AFRL-VS-TR-
2000-1106

LATCH-UP DETECTION AND CANCELLATION IN CMOS VLSI CIRCUITS

Philippe O. Pouliquen
Mark N. Martin

Padgett-Martin Technology, Inc
4222 Falls Road
Baltimore MD 21211

June 2000

Final Report

AIR FORCE RESEARCH LABORATORY
Space Vehicles Directorate
3550 Aberdeen Ave SE
AIR FORCE MATERIEL COMMAND
KIRTLAND AIR FORCE BASE, NM 87117-5776

AFRL-VS-TR-2000-1106

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data, does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report has been reviewed by the Public Affairs Office and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nationals.

If you change your address, wish to be removed from this mailing list, or your organization no longer employs the addressee, please notify AFRL/VSSE, 3550 Aberdeen Ave SE, Kirtland AFB, NM 87117-5776.

Do not return copies of this report unless contractual obligations or notice on a specific document requires its return.

This report has been approved for publication.

JOSEPH CHAVEZ
Project Manager

FOR THE COMMANDER

SCOTT D. BERGER, Lt Col, USAF
Chief, Space Electronics & Protection Branch

KIRT S. MOSER, DR-IV
Chief, Spacecraft Technology Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
20-06-2000	Final Technical Report	01-06-1999 - 20-06-2000

4. TITLE AND SUBTITLE
Latch-Up Detection and Cancellation in CMOS VLSI Circuits

5a. CONTRACT NUMBER
F29601-99-C-0069

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
65502F

6. AUTHOR(S)
Philippe O. Pouliquen
Mark N. Martin

5d. PROJECT NUMBER
3005

5e. TASK NUMBER
VO

5f. WORK UNIT NUMBER
CQ

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Padgett-Martin Technologies, Inc.
4222 Falls Road
Baltimore, MD 21211

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
3550 Aberdeen Ave. SE
Bldg 914 Rm 109
Kirtland AFB, NM 97117-5776

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/VSSE

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-VS-TR-2000-1106

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release, distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This report looks at the issues involved in designing integrated circuits in bulk-CMOS processes for radiation environments.

First, we describe how radiation can cause transistor threshold to change and leakage currents to increase, and how these effects are mitigated in high-density VLSI processes. Secondly, we describe how radiation can activate parasitic structures endemic to bulk-CMOS processes, causing damaging effect called latch-up. Thirdly, we describe how latch-up can be detected in an active circuit and cancelled before damage can occur - this is accompanied by successful experimental results with laser induced latch-up. Fourthly, we describe other integrated circuit fabrication technologies, which are naturally immune to latch-up. Finally, we conclude with recommendations regarding further research that would be needed to validate the concept of bulk-CMOS integrated circuits in radiation environments.

15. SUBJECT TERMS
Space environment, radiation, latch-up, CMOS design

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Joseph Chavez
U	U	U	Unlimited	66	19b. TELEPHONE NUMBER *(include area code)* (505) 846-6456

Contents

List of Figures

Acknowledgments

The research described in this technical report was supported by SBIR contract F29601-99-C-0069 "Latch-Up Detection and Cancellation in CMOS VLSI Circuits". This contract was awarded pursuant to the proposal submitted by Padgett-Martin Technology, Inc. (PMT) in response to SBIR 99.1 solicitation topic AF99-030: Micro-Latch-Up Characterization.

The research was performed by Mark N. Martin, Ph.D. (Electrical and Computer Engineering) and Philippe O. Pouliquen, Ph.D. (Electrical and Computer Engineering). Invaluable assistance was provided by Professor Andreas G. Andreou, Ph.D. (Electrical Engineering and Computer Science) of The Johns Hopkins University. Some of the experiments described herein were performed by Wesley Millard, M.S.E. (Electrical and Computer Engineering) at The Johns Hopkins University.

The project officer for this SBIR effort was Joe Chavez of the Air Force Research Laboratory (VSSE), Kirtland AFB, NM.

Although most of the data in this report was collected in this effort, the data for the figures in Chapter 2 were obtained from Dr. Mark N. Martin's Ph.D. dissertation (see [1]).

Abbreviations

AMI	American Microsystems, Inc. (integrated circuit manufacturer)
BJT	Bipolar junction transistor
BNL	Brookhaven National Laboratory
CMOS	Complimentary Metal-Oxide-Semiconductor
CSN	N-select (integrated circuit fabrication mask)
CSP	N-select (integrated circuit fabrication mask)
CWN	N-select (integrated circuit fabrication mask)
DI	Dielectric Isolation
FET	Field-Effect Transistor
GaAs	Gallium-Arsenide (atomic symbols)
HP	Hewlett-Packard (integrated circuit manufacturer)
MOS	Metal-Oxide-Semiconductor
NFET	N-channel Field-Effect Transistor
NMOS	N-channel Metal-Oxide-Semiconductor
PFET	P-channel Field-Effect Transistor
PMOS	P-channel Metal-Oxide-Semiconductor
polysilicon	Polycrystalline silicon
SCR	Silicon controlled rectifier
Si	Silicon (atomic symbol)
SOI	Silicon-On-Insulator
SOS	Silicon-On-Sapphire
TID	Total-ionizing dose
TMAH	TetraMethyl Ammonium Hydroxide (silicon etchant)

Symbols

A	Amp, unit of electric current
V	Volt, unit of electric potential
V_T	Thermal voltage, $\frac{kT}{q}$, where k is Bolztmann's constant, T is the temperature and q is the charge of a electron. $V_T = 25\text{mV}$ at room temperature.
W	Watt, unit of power

Chapter 1

Introduction

1.1 Motivation

The thrust of this research has been towards the implementation in standard commercial processes of integrated circuits for space vehicles. The object is to take advantage of the low cost and repeatability of mass produced integrated circuits to reduce the cost of the design and fabrication of electronics destined for space applications.

Although the main target is public and military space vehicles, reductions in cost would benefit the private sector as well, in the area of satellite construction, and in other areas in which the operating environment bears some resemblance to space (*e.g.* nuclear power plants, and medical radiology).

1.2 The problem

Ideally, integrated circuits built for use on earth would be usable in space and *vice-versa.* In practice however, integrated electronics destined for space must be designed to resist the deleterious effects of radiation. Although the added overhead of this so-called *radiation hardening* does not prevent the circuit from operating on earth, it does add a substantial cost to the design and the fabrication. Unfortunately, this added cost is not acceptable in commercial electronics which is an extremely cost-sensitive market.

The typical result of this incompatibility is that space electronics are produced at greater

expense and in smaller quantities than their earth electronics counterparts. Although there is always the option of using adequately shielded earth electronics in space, this is not an ideal solution: it still increases the cost over the earth electronics alone, and perhaps more importantly, it increases the weight of the resulting assembly.

1.3 Report organization

Apart from actual physical damage, radiation affects modern integrated circuits either by changing transistor characteristics or by activating parasitic structures or devices.

For instance, one typical radiation effect is transistor threshold voltage shift. Fortunately, as the integrated circuit process' minimum feature size is reduced, the transistor threshold voltage shift tends to become smaller also. Another typical radiation effect is edge leakage. Although edge leakage affects different processes to a varying degree, it can be eliminated in any process by selecting appropriate transistor geometries. These two effects and their mitigation are further described in Chapter 2.

Another radiation effect that occurs in bulk CMOS processes is latch-up. Latch-up is essentially a sudden catastrophic activation of a parasitic structure, which causes large, damaging currents to flow in the integrated circuit. It is described in greater detail in Chapter 3.

Because bulk CMOS is by far the most common integrated circuit process today, Chapter 4 describes the novel way that we have proposed for detecting the occurrence of latch-up and cancelling it.

Since latch-up has traditionally been avoided by using integrated circuit processes other than bulk CMOS, these alternative processes are reviewed in Chapter 5. They include NMOS, GaAs, DI SOI, and SOS, and they do not exhibit the parasitic structure which is at the root of latch-up.

This report ends with our conclusions and recommendations for future research in this area (Chapter 6).

Chapter 2

Radiation effects

2.1 The field-effect transistor

The transistor used in bulk CMOS (complimentary metal-oxide-semiconductor) processes is the field-effect transistor (FET). Its name derives from its method of operation, in which the application of an electric field creates a channel for current conduction. A schematic drawing of the cross-section through a typical N-channel FET is shown in Figure 2.1.

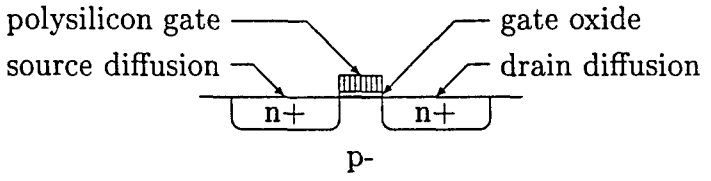

Figure 2.1: Schematic cross-section of an N-channel FET on a P-substrate.

The transistor is created from a crystalline silicon substrate which is selectively doped with other elements such as arsenic to form N regions or boron to form P regions. Polycrystalline silicon (polysilicon) is deposited on top of a thin oxide (the gate oxide) to form the gate (or control terminal) of the FET. It is from this oxide and the metallic nature of polysilicon that the term metal-oxide-semiconductor (MOS) is derived. If this oxide is omitted, the result is a junction FET (JFET) or a metal-semiconductor FET (MESFET).

The N-channel MOSFET (NMOS) shown in Figure 2.1 is formed from a lightly doped

P-substrate into which more heavily doped N regions of the source and drain are created by the diffusion of dopant elements into the substrate. P-channel MOSFETs (PMOS) can similarly be formed starting with an N-substrate and creating P regions for the source and drain terminals.

2.2 Self-aligned processes

Because the channel under the gate is formed by an electric field, it is important that the source and drain diffusion region be as close to the gate's edge as possible. This is achieved in a *self-aligned* process by using the gate polysilicon itself as the mask during the implantation of the source and drain.

If we take the transistor drawn in Figure 2.1 and view it from the top as shown in Figure 2.2, we see that it is essentially composed of two intersecting rectangles: the polysilicon gate and the so-called *active region.*

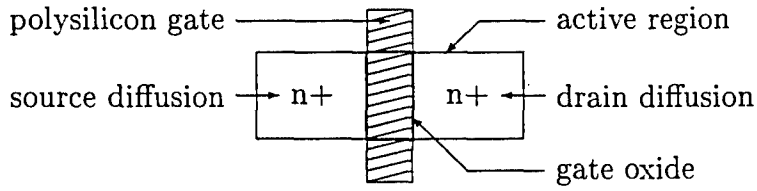

Figure 2.2: Schematic top view of an N-channel FET in a self-aligned process.

During the manufacturing process, the gate oxide and polysilicon gate are formed first. Then the source and drain dopants are implanted in the entire active region, including the region which overlaps with the middle portion of the polysilicon gate. This ensures that the source and drain regions begin and end exactly at the edge of the gate. The polysilicon is also always extended a safe amount beyond the active region, so that small misalignments in the fabrication process do not end up shorting the source to the drain.

2.3 Transistor characteristics

The channel is created under the gate when a sufficiently high potential is applied to the gate terminal. The minimum gate potential required to turn on the transistor in this way is called the *threshold voltage.*

Figure 2.3 shows the channel current at the drain in a PMOS fabricated in a 1.2μm process, for a -5 V bias at the drain. The threshold voltage is approximately -0.66 V for this transistor.

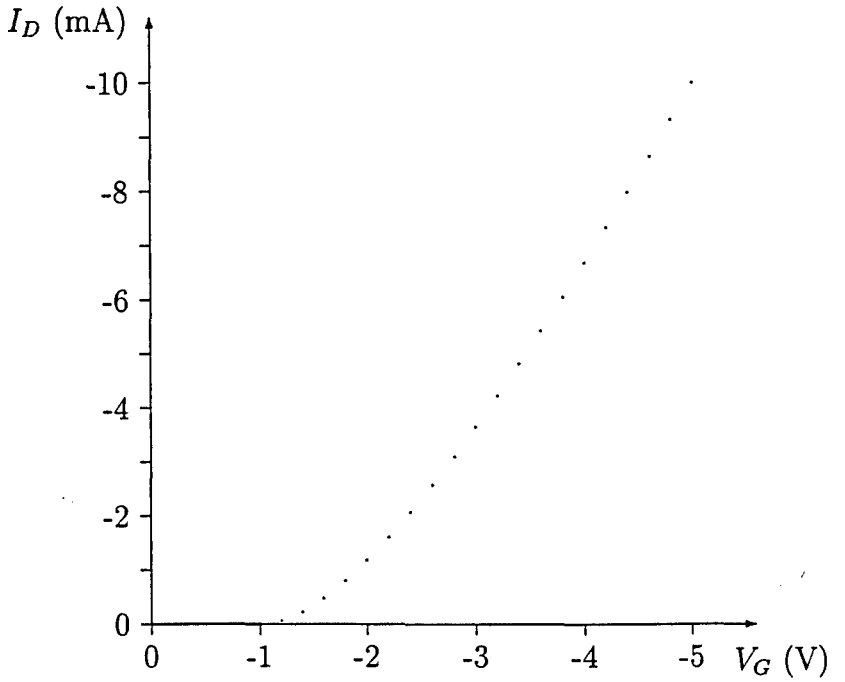

Figure 2.3: PMOS characteristics (1.2μm process, $V_D = -5$V).

However, the drain current is not completely zero below threshold. Rather, the characteristics change from an (ideally) square law behavior to an exponential behavior as shown in Figure 2.4. The threshold voltage is therefore usually computed by fitting transistor model equations to the empirical data.

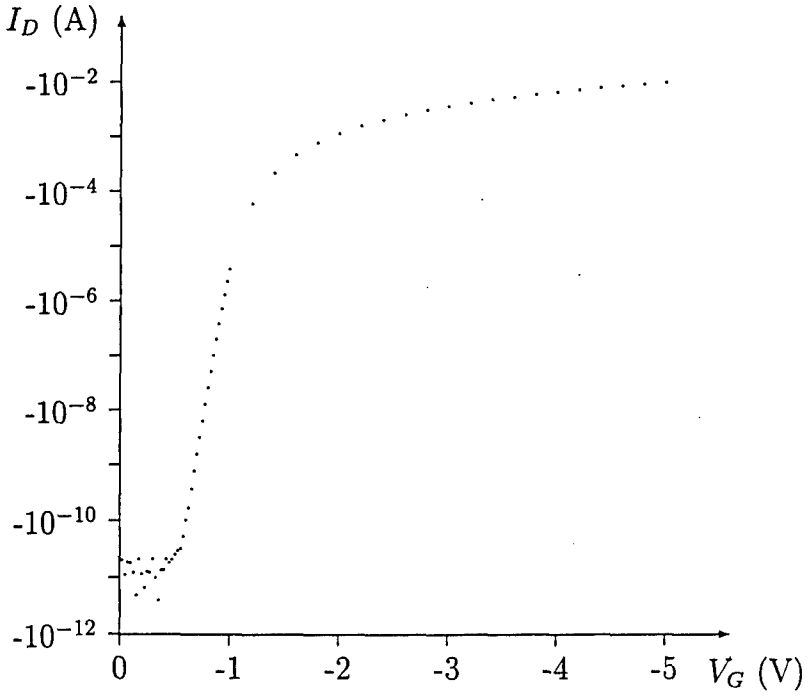

Figure 2.4: PMOS characteristics (semi-log) (1.2μm process, $V_D = -5$V).

