AN ABSTRACT OF THE THESIS OF

DAVID MING-SHIH TAO (Name) for the DOCTOR OF PHILOSOPHY (Degree)

in Electrical and Electronics Engineering (Major) presented on Nov 8, 1972 (Date)

Title: A COMPLEMENTARY MOS RANDOM-ACCESS-MEMORY CELL WITH DOUBLE-DIFFUSED MOS TRANSISTORS

Redacted for privacy

Abstract approved: James C. Looney

The use of double-diffused n-type MOS transistor (DN-MOS) in a complementary MOS random-access-memory (CMOS RAM) cell is the main objective of this investigation.

DN-MOS transistors and conventional p-channel MOS transistors on the same chip have been successfully fabricated. Process sequence effects on device threshold voltage and channel length are discussed. The CMOS RAM cell operating characteristics were studied as related to the characteristics of devices in the cell. Both hybrid and monolithic CMOS RAM cells with DN-MOS transistors were realized. The theoretical predictions of cell switching characteristics are experimentally justified.

The advantage of using DN-MOS transistors in a CMOS RAM cell is found to be that a faster switching cell can be built using the same area of silicon as compared to a regular complementary structure.

A COMPLEMENTARY MOS RANDOM-ACCESS-MEMORY CELL WITH DOUBLE-DIFFUSED MOS TRANSISTORS

by

David Ming-Shih Tao

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

June 1973

APPROVED:

Redacted for privacy

Associate Professor of Electrical and Electronics Engineering
in charge of major

Redacted for privacy

Head of Department of Electrical and Electronics Engineering

Redacted for privacy

Dean of Graduate School

Date thesis is presented Nov. 8 1972

Typed by Erma McClanathan for David Ming-Shih Tao

ACKNOWLEDGMENT

The author wishes to express his deep appreciation to Professor James C. Looney for his continued guidance and helpful assistance throughout the course of this study.

The Department of Electrical and Electronics Engineering has provided financial aid in the form of graduate assistantship which made this investigation possible.

Finally the author wishes to dedicate this thesis to his wife, Susan, for her patience and encouragement. Special thanks should go to Mrs. McClanathan for carefully preparing this final copy.

TABLE OF CONTENTS

I. Introduction 1

II. Double-Diffused and Conventional MOS Transistors 3

A. Device Structure 3
B. Device Operation 6
C. Processes and Considerations 12
1. Process Steps 13
2. Experimental Results 15
3. Threshold Voltage Considerations 22
a. Net Channel Doping Effect on Threshold Voltage 22
b. Impurity Profile Calculations 27
c. Impurity Profile Control on DN-MOS Transistor Threshold Voltage 30
d. Silicon Orientation Effect on Threshold Voltage 38
4. DN-MOS Transistor Channel Length Considerations -- Punch Through 39
5. Diffusion Schedule Considerations 46
6. Other Considerations 48

III. The Complementary MOS Random-Access-Memory Cell 50

A. The CMOS RAM Cell Circuit 52
B. Cell D.C. Operation 54
C. Cell Dynamic Operation 64
1. Cell Step Response 65
2. Cell Pulsed Response 69
D. Cell Power Dissipation 72

IV. DN-MOS in CMOS RAM Cell 74

A. Monolithic Structural Cross-Section 74
B. Advantages of DN-MOS Transistor in CMOS RAM Cell 76
C. Hybrid Realization 77
D. Monolithic Realization 84

V. Conclusion 87

Bibliography 89

Appendix I. Description of Process Steps 91

a. Field Oxidation 92
b. P-MOS Source and Drain Diffusion 94

c. DN-MOS Opening, p-Diffusion Followed by n^+-Diffusion 95
d. Gate Opening Followed by Gate Oxidation 96
e. Open Contacts to Source and Drain, Metalization and Excess Aluminum Removal 97

Appendix II. Numerical Data of I_{ds} versus V_{gs} for P-MOS Transistor and DN-MOS Transistor 98

Table A. I_{ds} versus V_{gs} for P-MOS Transistor 99

Table B. I_{ds} versus V_{gs} for DN-MOS Transistor 99

Appendix III. DN-MOS Transistor Impurity Profile Numerical Values 100

Table C. Numerical Values for the Plot of DN-MOS P-Region Impurity Profile 101

Table D. Numerical Values for the Plot of DN-MOS n^+-Region Impurity Profile 102

Table E. DN-MOS Transistor Net Impurity Profile Numerical Values 103

LIST OF FIGURES

Figure		Page
1.	Cross-Section of DN-MOS	4
2.	Impurity Profile of DN-MOS Structure	4
3.	Cross-Section of Conventional P-MOS	5
4.	Enhancement MOS Transistors in Operation	8
5.	V_{th} of Both P-Channel and N-Channel MOS Transistors versus $\lvert N_A - N_D \rvert$ and N_{ss}	11
6.	Major Steps of Device Fabrication Process	13
7.	Device Composite Layout	15
8.	Photomicrograph of Devices	16
9.	P-MOS Transistor Characteristics	17
10.	DN-MOS Transistor Characteristics	19
11.	P-MOS Transistor $\sqrt{\lvert I_{ds} \rvert}$ versus $\lvert -V_{gs} \rvert$	23
12.	DN-MOS Transistor $\sqrt{I_{ds}}$ versus V_{gs}	24
13.	DN-MOS Threshold Voltage (V_{tn}) versus Net Channel Doping $\lvert N_A - N_D \rvert$ Near Source End	26
14.	DN-MOS Transistor Impurity Profiles	31
15.	DN-MOS Net Impurity Profile	32
16.	R_s Effect on DN-MOS P-Region Impurity Profile	34
17.	Drive-In Time Effect on DN-MOS P-Region Impurity Profile	36
18.	Enlarged Cross-Section of DN-MOS Transistor with Defining Quantities	40
19.	Depletion Region Width Variation Effect on Channel Length of DN-MOS Transistor	45
20.	The Random-Access-Memory System Block Diagram	51
21.	CMOS RAM Cell Circuit	52

Figure		Page
22.	CMOS Storage Cell	54
23.	CMOS Inverter	56
24.	Inverter Operating Point Locus	57
25.	Inverter Voltage Transfer Curve. Input Voltage (V_Q) versus Output Voltage ($V_{\bar{Q}}$)	60
26.	Inverter Current (I_P,I_N) versus Input Voltage (V_Q)	61
27.	Inverter with Step Input	65
28.	Simplified Inverter Step Response	66
29.	Step Response of Cell	69
30.	Cell Pulsed Response	70
31.	CMOS Storage Cell with Nodal Capacitors	72
32.	Monolithic Storage Cell	74
33.	Hybrid Cell	77
34.	Cell Switching Test Schematic	78
35.	Cell Switching	79
36.	Enlarged Cell Switching	80
37.	T_F Testing Circuit	81
38.	Fall Time Response	82
39.	CMOS RAM Cell Composite	84
40.	Monolithic CMOS RAM Cell Photomicrograph	85

LIST OF APPENDIX FIGURES

A.	Accidental MOS Structure Between Two Diffused Pockets	93

LIST OF TABLES

Table		Page
1.	V_{tn} for Various $\lvert N_A - N_D \rvert$	25
2.	Numerical Values of Depletion Region Width and Effective Channel Length	44
3.	Diffusion Schedules	47
4.	Static Cell States	55
5.	Transistor Operating Condition versus Input Voltage (V_Q) of an Inverter	58

LIST OF APPENDIX TABLES

A.	$\sqrt{\lvert I_{ds} \rvert}$ versus $\lvert V_{gs} \rvert$ for P-MOS Transistor	99
B.	$\sqrt{I_{ds}}$ versus V_{gs} for DN-MOS Transistor	99
C.	Numerical Values for the Plot of DN-MOS P-Region Impurity Profile	101
D.	Numerical Values for the Plot of DN-MOS n^+-Region Impurity Profile	102
E.	DN-MOS Transistor Net Impurity Profile Numerical Values	103

A COMPLEMENTARY MOS RANDOM-ACCESS-MEMORY CELL WITH DOUBLE-DIFFUSED MOS TRANSISTORS

I. INTRODUCTION

There are four MOS transistors in a complementary MOS random-access-memory storage cell. This study investigates the use of a special kind of MOS transistor called "double-diffused MOS transistor" as two of the required four MOS transistors in implementing the cell. It is found that incorporating double-diffused MOS transistors into the complementary MOS random-access-memory cell will improve the switching speed and/or cell size without causing monolithic fabrication complexity, yet preserving the power saving nature which is inherent to the complementary structure.

The fabrication processes and the performance of this particular cell are the focal points of interest. Cell operation, power dissipation, switching speed, cell size and fabrication complexity are discussed in detail.

Experiments started with developing a process so that both conventional p-channel and double-diffused n-channel MOS transistors could exist on a single chip in enhancement mode. A hybrid prototype cell was then built by using the two types of MOS transistors mentioned. Finally a monolithic design was realized. Test results of both the individual devices involved and the cell are presented.

Semiconductor integrated circuit memory has just now

been widely used by industry to store data in a computing system. There are two categories of semiconductor integrated circuit memory: the bipolar memory and the MOS memory; this investigation belongs to the MOS memory category. Both p-channel type and n-channel type MOS transistors are employed in implementing the memory cell which is different from the regular configuration using only one type of MOS transistor; the word "complementary" is added to tell the difference.

Random-access-memory (RAM) is the structure where information can be written into the cell at random and stored, then read out later on. There is another name for it called "read-write memory" to distinguish it from the read-only memory (ROM) structure.

II. DOUBLE-DIFFUSED AND CONVENTIONAL MOS TRANSISTORS

The double-diffused n-type channel MOS transistor and conventional p-type channel MOS transistor are discussed alongside each other throughout this chapter. Investigation of the double-diffused MOS transistor as a discrete microwave device has been done by Hans J. Sigg et al. (19), while here the main objective is to explore this device as used together with conventional MOS transistors in monolithic complementary structure and the effect on circuit operating speed. Device characteristics related to circuit performance is emphasized in this chapter.

A. DEVICE STRUCTURE

Figure 1 shows the cross-section of a double-diffused n-type channel MOS transistor (DN-MOS). The channel of the DN-MOS transistor is laterally diffused through an oxide window opening and bounded by n^+ - p junction and p-n epitaxial junction. The p-diffusion into the n-epitaxial layer is followed by a n^+ - diffusion through the same oxide window.

The impurity profile of this structure is shown in Figure 2. The surface concentrations of n^+ and p impurity layer and their distribution shape are controlled by the fabrication process, which will be discussed later. Once X_{n+} and X_p are known then X_{n+s} and X_{ps} in Figure 1 can be

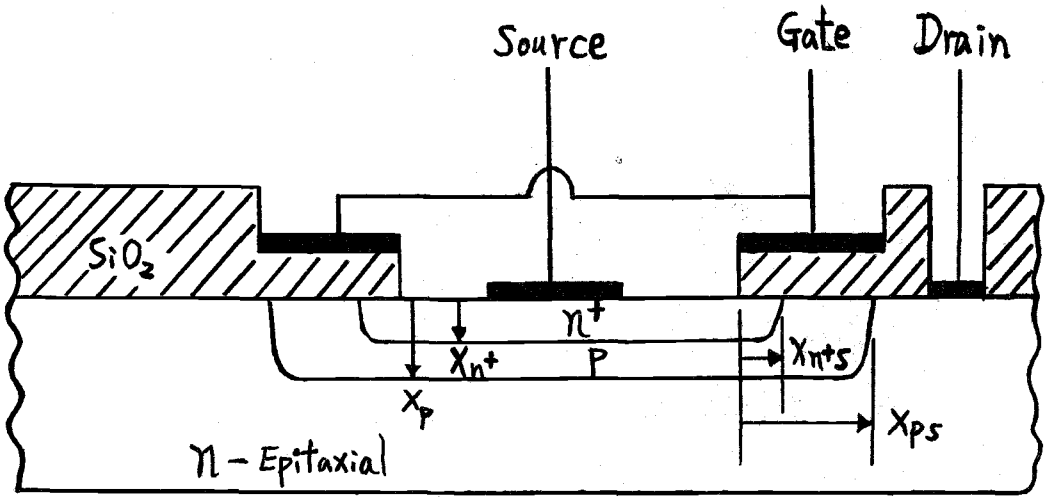


Figure 1. Cross-section of DN-MOS.

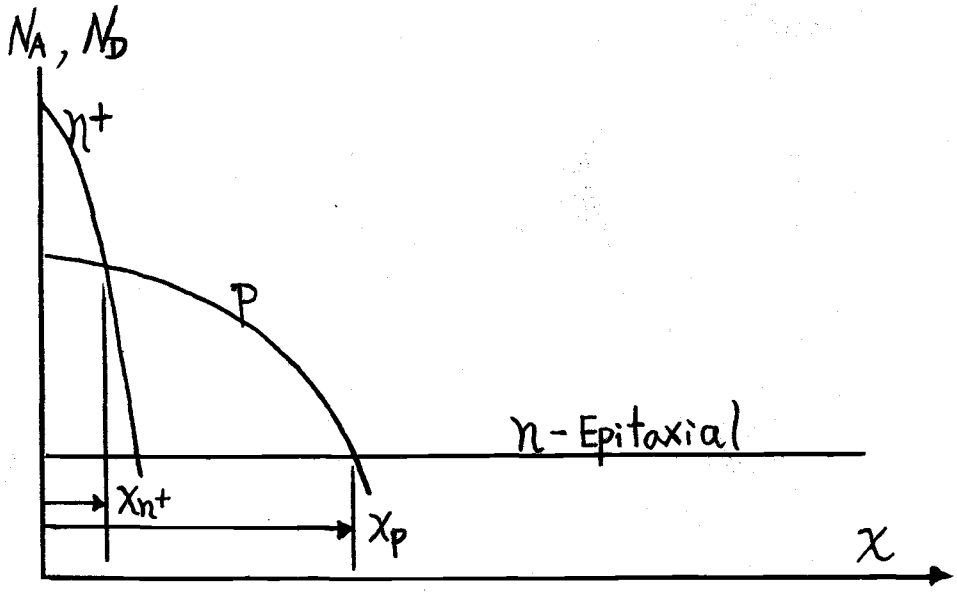


Figure 2. Impurity profile of DN-MOS structure.

found through the calculated results by D. P. Kennedy and R. R. O'Brien (13); the metallurgical channel length is $X_{ps} - X_{n+s}$. This is shown in a later section, II C 4, in detail. By using the double diffused technique channel length close to 1 μM (10^{-6} meter) can be realized while the conventional photolithography technique can only yield channel length of 5 μM at best.

Contacts to n^+ layer and n-epitaxial layer are provided to be source and drain contact respectively as in Figure 1. Gate thin oxides are grown after gate opening. Aluminum metal gate is evaporated on the thin oxide as shown in Figure 1.

Figure 3 gives the cross-section of a conventional p-type channel MOS transistor (p- MOS). The p-type source

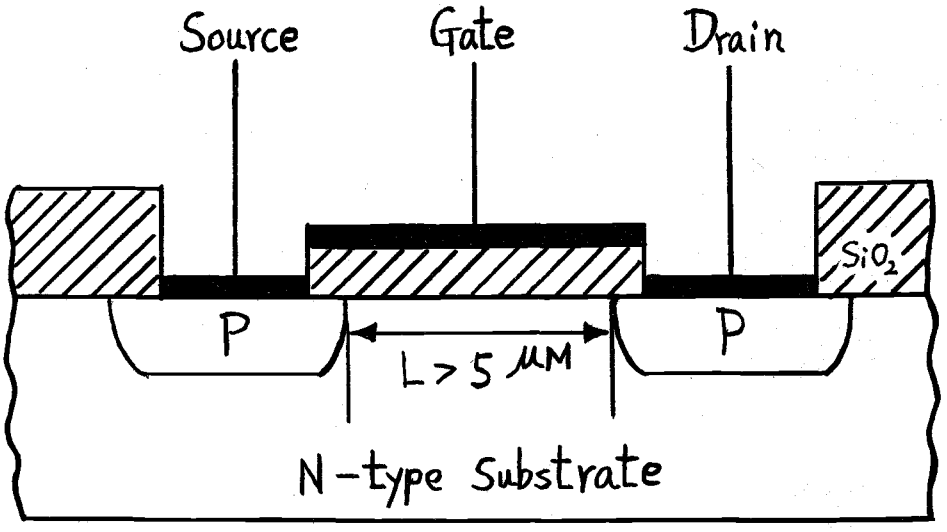


Figure 3. Cross-section of conventional p-MOS.

and drain region are diffused into uniformly doped n-type substrate, separated by at least 5 μM (5×10^{-6} meter) due to regular photolithography limit. Aluminum metal contacts are provided to source, drain and gate.

