32) houses CONSG and AGSGT (Figure 33) houses AGGSG. They represent, respectively, a conservative layout rule approach to silicon gate DIFMOS versus an aggressive layout rule approach to silicon gate DIFMOS. COSGT and AGSGT are functionally identical and allow the following:

1) Direct access to both injectors
2) Direct access to the column lines
3) Address transistors to the column line and both row lines
4) Sense transistor simulation for indirect determination of gate voltage on sense transistors in cell

All these device designs are capable of being implemented in array designs in that required voltages and currents can be switched by on-chip address transistors.

F. PROCESSING

The DIFMOS processing sequence including the implanted hole and electron injector implants is listed in Table VIIA. Table VIII compares the DIFMOS process to basic PMOS metal-gate, p-channel FAMOS, CMOS, and MNOS. DIFMOS is roughly equivalent to CMOS in processing complexity. Although it has one more mask level and processing step, the extra implant diffusions are much shorter than the single CMOS p-tank isolation diffusion, so the CMOS and DIFMOS process are roughly equivalent. Another way of comparing the DIFMOS process is to consider it as basic metal-gate p-channel MOS with three extra diffusion and patterning steps for the guard rings and injectors. In summary, DIFMOS has the processing complexity of CMOS, and it offers writable/erasable nonvolatile EAROM capability of MNOS. The simpler FAMOS process is not electrically erasable, and MNOS, while being electrically erasable, is a much more complicated process than DIFMOS.

The process can be modified to provide silicon gates by the insertion of three steps listed in Table VIIB. This is a simplified approach to permit evaluation of storage properties of oxidized poly-silicon gates. Using this modification, all MOS gates would be poly-silicon, and no aluminum gates would be functional unless further additional steps were added. Also, there would be no poly-silicon to silicon contacts. The poly-silicon would have to be doped during deposition if a low sheet resistance is desired.

G. TESTING

Within the time frame of this contract, five lots of material of the Fifth Experimental Series devices were processed. The first two lots of material were processed more or less simultaneously using the old CMOS-type 7.0 μm deep p-type isolation region as the hole injector diffusion. These first two lots of material showed that the smaller cell designs were inoperative due to lateral