2.4 Back gate effect

In an MOS transistor, the electric field at the channel induced by the gate depends on the substrate potential as well. The transistor characteristics shown in Figures 2.3 and 2.4 were taken with the source and substrate at the same potential. However, since typically many transistors coexist in the same substrate, it is likely that some of these will be operating with the source and substrate at different potentials.

The effect of the substrate potential (V_B) is shown in Figure 2.5: the threshold voltage appears to shift by approximately 0.2 V at a substrate potential of 1 V. This effect is called the *back gate effect* or *body effect.*

If the substrate is left unconnected, its potential may fluctuate leading to unpredictable transistor behavior. These fluctuations are limited only by the diode characteristics of the source-substrate P-N junction. Random transistor behavior due to missing or damaged substrate contacts is referred to as the *floating body effect.*

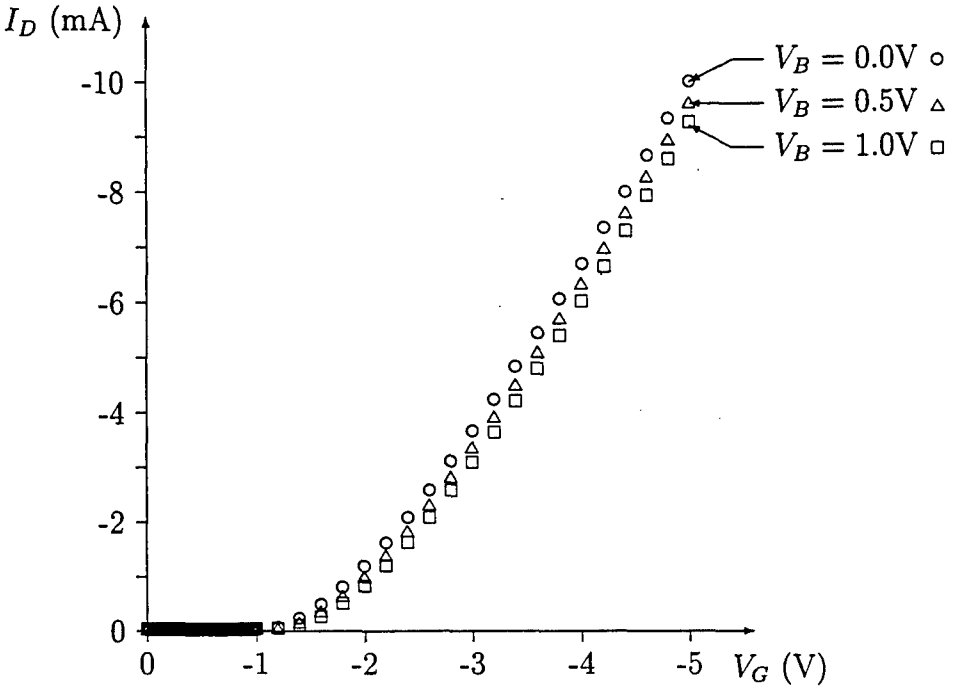

Figure 2.5: Back gate effect on PMOS characteristics (1.2μm process, $V_D = -5$V).

2.5 Gate oxide traps

The gate oxide is essential to the CMOS process to provide electrical isolation between the gate terminal and the source and drain terminals. However, is also one of the causes of radiation sensitivity in the FET.

Indeed, in the presence of ionizing radiation, an electron can be knocked out of its orbit around an atom. If this happens in the gate oxide, the electron is swept away by the gate potential, but the resulting hole moves more slowly. Moreover, the hole may get stuck in an oxide defect or *trap*, for extremely long periods of time.

Trapped charges in the oxide alter the vertical electric field in the silicon, which causes the threshold voltage to shift. This effect is shown in Figure 2.6 for the transistor from Figure 2.3 for radiation doses in the 0 to 100 kRad range: the threshold shifts by approximately 1.5 V at the 100 kRad dose.

However, as shown in Figure 2.7, a PMOS transistor fabricated in 0.5μm process shows a comparatively small threshold shift, even for much higher doses. This is typical of the smaller

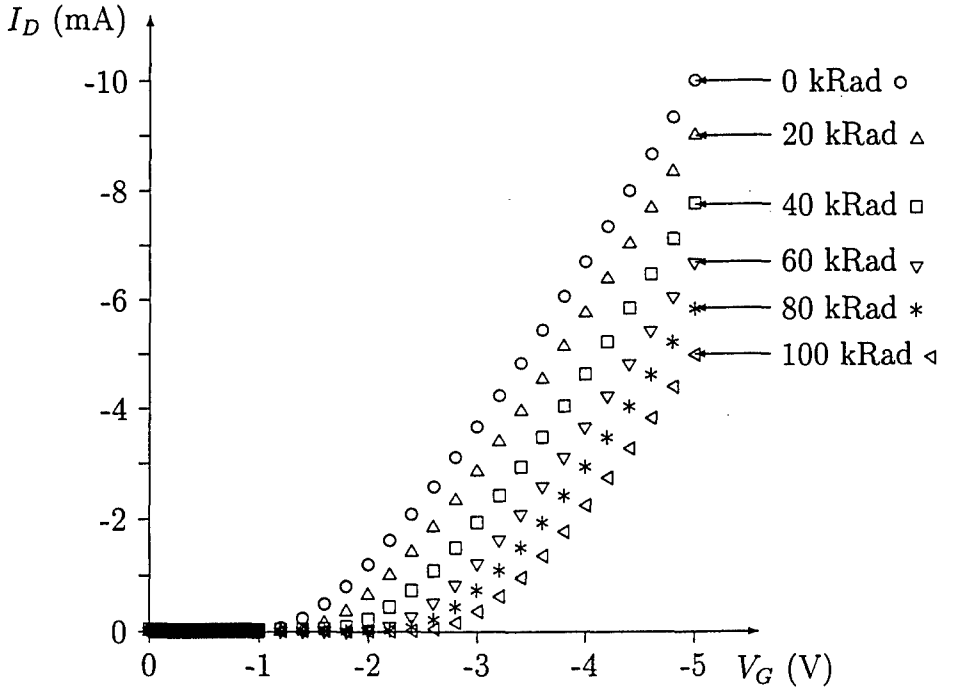

Figure 2.6: Radiation effect on PMOS characteristics (1.2μm process, $V_D = -5$V).

feature size processes: as the process feature size shrinks, the gate oxide thickness, width and length shrink also, which reduces the probability of ionizing radiation interactions within the rapidly shrinking gate oxide volume. In addition, as the gate oxide becomes thinner, integrated circuit manufacturers must continually improve its quality thereby reducing the density of defects.

2.6 Field oxide transistor

If we take the transistor drawn in Figure 2.2 and view a cross section through the gate as shown in Figure 2.8(a), we can see that the thin gate oxide only occurs in the active region. The oxide under the polysilicon gate extensions on either side of the transistor is much thicker, and is referred to as thick oxide or *field oxide.* Figure 2.8 also shows the location of the channel or threshold implant which is added to adjust the threshold voltage to the manufacturer's specifications.

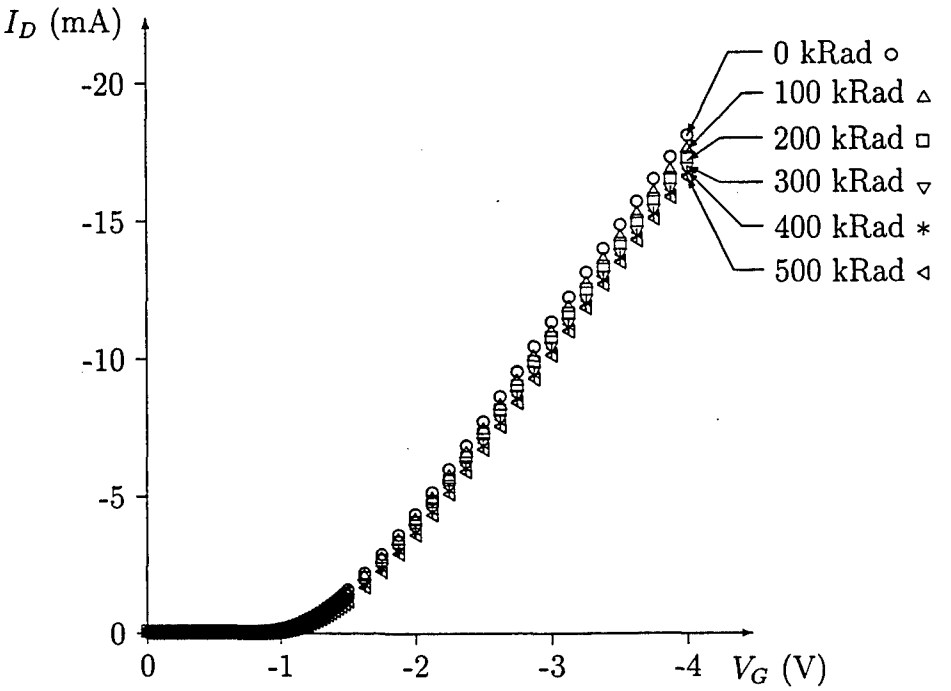

Figure 2.7: Radiation effect on PMOS characteristics (0.5μm process, $V_D = -2$V).

There is also a *field implant* or *channel stop* (not shown) which is placed outside the transistor's channel where the oxide is thick. Under normal operating conditions, the channel stop prevents a channel from forming outside the confines of the transistor. Unfortunately, in high radiation environments, this field implant is often insufficient to prevent leakage from occurring along the path shown in Figure 2.8(b).

The field oxide's quality with respect to the gate oxide accounts for most of the high sensitivity of the field oxide transistor's threshold to radiation. Indeed, the field oxide is generally of much poorer quality than the gate oxide, and therefore has a higher trap density. Additionally, the field oxide is thicker than the gate oxide, and this increased volume increases the probability of ionizing radiation interactions. The increased volume effect is somewhat mitigated by the fact that the trap locations are spread across a larger range of distances from the channel, and trapped charges further away will have less effect on the threshold.

Figure 2.9 shows the increased leakage in the characteristics for an NMOS transistor fabricated in 0.5μm process, as a result of the field-oxide transistor turning on at progressively

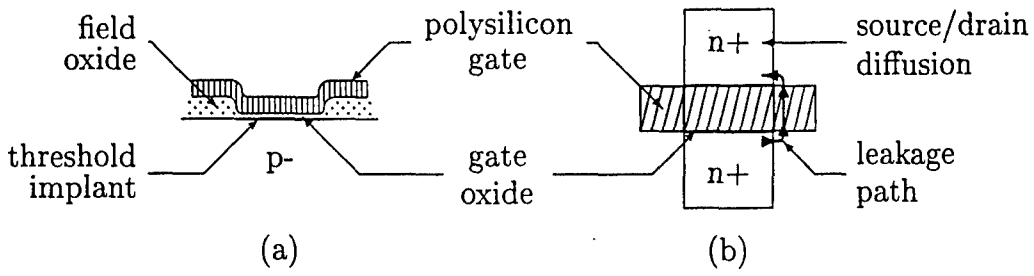

Figure 2.8: Schematic of an NFET on a P-substrate. (a) Cross-section through the gate. (b) Top view showing leakage path.

lower gate voltages. However, as shown in Figure 2.10, a PMOS transistor under the exact same conditions as the NMOS transistor of Figure 2.9 does not exhibit increasing leakage with radiation. The difference is due to the nature of the trapped charges: radiation causes holes to be trapped in the oxide of both NMOS and PMOS transistors alike, but the positive charges cause the NMOS transistors to pass more current and PMOS transistors to pass less current with increasing radiation dose. Furthermore, despite the low mobility of the holes, the positively charged gate in the NMOS transistor tends to push the holes closer to the channel, thereby exacerbating their effect, whereas the negatively charged gate in the PMOS transistor tends to pull the holes further away from the channel.

In the absence of specialized processes, the field oxide leakage is easily circumvented by the use of annular-shaped transistors. The top view of an annular transistor is shown in Figure 2.11. In this configuration, the only location of a parasitic field-oxide device is at the top, where the polysilicon gate extends beyond the active region. This stub is required to allow connectivity to the gate, as metal contacts to the polysilicon are generally not allowed on top of transistors. However, since the diffusion regions on either side of this stub are electrically connected anyway, any leakage between them is inconsequential.

The use of annular transistors does result in a significant circuit bloat compared to the use of minimum size transistors. For instance, in a Scalable CMOS process, we find that a single annular transistor consumes 228 squares, whereas a minimum size transistor consumes 48 squares. However, this factor of six is a worst case scenario: it does not take into account the minimum transistor separation (which reduces the ratio to 3.4375), the potential sharing

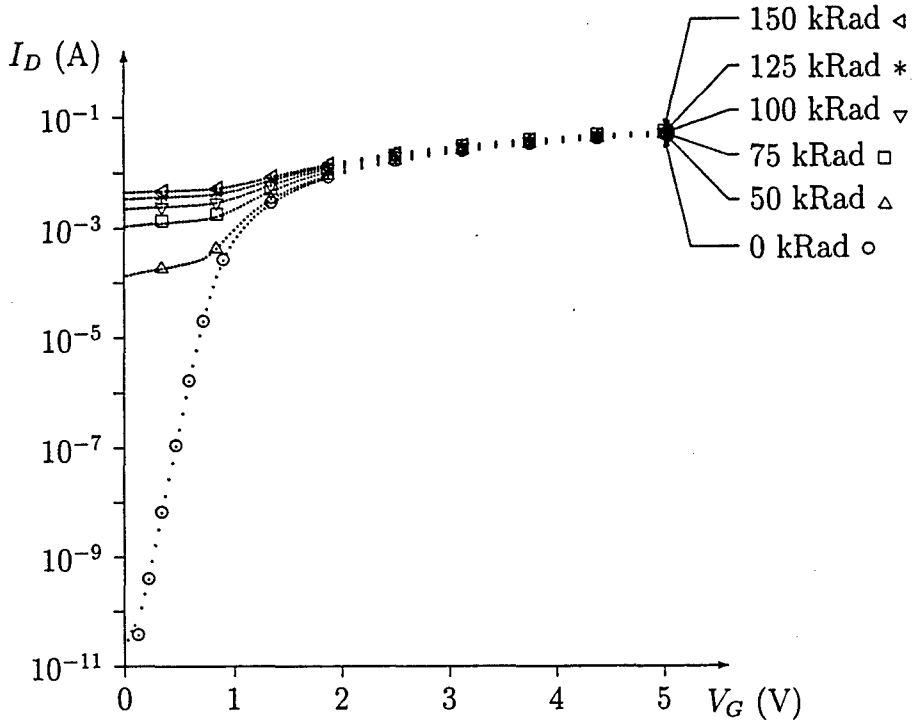

Figure 2.9: Radiation effect on NMOS characteristics (0.5μm process, $V_D = 2$V).