B. DEVICE OPERATION

The double-diffused MOS transistor differs from the conventional one in the channel doping; the former has a gradually decreasing channel doping level from source to drain as can be seen in Figures 1 and 2, while the latter has uniformly doped channel from source to drain as shown in Figure 3. The effect of this non-uniform channel doping on device threshold voltage has been studied by Chou (2), and it was found that the heavy impurity concentration near the source end of the channel determines the device threshold voltage. We assume the first order theory of conventional MOS equations (5) still holds for double-diffused MOS transistors (19); namely they are:

$$I_{ds} = \frac{\mu \epsilon_{ox} W}{2 t_{ox} L} \left[2(V_{gs} - V_{th}) V_{ds} - V_{ds}^2 \right] \qquad (1)$$

for $V_{ds} < V_{gs} - V_{th}$ (non-saturation region)

$$I_{ds} = \frac{\mu \epsilon_{ox} W}{2 t_{ox} L} (V_{gs} - V_{th})^2 \qquad (2)$$

for $V_{ds} \geq V_{gs} - V_{th}$ (saturation region)

$$I_{ds} = 0 \qquad (3)$$

for $V_{gs} < V_{th}$ (off region)

where:

I_{ds} = current from drain to source

μ = effective electron mobility for n-type MOS transistor; effective hole mobility for p-type MOS transistor

ϵ_{ox} = permitivity of silicon dioxide 1/3 pf/cm

t_{ox} = gate silicon dioxide thickness

W = MOS transistor channel width

L = MOS transistor channel length

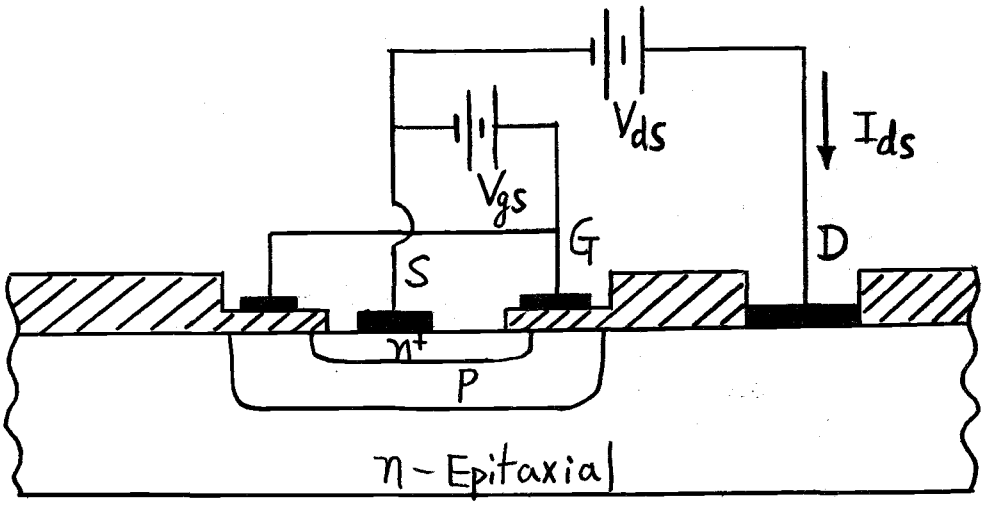
V_{gs} = voltage drop from gate to source

V_{th} = MOS transistor threshold voltage

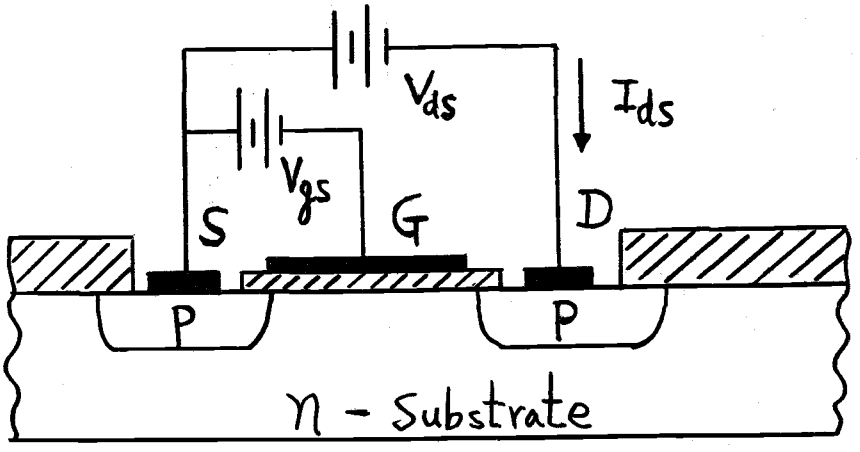
V_{ds} = voltage drop from drain to source

The above equations and definitions are for n-type channel MOS transistors. For p-type channel MOS transistor in enhancement mode quantities I_{ds}, V_{gs}, V_{th} and V_{ds} will all be negative, if their corresponding absolute values are used, equations (1), (2) and (3) still hold. Figure 4 gives a clear picture of the above statements, where two enhancement MOS transistors are used as examples.

Enhancement mode MOS transistor means it takes a positive V_{gs} to have I_{ds} flow for n-type channel MOS transistor, while it takes a negative V_{gs} to have I_{ds} flow for p-type channel MOS transistor. Some n-type channel MOS transistors will have I_{ds} flow even at zero V_{gs} or



(a) Enhancement DN-MOS transistor, I_{ds}, V_{gs}, V_{th}, and V_{ds} all positive.



(b) Enhancement conventional p-type channel MOS transistor, I_{ds}, V_{gs}, V_{th}, and V_{ds} all negative.

Figure 4. Enhancement MOS transistors in operation.

negative V_{gs}; they belong to depletion mode. All these will be obvious when threshold of device is discussed.

The threshold voltage (V_{th}) of the MOS transistor has been given by the following expression (12):

$$V_{th} = 2\phi_F + \phi_{ms} - \frac{qN_{ss}}{C_o} + \frac{Q_B}{C_o} \quad (4)$$

$$= V_{tn} \quad \text{(for n-type channel)}$$

$$V_{th} = 2\phi_F + \phi_{ms} - \frac{qN_{ss}}{C_o} - \frac{Q_B}{C_o} \quad (5)$$

$$= V_{tp} \quad \text{(for p-type channel)}$$

where:

ϕ_F = Fermi level associated with a given channel doping concentration

ϕ_{ms} = work function difference between gate metal aluminum and silicon bulk

q = electronic charge 1.6×10^{-19} coulomb

N_{ss} = surface states per cm^2

C_o = gate oxide capacitance per unit area

$= \epsilon_{ox}/t_{ox}$

Q_B = charge density per unit area within the surface depletion region in equilibrium

$$= \sqrt{2\epsilon_{si} q \left|N_A - N_D\right| \left|2\phi_F\right|}$$

ϵ_{si} = permitivity of silicon, 1.0632 pf/cm

$\left|N_A - N_D\right|$ = net channel impurity level at heavily doped source end for DN-MOS just N_D for conventional p-MOS

The value of ϕ_F is given as (16):

$$\phi_F = \frac{KT}{q} \ln \frac{|N_A - N_D|}{n_i} \qquad (6)$$

where:

K = Boltzmann's constant, 8.62×10^{-5} ev/Kelvin degree

T = temperature in Kelvin degree

n_i = intrinsic silicon carrier concentration, 1.6×10^{10} per cm^3 at room temperature

The numerical values of ϕ_{ms} given by Deal et al. (7) range from -0.4 volt to -0.1 volt for n-type silicon, and -0.8 volt to -1.1 volt for p-type silicon. The exact value of ϕ_{ms} depends on $|N_A - N_D|$ which runs from 10^{14} to 10^{18} per cm^3.

Assuming ϕ_{ms} = 0 for n-type silicon, -0.7 volt for p-type silicon, and t_{ox} = 1000 Å = 10^{-7} meter, equations (4) and (5) are plotted by Lin et al. (15) which is transferred as Figure 5 for reference.

The threshold V_{th} of an n-type channel MOS transistor can be either positive or negative depending on both $|N_A - N_D|$ and N_{ss} as shown in Figure 5. If V_{th} is positive for an n-type channel MOS transistor it is enhancement mode; it takes a more positive V_{gs} to have I_{ds} flow (see equations 1, 2, and 3). If V_{th} is negative for an n-type channel MOS transistor, it is depletion mode; a positive V_{gs} or less negative V_{gs} or zero V_{gs} will give I_{ds} flow

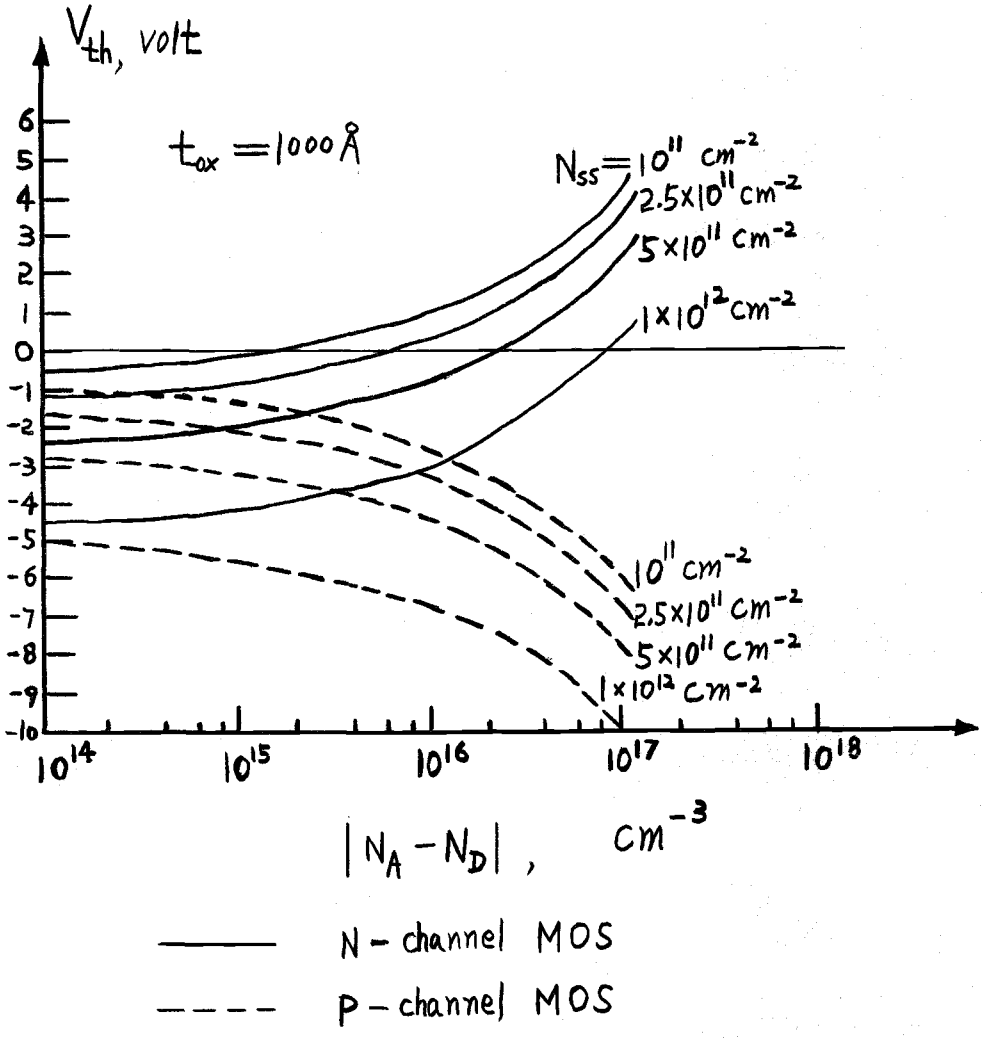


Figure 5. V_{th} of both p-channel and N-channel MOS transistors versus $|N_A - N_D|$ and N_{ss}. (Ref. 15)

(see equations 1, 2, and 3). The statement made earlier on pages 7 and 9 is justified.

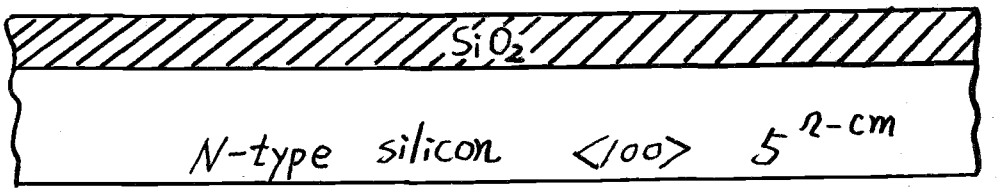
The threshold voltage of most practical p-type channel MOS transistors is negative as can be seen in Figure 5; a negative threshold voltage for the p-type channel MOS transistor means enhancement mode. The reason for this is that the only positive term in equation (5), $2\phi_F$, is usually small as compared to the sum of the other three negative terms.

C. PROCESSES AND CONSIDERATIONS

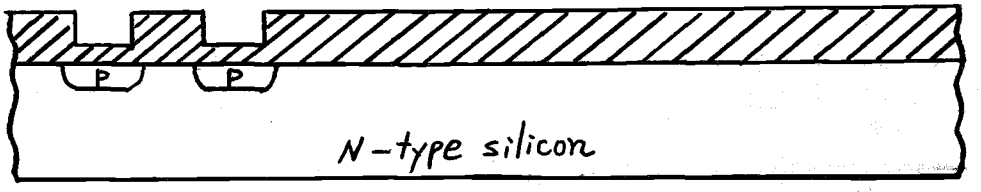
The objective of this section is devoted to the fabrication considerations of both double-diffused n-type channel MOS transistor (DN-MOS) and conventional p-type channel MOS transistor (p-MOS) monolithically on a chip. The complementary memory cell configuration requires both DN-MOS and p-MOS operating in enhancement mode. A lower threshold voltage (V_{th}) and shorter channel length will result in faster switching of the complementary memory cell. The circuit performance requirements will be clear when cell operation is discussed later; however, the requirements must be considered in process planning and fulfilled during processing.

1. Process Steps

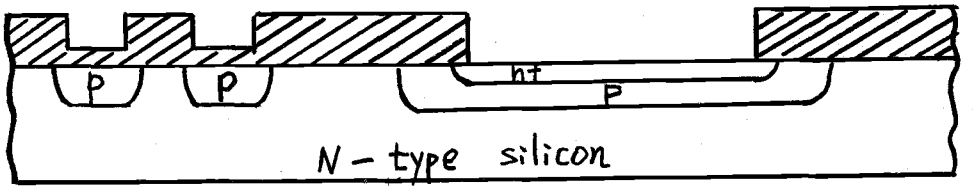
The fabrication sequence of the monolithic chip containing both DN-MOS and p-MOS is illustrated in Figure 6 below. Each step in Figure 6 is accompanied by the description in Appendix I.



a. Field oxidation.

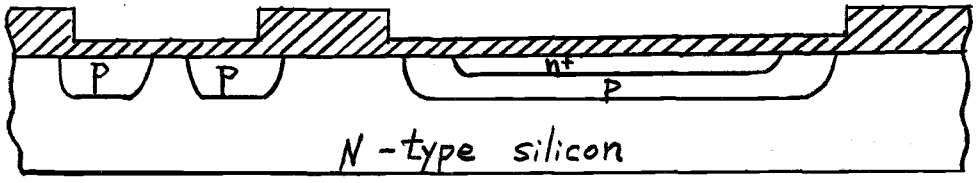


b. p-MOS source and drain diffusion.

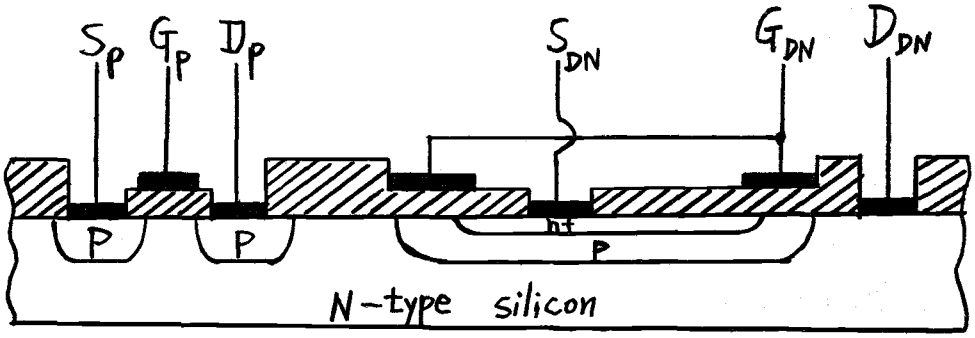


c. DN-MOS opening, p-diffusion followed by n^+-diffusion.

Figure 6a. Major steps of device fabrication process.



d. Gate opening, followed by gate oxidation.



e. Open Contacts to source and drain, metalization and excess aluminum removal.

Figure 6b. Major steps of device fabrication process.

2. Experimental Results

A composite layout was designed, as shown in Figure 7. It consists of a conventional p-MOS transistor, a DN-MOS transistor, two diffused resistors, and a testing MOS capacitor. The purpose of this part is to ensure that the process sequence stated in previous sections will yield good p-MOS and DN-MOS transistors.

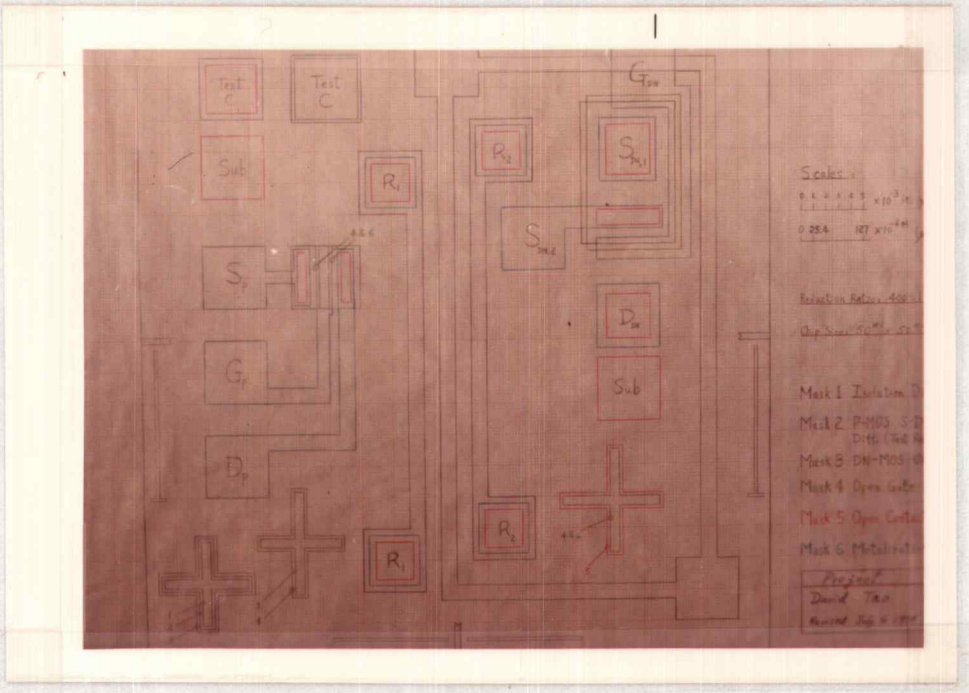


Figure 7. Device composite layout.

A set of masks are generated. Samples are processed as stated in section II C 1. Typical finished sample under the microscope looks like that shown in Figure 8.