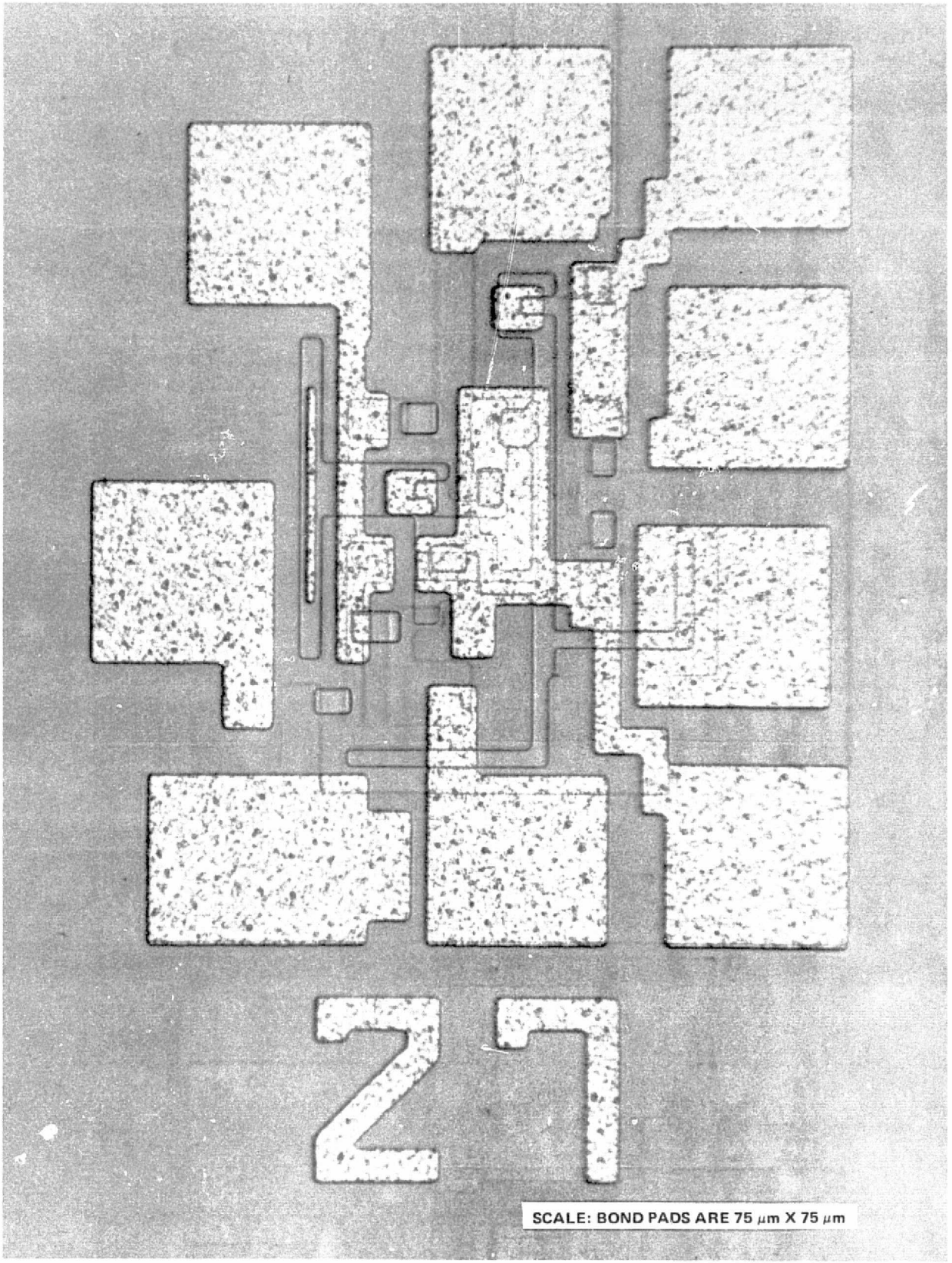


Figure 32. Fifth Experimental Series – CNSGT

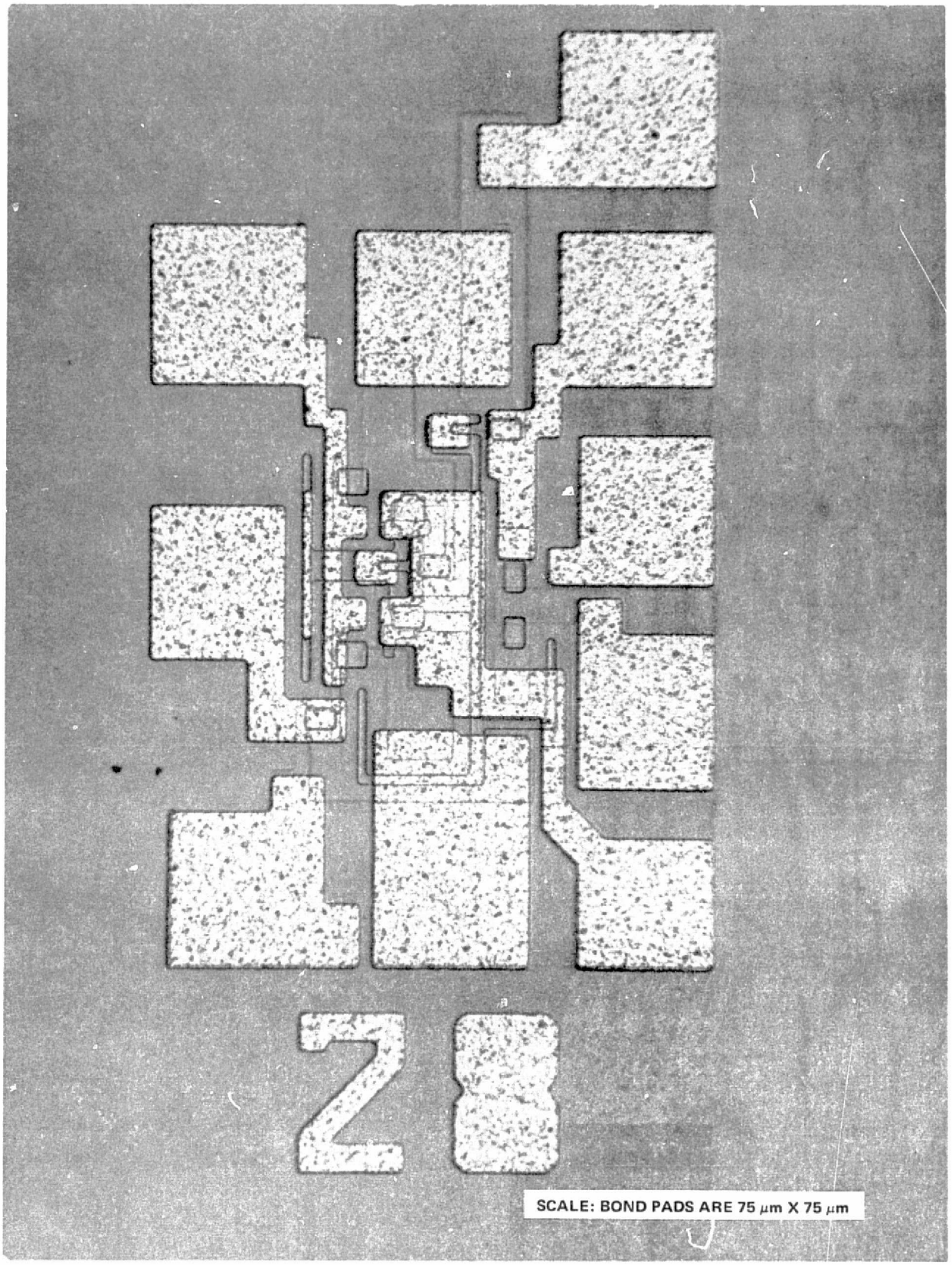


Figure 33. Fifth Experimental Series – AGSGT

TABLE VII. DIFMOS PROCESS SEQUENCE

A. Final metal-gate DIFMOS process sequence including implanted electron and hole injector diffusions.

Starting Material: 6 ohm-cm n-type silicon

*1. Field oxide growth
*2. Pattern p^{++} source and drains
*3. Diffuse p^{++} regions
4. Pattern n^{++} guard rings
5. Pattern n^{++} regions
*6. Pattern gate oxide regions
7. Pattern p^{+} hole injectors
8. Implant p^{+} hole injectors
9. Pattern n^{+} electron injectors
10. Implant n^{+} electron injectors
*11. Grow gate oxide
*12. Pattern metal-to -silicon contacts
*13. Deposit metallization
*14. Pattern metallization
*15. Deposit protective overcoat
*16. Pattern bond pad openings