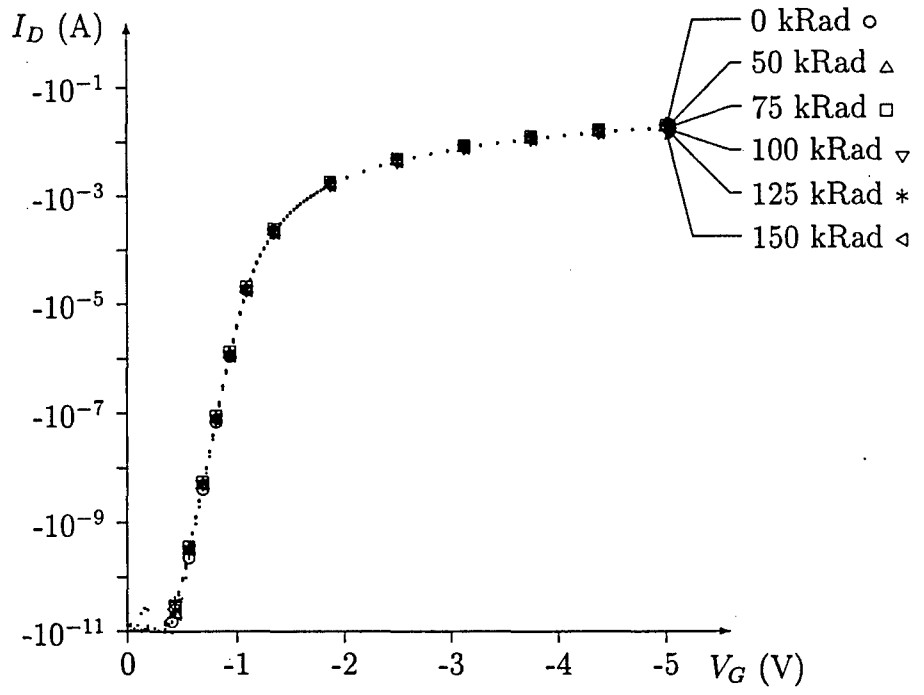

Figure 2.10: Radiation effect on PMOS characteristics (0.5μm process, $V_D = -2$V).

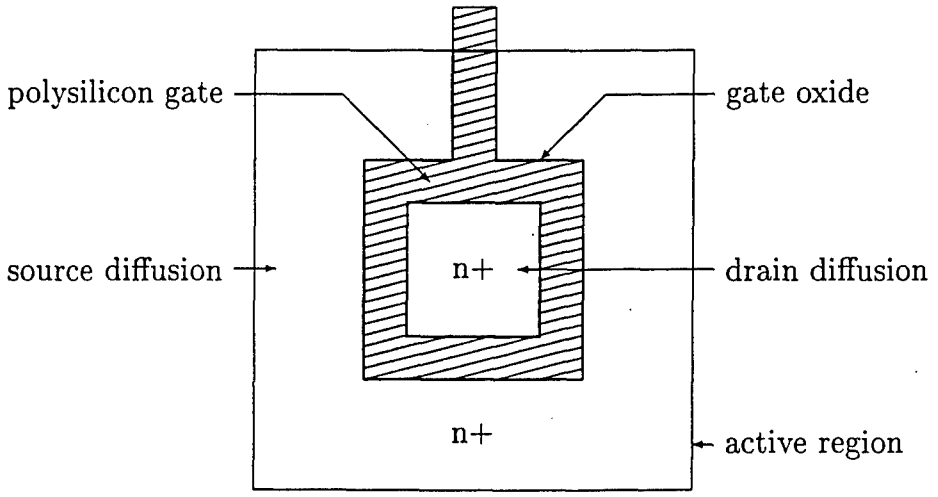

Figure 2.11: Schematic top view of an annular NFET.

of the external source or drain area, the approximately eight times greater width-to-length ratio, and the fact that many circuits are limited by the wiring density and not the transistor density. For practical circuits, the bloat factor ranges from 1.5 to 2.

2.7 Summary

Radiation induced threshold voltage shifts are a problem that is gradually going away with smaller process feature sizes (see Section 2.5). Radiation induced leakage has not been mitigated by smaller process feature sizes, but can be eliminated by careful geometric considerations at a cost in area (see Section 2.6).

Thus commercial processes gate oxides are becoming more and more suited to radiation environments, which could potentially significantly reduce the cost of space-based electronics.

Chapter 3

Latch-up

3.1 The bulk CMOS process

Although radiation affects all MOS transistors in the same way, latch-up is a problem specific to bulk CMOS (complimentary metal-oxide-semiconductor) processes. To understand why this is the case, we need to look at the way transistors are fabricated in a bulk CMOS process. Figure 3.1 is a schematic cross section of a typical CMOS process, showing two transistors, contacts and metalization.

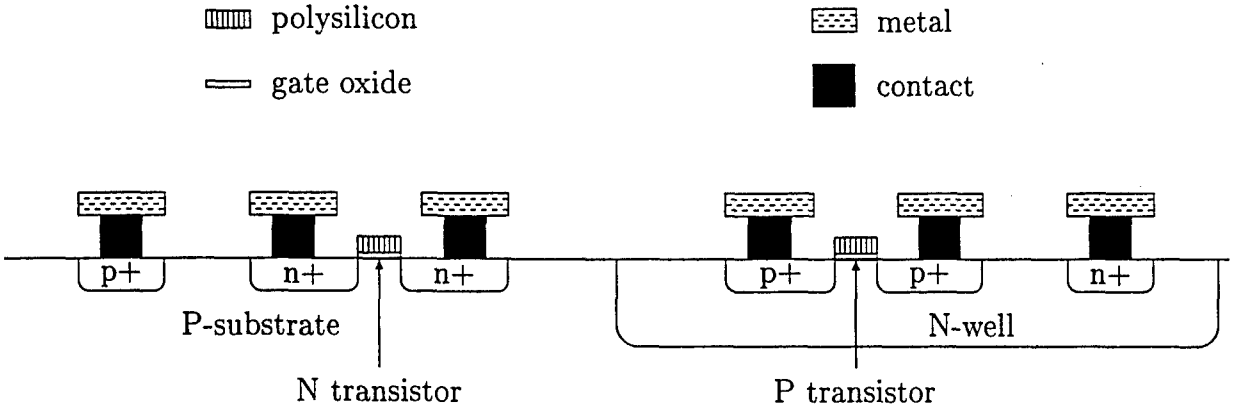

Figure 3.1: Schematic cross-section of an N-well bulk CMOS process.

The transistor on the left is an N-channel MOS (NMOS) transistor. It uses two heavily doped N-regions (n+) for its source and drain, and a lightly doped P-substrate (p-). A

separate heavily doped P-region (p+) to the left of the NMOS transistor is used to contact the substrate.

The transistor on the right is a P-channel MOS (PMOS) transistor. It uses two heavily doped P-regions (p+) for its source and drain, and a lightly doped N-substrate (n-). Because it is not possible for the substrate to be simultaneously n- and p-, the PMOS transistor is formed inside a restricted region called a *well.* A separate heavily doped N-region to the right of the PMOS transistor is used to contact this N-well.

The process described in Figure 3.1 is an N-well process. However, NMOS and PMOS transistors can also coexist by starting with an N-substrate and creating P-well regions for the NMOS transistors, resulting in a P-well process. It is also possible to start with an intrinsic (un-doped) substrate and form P-wells for NMOS transistors and N-wells for PMOS transistors. However, *bulk* CMOS processes refer only to those fabrication processes which start with lightly doped substrates and create wells for the complementary devices.

3.2 Silicon controlled rectifiers

A silicon controlled rectifier (SCR) is a four layer device formed of alternating P and N regions as shown in Figure 3.2(a). By slicing the device along the dotted line, it can be seen that the SCR is electrically equivalent to a PNP and an NPN bipolar junction transistors (BJT) wired back-to-back as in Figure 3.2(b): the top three layers form the PNP and the bottom three layers form the NPN.

The BJTs are connected in a positive current feedback loop, in which the collector current of one device feeds the base current of the other. Any current that flows through one transistor is amplified by the other and subsequently further amplified by the first one. Thus, when the product of the current gains (loop gain) of the two devices is greater than unity, the current flowing through the SCR grows in an unbounded fashion.

In practice however, the loop gain depends on the current through and the potential across the SCR. For instance, at very large currents, the current gain of a typical BJT drops significantly from its peak value. Furthermore, the emitter current in a BJT is exponentially related to the base-emitter voltage. Therefore, as the current in the SCR increases, the

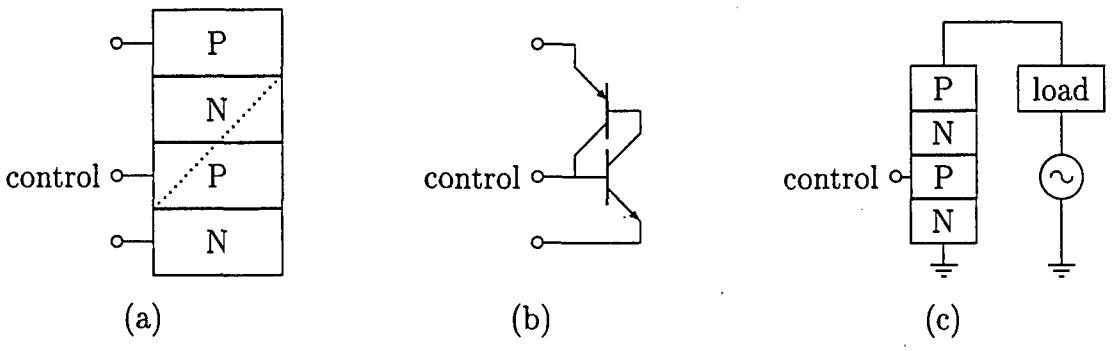

Figure 3.2: Silicon controlled rectifier. (a) Schematic. (b) Equivalent circuit. (c) Typical application.

base-emitter voltage of each BJT increases, resulting in a decrease in the collector-emitter voltage of each BJT, an effect which is further exacerbated by any series resistance in the SCR. Once the collector-emitter voltage drops below the saturation voltage, the current gain of the BJTs drops rapidly, and the SCR current will stabilize.

Since the current gain of a typical BJT also drops significantly from its peak value at very low currents, the SCR can also stabilize with no current flowing through it, other than leakage. Thus the SCR is typically bi-stable, with a high-current on state, and a zero-current off state.

In a typical SCR application, the SCR is connected in series between a high current load and an AC power source, as shown in Figure 3.2(c). Furthermore, a control terminal is added to the structure (shown in Figure 3.2(a) as attached to the bottom P layer), to switch the SCR from the off to the on state. When the control terminal is grounded, the SCR remains in the off state. However, when the polarity of the AC power source is positive, a small current can be injected into the SCR via the control terminal to cause the SCR to switch to the on state. The SCR will then remain on, with current flowing to the load, as long as the polarity of the AC power source remains positive. Since the SCR is a rectifying structure, no current will flow during the negative phase of the AC power source, unless back-to-back SCRs are used. Thus, by adjusting the turn-on delay of the SCR from the the beginning of the positive phase of the AC power source, a variable amount of power can be supplied to the high current load.

Once the SCR is in the on state, it is difficult to return to the off state. In theory, the SCR can be shut off by shunting all the collector current from the upper PNP transistor out via the control terminal, but in practice, the SCR current is usually so large that shunting it is impractical. Thus, the only other way to shut the SCR off again is to bring the potential across it to zero (or some negative value as in the AC SCR application).

3.3 Latch-up in bulk CMOS processes

The structure shown in Figure 3.1 bears more than a passing resemblance to an SCR: the P source/drain terminals of the PMOS, the N-well, the P-substrate, and the N source/drain terminals of the NMOS form a parasitic SCR, as shown in Figure 3.3.

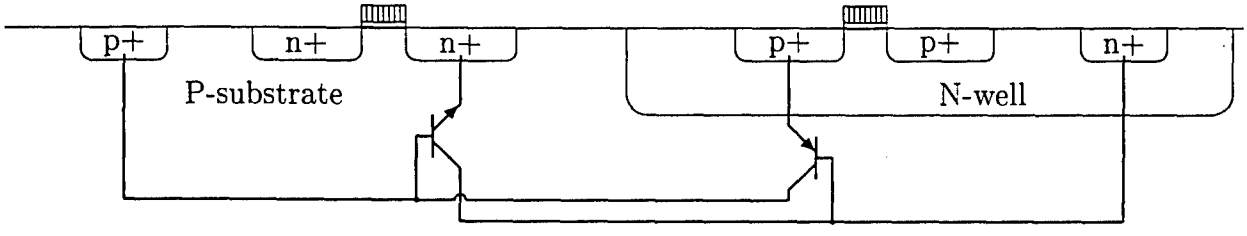

Figure 3.3: Parasitic SCR in a bulk CMOS process.

If the parasitic SCR structure turns on, the integrated circuit will draw a large amount of current from the power supply, possibly damaging the on-chip wiring or the P-N junctions of the transistors and SCR. This condition, called *latch-up,* is inhibited by grounding the P-substrate and connecting the N-well to the power supply. However, because the metal contacts to the silicon have some resistance associated with them, the possibility of a latch-up is not entirely eliminated (see Appendix A for a calculation of the limits on the contact resistance for latch-up prevention).