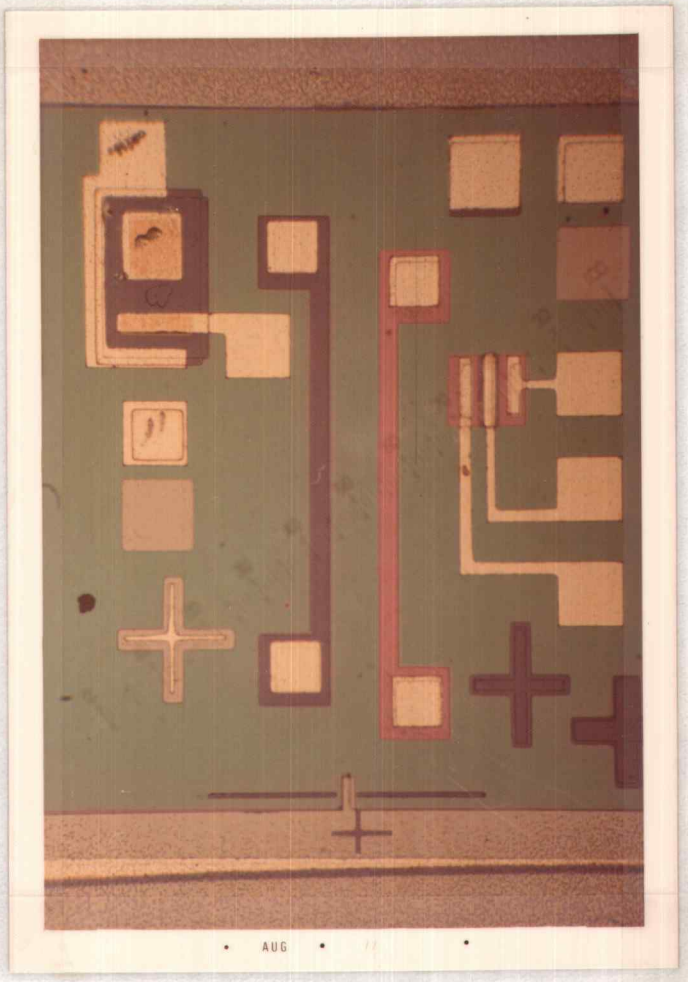
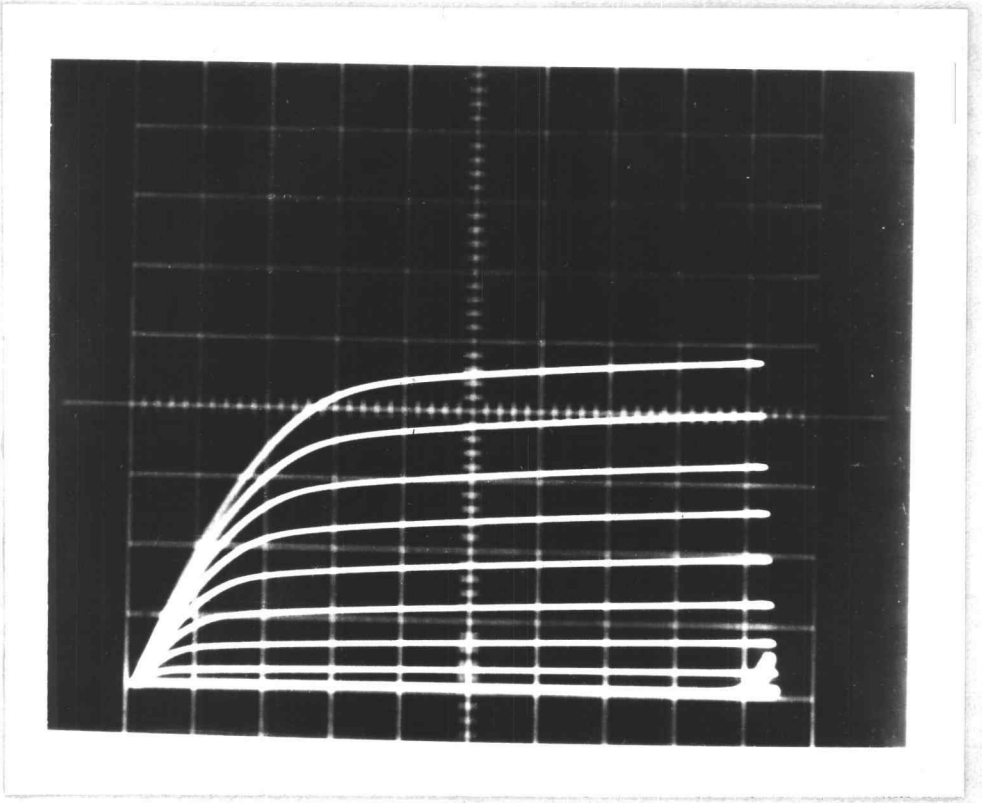


Figure 8. Photomicrograph of devices.

An individual chip is cut from the wafer, using a diamond scriber. A TO-5 header is used to encase the chip. Bonding is done by a Sonoweld ultrasonic bonder with 1×10^{-3} inch diameter aluminum wire.

Both p-MOS transistor and DN-MOS transistor characteristics are tested through the transistor curve tracer. Typical results are shown in Figures 9 and 10. Two pictures are shown for each type of transistor; gate bias

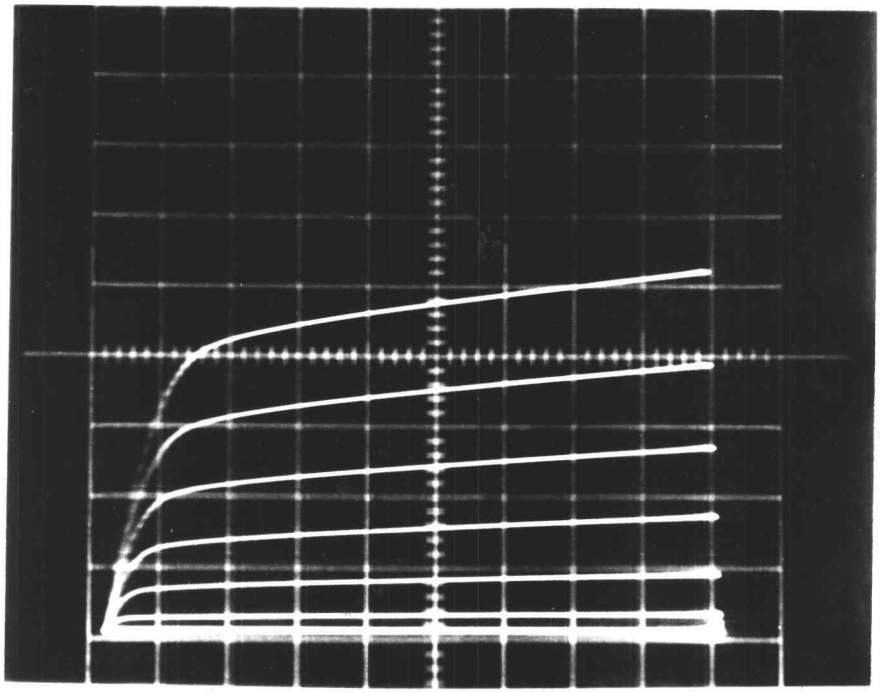


Vertical - I_{ds} : 2 milli-ampere per major division

Horizontal - V_{ds} : 5 volts per major division

Gate bias : 2 volts per step. Top curve V_{gs} = - 20 volts

Figure 9a. P-MOS transistor characteristics.

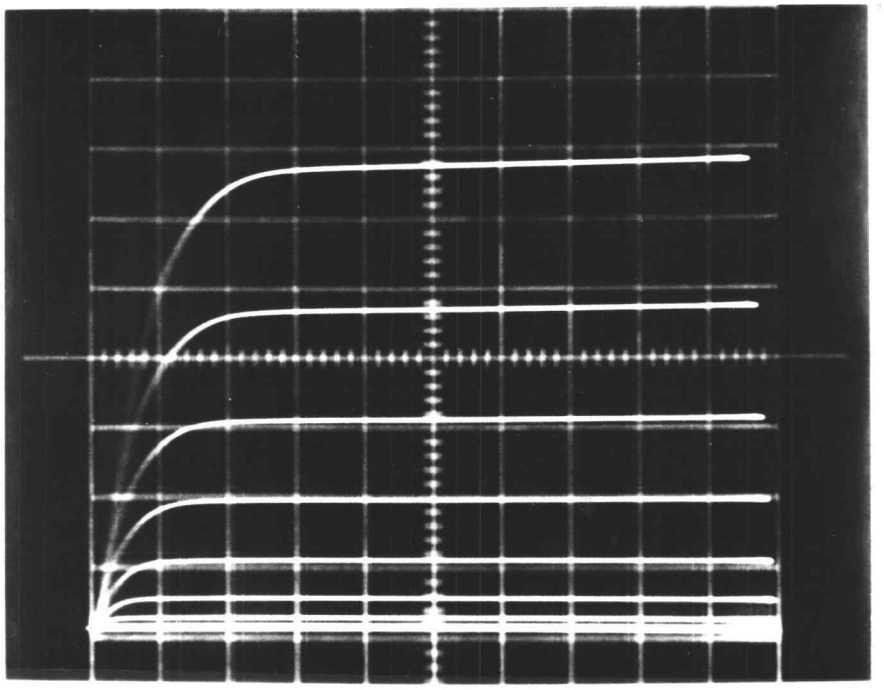


Vertical - I_{ds} : 500 micro-ampere per major division

Horizontal - V_{ds} : 5 volts per major division

Gate bias : 1 volt per step. Top curve V_{gs} = - 10 volts

Figure 9b. P-MOS transistor characteristics.

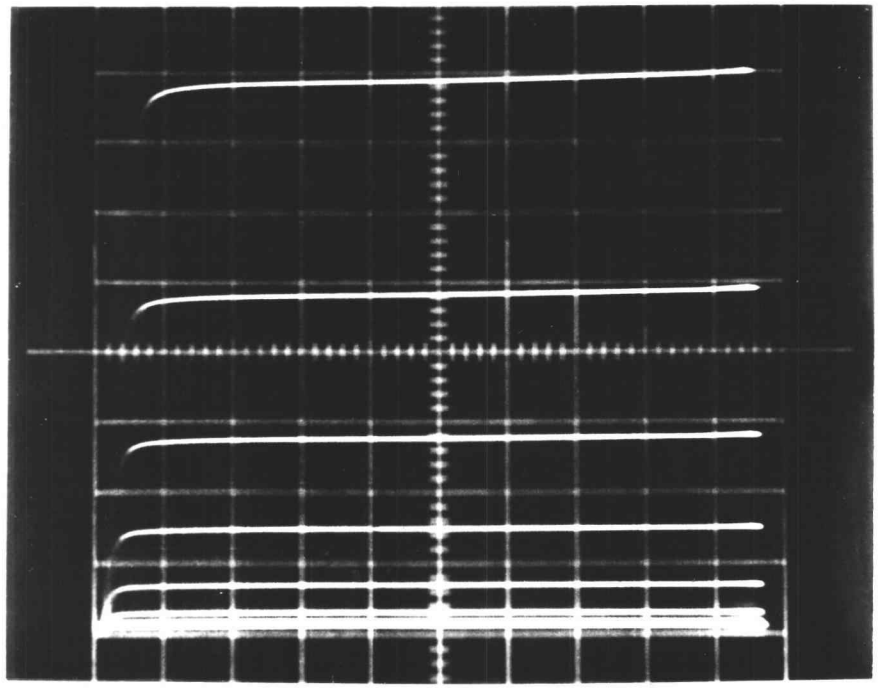


Vertical I_{ds} : 2 milli-ampere per major division

Horizontal V_{ds} : 5 volts per major division

Gate bias : 2 volts per step. Top curve V_{gs} = 20 volts

Figure 10a. DN-MOS transistor characteristics.



Vertical I_{ds} : 100 micro-ampere per major division

Horizontal V_{ds} : 5 volts per major division

Gate bias : 1 volt per step. Top curve V_{gs} = 10 volts

Figure 10b. DN-MOS transistor characteristics.

as high as 20 volts is shown. Drain voltage can reach beyond 40 volts without causing the junction between drain region and channel bulk to breakdown. Twenty volts is the highest voltage in the complementary memory cell configuration to be realized later by employing both p-MOS transistor and DN-MOS transistor, so it is safe as far as the voltage rating is concerned. However, if the n^+ -region of the DN-MOS transistor were used as drain and n-type silicon as source (refer to Figure 4a), then the breakdown voltage of the n^+-p junction would limit the voltage within the cell to less than ten volts, which may not be practical if the device threshold voltage is close to four volts because the biasing condition will cause very slow switching of the cell.

From equation (2) it has

$$\sqrt{I_{ds}} = \sqrt{\frac{\mu \epsilon_{ox} W}{2 t_{ox} L}} \quad (V_{gs} - V_{th}) \qquad (7)$$

$$\text{for } V_{ds} \geqslant V_{gs} - V_{th} \text{ (saturation region)}$$

In the linear plot of $\sqrt{I_{ds}}$ versus V_{gs}, V_{th} can be found from the point when $\sqrt{I_{ds}}$ is zero.

I_{ds} and corresponding V_{gs} values read from Figures 9 and 10 are tabulated in Tables A and B in Appendix II. $\sqrt{I_{ds}}$ is also listed for convenience. I_{ds} values are read when $-V_{ds}$ = 25 volts for the p-MOS transistor and V_{ds} = 25 volts for the DN-MOS transistor; it is the point when the transistors are in saturation.

Linear plots of $\sqrt{I_{ds}}$ versus V_{gs} for both p-MOS and DN-MOS transistors are shown in Figures 11 and 12. The extrapolated curve shows threshold voltage for p-MOS transistor V_{tp} = - 4 volts and threshold voltage for DN-MOS transistor V_{tn} = 4 volts.

3. Threshold Voltage Consideration

The threshold voltage of both p-MOS transistor and DN-MOS transistor is an important factor affecting the performance of the complementary MOS memory cell. Reference to equations (4), (5), and (6) shows that net channel doping $|N_A - N_D|$ determines values of ϕ_F, ϕ_{ms}, and Q_B. Its effect on V_{th} is observed first, and the calculation and control of $|N_A - N_D|$ is then discussed.

a. Net Channel Doping $N_A - N_D$ Effect on Threshold Voltage

For the p-MOS transistor, the channel is uniformly doped. $|N_A - N_D|$ is simply the doping level of the starting n-type silicon, 5 ohm-cm resistivity corresponding to a doping level N_D of 10^{15} per cm^3; hence the threshold voltage of p-MOS transistor V_{tp} can be calculated using equations (5) and (6) to be - 3.2 volts, assuming a surface state of 5×10^{11} per cm^2 and gate oxide thickness of 1200 Å (1.2×10^{-7} meter). An error of 0.8 volt is found between the calculated value and the experimental

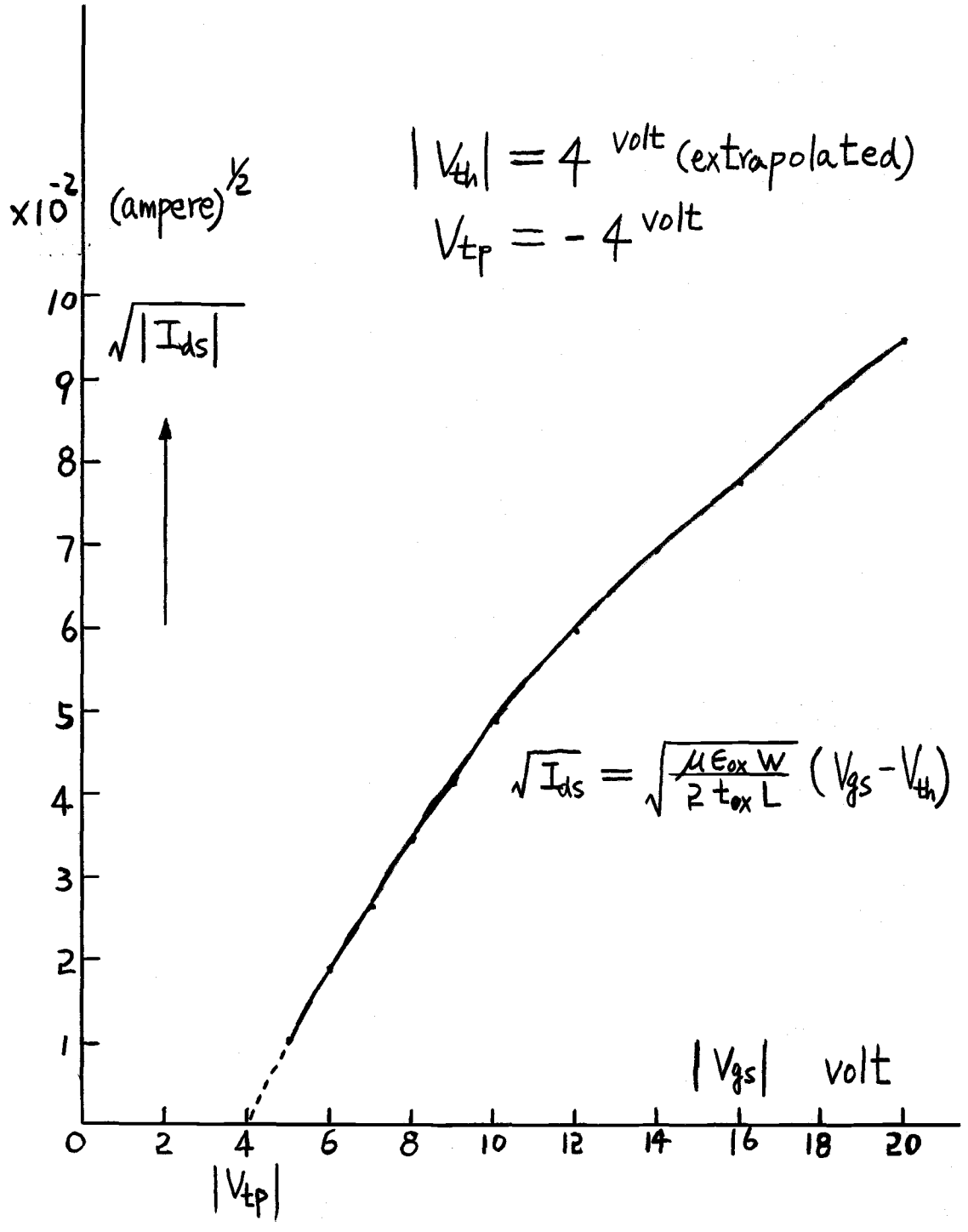


Figure 11. P-MOS transistor $\sqrt{|I_{ds}|}$ versus $|-V_{gs}|$.