* Indicates steps in basic metal-gate p-channel MOS process

B. Extra process steps to be inserted for poly-silicon gates

11a. Deposit poly-silicon
11b. Pattern poly-silicon
11c. Oxidize poly-silicon

TABLE VIII. DIFMOS PROCESS COMPARISON (Metal-Gate PMOS Basis)

	PMOS	FAMOS	CMOS	DIFMOS	MNOS
Number of mask levels	5	6	7	8	9
Number of diffusions	1	2	3	4	3
Number of major process steps	10	12	15	16	19
Nonstandard steps	None	None	None	None	EPI substrate, 2 nm tunnel oxide/nitride gate dielectric
Nonvolatile	No	Yes	No	Yes	Yes
Electrically erasable	—	No	—	Yes	Yes

diffusion of the 7.0 μm deep p-type diffusion spreading the hole injector into contact with adjacent portions of the cell. It was at this point that the process was modified by replacing the deep p-type diffusion with a shallow p-type implant just prior to growth of the gate oxide. Thus, both hole and electron injector breakdown voltages are set by diffusions which are patterned and implanted in successive steps. The next three lots of material proved that this improvement worked, and that the

small cell designs were fully functional. In fact, using the figure of merit in Equation (1), a relative improvement of over 10^8 was obtained from the new device designs STAGG and NOCON, compared to those in the Fourth Experimental Series.

There were some additional problems encountered. There was a coding error in the AGGNP cell design which caused the entire cell area to be covered with p^{++} diffusion. This was the only design or coding error discovered on this test bar, and it has since been corrected and a new mask made. However, the overall design-process-test cycle time has already overrun the original six months of this add-on portion of the contract. Therefore, we have insufficient time to process material with the corrected AGGNP mask or to work out a good silicon gate process in time to include the results in this final reprot. Therefore, test results are limited to the STAGG and NOCON cells which represent three of the seven individual device designs submitted. Except for possible improved storage characteristics on the silicon gate devices, characteristics of the other device designs will probably be nearly equivalent to the STAGG and NOCON results.

To summarize the testing results to date, all the cells which have been tested have worked extremely well in comparison to previous results. STAGG shows that both the 16.1 nm^2 and 9.7 nm^2 per bit cells work, and NOCON verifies that the 9.7 nm^2 per bit cells can function in an array. The smaller floating gates, together with more efficient injectors and better gate oxides, provide the 10^8 improvement in figure of merit that was previously mentioned.

1. Data Write, Erase, and Read Speed and Power

Figure 34 shows the time dependence of a STAGG-type device floating-gate voltage while the electron avalanche injector is operated. The avalanche voltage is 28.5 volts, with current drive to the injector diode set to 500 μA, resulting in 14.2 milliwatts of power dissipation in the electron injector. The floating-gate voltage begins to charge from 0 volts to a value approximately half the avalanche voltage at a rate of approximately -2.5×10^4 volts/second. In one millisecond, it reaches -9 volts; in ten milliseconds, -12 volts. This result was obtained on a device which had not been previously operated, so the oxide traps were empty, and the injectors operated with maximum efficiency. A device with filled traps would show a reduced charging rate.

Figure 35 shows the time dependence of the STAGG cell floating-gate voltage during the operation of the hole injector. The initial voltage on the floating gate was set to -40 volts by an external probe, which was removed from physical contact to the floating gate prior to starting the avalanche. Because a negative electric field is required for hole injection, the hole injector diode can only partially discharge the floating gate to a minimum voltage level; this minimum level is the