Latch-up can also be inhibited by fabricating the integrated circuit so that the product of the current gains of the two bipolar transistors in the SCR is less than unity: this inhibits the positive feedback loop that enables the SCR to remain on after it has been triggered. (See Appendix A.)

3.4 SCR characteristics

The SCR characteristics can be measured for low power supply voltages, where the current in not sufficiently large to cause permanent damage. These characteristics are shown in Figures 3.4 and 3.5.

Figure 3.4 shows that the current is principally exponential: this is due to the BJT's exponential relationship between the base-emitter voltage and the collector current.

Figure 3.5 shows that the P-substrate voltage ($V_{\text{P-substrate}}$) is close to the power supply (V_{SCR}), and that the N-well voltage ($V_{\text{N-well}}$) is close to ground. This is somewhat surprising because we would not normally expect the P-substrate voltage to exceed the N-well voltage. Indeed, a BJT's maximum current gain occurs in the *active* region, which is the region for which the emitter-base junction is forward biased and the collector-base junction is reverse biased. However, the BJTs in Figure 3.5 are operating in the *saturation* region, which is the region for which both the emitter-base junction and the collector-base junction are forward biased (see [2]).

This apparent anomaly is explained by looking at the BJT's characteristics, shown in Figures 3.6 and 3.7: the collector current levels off quite rapidly, within about 100 mV.

3.5 Latch-up current limits

As noted in Section 3.2, the positive feedback connection of the BJTs causes the current to increase until the feedback is limited by another mechanism. In the case of the bulk CMOS SCR, the P-substrate voltage cannot increase beyond the supply voltage, and the N-well voltage cannot decrease below ground.

However, the P-substrate voltage never quite reaches the supply voltage, and the N-well voltage never quite reaches ground. This is because the current gain of the BJT in the saturation region depends on the collector voltage: below about 200 mV collector-emitter bias, a portion of the base current begins to flow to the collector. If the collector-emitter bias is further reduced, the base current eventually dominate, and the collector current direction is reversed (as shown at the left edge of the graphs in Figures 3.6 and 3.7). The effect on

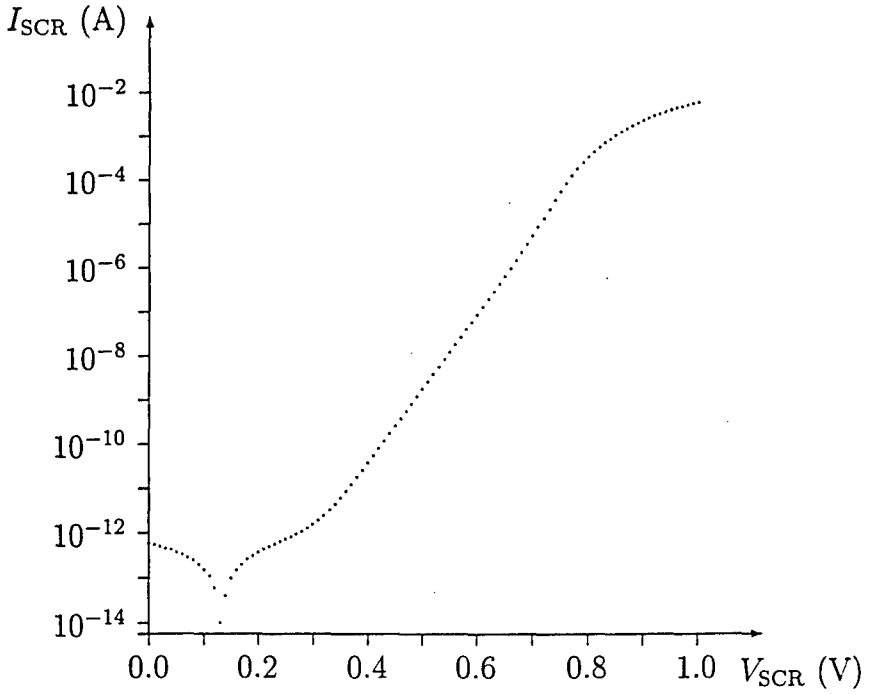

Figure 3.4: Bulk CMOS SCR current versus power supply voltage.

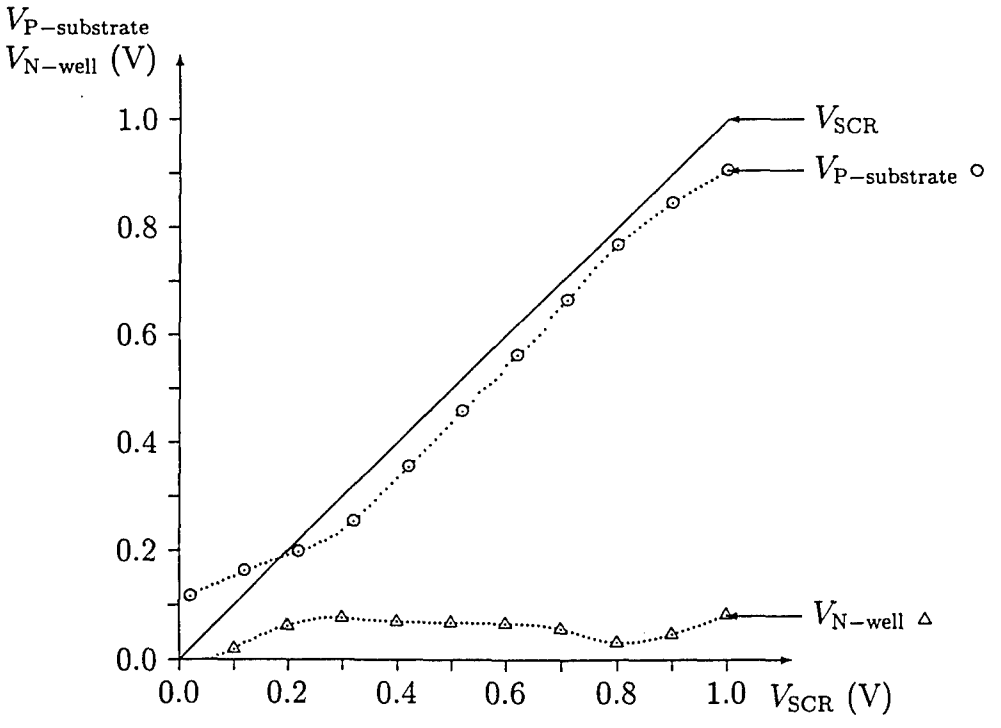

Figure 3.5: Bulk CMOS SCR internal voltages versus power supply voltage.

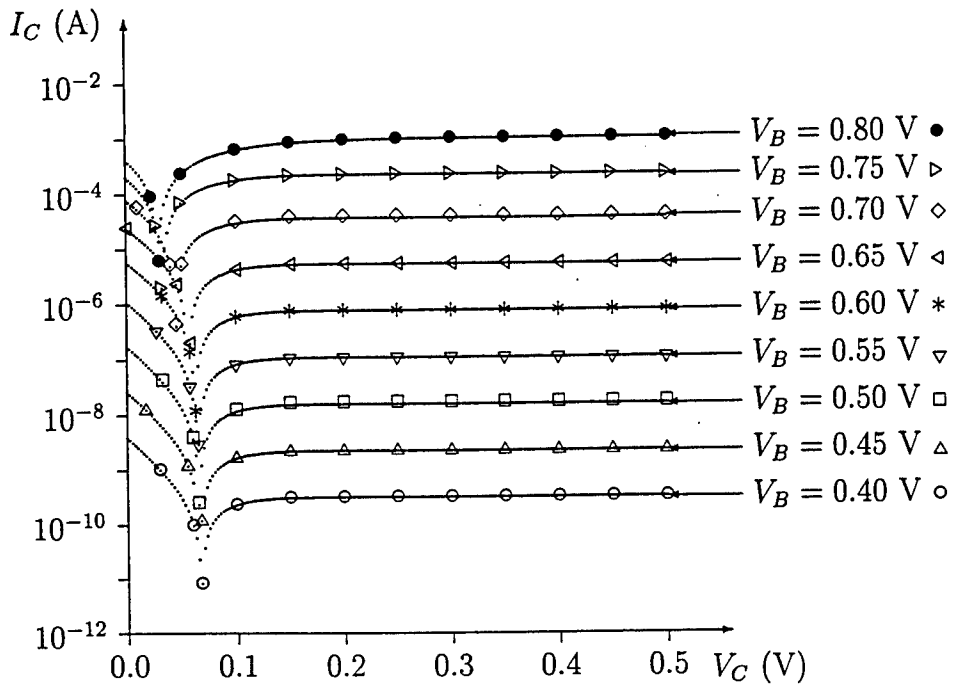

Figure 3.6: Bulk CMOS NPN BJT characteristics.

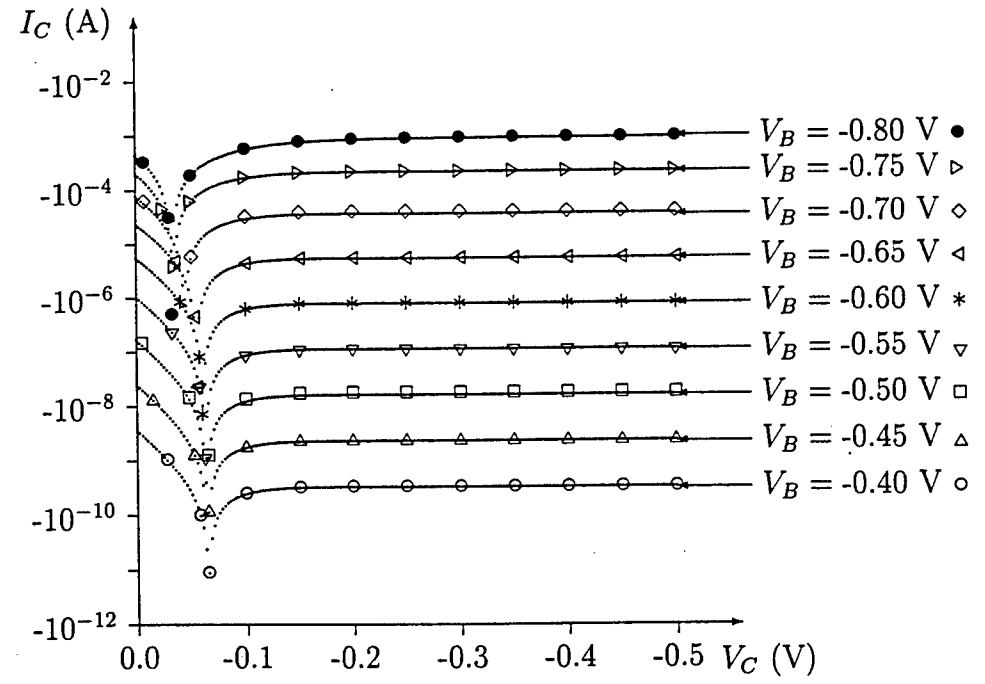

Figure 3.7: Bulk CMOS PNP BJT characteristics.

the current gain is quite prominent, as shown in Figure 3.8 and 3.9. These Figures also show that the current gain saturates much more slowly at large base currents, and that the current gain is largest for intermediate base currents.

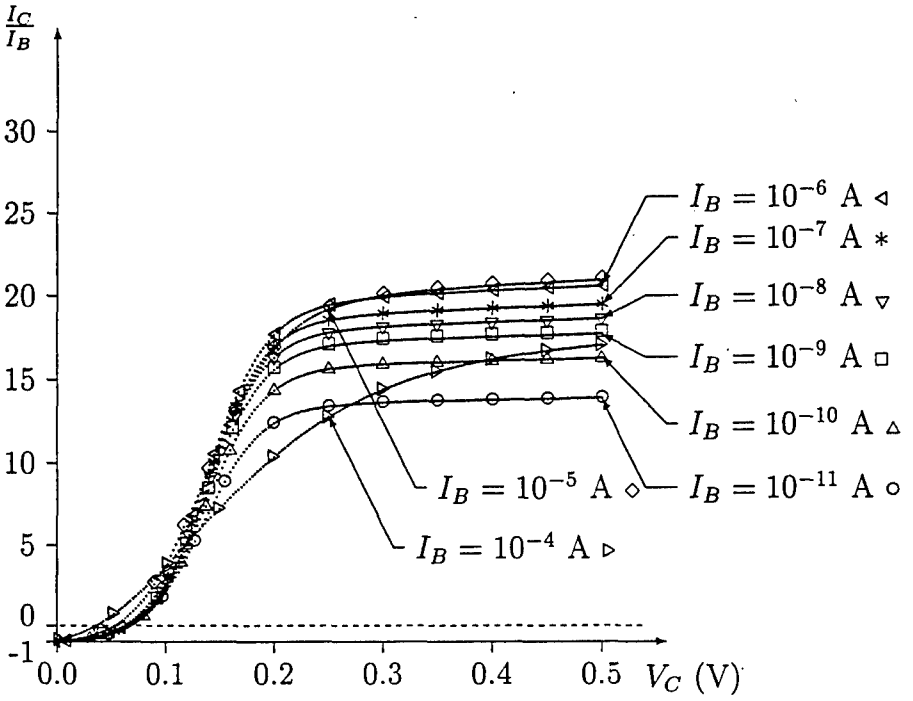

Figure 3.8: Bulk CMOS NPN BJT current gain.

3.6 SCR breakdown

Figure 3.10 shows the evolution of the current in two SCRs when the power supply is slowly increased to 5 V. Note that one of the devices is completely damaged (the current drops to a negligible value), whereas the other device is only partially damaged (the current drops only a little).