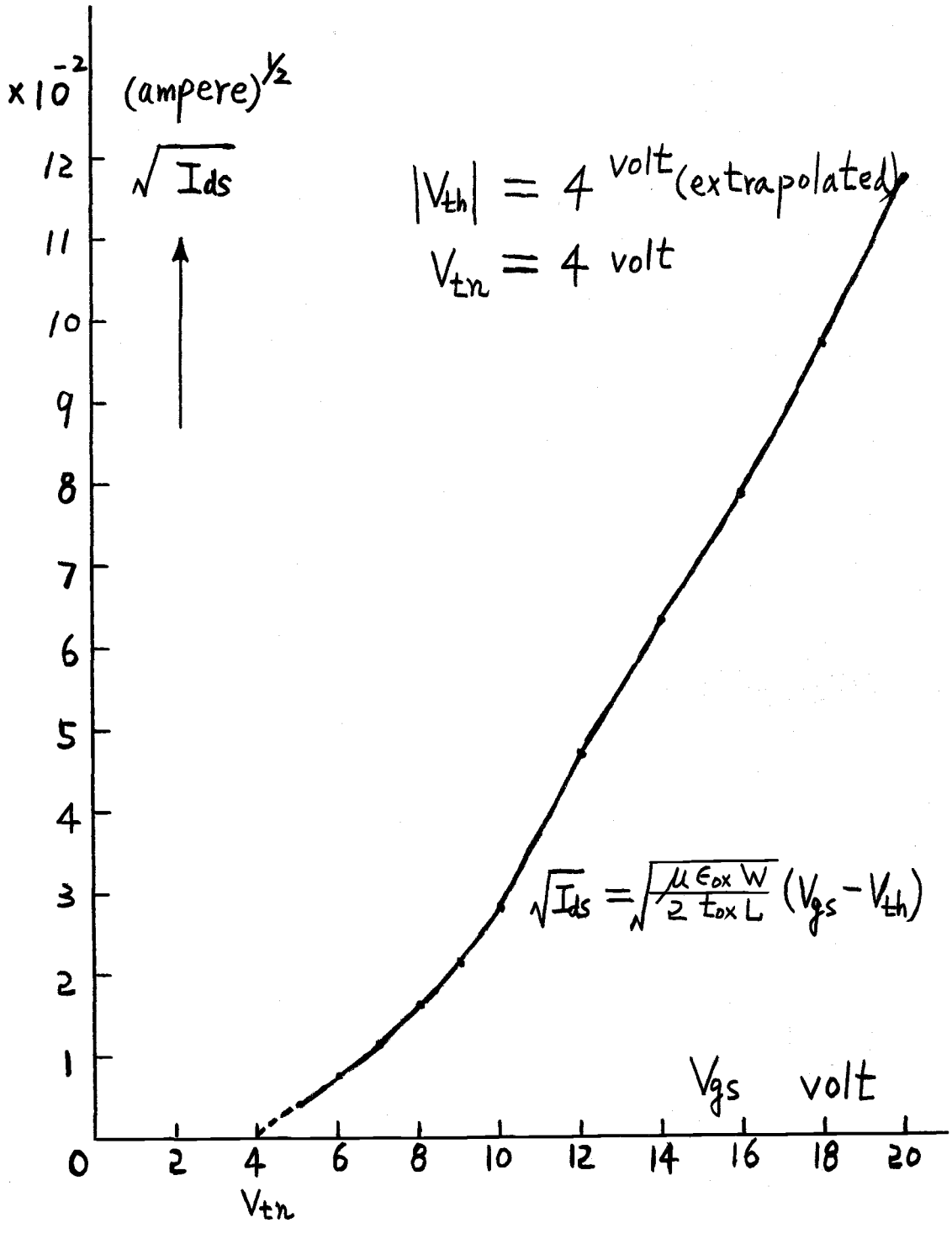


Figure 12. DN-MOS transistor $\sqrt{I_{ds}}$ versus V_{gs}.

extrapolated value in Figure 11. Error source could very well come from the estimation of $|N_A - N_D|$. The lot of silicon wafers used has a specified resistivity of 1 to 5 ohm-cm, which means $|N_A - N_D|$ has a range between 5×10^{15} to 10^{15} per cm^3 respectively. If $|N_A - N_D|$ is 5×10^{15} per cm^3, then the calculated V_{tp} would be - 4 volts.

For the DN-MOS transistor, the heavy net channel doping near the source end determines threshold voltage V_{tn} (2). Figure 5 shows $|N_A - N_D|$ up to 10^{17} per cm^3 only, while the range of interest to the particular process sequence stated in Appendix I is between 10^{17} and 10^{18} per cm^3. V_{tn} for various $|N_A - N_D|$ between 10^{17} and 10^{18} per cm^3 is calculated using equations (4) and (6), assuming 5×10^{11} per cm^2 surface states and 1200 Å (1.2×10^{-7} meter) gate oxide thickness. Φ_{ms} used is -1.1 volt from reference (7). The results are listed in Table 1 below and plotted in Figure 13.

Table 1. V_{tn} for Various $|N_A - N_D|$

$\lvert N_A - N_D \rvert$ (cm^{-3})	10^{17}	2×10^{17}	4×10^{17}	6×10^{17}	8×10^{17}	10^{18}
V_{tn} (volt)	2.88	5.50	9.30	12.32	14.94	17.15

It can be seen in Figure 13 that V_{tn} goes from 2.88 volts to 17.15 volts as net channel doping varies. The control

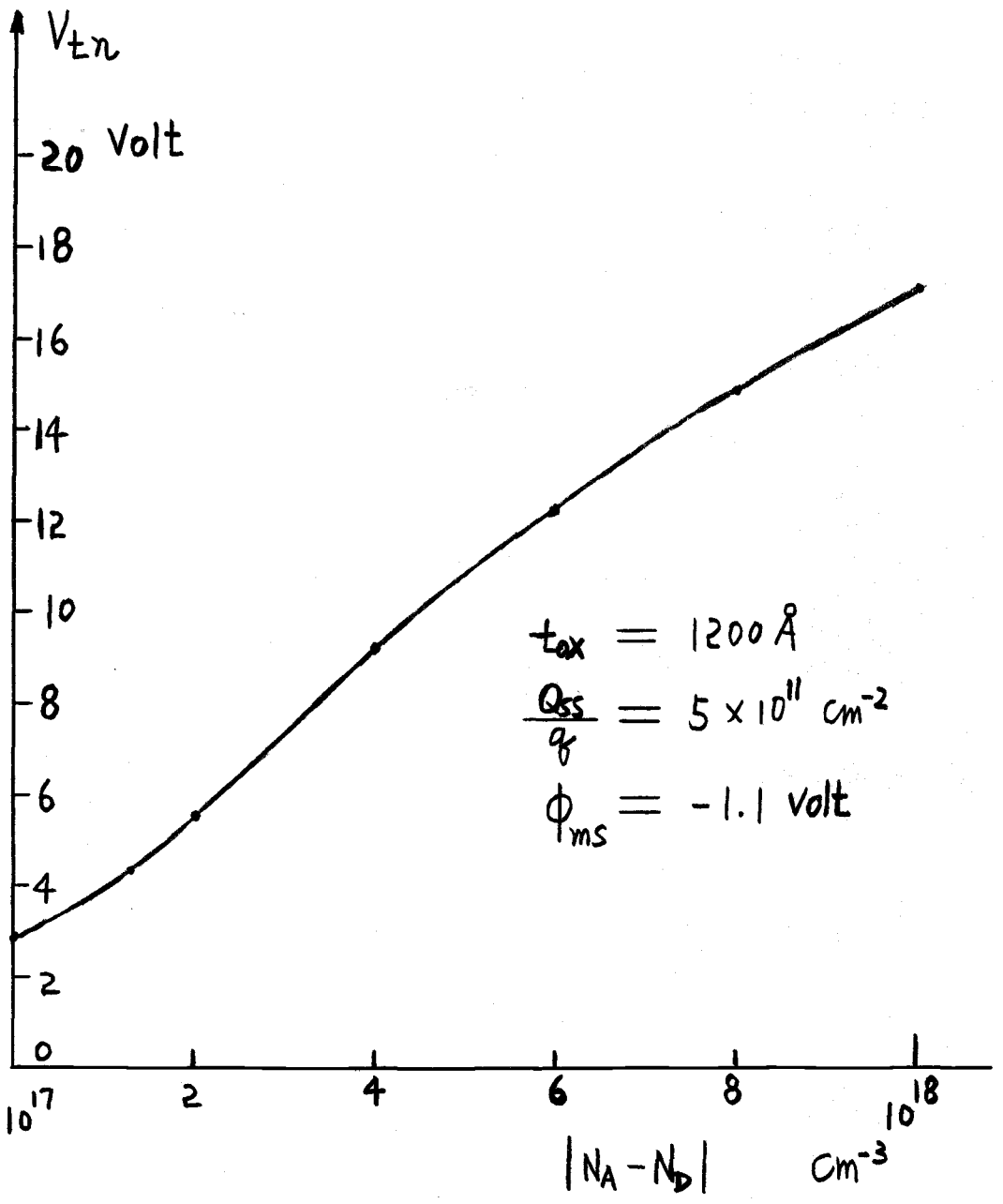


Figure 13. DN-MOS threshold voltage (V_{tn}) versus net channel doping $|N_A - N_D|$ near source end.

of V_{tn} is therefore dependent on the impurity profile of the DN-MOS transistor.

b. Impurity Profile Calculations

There are two kinds of impurity profile distribution involved in the processing sequence stated in section II C 1. A predeposition giving a sheet resistance followed by a long drive-in period results in Gaussian distribution, which has a limited-source boundary condition. The p-MOS transistor source and drain region and the DN-MOS transistor p-region belong to the Gaussian distribution. In a later section when the memory cell is realized, the isolation diffusion belongs to this kind also. The DN-MOS transistor n^+-region is fabricated by one step diffusion as stated in Appendix I C, which has a constant-source boundary condition. The impurity distribution profile is a complementary error function (erfc); this is the second kind.

For p-type Gaussian distribution, it has (22):

$$N_A(X) = \frac{Q}{\sqrt{\pi D_p t}} e^{-x^2/4D_p t} \tag{8}$$

$$Q = \frac{1}{R_s q \mu} \tag{9}$$

where:

N_A = acceptor impurity atoms per unit volume, cm^{-3}.

Q = impurity concentration in atoms per unit area, cm^{-2}.

D_p = diffusion coefficient, or diffusivity, for boron in silicon at 1100°C. D_p is 0.09 $(\mu M)^2$ per hour or 0.09×10^{-8} cm^2 per hour (12).

t = drive-in diffusion time in hour.

X = depth into the surface of silicon in micrometer μM.

R_s = sheet resistance of silicon surface after predeposition, in ohms per square.

q = electronic charge, 1.6×10^{-19} coulomb.

μ = mobility of holes, 480 cm^2 per volt-sec.

For n-type complementary error function distribution (erfc) it has (22):

$$N_D(X) = N_o \operatorname{erfc} \frac{X}{2\sqrt{D_n t}} \qquad (10)$$

where:

N_D = donor impurity atoms per unit volume, cm^{-3}.

N_o = surface concentration in atoms per cm^3, independent of diffusion time. 10^{21} per cm^3 for phosphorus in silicon at 1000°C, determined by solid solubility (22).

D_n = diffusion coefficient of phosphorus into silicon, 0.09 $(\mu M)^2$ per hour at 1000°C (12).

Consider the DN-MOS transistor p-region first. It has a sheet resistance of 150 ohms per square, and an

over-all effective drive-in time of six hours (refer to Appendix I c). Equation (9) gives:

$$Q = \frac{1}{(150)(1.6\times10^{-19})(480)}$$

and

$$Q = 8.7 \times 10^{13}\ cm^{-2} \tag{11}$$

Substituting equation (11) into equation (8) yields:

$$N_A(X) = \frac{8.7 \times 10^{13}}{\sqrt{(3.14)(0.09\times10^{-8})\ 6}}\ e^{\frac{-X^2}{4(0.09)\ 6}}$$

and

$$N_A(X) = 6.7 \times 10^{17}\ e^{\frac{-X^2}{2.16}} \tag{12}$$

The plot of equation (12) is the dotted line in Figure 14, based on numerical values listed in Table C, Appendix III.

The value of X_p can be calculated by solving $N_A(X_p) = 10^{15}$ from equation (12) as follows:

$$10^{15} = 6.7 \times 10^{17}\ e^{\frac{-X_p^2}{2.16}} \tag{13}$$

then

$$X_p^2 = 2.16 \ln 670 = 2.16 \left[\ln 6.7 + 2 \ln 10\right]$$

and

$$X_p = 3.74\ \mu M = 3.74 \times 10^{-6}\ \text{meter}$$

For the n^+-region of DN-MOS transistor, it has a complementary error function distribution, the phosphorus diffusion is done at 1000°C for one hour. Thus

equation (10) gives:

$$N_D(X) = 10^{21} \text{ erfc } \frac{X}{2\sqrt{(0.09)(1)}}$$

$$= 10^{21} \text{ erfc } \frac{X}{0.6} \qquad (14)$$

The plot of equation (14), also shown in Figure 14, is based on the numerical values listed in Table D, Appendix III. Error function values are obtained from a reference table (18). The value of X_{n+} is read from Figure 14 to be 1.55×10^{-6} meter. This graphical solution is much easier than setting equation (12) equal to equation (14) and trying to solve for X_{n+} analytically.

Net impurity doping profile for the DN-MOS transistor is plotted in Figure 15 based on numerical values listed in Table E, Appendix III. The channel has the heaviest doping at point A in Figure 15 near the source end. The value of this high doping level is found to be 1.3×10^{17} atoms per cm^3 from Table E, Appendix III.

Based on the value of 1.3×10^{17} cm^{-3} as $|N_A - N_D|$ and from Figure 13, the value of V_{tn} should be 4.4 volts, which is very close to the experimental extrapolated result of four volts in Figure 12.

c. Impurity Profile Control on DN-MOS Transistor Threshold Voltage

The threshold voltage of DN-MOS transistor V_{tn} is dependent on the impurity profile of the p-region near the

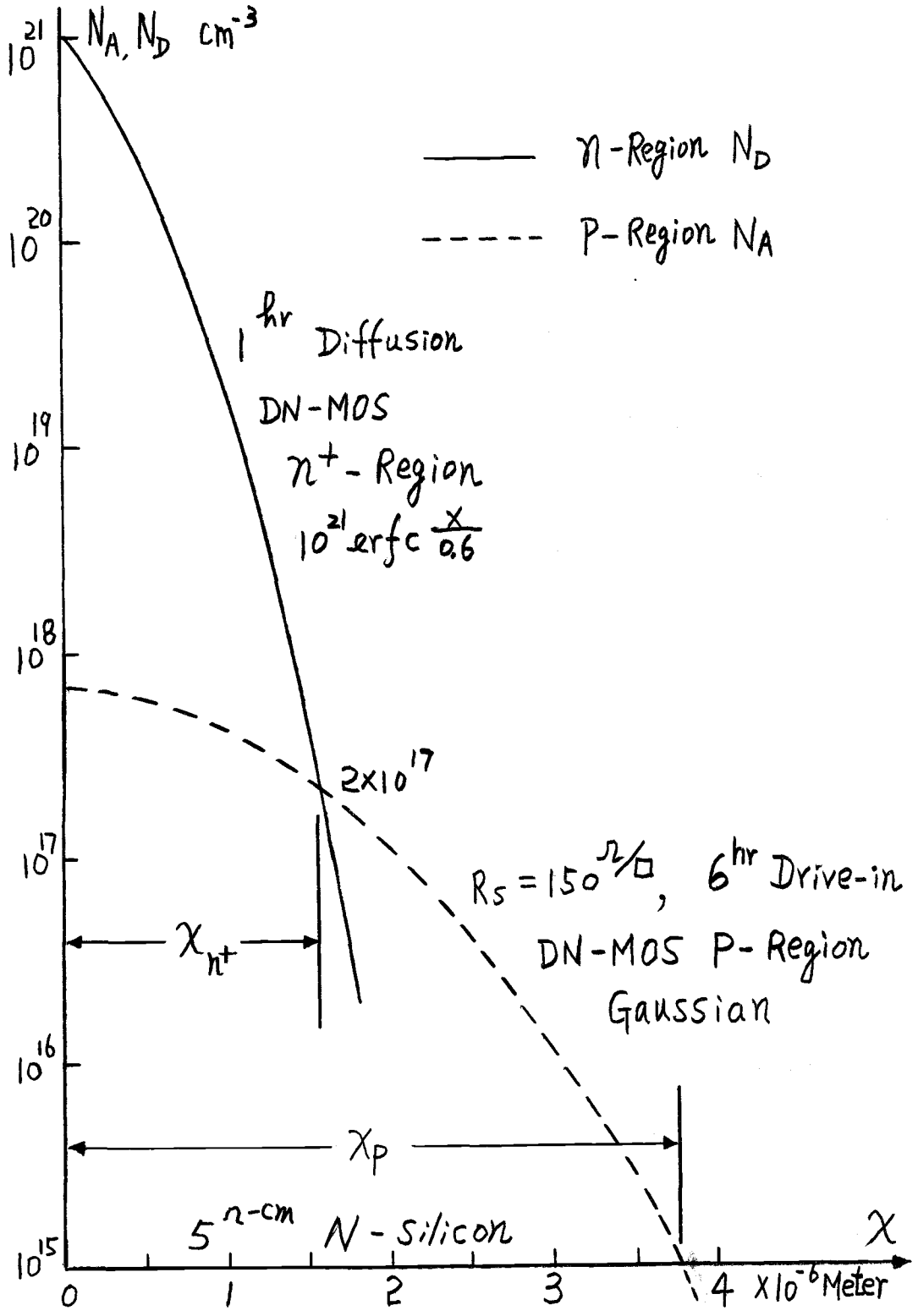


Figure 14. DN-MOS transistor impurity profiles.

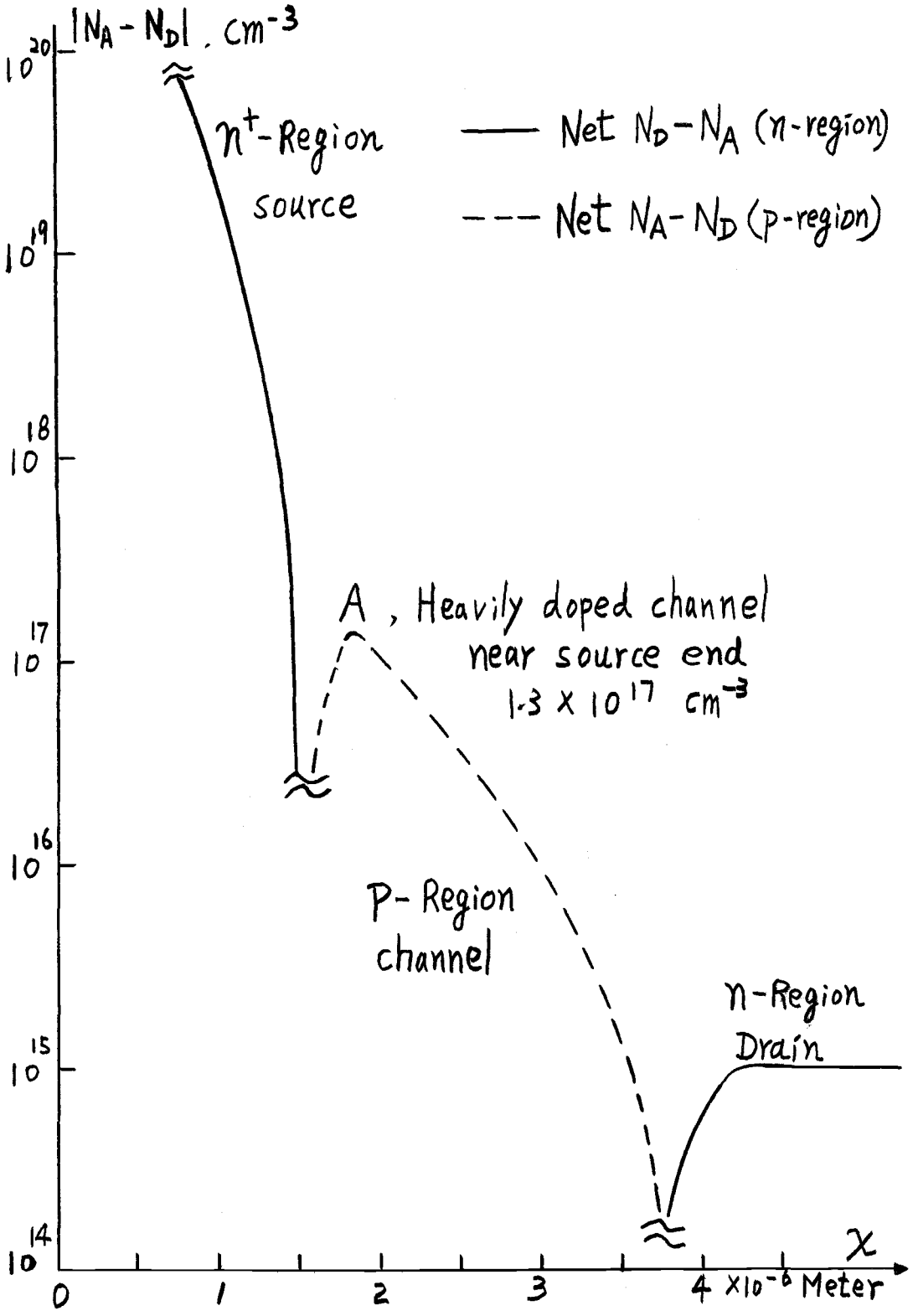


Figure 15. DN-MOS net impurity profile.

n^+-p junction as point A in Figure 15.