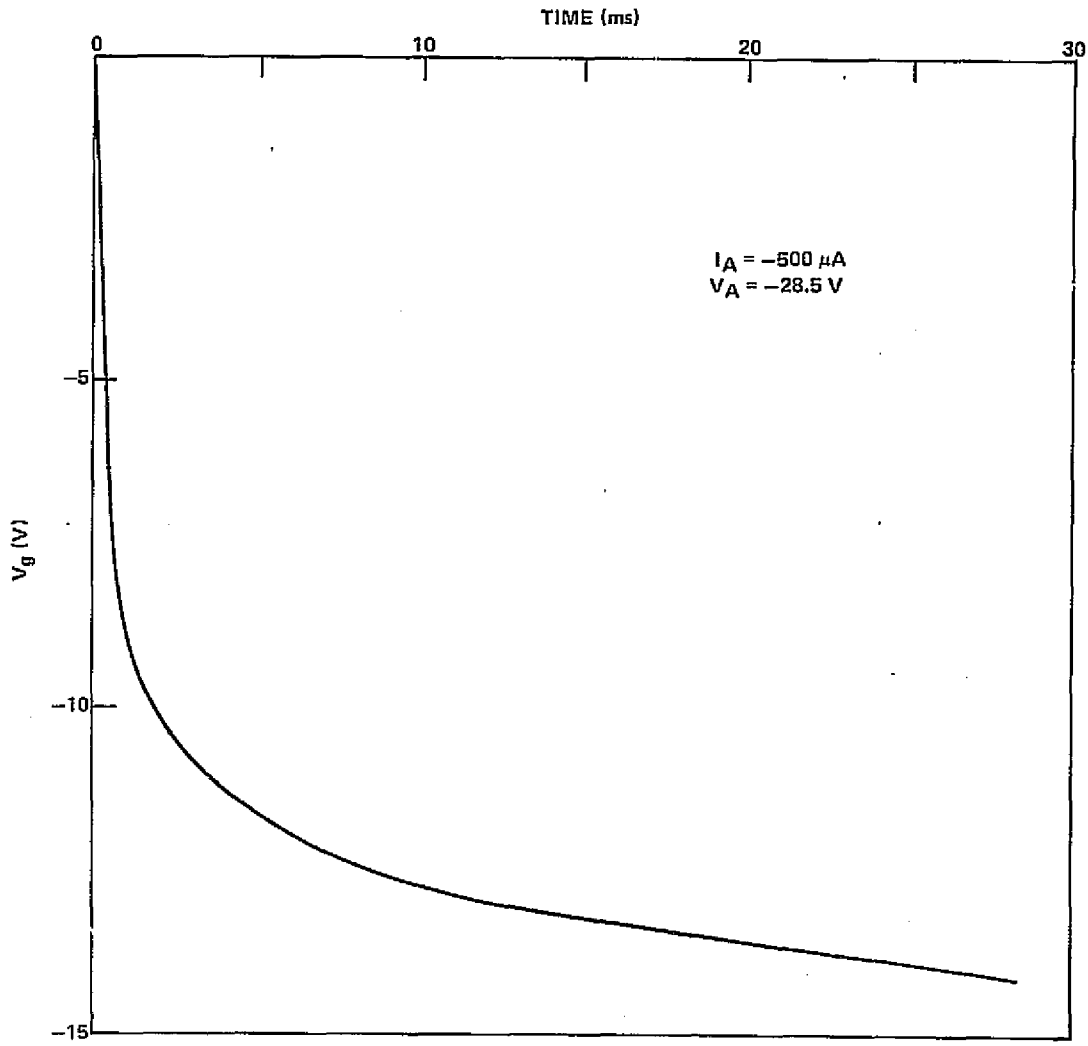


Figure 34. Time Dependence of Floating Gate Voltage During Avalanche of the $p^{++}n^{+}$ Electron Injector

threshold voltage for hole injection, V_{htx}. Devices in the Fifth Experimental Series exhibited V_{htx} as low as −14.5 volts. The initial discharge rate of the floating-gate voltage in Figure 35 is over 9×10^4 V/cm. Within one millisecond, it has discharged to −24 volts; in ten milliseconds, −18 volts. The device in Figure 35 asymptotically approaches $V_{htx} = -14.5$ as the avalanche continues. There has been some evidence that V_{htx} is related to the hole injector avalanche voltage. Experiments comparing hole injector avalanche voltages of −25 volts and −15 volts indicate that V_{htx} is essentially equal to the avalanche voltage. This effect may be coincidental, and consequently is being investigated further.