These results show that it is not always possible to distinguish a damaged SCR from an intact SCR. However, it is possible to compare the SCR's characteristics for low power supply voltages before and after an experiment to determine if the SCR has been damaged or not. As an example, the before and after characteristics of the second SCR are shown in

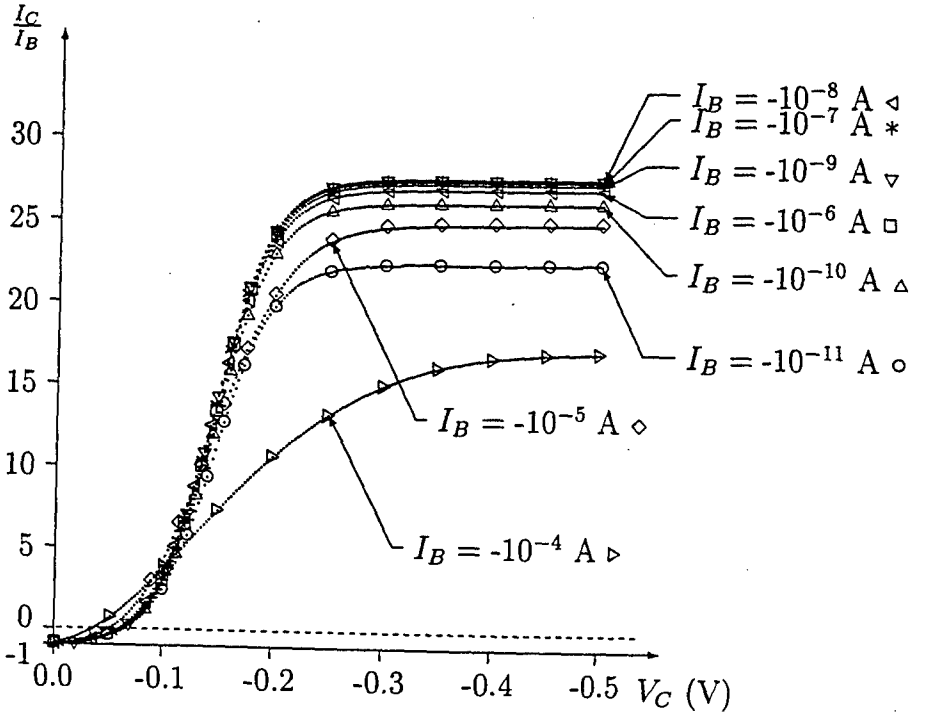

Figure 3.9: Bulk CMOS PNP BJT current gain.

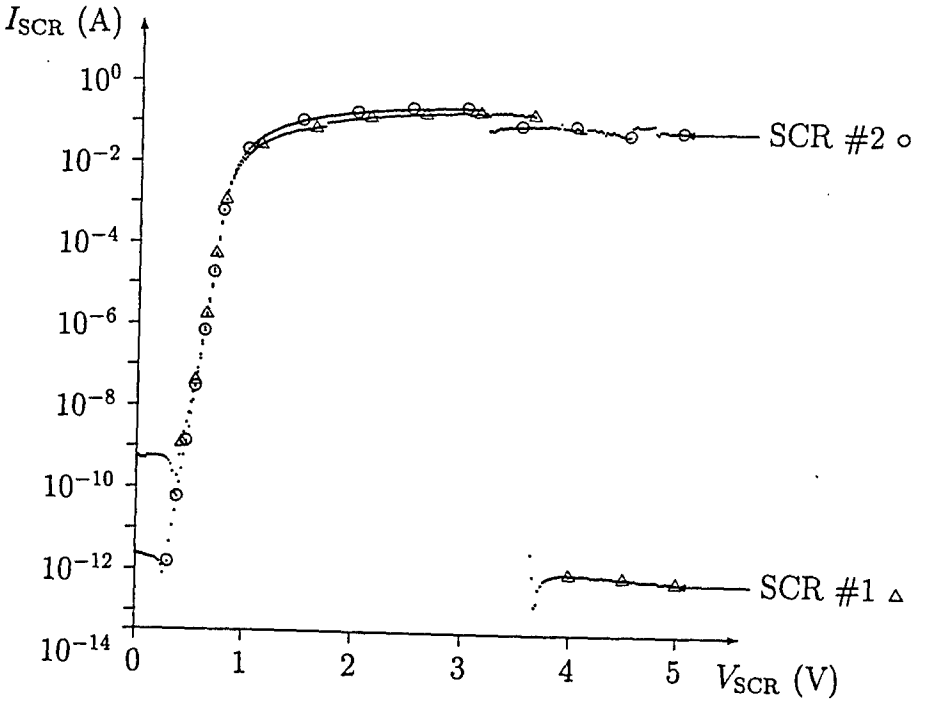

Figure 3.10: SCR breakdown evolution.

Figure 3.11.

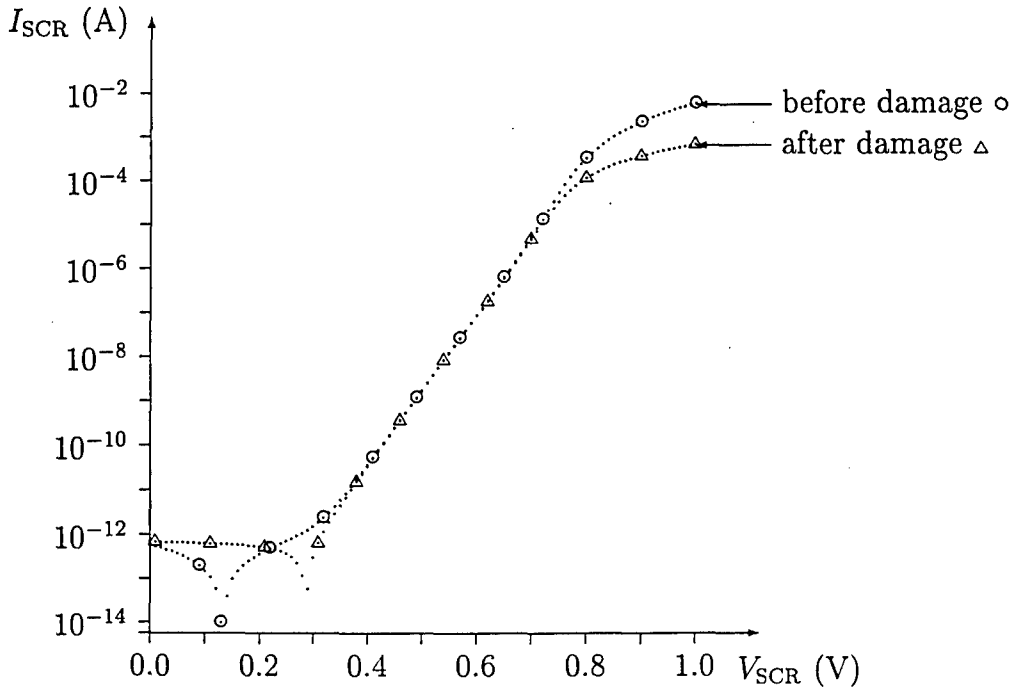

Figure 3.11: Before and after characteristics for SCR #2.

3.7 Summary

Latch-up is a problem specific to bulk CMOS processes. Its occurrence can be quite damaging, although it is not always possible to readily observe the damage (see Section 3.6). Latch-up in a CMOS integrated circuit can be characterized by a large current draw from the power supply, and by the P-substrate voltage approaching the power supply while the N-well voltage approaches ground (see Section 3.4).

Chapter 4

Latch-up detection and cancellation

4.1 Latch-up detection

Of the distinguishing features of latch-up described in Section 3.4, current monitoring of an integrated circuit is the most impractical to use: by the time the latch-up current becomes sufficiently large to detect, the damage may have already occurred.

Furthermore, the latch-up current may only represent a small fraction of the normal overall integrated circuit current consumption. This can easily occur if the latch-up condition is limited to a single well. Such a *micro-latch-up* condition will produce a current which is dependent on geometric factors such as the size of the well and the number of transistors in the well. Nevertheless, a micro-latch-up in a vital portion of the integrated circuit may incapacitate the chip as a whole.

A better approach is to monitor the substrate and/or well voltages. These signals provide an early warning of an imminent latch-up, because the base-emitter junctions of the BJTs in the SCR structure can be allowed to rise as much as 0.5 V without a significant amount of current flow.

4.2 Latch-up cancellation

One approach to latch-up cancellation is to inject current into the well and the substrate. Although this approach works, the amount of current required is roughly equal to the latch-

up current itself. Therefore, it is unlikely that this approach can be made to work on an integrated circuit without additional, higher-voltage power supplies.

The only remaining practical method of eliminating a latch-up condition is to power down the affected circuit and then power it back up again. This can be achieved automatically by supplying the integrated circuit with a transistor acting as a power switch. Unfortunately, powering down the entire integrated circuit as a whole may not be a practical option. However, by monitoring each well voltage separately, we can precisely identify the location of a micro-latch-up event, and thereby limit the corrective action to the immediately affected areas.

4.3 Integrated circuit requirements

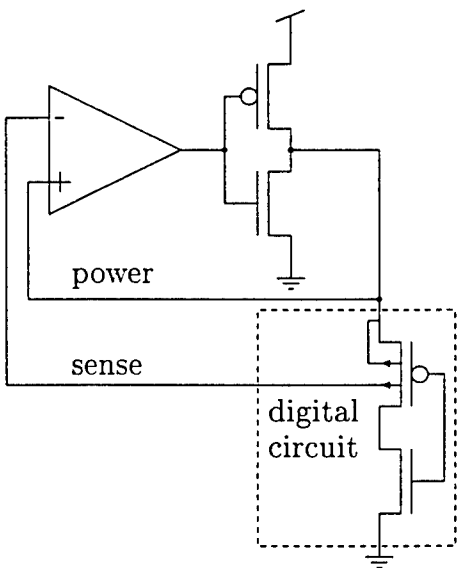

Figure 4.1: Latch-up detection and cancellation circuit.

Given an N-well bulk CMOS process, Figure 4.1 shows the components required to detect and cancel latch-up in a digital circuit. A simple digital inverter is used as a sample digital circuit to illustrate the connections. The required components are as follows:

First, each N-well must have a sense point to monitor its potential. Unfortunately, since the N-well potential must be compared to the power supply to detect latch-up, the existing

N-well contact cannot be used as a sense point because it is already connected to the power supply specifically to inhibit latch-up. A separate *sense contact* must therefore be supplied for each well. This is why the PMOS transistor in the digital circuit shown in Figure 4.1 is drawn with two N-well terminals.

Second, a differential sense amplifier is needed to compare the N-well voltage to the power supply. If the N-well voltage begins to dip below the power supply voltage, this indicates that the parasitic PNP bipolar transistor is turning on, and that latch-up is imminent. The sense amplifier's inverting input is connected to the sense contact and the non-inverting input is connected to the digital circuit's local power supply.

Third, a PMOS power switch is needed to turn off the power to the digital circuit to cancel a latch-up condition. When the sense amplifier detects latch-up, it turns the PMOS power switch off until latch-up is eliminated.

As shown in Figure 4.1, the sense amplifier is also connected to an NMOS transistor. This additional transistor is turned on when latch-up is detected to absorb any leakage currents from the PMOS transistor (in its off state) which might otherwise be sufficient to keep the SCR latched-up.

The sense amplifier operates as follows. Under normal operating conditions, both the sense and the power inputs to the sense amplifier are at the power supply rail. When the digital circuit latches-up, the N-well voltage will drop, causing the sense amplifier's output to go high. This shuts the power off so that both the sense and the power inputs start to fall to ground. The sense input will remain below the power input as long as the SCR is latched-up, until both inputs reach ground. At that point, since the two input have become equal again, the sense amplifier's output will go low, restoring power to the digital circuit.

4.4 Sense amplifier characteristics

Since the sense amplifier must operate properly with both inputs at the power supply and with both inputs at ground, a *rail-to-rail* differential amplifier is needed. Furthermore, since the output must be low whenever the inputs are equal, the sense amplifier must have a small positive input offset voltage.

After testing a number of differential transconductance amplifiers, we have selected the rail-to-rail wide-range transconductance amplifier shown in Figure 4.2 for use as the sense amplifier.

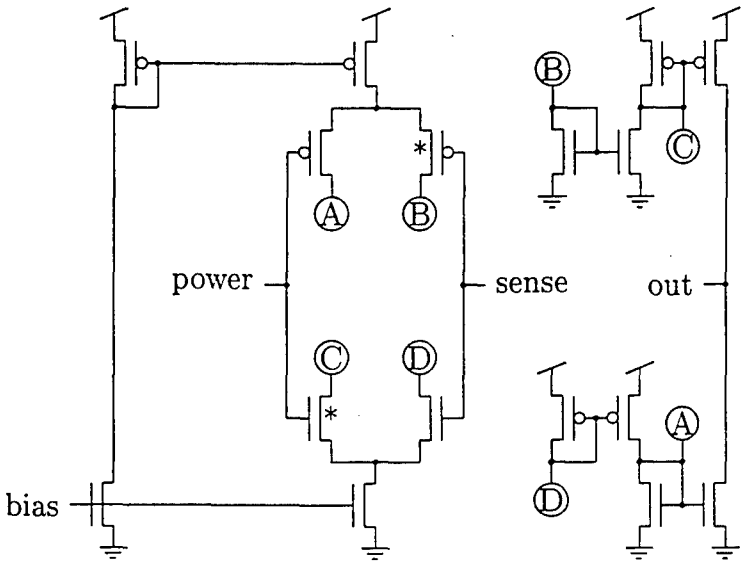

Figure 4.2: Rail-to-rail wide-range transconductance amplifier.

This sense amplifier is composed of two differential pairs. The NMOS differential pair can operate close to the power supply rail, but not near ground, whereas the PMOS differential pair can operate close to ground but not near the power supply rail. By combining the currents from both differential pairs using four current mirrors, we get the rail-to-rail operation that we need (although the transconductance will drop by half when the inputs are near either rail).

The input offset is provided by an *inverted select* transistor (marked with an *) in each differential pair. The inverted select transistors (described in more detail in Appendix B) produce approximately 30 mV of input offset independent of the bias conditions.

An additional current mirror supplies the bias current for the PMOS differential pair by copying the bias current for the NMOS differential pair.

4.5 Latch-up test setup

In order to test the proposed circuit, we fabricated an SCR structure in a 0.5μm process. The SCR is composed of four parallel bands of diffusion, 1.2μm wide, 104.4μm long and 3.6μm apart. The die containing the SCR was glued to a PC board, and the SCR was wire-bonded to copper pads on the PC board. The board was mounted in a fixture with a 4.15 mW, 660 nm visible laser diode module.

When turned on, the laser produced approximately 20μA of photo-current in the reverse-biased N-well to P-substrate junction of the SCR. This allowed us to simulate a contact resistance as low as 100KΩ and still trigger the SCR with a laser pulse.