Referring to equations (8) and (9), both sheet resistance R_s and drive-in time t are the two parameters which can be adjusted to fit a designer's need. The boron diffusivity D_p can be varied depending on the drive-in temperature (practical value is from 1000°C to 1200°C). The effect of varying D_p on N_A (X) is the same as that of varying t on N_A (X); this can be seen from equation (8). In the diffusion practice it is desirable to plan on varying t rather than changing diffusion furnace temperature.

The effect of changing sheet resistance R_s after predeposition on p-region profile is shown in Figure 16, where three different R_s values are shown assuming same six hours drive-in at 1100°C. N_A (X) is linearly proportional to the reciprocal of R_s, as can be seen from equations (8) and (9). Figure 16 shows this effect clearly. If the same n^+-region diffusion is used, then the dotted line in Figure 16 represents the n^+-region impurity profile. The point A with the highest net channel doping would move up approximately to A' or down to A'' in Figure 16, depending on sheet resistance. A' has a $|N_A - N_D|$ larger than 10^{18} per cm^3, which gives a V_{tn} larger than 17 volts (see Figure 13). Point A'' has a $|N_A - N_D|$ close to 10^{16} per cm^3, which may make the DN-MOS operating in depletion mode with a negative V_{tn} (refer to Figure 5). As R_s goes from 15 ohms per square to 1500 ohms per square, V_{tn} goes from above 17

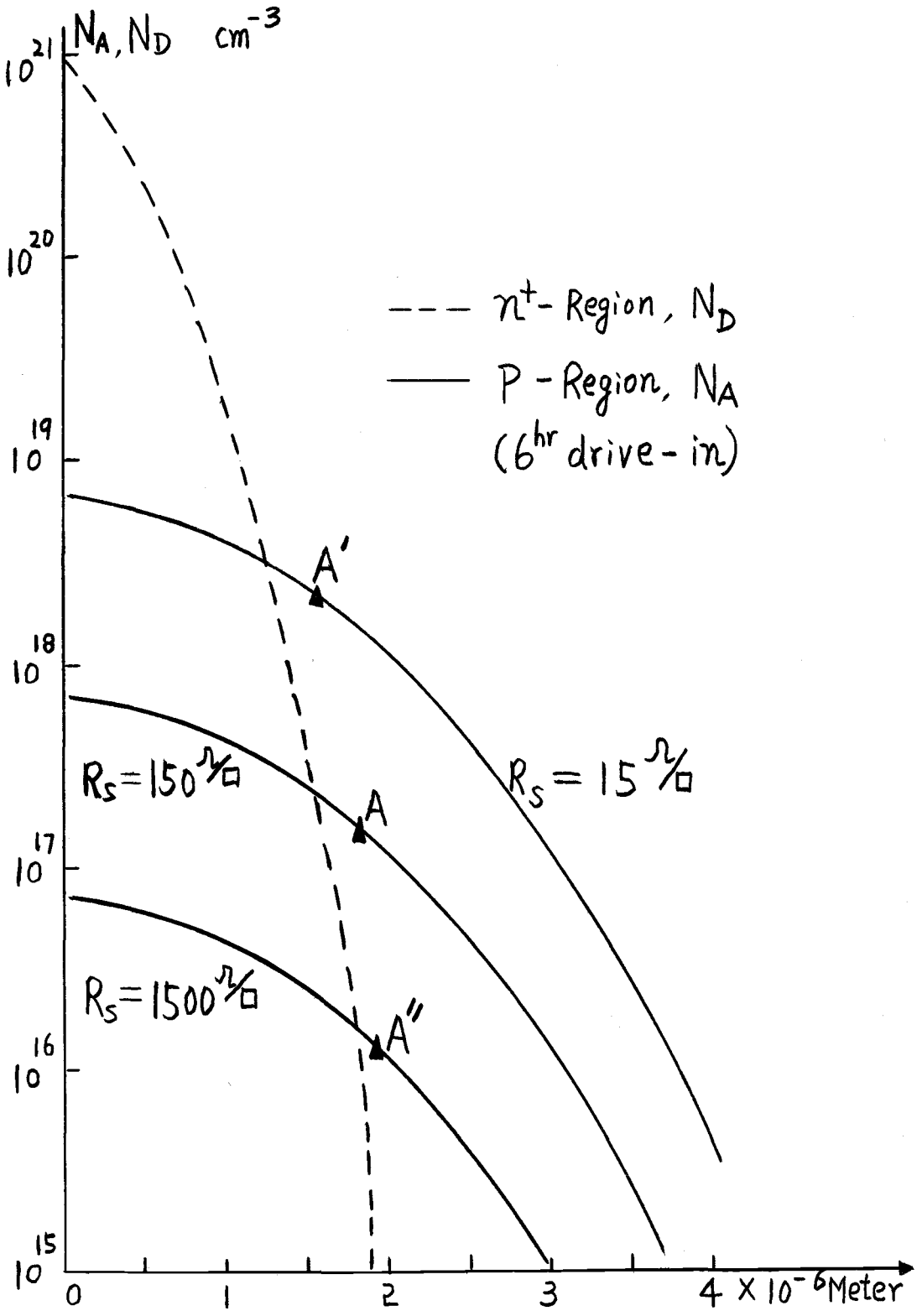


Figure 16. R_S effect on DN-MOS p-region impurity profile.

volts to a negative value. It is then concluded that R_s affects V_{tn} drastically.

The other diffusion parameter is the drive-in time t. From equations (8) and (9), assuming R_s = 150 ohms per square constant and t = six hours, ten hours or 48 hours respectively, a family of curves is plotted in Figure 17 according to the following equations*:

$$N_A(X) = 6.7 \times 10^{17} \, e^{\frac{-X^2}{2.16}} \quad \text{(for t = 6 hours)} \tag{15}$$

$$N_A(X) = 5.19 \times 10^{17} \, e^{\frac{-X^2}{3.6}} \quad \text{(for t = 10 hours)} \tag{16}$$

$$N_A(X) = 2.37 \times 10^{17} \, e^{\frac{-X^2}{17.28}} \quad \text{(for t = 48 hours)} \tag{17}$$

Also in Figure 17 the n^+-region impurity profile is shown in dotted line assuming the same n^+ diffusion as before. It can be seen that N_D of the n^+-region is a very steep function with respect to X within the vicinity where point A locates. As the drive-in time of the DN-MOS p-region changes from six hours to ten hours or to 48 hours, the point A where net channel doping $|N_A - N_D|$ is highest

*Equation (15) is the same as equation (12), copied for convenience.

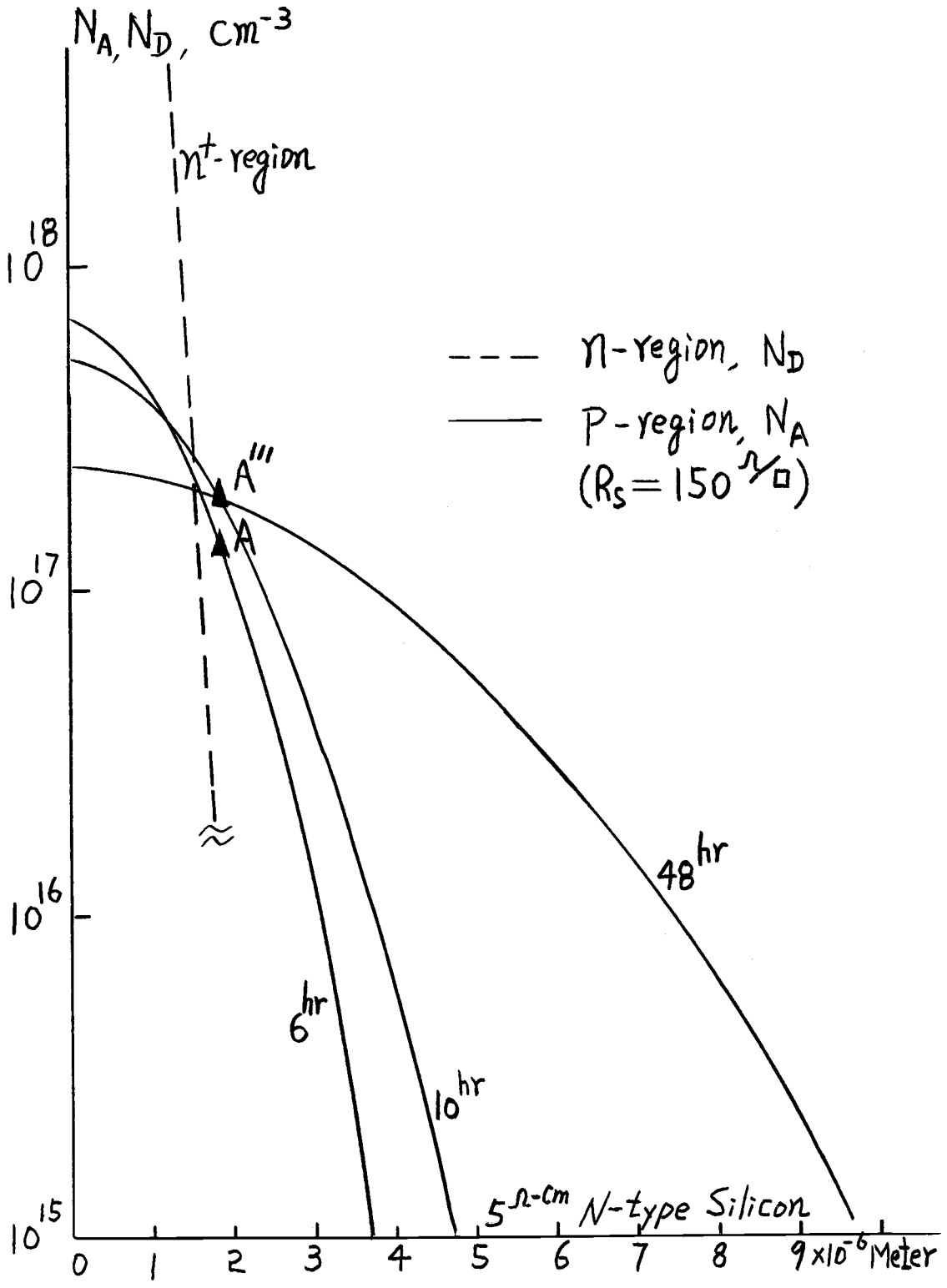


Figure 17. Drive-in time effect on DN-MOS p-region impurity profile.

moves approximately to point A''' for both ten hour and 48 hour cases, as shown in Figure 17. Corresponding new $|N_A - N_D|$ is then 2 x 10^{17} per cm^3; it does not deviate much from the 1.3 x 10^{17} per cm^3 for point A before. Of course V_{tn} still changes to 5.5 volts as listed in Table 1 as compared to 4.4 volts for point A before. A change of 1.1 volt in V_{tn} results from drastically changing p-region drive-in time from six hours to ten hours or even to 48 hours.

Because of the insensitive nature of the Gaussian (p-region) distribution with respect to t near the small X end, and the sensitive nature of the complementary error function (n^+-region) distribution with respect to X near the tailing end, it is concluded that varying t will not change $|N_A - N_D|$ and thence V_{tn}.

From Figure 17 it is seen that the junction depth of p-region varies drastically as the drive-in time increases. Ten hours drive-in gives a junction depth of 4.7 x 10^{-6} meter, which is the p-MOS transistor source and drain region. Forty-eight hours drive-in gives a junction depth of 9.7 x 10^{-6} meter, which can be used for isolation diffusion when a memory cell is realized on an 8 x 10^{-6} meter thick n-type epitaxial layer on p-type substrate. These junction depths can be calculated using equations (16) and (17) in the same manner as X_p was calculated in equation (13).

d. Silicon Orientation Effect on Threshold Voltage

The surface state charge density N_{ss} (refer to equations 4 and 5) of thermally oxidized gate silicon has a smaller value for (100)* oriented silicon than (111)* oriented silicon (6).

For a p-MOS transistor this means (100) oriented silicon will yield a smaller threshold voltage $|V_{tp}|$ than (111) oriented silicon. This is because the third term in equation (5) will have a smaller absolute value for (100) silicon and usually the third term is a major contributing term for V_{tp}. The dotted lines in Figure 5 also illustrate this phenomenon.

For a DN-MOS transistor, less N_{ss} means equation (4) will have a better chance to yield a positive V_{tn}. In other words, a DN-MOS transistor is likely to operate in enhancement mode. This is because the third term in equation (4) is a major negative term. The solid lines in Figure 5 demonstrate this fact clearly.

Since the DN-MOS transistor must operate in enhancement mode in the complementary memory cell discussed later, and it is desirable to have low absolute value V_{tp} for p-MOS transistors used in the complementary memory cell, (100) orientation is therefore much preferred over (111) orientation.

*Miller indices for crystal orientation.

4. DN-MOS Transistor Channel Length Consideration -- Punch Through

An enlarged cross-section view of the DN-MOS transistor with definition of quantities used in channel length calculation is shown in Figure 18. Figure 14 is also inserted to indicate the correlation of quantities in the impurity profile diagram. Quantities in Figure 18 have the following definitions:

X_{n+} = junction depth of the n^+-p junction between n^+-region and p-region

X_p = junction depth of the p-n junction between p-region and n-type silicon

X_{n+s} = lateral junction depth of the n^+-p junction

X_{ps} = lateral junction depth of the p-n junction

a_1 = depletion width in the p-region side of the n^+-p junction

a_2 = depletion width in the p-region side of the p-n junction

The values of X_p and X_{n+} were found in section II C 3b to be 3.74×10^{-6} meter and 1.55×10^{-6} meter respectively.

In order to find X_{ps}, the p-region impurity distribution function in equation (12) is considered again. The surface impurity level is found by setting X = 0:

$$N_A(0) = 6.7 \times 10^{17} \text{ cm}^{-3} \qquad (18)$$

The background to be diffused in has an impurity level

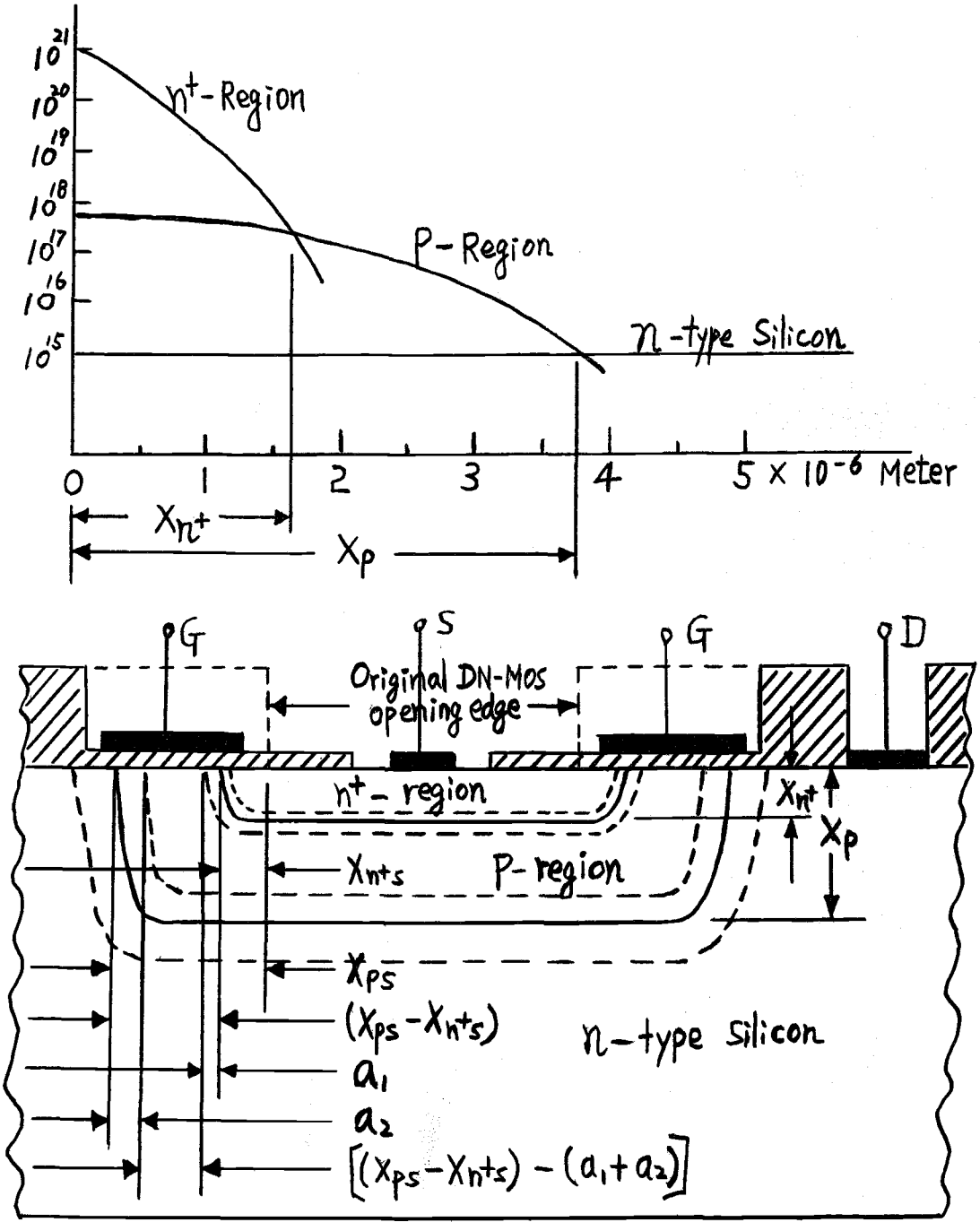


Figure 18. Enlarged cross-section of DN-MOS transistor with defining quantities.

of 10^{15} per cm^3. Considering the concentration ratio

$$\frac{10^{15}}{6.7 \times 10^{17}} = 0.00149 \tag{19}$$

as a parameter, from the Gaussian distribution contours calculated by Kennedy *et al.* (13), it is found that:

$$\frac{X_{ps}}{X_p} = \frac{2.1}{2.6} \tag{20}$$

Equation (20) gives X_{ps} as

$$X_{ps} = (3.74)\left(\frac{2.1}{2.6}\right) \times 10^{-6}$$

$$= 3.02 \times 10^{-6} \text{ meter} \tag{21}$$

For the value of X_{n+s}, the surface concentration of n^+-region is found from equation (14) by setting X = 0

$$N_D(0) = 10^{21} \text{ erfc } 0$$

$$= 10^{21} \text{ cm}^{-3} \tag{22}$$

The p-region background to be diffused in has a doping level of 2×10^{17} per cm^3 (refer to Figure 14). Considering the concentration ratio

$$\frac{2 \times 10^{17}}{10^{21}} = 0.0002 \tag{23}$$

as a parameter, from the complementary error function distribution contours calculated by Kennedy *et al.* (13), it is found that

$$\frac{X_{n+s}}{X_{n+}} = \frac{2.3}{2.7} \tag{24}$$

Equation (24) gives X_{n+s} as

$$X_{n+s} = (1.55)\left(\frac{2.3}{2.7}\right) \times 10^{-6}$$

$$= 1.32 \times 10^{-6} \text{ meter} \qquad (25)$$

The metallurgical channel length should be $(X_{ps} - X_{n+s})$, which is 1.7×10^{-6} meter.