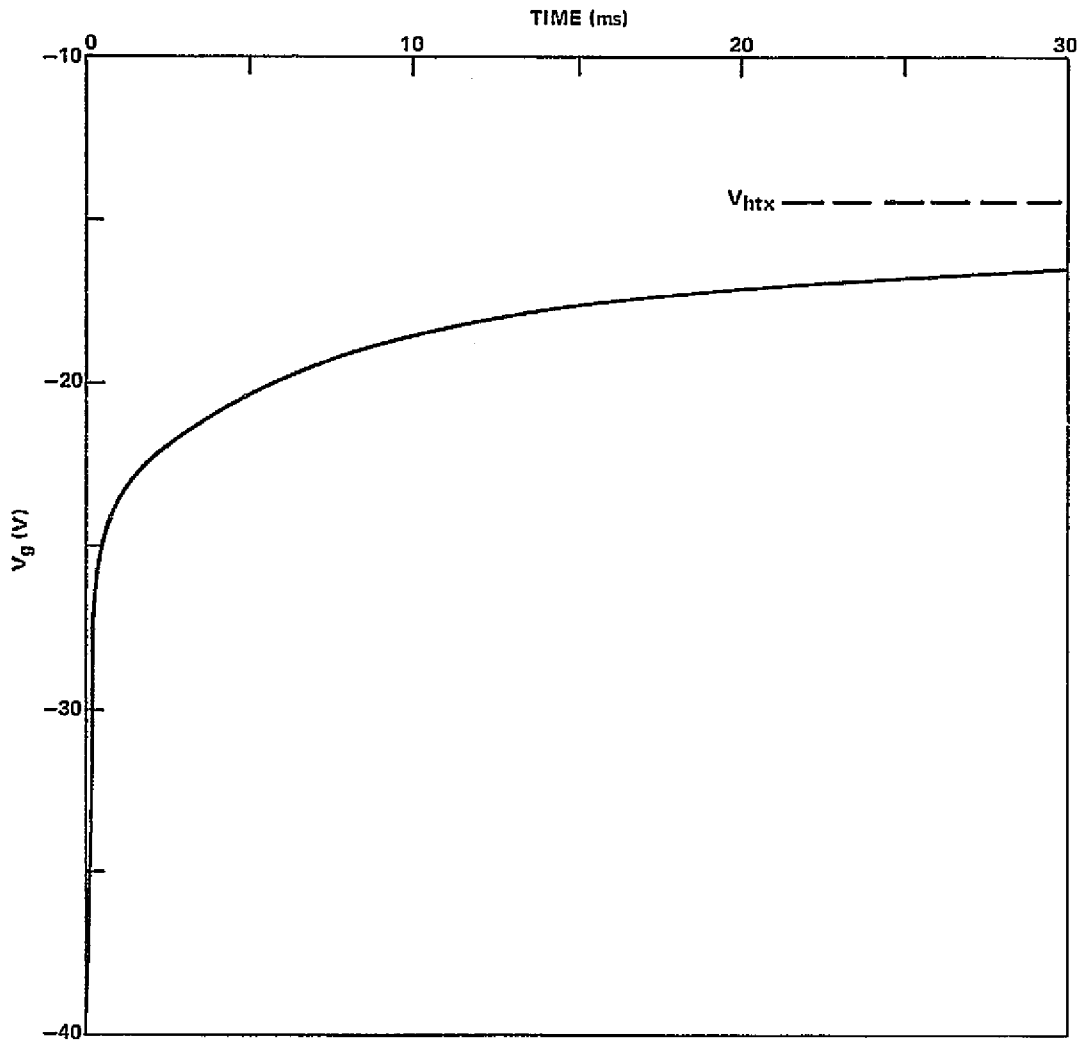


Figure 35. Time Dependence of Floating Gate Voltage During Avalanche of a p^+/n^{++} Hole Injector Junction $I_A = -500\ \mu A$ at −14.9 volts

Power consumption during the erase operation takes piace in two ways, (1) direct power dissipated in the avalanche junc ion, and (2) power required to operate the bootstrap capacitor. The avalanche power is 500 μA X 14.9 volts = 7.45 milliwatts. The power required to charge the bootstrap capacitor is negligible compared to the avalanche power, because it requires only a transient current.

Figure 36 shows the floating-gate voltage of a STAGG-type cell during a write and an erase operation. A 50-ms avalanche pulse applied to the electron injector results in the charging of the floating gate to about 14 volts. A similar pulse applied to the hole injector, together with a −40-volt