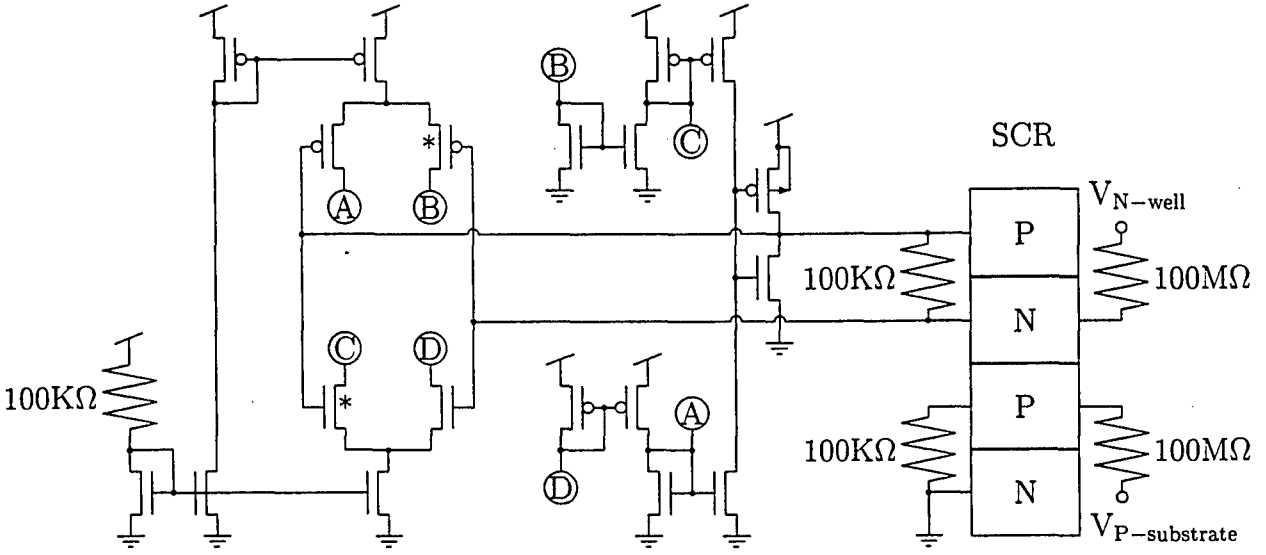

Figure 4.3: Latch-up test setup with sense amplifier, power switch, SCR and biasing resistors.

The full test setup is shown in Figure 4.3. The sense amplifier was constructed of discrete annular transistors (as described in Section 2.6) fabricated in a 1.2μm process. The bias current was provided by a 100KΩ resistor to the power supply, resulting in approximately 50μA of bias current.

Two 100KΩ resistors were attached to the SCR, one each between the top and bottom P-N junctions, to simulate the contact resistance. One 100MΩ resistor was connected to each of the internal layers to allow the monitoring of the internal voltages during the test: these 100MΩ resistors where connected to an oscilloscope with 1MΩ input impedance, resulting

in a 100 fold attenuation of the signals, but preventing the oscilloscope from affecting the SCR's operation.

4.6 Test protocol

The laser was pulsed on and off with a computer controlled square wave to periodically trigger the SCR. The on duration of the laser module supply voltage was 15 ms and the off duration was 5 ms. Note however, that the laser does not turn on instantly when the power is applied to its module.

Although the PMOS power switch in Figure 4.3 was initially unable to supply enough current to allow the SCR to be damaged by latch-up, additional transistors were added in parallel up to a total of six. The SCR was characterized before the experiment, and again after the addition of each extra transistor. No damage to the SCR was detected.

After the last test, the sense input of the sense amplifier was tied to the power supply to disable the latch-up detection. Then the circuit was tested again, and the SCR was characterized to confirm that it had been obviously damaged.

4.7 Results

Typical waveforms for the laser control signal, the SCR power supply, and the SCR internal voltages are shown in Figures 4.4 and 4.5.

In Figure 4.4, the laser module is turned on at the 2 ms time mark, and the sense amplifier shuts off the power approximately 5 ms later.

Figure 4.5 shows that no detectable rise in the P-substrate potential occurs before the sense amplifier shuts off the power. Directly measuring the SCR current confirms that no detectable current flows through the SCR before the sense amplifier reacts to prevent damaging latch-up conditions. Note that while the power is off, the P-substrate potential rises by about 0.3 V and the N-well potential goes to about -0.4 V: this is due to the photo-voltaic effect of the laser on the SCR, and prevents the sense amplifier from restoring power until the laser is turned off.

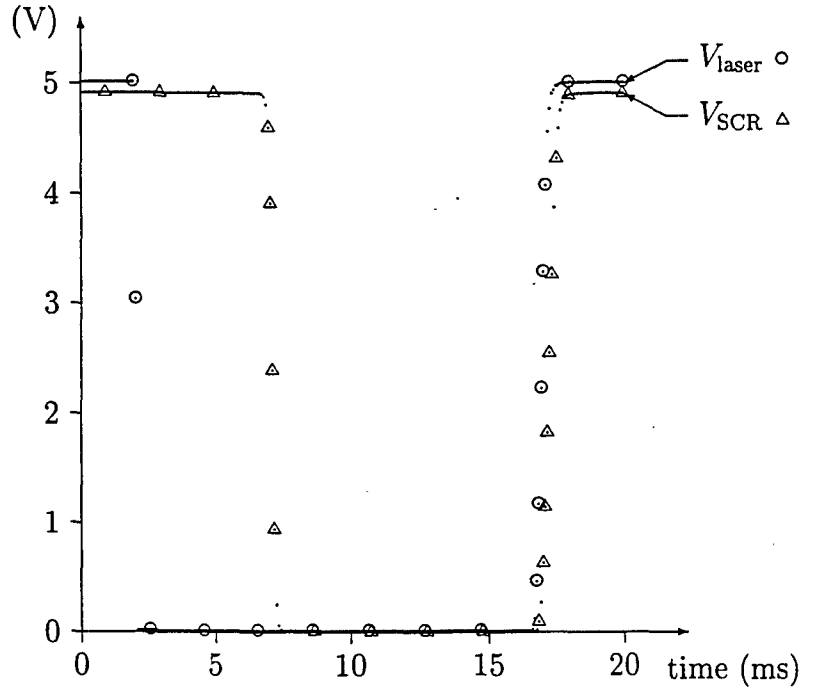

Figure 4.4: Laser control signal and SCR power supply. Laser is on when V_{laser} is low.

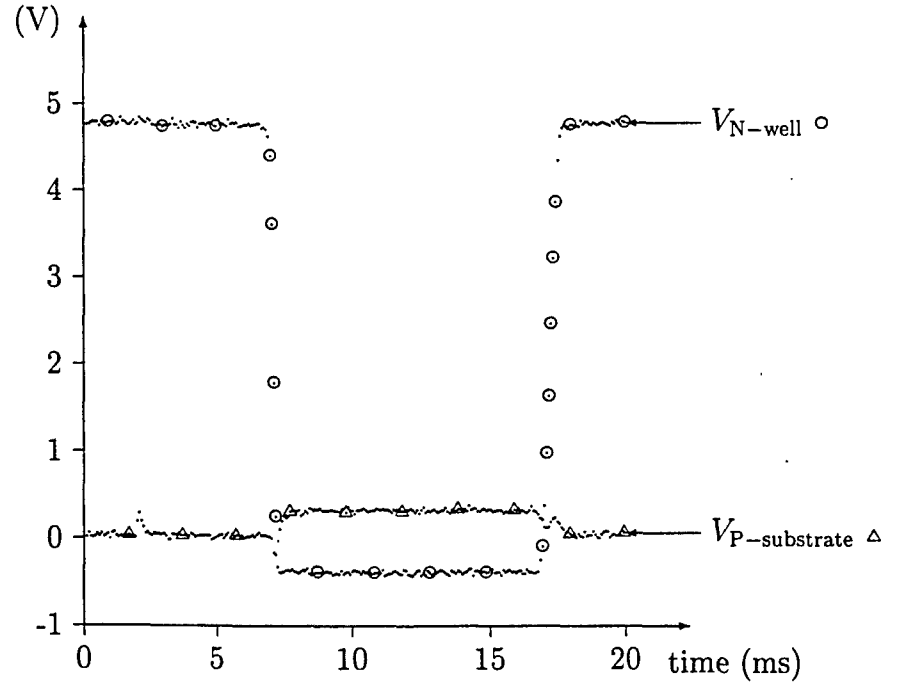

Figure 4.5: P-substrate and N-well potentials during test.

Figure 4.6 shows the SCR characteristics at the beginning and in between test runs. Figure 4.6 also shows the SCR characteristics after intentionally damaging the device by connecting the sense input of the sense amplifier to the power supply.

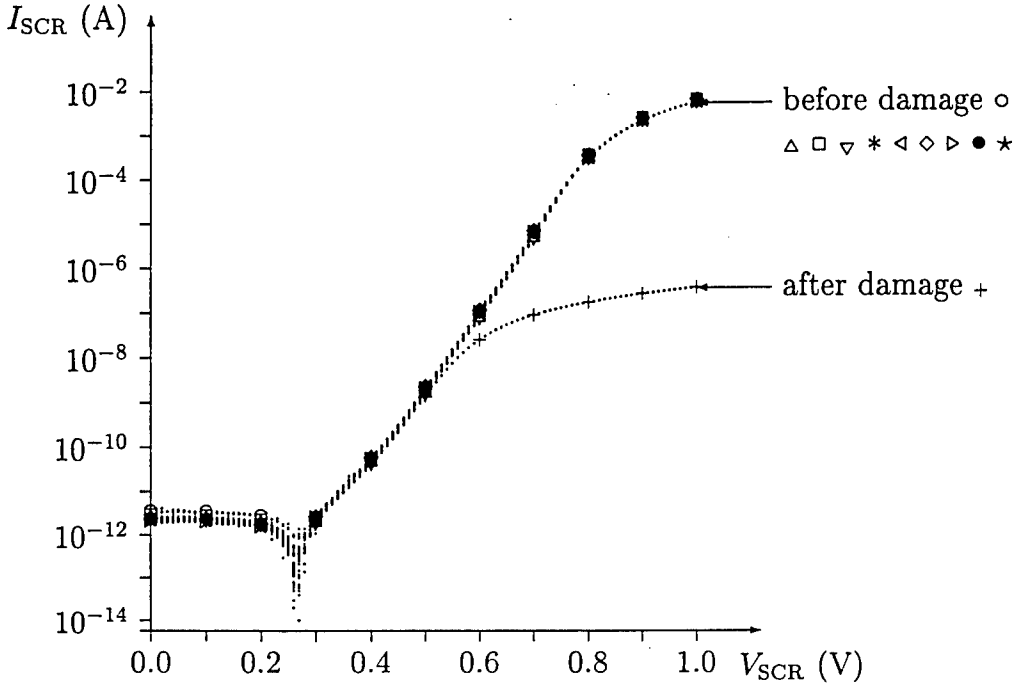

Figure 4.6: SCR characteristics before, between and after tests.

4.8 Limitations

The first principal limitation of the proposed latch-up detection and cancellation approach is that the sense amplifier circuit and power switch are themselves susceptible to latch-up.

Fortunately, there are some techniques which can be leveraged to reduce the latch-up probability (see Appendix A), as well as post-fabrication chemical etching techniques to isolate the N-well(s) from the substrate. These methods all consume a large area on a chip, but can still be practical for the small number of sense amplifiers and power switches envisioned for an integrated circuit.

The second principal limitation is that the circuitry affected by the latch-up cannot

operate when the power supply is shut off. Thus, minimizing the probability of a latch-up event, and providing redundant circuits are a necessity if the integrated circuit is to function normally through a latch-up event.

4.9 Summary

Latch-up detection and cancellation by means of on-board sense amplifiers and power switches appears to be a viable way of protecting an integrated circuit from the damaging effects of latch-up. Furthermore, by localizing the occurrence of latch-up, we can limit the temporary effects of latch-up cancellation to a small part of the overall integrated circuit, and employ redundant circuitry to continuously maintain the integrated circuit's functionality.

Chapter 5

Competing technologies

Unlike the methods proposed in this report, current methods for handling latch-up issues all involve specialized fabrication processes in which latch-up is known not to occur.

There are two principle categories for these processes: the enhancement/depletion technologies which avoid latch-up by not providing complimentary devices, and the twin-tub-like technologies which avoid latch-up by electrically isolating the complimentary devices from one-another.

5.1 Enhancement/Depletion technologies

There are two common types of enhancement/depletion integrated circuit processes: those made from silicon (Si), and those made from gallium-arsenide (GaAs).

5.1.1 Silicon

Before the advent of the current CMOS technologies, silicon integrated circuits based on the field-effect transistor (FET) were manufactured in NMOS (N-channel metal-oxide-semiconductor) technologies.

As its name implies, NMOS processes does not have any complimentary devices, only N-channel transistors. However, these transistors can be fabricated with two (or more) different threshold voltages. Typically, for digital processes, one device has a high, positive threshold

voltage so that it can be turned on or off depending on whether the transistor's gate terminal is at the upper or lower power supply rail, and the other device has a zero or slightly negative threshold voltage, so that it behaves like a current source when the gate and source are tied together.

The first type of device is referred to as an *enhancement mode* device, and is the same as the type used in CMOS processes. The other device is referred to as a *depletion mode* device and is typically used in place of a passive load in digital gates (where the P-channel transistor would normally go in a CMOS process) and is often fabricated by simply omitting the channel implant step in the fabrication.

Since the NMOS process does not have P-channel transistors, and since the depletion mode device can be fabricated along with the enhancement mode device without additional steps, the fabrication of an NMOS integrated circuit is simpler and therefore cheaper than the fabrication of a CMOS integrated circuit. (This is part of the reason why NMOS predates CMOS.)

The disadvantages of NMOS processes are principally the inability to make wired-NAND structures (due to the lack of an active pull-up device), the static power dissipation of any digital gate whose output is low, and the lower operating speed. These disadvantages have all but driven NMOS processes from the current marketplace.

5.1.2 Gallium-arsenide

Typical GaAs processes also use only enhancement and depletion mode N-channel transistors. Although the operation of GaAs integrated circuits is similar to the silicon NMOS counterpart, there are two distinctions: the electron mobility is higher in GaAs, and GaAs processes usually use MESFETs (metal-semiconductor field-effect transistors, see Section 2.1).

The higher mobility of GaAs allows the transistors to switch faster than similarly sized silicon transistors. However this advantage is substantially diminished by the fact that contemporary silicon processes typically use smaller lithographic feature size, which results in horizontal electric fields between the source and the drain in excess of $10^4 \frac{V}{cm}$. In the presence of such high field strengths, the carrier velocity in silicon exceeds that in GaAs resulting in a faster device for a fixed channel length and drain to source voltage. Furthermore, in a

MESFET, the gate terminal is a metal-semiconductor rectifying contact to the channel, so that the transistor's gate conducts current when the gate to source voltage is positive. This typically leads to increased static power dissipation in GaAs VLSI circuits.