The values of depletion widths a_1 and a_2 depend on the biasing condition of the DN-MOS transistor and the impurity distribution near n^+-p junction and p-n junction. If $(X_{ps} - X_{n+s})$ is equal to or smaller than $(a_1 + a_2)$, then the depletion regions occupy the metallurgical channel, the DN-MOS transistor is accidentally "on" all the time. There will be no "off" state, which is called "punch through." Frohman-Bentchkowsky *et al.* (11) reported this phenomenon for a lightly doped short channel device.

The worst case to consider possible punch through is when the transistor is supposed to be in "off" state, while no drain current is drawn, and the large drain voltage V_{ds} is reverse biasing the drain p-n junction giving a large a_2.

Assume V_{ds} is 20 volts, which is the highest d.c. voltage in the memory cell to be discussed later. Consider the following parameters

$$\frac{\text{(junction bias)}}{\text{(background doping)}} = \frac{20^{v}}{10^{15}\ \text{cm}^{-3}}$$

$$= 2 \times 10^{-14}\ \text{v - cm}^3 \qquad (26)$$

$$\frac{\text{(background doping)}}{\text{(surface doping)}} = \frac{10^{15}}{6.7 \times 10^{17}}$$

$$= 1.49 \times 10^{-3} \qquad (27)$$

and a Gaussian junction depth of 3.74×10^{-4} cm, the total depletion width and a_2 can be found from the calculated results by Lawrence and Warner (14) to be:

(drain junction total depletion width)

$= 5 \times 10^{-4}$ cm

$= 5 \times 10^{-6}$ meter

and

$a_2 = (0.17)(5 \times 10^{-4}) = 8.5 \times 10^{-5}$ cm

$= 0.85 \times 10^{-6}$ meter

A similar approach is used to find a_1. Consider 0.7 volt n^+-p junction built-in voltage as junction bias, 2×10^{17} per cm^3 as background doping, 10^{21} per cm^3 as surface doping, and a complementary error function junction depth of 1.55×10^{-4} cm. These new parameters yield:

(source junction total depletion width)

$= 0.14 \times 10^{-6}$ meter

and

$a_1 = 0.08 \times 10^{-6}$ meter

Then $(a_1 + a_2)$ gives 0.93×10^{-6} meter which is smaller than $(X_{ps} - X_{n+s}) = 1.7 \times 10^{-6}$ meter. The DN-MOS transistor will not punch through under 20 volts bias; an effective channel length of 0.77×10^{-6} meter still exists.

The numerical values involved so far are summarized in Table 2. The depletion widths and effective channel length, when the DN-MOS transistor is in "on" state, are also listed in Table 2. During the calculation built-in voltages are considered as the junction bias for both source and drain junctions.

Table 2. Numerical Values of Depletion Region Width and Effective Channel Length

DN-MOS	State	OFF $V_{ds} = 20^{v}$	ON $V_{ds} = 0^{v}$
source n^+-p junction	total depletion width, cm	0.14×10^{-4}	0.14×10^{-4}
	a_1, cm	0.08×10^{-4}	0.08×10^{-4}
drain p-n junction	total depletion width, cm	5×10^{-4}	1.3×10^{-4}
	a_2, cm	0.85×10^{-4}	0.44×10^{-4}
effective channel length $(X_{ps}-X_{n+s}) - (a_1+a_2)$ cm		0.77×10^{-4}	1.18×10^{-4}

The depletion region width variation effect on DN-MOS transistor effective channel length is shown in Figure 19 with the numerical values in Table 2 as reference. Two approximations are made during the calculation process: the first one is that the p-region impurity doping level is a constant of 2×10^{17} per cm^3 near ($\pm 0.1 \times 10^{-6}$ meter, where depletion region extends) the source n^+-p junction; the second one is that the depletion width does not change

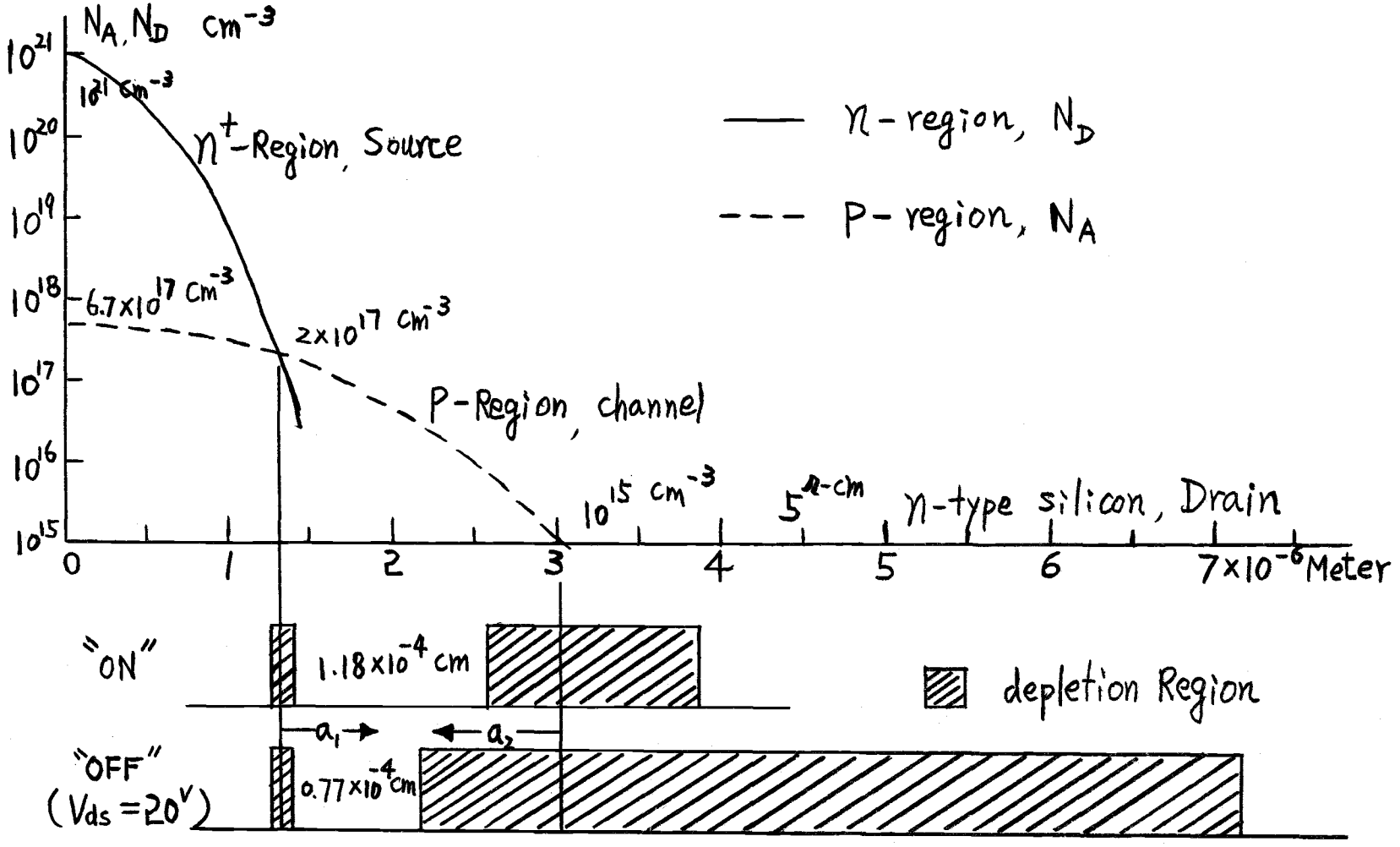


Figure 19. Depletion region width variation effect on channel length of DN-MOS transistor.

in lateral cases.

Figure 19 indicates that effective channel length varies from 0.77×10^{-6} meter to 1.18×10^{-6} meter as the DN-MOS transistor switches from "off" state to "on" state.

It has been shown that the channel length of the DN-MOS transistor depends on biasing condition and impurity profile distribution. Section II C 3c shows that the impurity profile also controls the threshold voltage of the DN-MOS transistor. It is then concluded that the impurity profile determines both threshold voltage and channel length of the DN-MOS transistor.

5. Diffusion Schedule Considerations

There are three diffused regions under consideration: p-MOS transistor source and drain region, DN-MOS transistor p-region, and DN-MOS transistor n^+-region. Because diffusions are performed consecutively, the heat cycle experienced must be accumulatively considered for each region. Table 3 shows the schedule summarized from Appendix I.

In planning this schedule, several conditions must be met. The first condition is that the total time experienced by DN-MOS p-region and DN-MOS n^+-region affect their impurity profiles, which in turn control the channel length and the threshold voltage of the DN-MOS transistor. They must be properly arranged to avoid punch through and/or too high a threshold voltage.

Table 3. Diffusion Schedules

diffused region	p-MOS source and drain	DN-MOS p-region	DN-MOS n^+-region
section describing diffusion	II C 1b	II C 1c	II C 1c
diffusion type	Gaussian	Gaussian	erfc
diffusion time	4 hr	5 hr	1 hr
total time experienced	10 hr	6 hr	1 hr

In the determination of DN-MOS n^+-region diffusion time, it is assumed that the DN-MOS p-region has been chosen (within limits stated later in section II C 6) as the dotted line in Figure 14, and the n^+-p source junction occurs at $X = 1.55 \times 10^{-6}$ meter $N(X) = 2 \times 10^{17}$ per cm^3 point. This choice determines both the threshold voltage and channel length of the DN-MOS transistor. Set $N_D(X)$ in equation (10) equals to 2×10^{17} per cm^3, and X equals to 1.55×10^{-6} meter to solve for t; t is found to be one hour for DN-MOS n^+-region diffusion.

Every planned schedule must be checked according to the same approach as in sections II C 4 and II C 3c, for punch through and threshold voltage.

The second condition is that since different regions are diffused consecutively, there must be an appreciable difference between the total time experienced by each

region, so that it is possible to schedule diffusion time for each step.

The third condition is that the drive-in time for p-MOS source and drain region must be long enough (four hours in Table 3) so that thick enough oxide can be grown during drive-in to protect the region against the following p and n^+ diffusions for the DN-MOS transistor.

Too long a drive-in time for p-MOS source and drain region will give too deep a junction depth (refer to Figure 17). The lateral diffused source and drain region will decrease the channel length as seen from the diffusion mask, causing large Miller's capacitance between gate and source (or drain).

When isolation diffusion is needed, a total drive-in time of 48 hours will safely isolate a 5 ohm-cm resistivity and 8.5 to 9 x 10^{-6} meter thick n-type silicon layer (refer to Figure 17). This isolation could be done before p-MOS source and drain region which has a ten hour total drive-in time.

6. Other Considerations

During the drive-in diffusion of the DN-MOS p-region the atmosphere must be in inert gas as described in Appendix I c. Nitrogen is used to prevent any oxidation so that the original DN-MOS opening edge (see dotted line in Figure 18) can be preserved as mask for the following

n^+ diffusion.

The doping level of point A in Figure 15 has an upper limit of 4.5×10^{17} per cm^3 if threshold voltage of the DN-MOS transistor, V_{tn}, is limited to a maximum value of ten volts (see Figure 13). There is a lower limit for the doping level of point A to be 4.5×10^{16} per cm^3 so that V_{tn} can be positive, then the DN-MOS transistor operates in enhancement mode. The impurity level corresponding to point A in Figure 15 is determined mostly by the p-region impurity profile. The above mentioned limits serve as limits for the DN-MOS p-region impurity level near n^+-p junction as well.

The 5 ohm-cm starting material affects the threshold voltage of p-MOS transistor V_{tp} (see Figure 5). It also affects the DN-MOS transistor drain p-n junction breakdown voltage; for n-type silicon having 1 to 25 ohm-cm resistivity (5×10^{15} to 2×10^{14} per cm^3) 3.74×10^{-6} meter Gaussian junction depth, the breakdown voltage is well above 20 volts (reference 20), so that the d.c. bias in the memory cell circuit discussed later can operate safely on the DN-MOS transistor.

III. THE COMPLEMENTARY MOS RANDOM ACCESS MEMORY CELL

The complementary MOS random access memory (CMOS RAM) cell to be discussed can operate from static (d.c.) to high speed (a few megahertz). The cell offers a non-destructive read out process; reading need not be followed by a write operation to restore memory contents. Each cell is directly addressed within the memory matrix of cells through its associated selecting MOS transistors. The addressing process is random, because each cell has its own combination of "on" selecting transistors.

Figure 20 shows the functional blocks of the memory system. Binary addressing signals are decoded through x or y decoder circuitry to select an x and a y line; the selected x and y lines determine the particular cell in the memory matrix where data can be written in or read out through the read/write circuitry. For an n x n memory matrix, there are n x-lines and n y-lines, while binary input to both x and y decoder should have $\log_2 n$ lines.

Cells within the memory matrix are all identical. A typical cell_{ij} is investigated in the following sections.

As shown in Figure 21, the cell consists of eight MOS transistors; four (Q_5 Q_6 Q_7 Q_8) are selecting transistors, and the other four (Q_1 Q_2 Q_3 Q_4) form a bistable storage flip-flop. Only Q_3 and Q_4 are n-type channel MOS

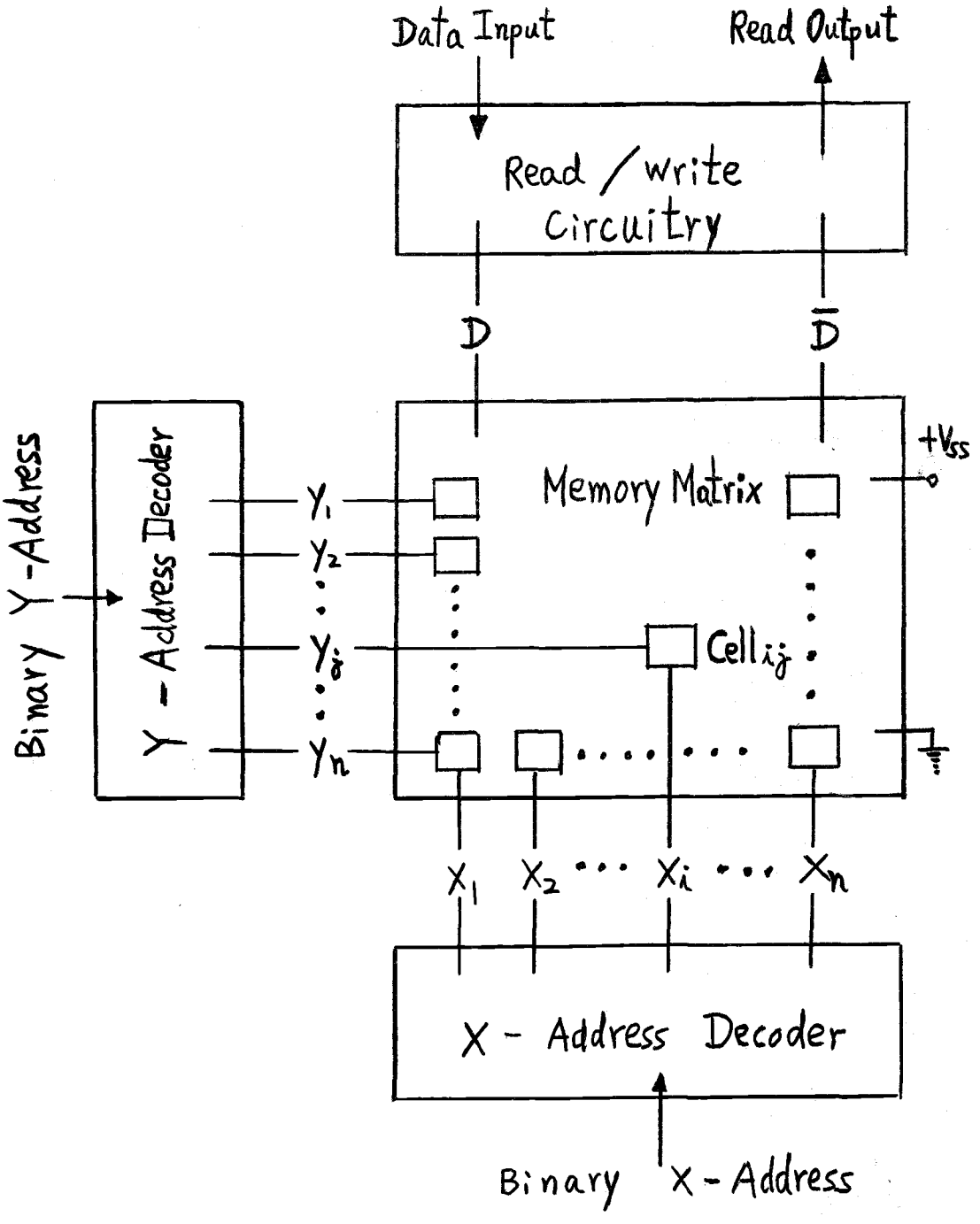
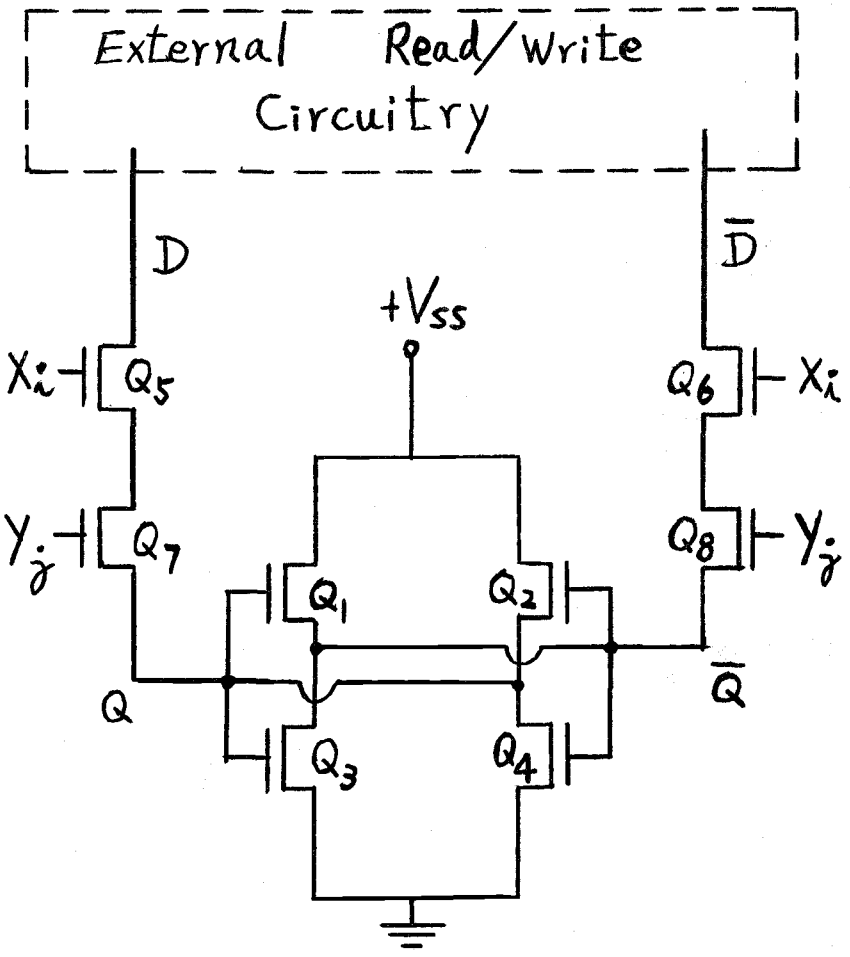


Figure 20. The random-access-memory system block diagram.