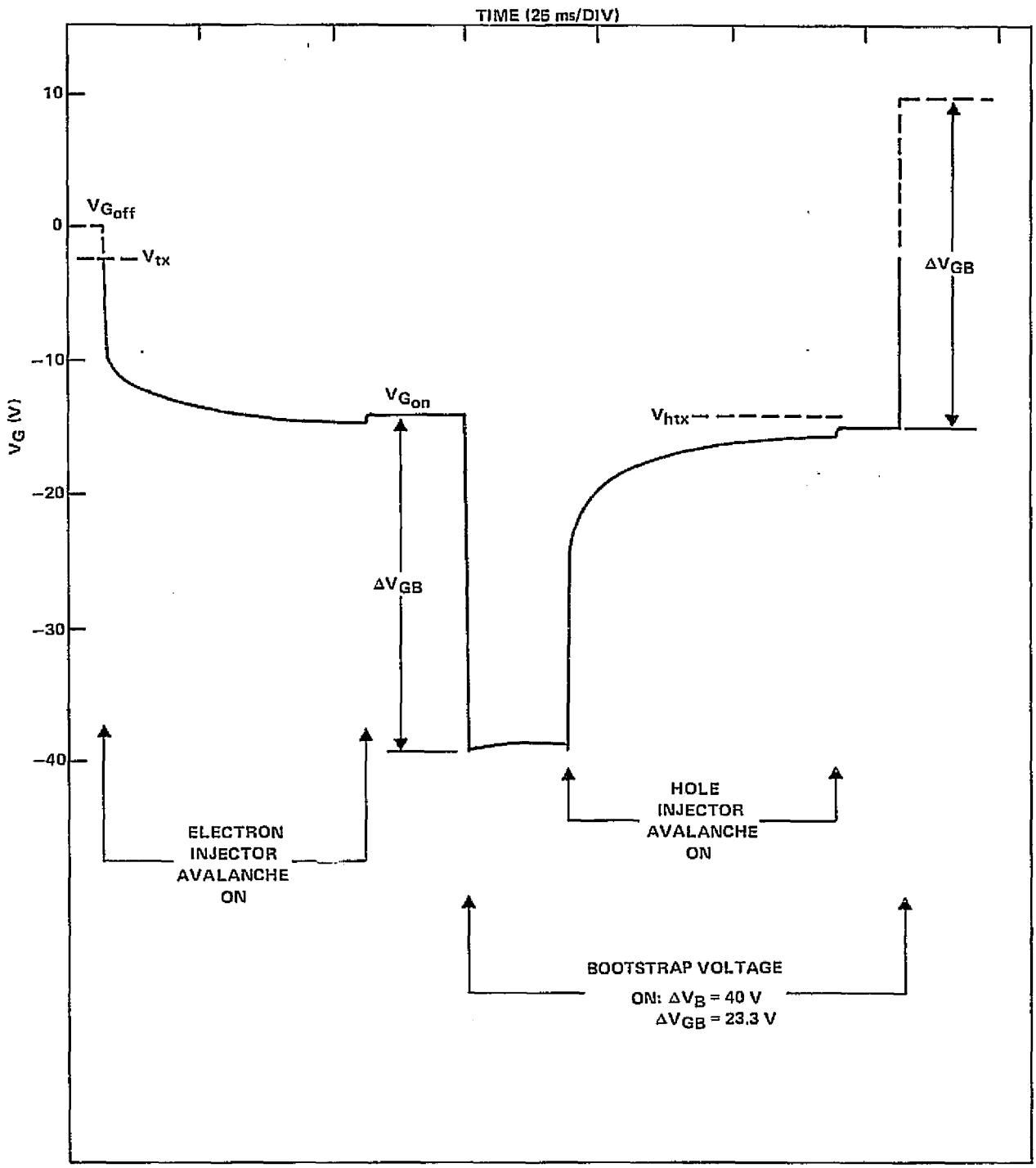


Figure 36. Time Dependence of Floating Gate Voltage During Writing and Erasing

bootstrap voltage pulse, results in the discharge of the floating-gate voltage through zero volts up to about +10 volts. This behavior was measured on a fresh device. As will be shown later, this window closes as the device is repeatedly cycled through write and erase operations.

In Figure 36, the floating-gate voltage is measured by using the sense transistor as a built-in electrometer. By suitable calibration, the sense transistor can provide a continuous readout of the floating-gate voltage during the writing and erasing cycles.

In array applications, important considerations are the read speed and power. For purposes of reading, the DIFMOS nonvolatile memory devices behave like the fixed-program memory devices in a p-channel ROM. Reading itself is a passive operation and requires that no power be dissipated within the DIFMOS memory device. Therefore, reading specifications and performance will be like a standard MOS ROM: read-access time for a given word or bit is in the microsecond range. The speed is strictly a function of the address and decoding circuitry of the design, and is not limited by the DIFMOS device itself. Likewise, the power per bit will be that power consumed in the peripheral circuits used for decoding and addressing.

2. Data Storage Without Applied Power

Figure 37 shows the effect of data storage without power. In this test, a floating-gate device is charged with a single programming pulse of 40 to 100 ms. The worst case for charge storage is charge loss from a gate which has been written. (There are cases where severe surface leakage results in the charging of one floating-gate by leakage from an adjacent gate; but such cases represent failures in the assembly process, and are not of interest in determining the storage characteristics.)

The decay rate of charge from the floating-gate obeys a log time dependence given by

$$V/V_0 = Q/Q_0 = D \log (t/t_0) \qquad (6)$$

where V_0 and Q_0 are voltage and charge at t_0 = 1 second, and D is the rate of decay per decade of time. At room temperature, D is approximately 0.06%; at 80°C, it is approximately 1%, and it increases to about 10% at 125°C. This provides an activation energy of 0.51 eV for the temperature dependence of charge loss. Charge loss in the aluminum-gate DIFMOS device is predominantly surface leakage; it is very sensitive to the type and quality of the overcoat oxide. Presently, achieving sufficiently low charge leakage requires the DIFMOS devices to be sealed in dry nitrogen within hermetic packages.