However, gate current does have some benefits. For instance, in GaAs Direct-Coupled FET Logic (DCFL), each digital logic gate is composed of a single depletion mode output load transistor, and a network of multiple enhancement mode input transistors (in parallel and serial combinations). In this arrangement, when the gate's output is low, the current from the load transistor flows through the network of input transistors, and when the gate's output is high, the current from the load transistor flows into the gate of the input transistors in the next DCFL gate. Thus, the DCFL gate's current consumption remains constant regardless of the state of the gate's output (reducing ground bounce), and the voltage swing at the output is limited. Consequently, a DCFL gate can be operated faster than a comparable NMOS process digital logic gate.

Nevertheless, the high power dissipation of GaAs as well as its high fabrication costs have overwhelmed its speed advantage over silicon CMOS for most applications. Therefore, although GaAs has been used in the past as the basis for high speed digital logic families (*e.g.* for super-computers), it is currently relegated to the high-end RF circuits area, where power dissipation is dominated by its dynamic rather than its static component.

5.2 Twin tub technologies

Twin tub or twin well technologies, beginning with DI (Dielectric Isolation) and continuing with SOI (Silicon-On-Insulator), are processes that have been around for a while and have found extensive use in analog and high-voltage integrated circuits, mainly in the telephone company and automatic test equipment industry. Within the SOI category is also SOS (Silicon-On-Sapphire), in which the underlying support substrate is a sapphire wafer instead of a silicon wafer.

5.2.1 Dielectric Isolation and Silicon-On-Insulator

The early objectives of DI and SOI were to allow high temperatures and very high voltages (100 to 600 Volts) – but not necessarily high currents – on an integrated circuit by isolating transistors from one another and confining the electric fields from the high voltages to the area immediately enclosing each transistor. Currently however, SOI has become particularly attractive to the VLSI community for its high packing density and low parasitic capacitance in sub-micron designs.

Both DI and SOI are formed of silicon islands or tubs, each containing a single transistor. Each tub is isolated from its neighbors and the silicon substrate by a layer of silicon dioxide. The thickness of the silicon dioxide and of the silicon film from which the islands are formed varies according to the application. The main purpose of the silicon islands is to confine dopants during drain/source implantation and to confine the electric fields during device operation. As an added benefit, the drain/source to substrate leakage current and capacitance of the bulk CMOS process are eliminated, which permit increased operating temperatures and speeds.

Unfortunately, although the thorough isolation via silicon dioxide is a necessary feature of DI and SOI, it is also an impediment to heat dissipation due to the poor thermal conductivity of silicon dioxide. Higher power dissipation can be achieved by thinning the insulating layer between the transistor and the substrate (the buried oxide), but only at the cost of increased capacitive coupling between the transistor and the substrate. Although the higher permissible operating temperatures somewhat offsets the poor power dissipation, local or self-heating can cause somewhat erratic transistor characteristics, which must be taken into account during the design process.

The key to the rapid adoption of SOI by former bulk CMOS manufacturers is the increased chip functionality without the cost of major process equipment changes such as higher resolution lithographic tools. The lack of a well to substrate junction permits much closer spacing of complimentary devices, and high-performance digital processes tend to omit body contacts in favor of even higher transistor density. However, for moderately thin silicon films, the lack of body contacts results in the so-called *kink effect*, in which a variation in the

transistor characteristics is observed due to random charge accumulation in the body (see Section 2.4). The kink effect can be eliminated by making the silicon film thinner than the depletion region under the channel, such as in Fully Depleted SOI (FDSOI) or Thin Body SOI However, the film doping, channel implant and threshold voltage must be carefully balanced to maintain full depletion over the transistor's entire operating range.

The transistor's self-heating and kink effects are the main challenges to SOI design, due to their dependence on the device's somewhat unpredictable past operation history. Nevertheless, the eventual switch of bulk CMOS foundries to SOI currently seems inevitable, and there is some evidence that SOI will be well suited to radiation environments due to the absence of latch-up, and the small transistor radiation interaction volumes.

5.2.2 Silicon-On-Sapphire

Silicon-On-Sapphire (SOS) was invented in an attempt to eliminate the thermal barrier caused by the buried oxide layer in SOI. Since sapphire is an electrical insulator, the silicon film can be placed directly onto the sapphire, and the wafer can then be processed just like an SOI wafer. Although sapphire's thermal conductivity ($25\frac{W}{m\cdot°K}$ at 100°C) is lower than silicon($109\frac{W}{m\cdot°K}$ at 100°C), it is much higher than silicon dioxide, which effectively eliminates any noticeable self-heating effects. Moreover, sapphire is transparent to visible light and infrared, making it an ideal choice for opto-electronic devices.

The difficulty with SOS lies in achieving a good interface between the sapphire and the silicon. For very thin silicon films, this interface must be almost as free of defects as the interface between the silicon and the gate oxide. The silicon dioxide to silicon interface is comparatively well understood, since it is present in every MOS transistor, and has facilitated the transition to SOI. In contrast, early SOS integrated circuits were plagued by low transistor yields due to the poor interface at the sapphire substrate. Consequently, SOS has been restricted to high-speed but low transistor count circuits such as RF components.

5.3 Summary

Each competing technologies (NMOS, GaAs, SI, SOI, SOS) has its share of advantages (latch-up and/or radiation immunity) and drawbacks (typically power dissipation and/or cost). Although none of these technologies is too common today, the adoption of any of them, by DRAM or microprocessor manufacturers for instance, could significantly change their popularity.

Chapter 6

Conclusion and recommendations

6.1 Conclusion

We have found that latch-up can indeed be detected by observing the well potential, and that latch-up can be subsequently cancelled by shutting down the power to the latched-up portion of the integrated circuit. We have verified that the latch-up detection is both early enough and fast enough to allow the shutdown to complete before the integrated circuit can be damaged.

However, we have not tested the latch-up detection and cancellation components in a radiation environment, nor have we attempted to trigger latch-up with high energy radiation.

Furthermore, the circuits were tested at a rather low repetition rate. It is conceivable, that if an integrated circuit receives too much latch-up inducing stimulation, the power will remain off, and no useful work will be performed.

Finally, the tests were performed using circuits constructed out of discrete components (although fabricated on a single integrated circuit). Scaling of the circuitry so that all the circuitry and wiring is on a single chip may expose unforeseen parasitic capacitance and inductance effects.

6.2 Recommendations

Addressing the shortcomings of this work will require a larger research effort in the following areas.

6.2.1 Process variation

Integrated circuit characteristic will vary from one manufacturer to another (*e.g* HP and AMI), from one process to another from the same manufacturer (*e.g* AMI 1.2μm and 0.5μm processes), and even from run to run of the same process.

However, all the parts used for the tests came from the same run. A thorough investigation of the effectiveness of the proposed technique would require having the same circuits fabricated on multiple runs and processes, and by different manufacturers: we may discover that some of the variations that we encounter may render the latch-up detection and cancellation circuits unusable.

6.2.2 Radiation exposure

To reduce the cost of testing the latch-up detection and cancellation circuits, the test integrated circuit was manufactured in an older technology which is known to be susceptible to large threshold shifts in the presence of radiation (as shown in Chapter 2).

These circuits need to be re-fabricated in a more modern, sub-micron process, which are typically much less sensitive to radiation. The chips should then be exposed to radiation to determine whether or not the circuits still operate properly, and what level of total-ionizing dose (TID) can be withstood.

6.2.3 Radiation induced latch-up

Full testing of the latch-up detection and cancellation technique also needs to be performed under more realistic conditions. Although photo-activation of the SCR is useful as a testing method, a circuit in a real space application will be exposed to photons of much higher energies, along with other high energy particles (electrons, protons, beta particles, *etc.*).

To this end, some testing needs to be performed at a particle beam facility, such as Brookhaven National Laboratory (BNL). At the BNL facility, devices can be placed in the path of a high energy particles of various mass to trigger the latch-up condition. This would result in a more realistic test, because the particle energies are closer to what one would expect in nature, and an accelerated test rate, because the particle densities are higher.

6.2.4 Redundancy

The circuits protected by the latch-up detection and cancellation technique are not functional during a latch-up cancellation cycle due to the lack of power. If the latch-up event is sufficiently rare, it may be possible to restart the interrupted work from scratch. Some integrated circuits, such a microprocessors, may be able to recover gracefully from more frequent interruptions by using some sort of work check-pointing method. However, for the remaining integrated circuits, some provision must be made for redundant circuitry to take over.

Circuit redundancy methods need to developed in order to produce useful integrated circuits. A digital circuit could use a simple voting scheme, or a more sophisticated approach, such as using the sense amplifier output to decide whether or not a circuit's output is valid.

6.2.5 Complex digital circuits

The testing of the latch-up detection and cancellation technique was performed with an SCR rather than with a digital circuit, because no sacrificial digital circuit containing the required sense contacts was available in sufficiently large quantities for testing.

Although some effort was made to gauge the effect of a running digital circuit on the latch-up detection and cancellation technique, no conclusive result was obtained, and therefore the data has not been included in this report.

A "full test" of the technique is therefore desirable. Ideally, this should include a fairly large digital circuit with redundant circuitry, tested under radiation (both TID and high-energy), in which the digital circuit can be verified to be computing properly throughout the test, while portions of the circuit exhibit micro-latch-up.

Bibliography

[1] Mark N. Martin. *Integrated Circuit Design Considerations for Spacecraft VLSI Implemented in Standard CMOS Processes.* Ph.D. Dissertation, The Johns Hopkins University, Department of Electrical and Computer Engineering, Baltimore, MD, 2000.

[2] Adel S. Sedra and Kenneth C. Smith. *Microelectronic Circuits,* third edition. Saunders College Publishing, 1991. ISBN 0-03-051648-X.

[3] Ronald R. Troutman. *Latchup in CMOS Technology.* Kluwer Academic Publishers, 1986. ISBN 0-89838-215-7.

[4] Neil Weste and Kamran Eshraghian. *Principles of CMOS VLSI Design.* (First Edition) Addison-Wesley Publishing Company, 1985. ISBN 0-201-08222-5.

[5] Carver Mead. *Analog VLSI and Neural Systems.* Addison-Wesley Publishing Company, 1989. ISBN 0-201-05992-4.

[6] Stephen A. Campbell. *The Science and Engineering of Microelectronic Fabrication.* Oxford University Press, 1989. ISBN 0-19-510508-7.

[7] Sorin Cristoloveanu, SOI: A Metamorphosis of Silicon. *IEEE Circuits & Devices.* Vol. 15, No. 1, January 1999. ISSN 8755-3996.

[8] John Y. Chen. *CMOS Devices and Technology for VLSI,* Chapter 8. Prentice Hall, 1990. ISBN 0-13-138082-6.

Appendix A

SCR latch-up limits

A.1 Effect of contact resistance

The equivalent circuit of the parasitic SCR with the contact resistances is shown in Figure A.1.

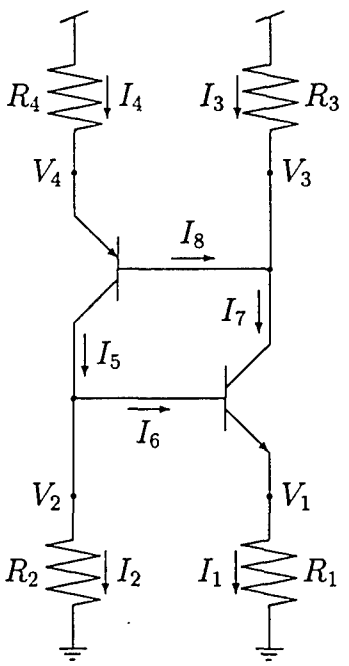

Figure A.1: Equivalent circuit for an SCR in a bulk CMOS process with contact resistances.

For latch-up to occur, the loop gain around the BJTs must be greater than unity. Using

the following equations for the bipolar transistors

$$I_7 = I_{\text{npn}}\left(e^{\frac{V_2 - V_1}{V_T}} - 1\right)$$

$$I_7 = \beta_{\text{npn}} I_6$$

we can compute the relationship between I_5 and I_7 due to the NPN transistor:

$$V_1 = I_1 R_1 = (I_6 + I_7)R_1 = \frac{\beta_{\text{npn}} + 1}{\beta_{\text{npn}}} I_7 R_1$$

and

$$V_2 = I_2 R_2 = (I_5 - I_6)R_2 = I_5 R_2 - \frac{1}{\beta_{\text{npn}}} I_7 R_2$$

and

$$V_2 - V_1 = V_T \ln\left(\frac{I_7}{I_{\text{npn}}} + 1\right)$$

therefore

$$I_5 R_2 - \frac{1}{\beta_{\text{npn}}} I_7 R_2 - \frac{\beta_{\text{npn}} + 1}{\beta_{\text{npn}}} I_7 R_1 = V_T \ln\left(\frac{I_7}{I_{\text{npn}}} + 1\right)$$

or equivalently

$$\frac{I_5}{I_7} = \left(\frac{1}{\beta_{\text{npn}}} + \frac{\beta_{\text{npn}} + 1}{\beta_{\text{npn}}} \frac{R_1}{R_2}\right) + \frac{V_T}{I_7 R_2} \ln\left(\frac{I_7}{I_{\text{npn}}} + 1\right)$$

To ensure a large enough loop gain, we want to make $\frac{I_5}{I_7}$ smaller than unity, or at least smaller than α, where α is the gain for the PNP.

Since the ln function is sub-linear, the logarithm term can be made as small as we want by making I_7 suitably large. Therefore, the term to concentrate on is the reverse gain R_{npn}

$$R_{\text{npn}} = \frac{1}{\beta_{\text{npn}}} + \frac{\beta_{\text{npn}} + 1}{\beta_{\text{npn}}} \frac{R_1}{R_2} = \frac{R_2 + (\beta_{\text{npn}} + 1)R_1}{\beta_{\text{npn}} R_2}$$

A similar reverse gain can be computed for the PNP

$$R_{\text{pnp}} = \frac{R_3 + (\beta_{\text{pnp}} + 1)R_4}{\beta_{\text{pnp}} R_3}$$

The conditions under which latch-up can occur, can now be expressed as

$$R_{\text{npn}} R_{\text{pnp}} < 1$$

or equivalently

$$\frac{R_2 + (\beta_{\text{npn}} + 1)R_1}{\beta_{\text{npn}} R_2} \frac{R_3 + (\beta_{\text{pnp}} + 1)R_4}{\beta_{\text{pnp}} R_3} < 1 \qquad \text{(A.1)}$$

A.2 Reducing BJT gain

Equation A.1 indicates that if we can make the BJT current gain small enough, then the product of the reverse gains can be made larger than one.