A. THE CMOS RAM CELL CIRCUIT



Q_5 Q_7 Q_1 Q_2 Q_8 Q_6	p-MOS transistors
Q_3 Q_4	n-MOS transistors
Q_5 Q_6 Q_7 Q_8	selecting transistors
Q_1 Q_2 Q_3 Q_4	storage flip-flop

Figure 21. CMOS RAM cell circuit.

transistors; the remaining six are p-type channel MOS transistors.

Two possible circuit configurations for external read/write circuit can be found in a Texas Instrument application report (17).

The selecting process is accomplished by negative going voltage step from V_{ss} to 0 reaching node x_i and y_j in Figure 21. If writing of "1" is performed at this moment, the external read/write circuit will provide + V_{ss} at node D, which pulls node Q to V_{ss} through Q_5 and Q_7; the writing of "0" is done the same way only through $\bar{D}$ $\bar{Q}$ branch instead. Once the information is written in the storage flip-flop (Q_1 Q_2 Q_3 Q_4), it will be memorized by the bistable nature of the flip-flop in terms of nodal voltage at Q and $\bar{Q}$. If reading of "1" in the cell is performed, the high voltage at node Q ($V_Q = V_{ss}$) will cause a current flow in branch Q D, which will then be detected by a sense amplifier in the read/write circuitry in Figure 21. The reading of "0" is done the same way through current flow in $\bar{Q}$ $\bar{D}$ branch, but the sense amplifier will have a different response due to its input voltage polarity. Both node x_i and y_j must be at zero ground potential ($cell_{ij}$ being selected from the matrix) when writing or reading is taking place. During stand-by situation, at least one of x_i or y_j must be at V_{ss} turning off the corresponding selecting transistor isolating the storage

flip-flop from external read/write circuit. The coincident select method is the same as that used by Friedrich (10).

B. CELL D.C. OPERATION

Q_5 and Q_7 or Q_6 and Q_8 in Figure 21 only provide the necessary path to the read/write circuitry when needed. The storage of information is essentially performed by Q_1, Q_2, Q_3, Q_4 in Figure 21. The Q_1, Q_2, Q_3 and Q_4 portion is redrawn in Figure 22 below.

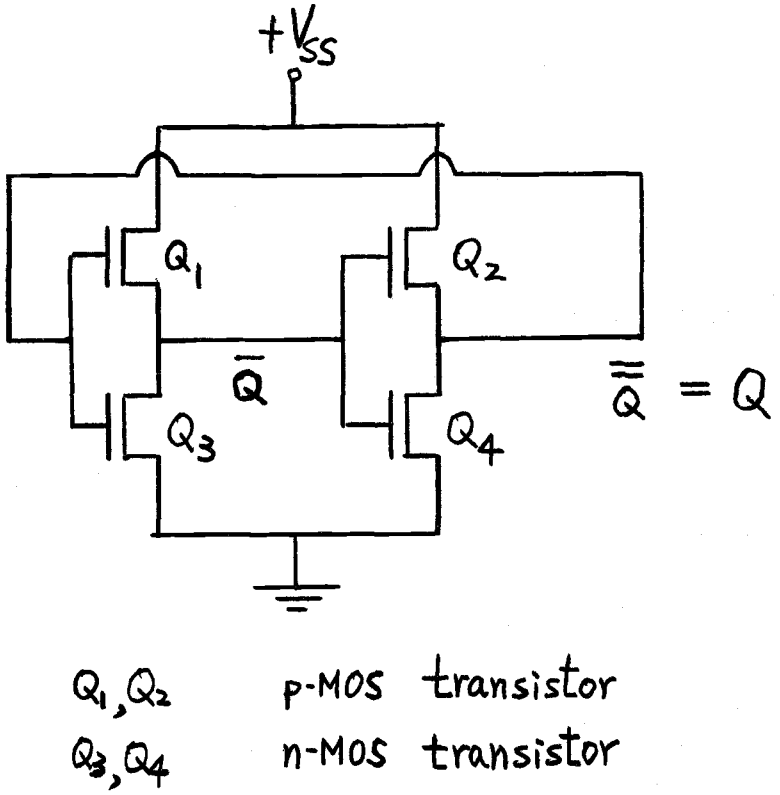


Figure 22. CMOS storage cell.

The d.c. operation of the CMOS storage cell can be best stated in a tabled form:

Table 4. Static Cell States

cell state	Q_1	Q_2	Q_3	Q_4	V_Q	$V_{\bar{Q}}$
"1"	OFF	ON	ON	OFF	$+V_{ss}$	0
"0"	ON	OFF	OFF	ON	0	$+V_{ss}$

It is noted that in either state "1" or "0", no current flows (except leakage current of the "off" device) between the d.c. power supply + V_{ss} and ground, which is the power saving nature of complementary MOS memory.

Close examination of the configuration in Figure 22 reveals that the cell is two inverters in cascade with the output of the second inverter fed back to the input of the first inverter to ensure bistable of the cell when input is removed. It is now imperative to study the static operation of the complementary MOS inverter so that how the cell changes state can be understood. Figure 23 shows the first inverter which is composed of Q_1 and Q_3.

The p-MOS transistor Q_1 can be thought as load to the n-MOS transistor Q_3. As V_Q increases from zero to + V_{ss}, $V_{\bar{Q}}$ decreases from + V_{ss} to zero. The operating point corresponding to different V_Q value (0_{VQ}) is shown in Figure 24. Transistor drain characteristics are, from Figure 10, dotted line for n-MOS transistor Q_3 and solid line for load p-MOS transistor Q_1. V_{ss} is assumed to be 20 volts. It is noted that Q_3 goes from off to

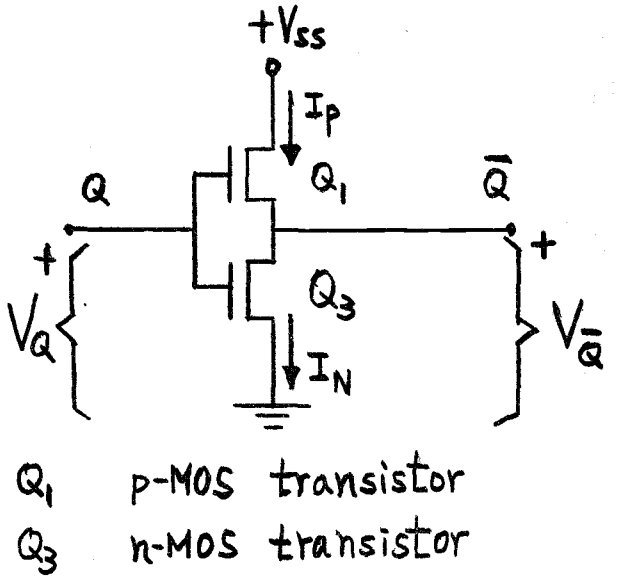


Figure 23. CMOS inverter.

saturation then to non-saturation, and Q_1 goes from non-saturation to saturation then to off as V_Q increases. Based on these different operating conditions of Q_1 and Q_3, five different regions are listed in Table 5. The governing inequalities in Table 5 are based on the conditions given in equations (1), (2), and (3), with the understanding that for Q_1 $|V_{gs}| = V_{ss} - V_Q$ and $|V_{ds}| = V_{ss} - V_{\bar{Q}}$ while for Q_3 $|V_{gs}| = V_Q$ and $|V_{ds}| = V_{\bar{Q}}$. The boundary between regions is determined from two relating inequalities; for instance, in region I ($V_Q < V_{th}$) while in region II ($V_Q > V_{th}$) the boundary dividing regions I and II must be at ($V_Q = V_{th}$) etc. Results are shown in the scale below Table 5.

The inverter voltage transfer curve is shown in

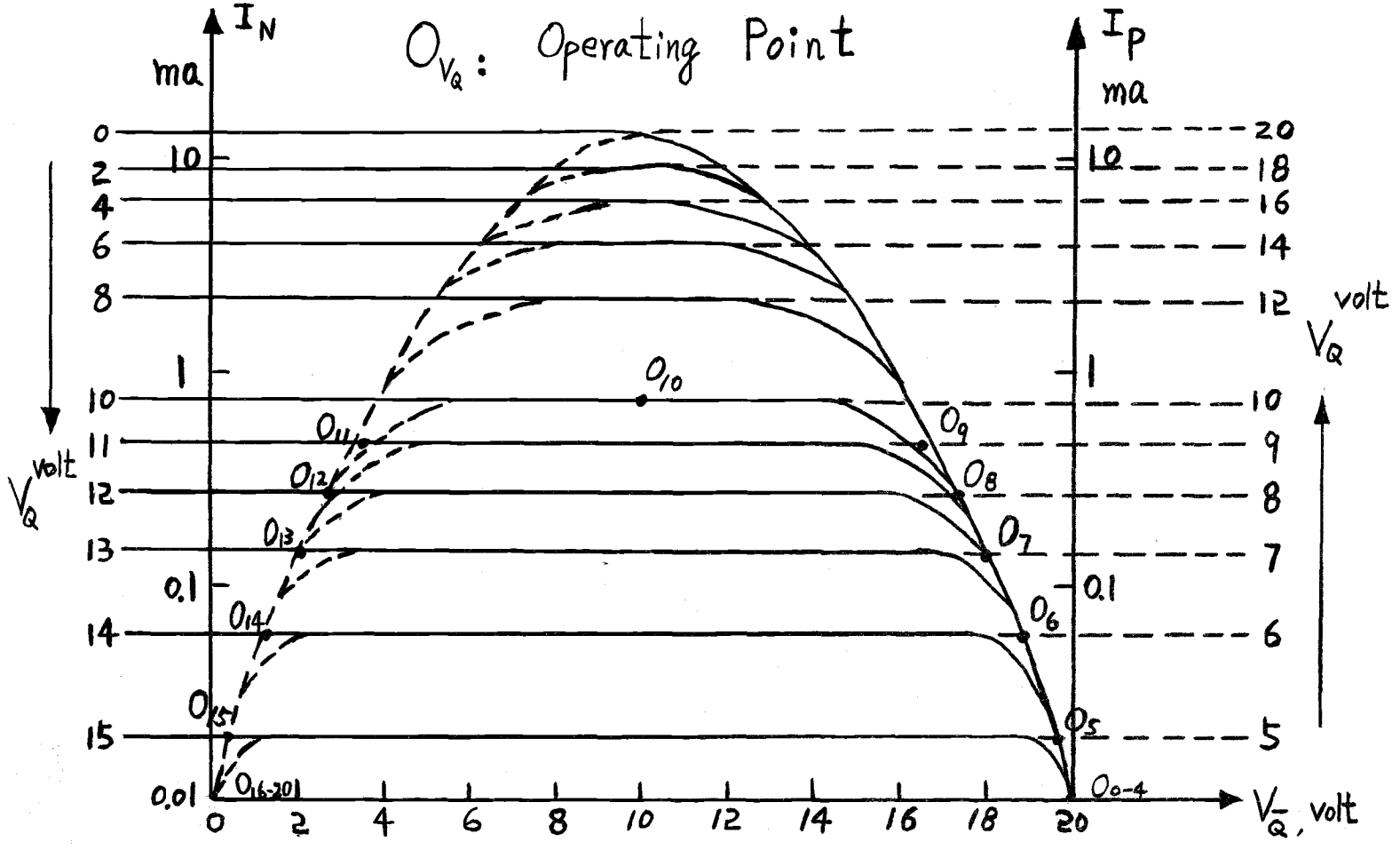


Figure 24. Inverter operating point locus.

Table 5. Transistor Operating Condition versus Input Voltage (V_Q) of an Inverter

Transistor \ Region	I	II	III	IV	V
N-MOS (Q_3)	(OFF) ($V_Q < V_{tn}$)	(ON) ($V_Q > V_{tn}$) (Saturated) ($V_{\bar{Q}} > V_Q - V_{tn}$) ($V_Q < V_{\bar{Q}} + V_{tn}$)		(ON) ($V_Q > V_{tn}$) (Nonsaturated) ($V_{\bar{Q}} < V_Q - V_{tn}$) ($V_Q > V_{\bar{Q}} + V_{tn}$)	
P-MOS (Q_1)	(ON) ($V_{ss}-V_Q > \lvert V_{tp}\rvert$) ($V_Q < V_{ss} - \lvert V_{tp}\rvert$) (Nonsaturated) ($V_{ss}-V_{\bar{Q}} < V_{ss}-V_Q - \lvert V_{tp}\rvert$) ($V_Q < V_{\bar{Q}} - \lvert V_{tp}\rvert$)		(ON) ($V_{ss}-V_Q > \lvert V_{tp}\rvert$) ($V_Q < V_{ss} - \lvert V_{tp}\rvert$) (Saturated) ($V_{ss}-V_{\bar{Q}} > V_{ss}-V_Q - \lvert V_{tp}\rvert$) ($V_Q > V_{\bar{Q}} - \lvert V_{tp}\rvert$)	(OFF) ($V_{ss}-V_Q < \lvert V_{tp}\rvert$) ($V_Q > V_{ss} - \lvert V_{tp}\rvert$)	

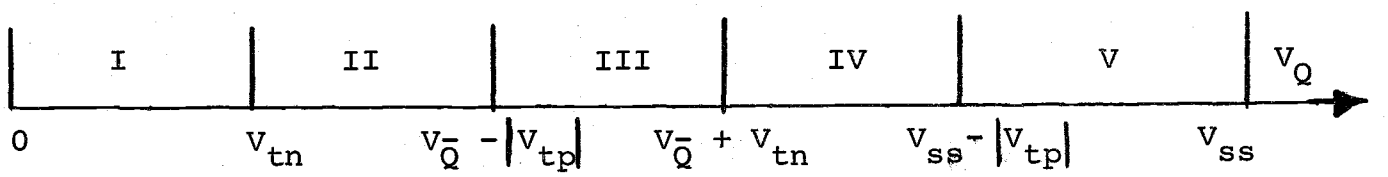


Figure 25, and the inverter current versus input voltage curve is shown in Figure 26. All data are taken from Figure 24.

From Figure 25, the slope ($dV_{\bar{Q}}/dV_Q$) is zero in region I and V, while in region III it has a very large negative value. This means the inverter has a high gain region in the center separating two regions with zero gain, which is an ideal characteristic based on input voltage V_Q noise consideration.

The value of input voltage V_Q^* when the inverter operates in high gain region III is called the transition voltage. For better noise immunity, V_Q^* is designed to be 1/2 V_{ss} (see Figure 25). A condition relating device constant and biasing voltage can be found according to this design criteria.