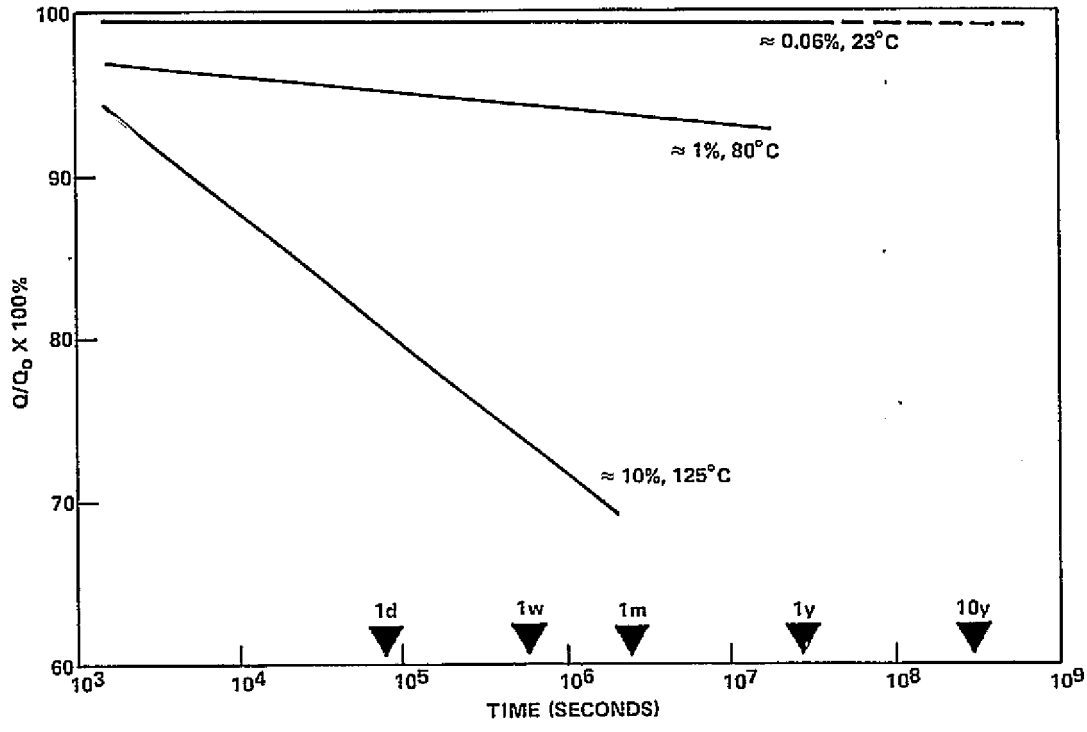


Figure 37. Temperature Dependence of Charge Storage

3. Data Degradation Due to Read Cycles

Reading is a passive operation in the DIFMOS cell in that the state of the floating-gate voltage is measured by measuring the channel conductivity of the sensing transistor. Therefore, reading has no effect on the floating-gate voltage. Our measurements indicate that the device can be read continuously forever without affecting stored data, or that over 10^6 read cycles (of arbitrary duration) can take place without disturbing the stored data. Unlike MNOS, DIFMOS has a well defined threshold voltage (the hole and electron avalanche voltages) for memory storage, any applied voltage less than this value can be utilized in the read operations without danger of affecting the stored data.

4. Device Degradation Due to Write/Erase Cycling

Figure 38 shows an envelope of the DIFMOS floating-gate voltages during write/erase cycling as a function of the number of cycles. The window amplitude initially is almost 14 volts, and it decays with the number of cycles. Note that voltages following the erase operations are clamped to V_{tx} of the sensing device, and values more positive than V_{tx} cannot be directly measured. At around 10^4 write/erase cycles, the device no longer can be turned completely off, and the window completely disappears around 10^5 cycles. This envelope was measured using 100 ms at 500 μA avalanche pulses for both writing and erasing, and a −40 volt bootstrap voltage pulse during erasing. At any point, the window can be widened by increasing either the avalanche current or time. Accordingly, the maximum number of cycles can be increased by increasing the avalanche time; but specification of a fixed avalanche time also defines the maximum number of cycles before closure of the window is attained.

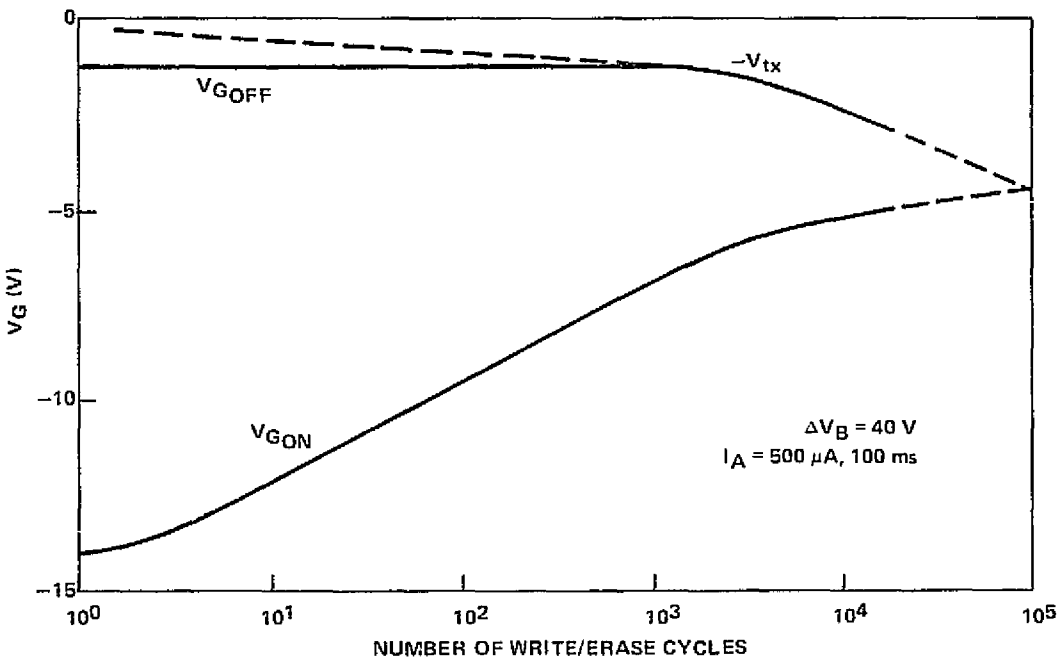


Figure 38. Envelope of On/Off Window versus Number of Cycles

SECTION VII

CONCLUSIONS AND SUMMARY

DIFMOS is a technology for building electrically writable and erasable nonvolatile semiconductor memory circuits suitable for read-mostly circuit applications. It features a metal-gate PMOS process with all standard processing steps. Nonvolatile data storage is achieved using the floating-gate technique with special injector diodes in each bit for charging and discharging the floating gate by avalanche injection of electrons and holes, respectively.

DIFMOS and FAMOS both utilize avalanche injection of electrons to charge floating gates for programming. FAMOS is not electrically erasable, however DIFMOS is erasable. DIFMOS provides the full electrically writable/erasable capability of MNOS technology without the difficult process technology associated with the MNOS tunneling oxide/nitride gate dielectric. The achievement of electrically erasable EPROM circuits from a PMOS process with three extra diffusions is the main feature of DIFMOS.