In a BJT, charge carriers from the emitter travel through the base to the collector. Consequently, the current gain is limited by the amount of carrier recombination in the base, and for this reason, bipolar transistors are generally designed to have a thin base. Conversely, we can artificially reduce the BJT's gain by thickening the base.

Since the PNP transistor is mainly a vertical structure, the base thickness is simply the depth of the N-well (as shown in Figure 3.3), and can only be adjusted by the integrated circuit fabrication facility. The current gain of this parasitic device is therefore fixed, and can be as high as 27.7 as shown in Figure 3.9, or as high as 100 in other CMOS processes.

On the other hand, the NPN transistor is mainly a lateral structure, for which the base thickness is the distance between the N-diffusion emitter (to the right of the NMOS transistor in Figure 3.3) and the N-well. The measured the gain of the NPN transistor for various distances of the emitter to the N-well as a function of base current is shown in Figure A.2 for a 1.2μm process.

In Figure A.2, the top five curves are for the emitter at 16.2, 21.0, 25.8, 30.6, and 35.4 μm from the N-well. The bottom two curves are for the emitter at 85.8 and 90.6 μm from the N-well. As clearly be seen, the current gain does drop significantly as the distance increases.

However, the return diminishes steadily, so that even when we place the emitter very far away, the current gain still peaks at 0.6. Therefore, given the large gain of the PNP device, separation alone is unlikely to prevent the occurrence of latch-up.

A.3 Reducing contact resistance

Equation A.1 also indicates that if we can make R_2 smaller than R_1 and/or R_3 smaller than R_4, then the product of the reverse gains can be made larger than one. What this means is that the P-substrate and N-well contacts should have a lower resistance than the other contacts.

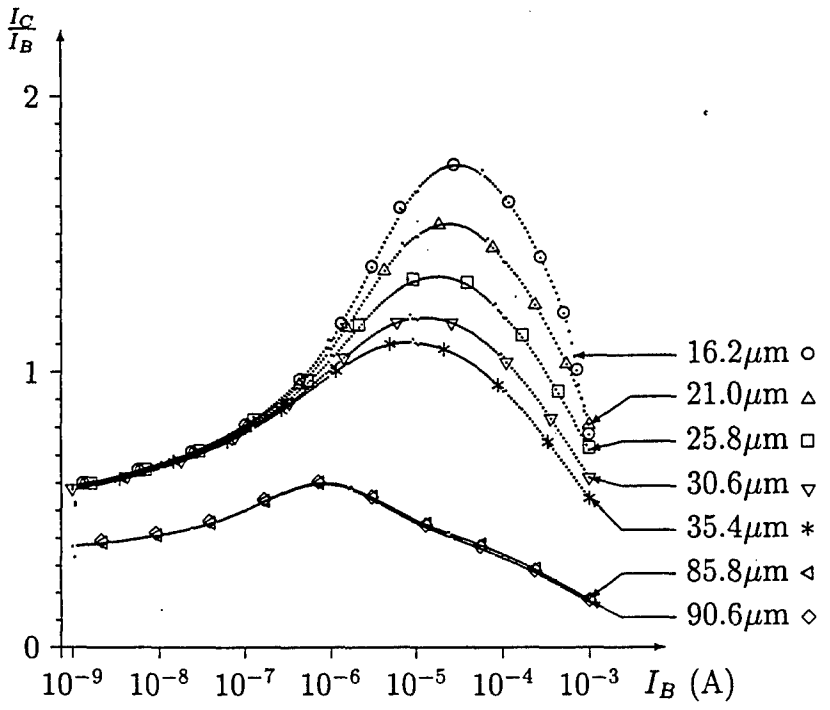

Figure A.2: NPN current gain dependence on geometry.

Although only the integrated circuit foundry can adjust contact resistances, we can also reduce the reduce the effective P-substrate and N-well contact resistance by using many contacts in parallel. This is typically achieved by placing a ring of contacts around each transistor.

A.4 Back gate

Another method that is commonly used to reduce the probability of latch-up is to bias the P-substrate below ground and the N-well above the power supply. Although this changes Equation A.1, it does not eliminate the dependence on the BJT current gain or the contact resistance. Furthermore, it has the undesirable effect of increasing the MOS transistor thresholds as shown in Figure 2.5.

A.5 Chemical etch

A final method that we have considered is to isolate the PMOS power transistors after fabrication using a chemical etch procedure. For instance, it is now possible to etch an integrated circuit die in tetramethyl ammonium hydroxide (TMAH) to suspend a N-well and isolate it from the P-substrate.

TMAH is an anisotropic etchant which will not etch an N-well, provided that bias voltages are maintained between the N-well, the P-substrate and the TMAH solution during the etching process. The result of a TMAH etch is the production of an inverted pyramidal pit under the suspended N-well.

Appendix B

Inverted select devices

B.1 Differential amplifier input offset

Real differential amplifiers are rarely perfect: they have finite gain, they have a non-zero common mode gain, and they have a non-zero input offsets. The input offset is simply the input differential signal required to get the amplifier's output to be midway between the power rails.

Usually, differential amplifiers are designed to have as small an input offset as possible. However, for the sense amplifier application of Chapter 4, a small predictable input offset is necessary.

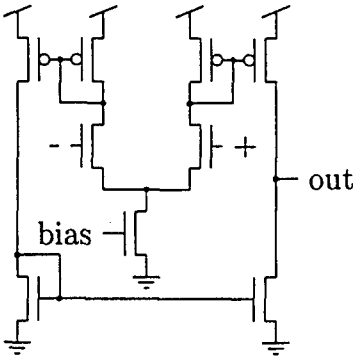

Figure B.1: Differential amplifier with positive input offset.

Producing an input offset can be achieved in several ways. For instance, in Figure B.1,

the input offset is produced by the mismatch between the input and output voltages of the current mirror: since the output voltage is almost always greater, the output current is always slightly larger than the input current. Thus, although the differential pair produces the same current on either side when the inputs are equal, the inverting input's current gets mirrored once more than the non-inverting input's current. Thus, at the output, the pull-down current is always larger than the pull-up current (for equal inputs), resulting in a low output signal.

Similar current mismatches can be obtained by intentionally changing the current mirror or differential pair transistor ratios. However, this results in an offset voltage whose magnitude is dependent on the bias current and the temperature.

B.2 Inverted select transistors

In order to get an invariant input offset voltage, one possible approach is to use floating gate transistors. However, floating gate transistors do not operate reliably in radiation environments. Another approach, which yields limited offset values, is to use *inverted select* transistors.

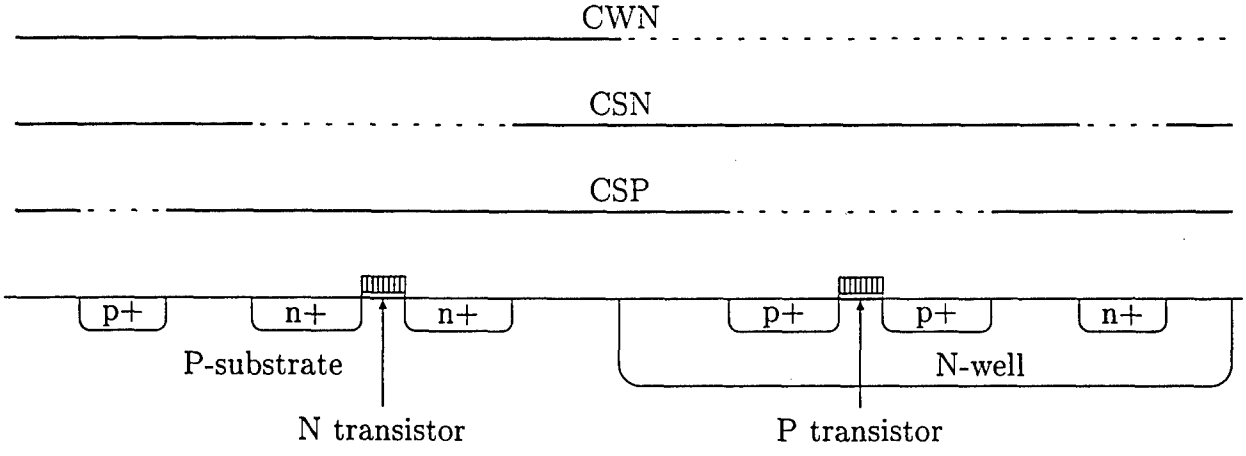

Figure B.2: Doping masks for an N-well bulk CMOS process.

Inverted select refers to the masks used to dope the silicon during the integrated circuit manufacture. Figure B.2 shows three of these masks for an N-well bulk CMOS process.

The CWN mask is used to create the N-well: where the mask is open, the dopant ions are permitted to diffuse into the substrate to form the N-well. Similarly, the CSN (N-select) mask marks the N-diffusion areas, and the CSP (P-select) mask marks the P-diffusion areas. Note that the select masks are normally open over the transistors as well as their source and drain area: this causes the polysilicon gate to be doped as well. If we switch the select mask over the polysilicon gate, the source and drain areas remain the same, but the polysilicon gets doped differently. The result is a slight change in the transistor's threshold voltage, as shown in Figure B.3 for the annular transistors used in Chapter 4.

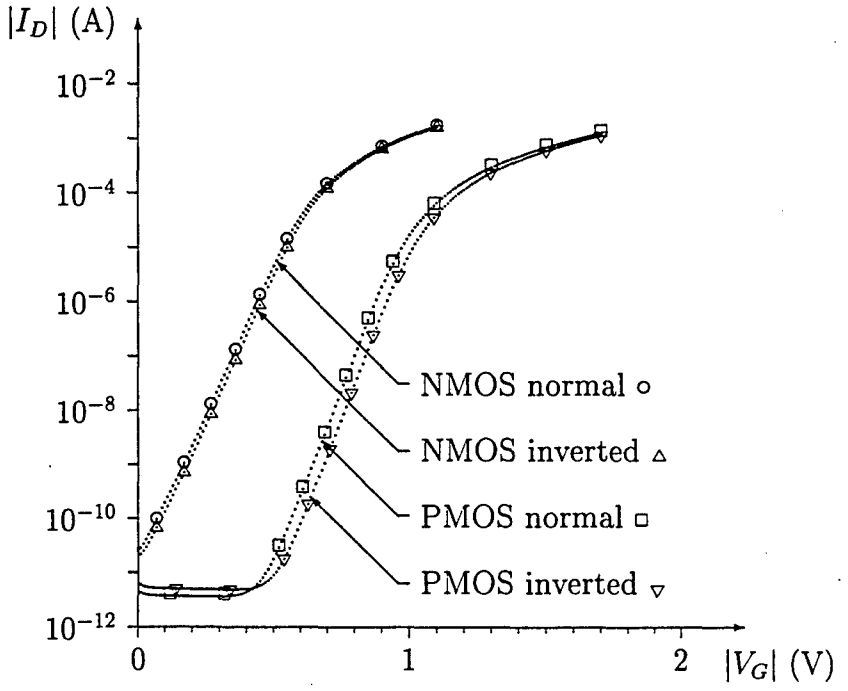

Figure B.3: Comparative inverted select MOS characteristics.

Figure B.4 shows how the transistor is constructed.

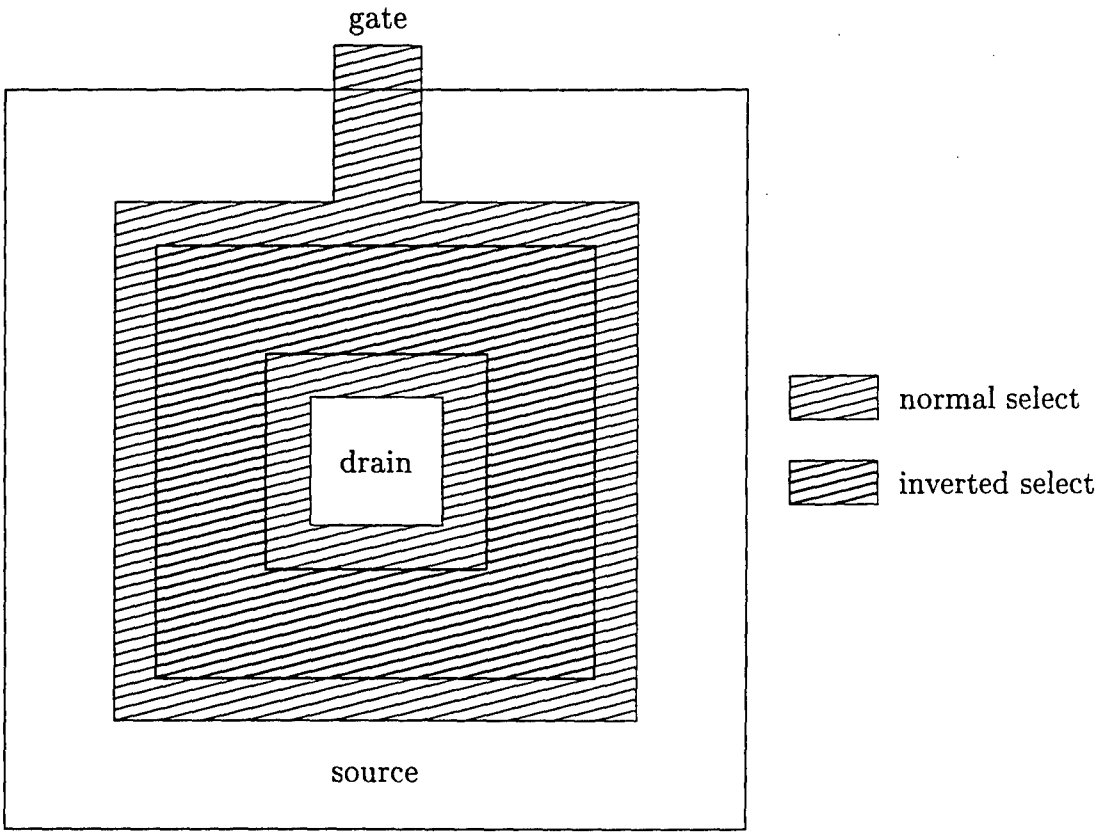

Figure B.4: Top view of annular inverted select transistor.

DISTRIBUTION LIST

DTIC/OCP 8725 John J. Kingman Rd, Suite 0944 Ft Belvoir, VA 22060-6218	1 cy
AFRL/VSIL Kirtland AFB, NM 87117-5776	2 cys
AFRL/VSIH Kirtland AFB, NM 87117-5776	1 cy
Padgett-Martin Technology 4222 Falls Road Baltimore MD 21211	1 cy
Official Record Copy AFRL/VSSE/Joseph Chavez	3 cys