Like other nonvolatile semiconductor memory technologies, DIFMOS is limited by a slow programming time relative to the read time, a finite number of times it can be reprogrammed, and a comparatively large cell size. With present cell designs in the 9.7 nm^2 to 16 nm^2 per bit size range, DIFMOS can provide circuits of 1K to 2K bit capacity. However, the ease with which DIFMOS can be used in system designs makes it attractive for many applications such as calculator and microprocessor program storage. Standard processing, electrically erasable by row or by bit, fully decodable without substrate isolation, unlimited reading capability, and a variety of possible organizations tend to offset some of the disadvantages. Additional reductions in cell size and improvements in device performance and circuit density will be realized as the floating-gate, avalanche-injected, nonvolatile semiconductor memories continue to evolve.

SECTION VIII

REFERENCES

1. Frohman-Bentchkowsky, D., "A Fully-Decoded 2048-Bit Electrically Programmable MOS-ROM," *1971 ISSCC Digest of Technical Papers*, 80-81.

2. J.T. Wallmark and J.H. Scott, Jr., "Switching and Storage Characteristics of MIS Memory Transistors," *RCA Rev.*, 30, 335 (1969).

3. Y. Tarui, et. al., "Proposal of Electrically Reprogrammable Non-Volatile Semiconductor Memory," *2nd Conference on SSS.D.*, Tokyo (September, 1971).

4. E.H. Nicollian, A. Goetzberger and C.N. Berglund, "Avalanche Injection Currents and Charging Phenomena in Thermal SiO^2," *Appl. Phys. Letters*, 15, 174 (1969).

5. E.H. Snow, A.S. Grove and D.J. Fitzgerald, "Effects of Ionizing Radiation on Oxidized Silicon Surfaces and Planar Devices," *Proc. IEEE*, 55, 1168 (1967).

6. R. Williams, "Photoemission of Electrons from Silicon into Silicon Dioxide," *Phys. Rev.*, 140, A569 (1965).

7. S.M. Hu, D.R. Kerr and L.V. Gregor, "Evidence of Hole Injection and Trapping in Silicon Nitride Films Prepared by Reactive Sputtering," *Appl. Phys. Letters*, 10, 97 (1967).

8. J.F. Verwey, "Hole Currents in Thermally Grown SiO_2," *J. Appl. Phys.*, 43, 2273 (1972).

9. C. Bulucea, A. Rusa and C. Postolache, "Surface Breakdown in Silicon Planar Junctions – A Computer-Aided Experimental Determination of the Critical Field," *Solid-State Elecronics*, 17, 881 (1974).

10. A.S. Grove, O. Leistiko, Jr. and W.W. Hooper, *IEEE Trans. Electron Devices*, ED-14, 157 (1967).

11. C. Bulucea, "Avalanche Injection into the Oxide in Silicon Gate-Controlled Devices – I Theory," *Solid-State Electronics* 18, 363 (1975).

12. C. Bulucea, "Avalanche Injection into the Oxide in Silicon Gate-Controlled Devices – II Experimental Results," *Solid-State Electonics* 18, 381 (1975).

APPENDIX

LIST OF TEST EQUIPMENT USED TO PERFORM TASKS VI AND X

Number	Item	Model No.	Manufacturer
1	X-Y Recorder	7004 B	Hewlett-Packard
	DC Preamplifier	17171A	Hewlett-Packard
	Time Base	17172A	Hewlett-Packard
	Filter	17175A	Hewlett-Packard
2	Analytical Probe Station	S-1000	Signatone
3	Counter, Universal	DC503	Tektronix
4	Digital Multiplier	DM501	Tektronix
5	Operational Amplifier	AM501	Tektronix
6	Power Module	TM504	Tektronix
7	Curve Tracer	576	Tektronix
8	Oscilloscope	7904	Tektronix
	Dual Trace Amplifier	7A26	Tektronix
	Digital Delay	7D11	Tektronix
	Dual Time Base	7B92	Tektronix
9	Function Generator	FG501	Tektronix
10	Power Supply	6215 A	Hewlett-Packard
11	Power Supply	6202 A	Harrison Labs
12	Digital Electrometer	616	Keithley
13	Multi-Function Meter	3450 A	Hewlett-Packard
14	Multimeter	630-NA	Triplett
15	Power Supply	6205 B	Harrison
16	Picoammeter	HP-425A	Hewlett-Packard
17	Power Supply	6218 A	Hewlett-Packard
18	Display	34740A	Hewlett-Packard
	DC Voltmeter	34701 A	Hewlett-Packard
19	Temperature Chamber	SD 60-1	Statham
20	Function Generator	501	Ailtech

ACKNOWLEDGMENT

The authors would like to acknowledge the assistance of the following people for their part in the development of DIFMOS: Dr. B. G. Carbajal for his belief and support; Dr. H. R. Plumlee for technical assistance; Joel Stanford for computer-aided design; Irene York for data collection; Dr. J. A. Reynolds and Joy Cookston for processing; and Dr. Archie Fripp of NASA Langley for his interest and support.