First, the set of MOS equations for both p-channel Q_1 and n-channel Q_3 are rewritten below, referring back to equations (1), (2), (3) and Figure 23 for convenience. For p-channel Q_1 we have:

$$I_p = |I_{ds}| \tag{28}$$

$$= K_p \left[2(|V_Q - V_{ss}| - |V_{tp}|)(V_{\bar{Q}} - V_{ss}) - (|V_{\bar{Q}} - V_{ss}|)^2 \right.$$

$$\text{for } |V_{\bar{Q}} - V_{ss}| < |V_Q - V_{ss}| - |V_{tp}|, \text{ non-saturation region}$$

$$I_p = K_p(|V_Q - V_{ss}| - |V_{tp}|)^2 \tag{29}$$

$$\text{for } |V_{\bar{Q}} - V_{ss}| \geq |V_Q - V_{ss}| - |V_{tp}|, \text{ saturation region}$$

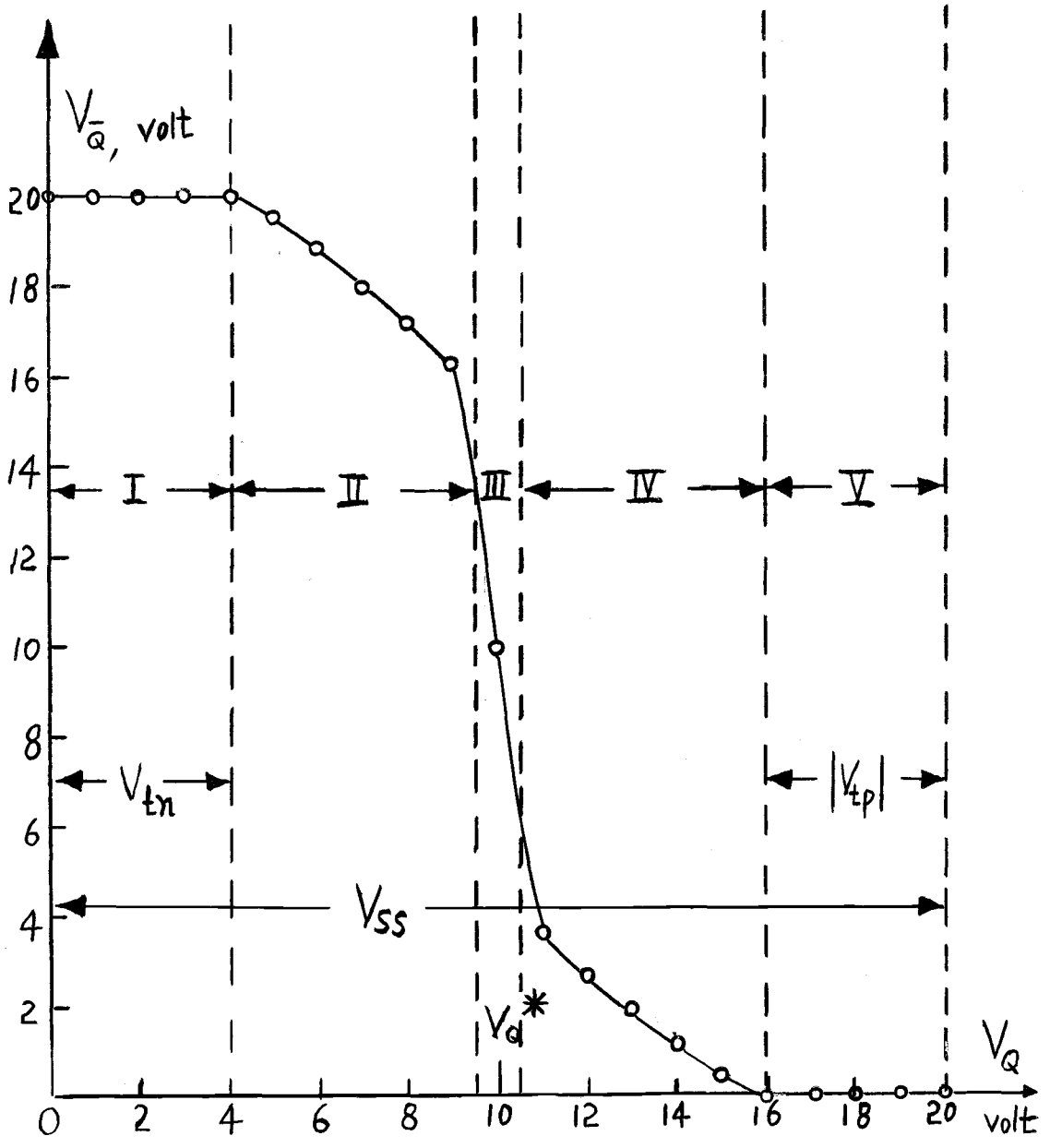


Figure 25. Inverter voltage transfer curve. Input voltage (V_Q) versus output voltage ($V_{\bar{Q}}$).

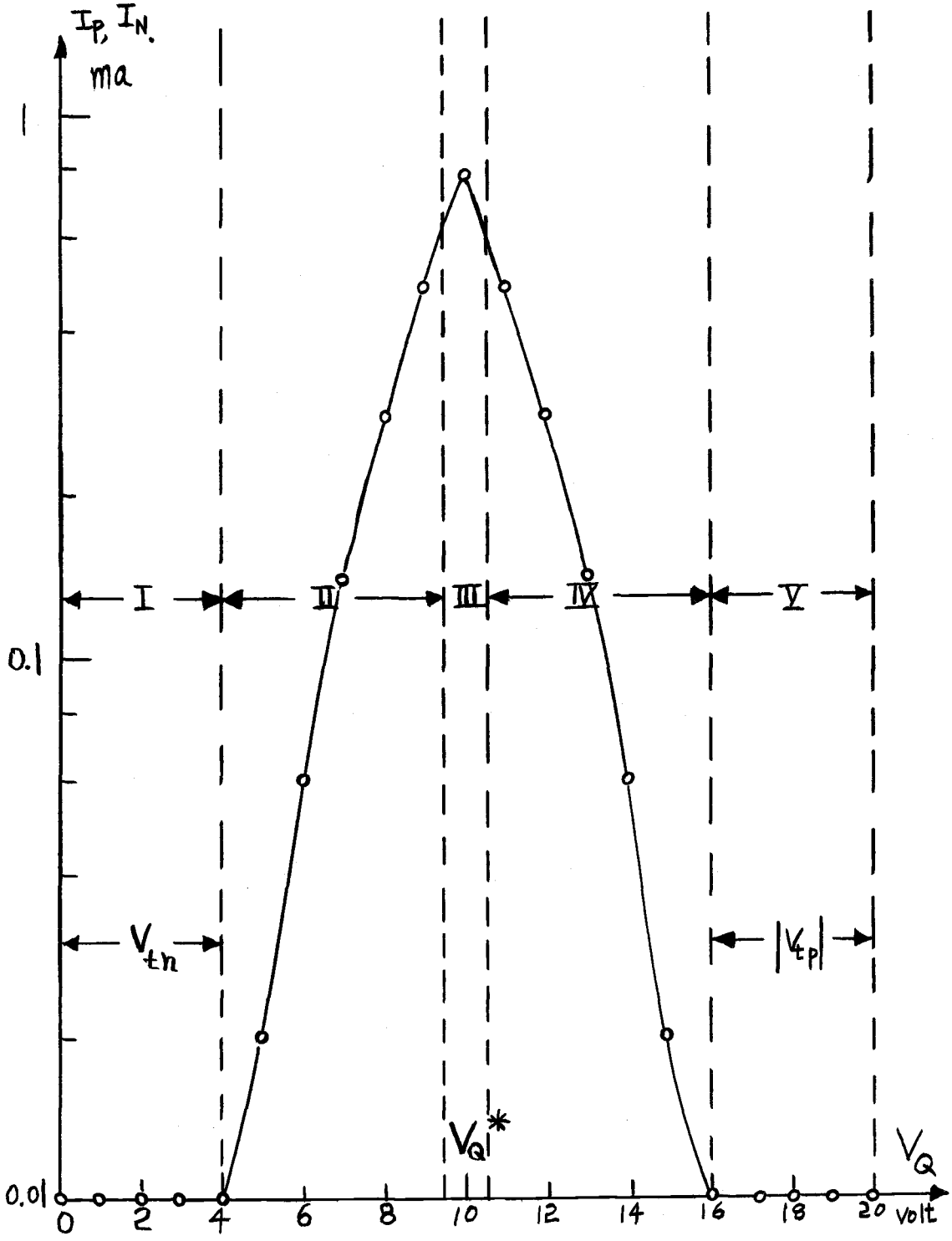


Figure 26. Inverter current (I_p,I_N) versus input voltage (V_Q).

$$I_p = 0 \tag{30}$$

for $|V_Q - V_{ss}| < |V_{tp}|$, off region

For n-channel Q_3 we have:

$$I_N = I_{ds} = K_N \left[2(V_Q - V_{tn})V_{\bar{Q}} - V_{\bar{Q}}^2 \right] \tag{31}$$

for $V_{\bar{Q}} < V_Q - V_{tn}$, non-saturation region

$$I_N = K_N (V_Q - V_{tn})^2 \tag{32}$$

for $V_{\bar{Q}} \geq V_Q - V_{tn}$, saturation region

$$I_N = 0 \tag{33}$$

for $V_Q < V_{tn}$, off region

where:

$$K_p = \frac{\bar{\mu}_h \epsilon_{ox} W_p}{2 t_{ox} L_p} \tag{34}$$

$$K_N = \frac{\bar{\mu}_e \epsilon_{ox} W_N}{2 t_{ox} L_N} \tag{35}$$

K_p and K_N are the MOS transistor conduction constants.

In region III, both Q_1 and Q_3 are in saturation (see Table 5). Equations (29) and (32) yield:

$$I_{ds} = K_p(|V_Q^* - V_{ss}| - |V_{tp}|)^2 = K_N(V_Q^* - V_{tn})^2$$

namely

$$\left(\frac{K_p}{K_N}\right)\left(\left|\frac{V_Q^*}{V_{ss}} - 1\right| - \frac{|V_{tp}|}{V_{ss}}\right)^2 = \left(\frac{V_Q^*}{V_{ss}} - \frac{V_{tn}}{V_{ss}}\right)^2 \tag{36}$$

Define:

$$\beta = \frac{K_p}{K_N} \tag{37}$$

$$\alpha_n = \frac{V_{th}}{V_{ss}} \tag{38}$$

$$\alpha_p = \frac{|V_{tp}|}{V_{ss}} \tag{39}$$

Equation (36) becomes:

$$\sqrt{\beta}(0.5 - \alpha_p) = (0.5 - \alpha_n) \tag{40}$$

where $V_Q^* = \frac{1}{2} V_{ss}$ was assumed.

Equation (40) represents a family of straight lines in the α_n, α_p plane with slope $\sqrt{\beta}$ as parameters. Rapp (21) and Cobbold (3) gave the curve in their works. An easy way to satisfy equation (40) is when $V_{tn} = |V_{tp}|$ and $K_p = K_N$. An interesting condition is that if $\alpha_p = \alpha_n = 0.5$ equation (40) will always hold independent of β value, but this will give a relatively slow dynamic time response as stated by Burns (1).

So far, node voltage V_Q is considered to be the input to the inverter composed of Q_1 and Q_3. $V_{\bar{Q}}$ is considered to be the response, which is exactly the case when writing of "1" is implemented by the cell. The writing of "0" can be analyzed by considering $V_{\bar{Q}}$ as input to inverter Q_2, Q_4 (see Figure 22) and V_Q as output response, which leads to the same analysis.

During the "read" operation, no change of state is happening in the cell (Q_1, Q_2, Q_3, Q_4 remain in their original state). Only the side of the cell with high

voltage (+ V_{ss}) will cause current flow in either digit line D or $\bar{D}$ depending on the state of the cell. If the state is "1", $V_Q = V_{ss}$ produces current in digit line D, otherwise digit line $\bar{D}$ will have current flow (see Table 4 and Figure 21).

C. CELL DYNAMIC OPERATION

Two different cases will be considered when the cell is in dynamic operation.

The first case is when a step voltage $\left(V_{ss}\ u(t)\right)$ is reaching one side (either node Q or node $\bar{Q}$ in Figure 22) of the cell, the other side is having a voltage change (from V_{ss} to zero) as a response. This is the situation when writing of either "1" or "0" is taking place in the cell. How long it takes for the response to change from 0.9 V_{ss} to 0.1 V_{ss} is calculated in detail. This amount of time is the minimum time required for a "write" operation. If the input voltage is a time-dependent function, then an analytical closed form solution for the response voltage is quite difficult. Feller (9) used a lumped model to simulate the response through computer analysis, which is a practical way of predicting the response. However, the step response analysis will yield some insight of the factors affecting the write time.

The second case considers the pulsed response of the cell. The maximum switching rate of the cell is determined

from the inverter pair delay time.

1. Cell Step Response

In Figure 22 assume $V_Q(t) = V_{ss}u(t)$, and $V_{\bar{Q}}(0) = V_{ss}$. The inverter in Figure 23 is redrawn in Figure 27 with nodal capacitor between node $\bar{Q}$ and ground shown. In the actual physical structure the nodal capacitor is the output capacitor of Q_1 and Q_3 (between their source and drain) plus the input capacitance of Q_2 and Q_4 (between their gate and source).

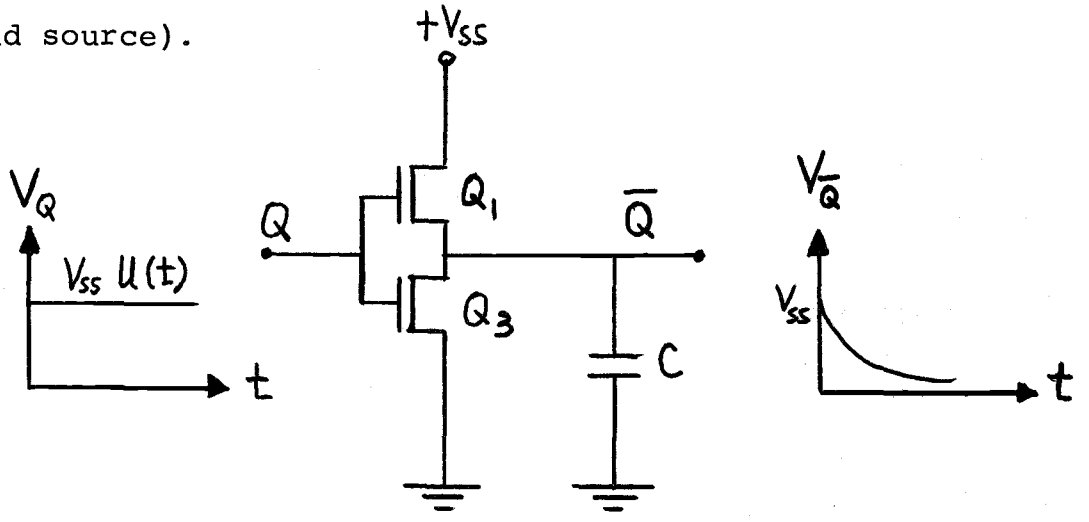


Figure 27. Inverter with step input

The step voltage will turn off Q_1, leaving Q_3 on as the only discharging path for capacitor C. The situation is further simplified as shown in Figure 28.

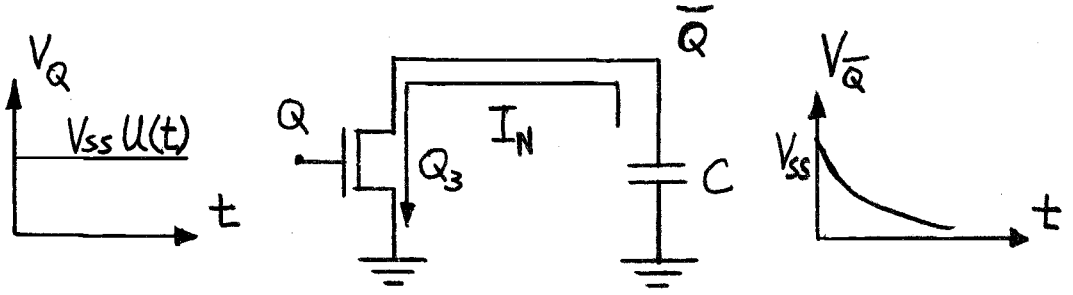


Figure 28. Simplified inverter step response.

As the discharge process continues, $V_{\bar{Q}}$ is decreasing from V_{ss} towards zero. Q_3 operates from saturation ($V_{\bar{Q}} \geq V_{ss} - V_{tn}$) to non-saturation ($V_{\bar{Q}} < V_{ss} - V_{tn}$). Due to changing biasing condition, Q_3 acts like a current valve regulating the discharge current I_N.

For saturation region, it has:

$$- C \frac{dV_{\bar{Q}}}{dt} = K_N (V_{ss} - V_{tn})^2 \tag{41}$$

where equation (32) was referred to. Solving equation (41) for $V_{\bar{Q}}(t)$ and employing initial condition $V_{\bar{Q}}(0) = V_{ss}$ yields:

$$V_{\bar{Q}}(t) = V_{ss} - \frac{K_N(V_{ss}-V_{tn})^2}{C} t \tag{42}$$

$$\text{for } (V_{\bar{Q}} \geq V_{ss} - V_{tn})$$

Equation (42) indicates a linear decay of $V_{\bar{Q}}(t)$ before t_o,

where

$$V_{\bar{Q}}(t_o) = V_{ss} - V_{tn}$$

$$= V_{ss} - \frac{K_N(V_{ss}-V_{tn})^2}{C} t_o \tag{43}$$

hence

$$t_o = \frac{C\, V_{tn}}{K_N(V_{ss}-V_{tn})^2} \tag{44}$$

From t_o and on the binding equation becomes:

$$-C \frac{dV_{\bar{Q}}}{dt'} = K_N \left[2(V_{ss}-V_{tn})\, V_{\bar{Q}} - V_{\bar{Q}}^2 \right] \tag{45}$$

$$\text{for } (V_{\bar{Q}} < V_{ss}-V_{tn}) \text{ and } t' = (t-t_o) \geq o$$

Solving equation (45) for $V_{\bar{Q}}(t')$ is quite difficult; however, the time for $V_{\bar{Q}}$ decreases to v, namely t'(v) can be solved from equation (45) as:

$$t'(v) = - \int_{V_{ss}-V_{tn}}^{v} \frac{C\, d\, V_{\bar{Q}}}{K_N \left[2(V_{ss}-V_{tn})V_{\bar{Q}} - V_{\bar{Q}}^2 \right]}$$

$$= \left(\frac{C}{K_N}\right)\left(\frac{1}{V_{ss}-V_{tn}}\right) \int_o^{V_{ss}-V_{tn}-v} \frac{(V_{ss}-V_{tn})\, d(V_{ss}-V_{tn}-V_{\bar{Q}})}{(V_{ss}-V_{tn})^2 - (V_{ss}-V_{tn}-V_{\bar{Q}})^2}$$

$$= \frac{C}{K_N(V_{ss}-V_{tn})} \tanh^{-1} \frac{V_{ss} - V_{tn} - v}{V_{ss} - V_{tn}} \tag{46}$$

where Dwight's integration table (8) was used. Equation (46) is rearranged to be:

$$v(t') = (V_{ss}-V_{tn}) \left[1 - \tanh \frac{K_N(V_{ss}-V_{tn})\, t'}{C} \right] \tag{47}$$

hence

$$V_{\bar{Q}}(t) = (V_{ss}-V_{tn})\left[1-\tanh\frac{K_N(V_{ss}-V_{tn})(t-t_o)}{C}\right] \qquad (48)$$

$$\text{for } t \geqslant t_o$$

Equations (42) and (48) can be written as (refer to equation 38)

$$\frac{V_{\bar{Q}}(t)}{V_{ss}} = 1 - \frac{(1-\alpha_n)^2}{T_N}\,t \qquad (49)$$

$$\text{for } t < t_o$$

and

$$\frac{V_{\bar{Q}}(t)}{V_{ss}} = (1-\alpha_n)\left[1-\tanh\frac{(1-\alpha_n)(t-t_o)}{T_N}\right] \qquad (50)$$

$$\text{for } t \geqslant t_o$$

where defining:

$$T_N = \frac{C}{K_N\,V_{ss}} \qquad (51)$$

Let equation (49) equal to 0.9 solve for $t_{0.9}$ and equation (50) equal to 0.1 solve for $t_{0.1}$; the fall time of $V_{\bar{Q}}$ can be expressed as:

$$T_F = t_{0.1} - t_{0.9}$$

$$= T_N\left[\frac{\alpha_n-0.1}{(1-\alpha_n)^2} + \frac{\tanh^{-1}\left(\frac{0.9-\alpha_n}{1-\alpha_n}\right)}{1-\alpha_n}\right] \qquad (52)$$

Equation (52) has exactly the same expression given by Burns (1). Figure 29 gives a clear picture of the quantities calculated before.

So far V_Q has been considered input to the cell and

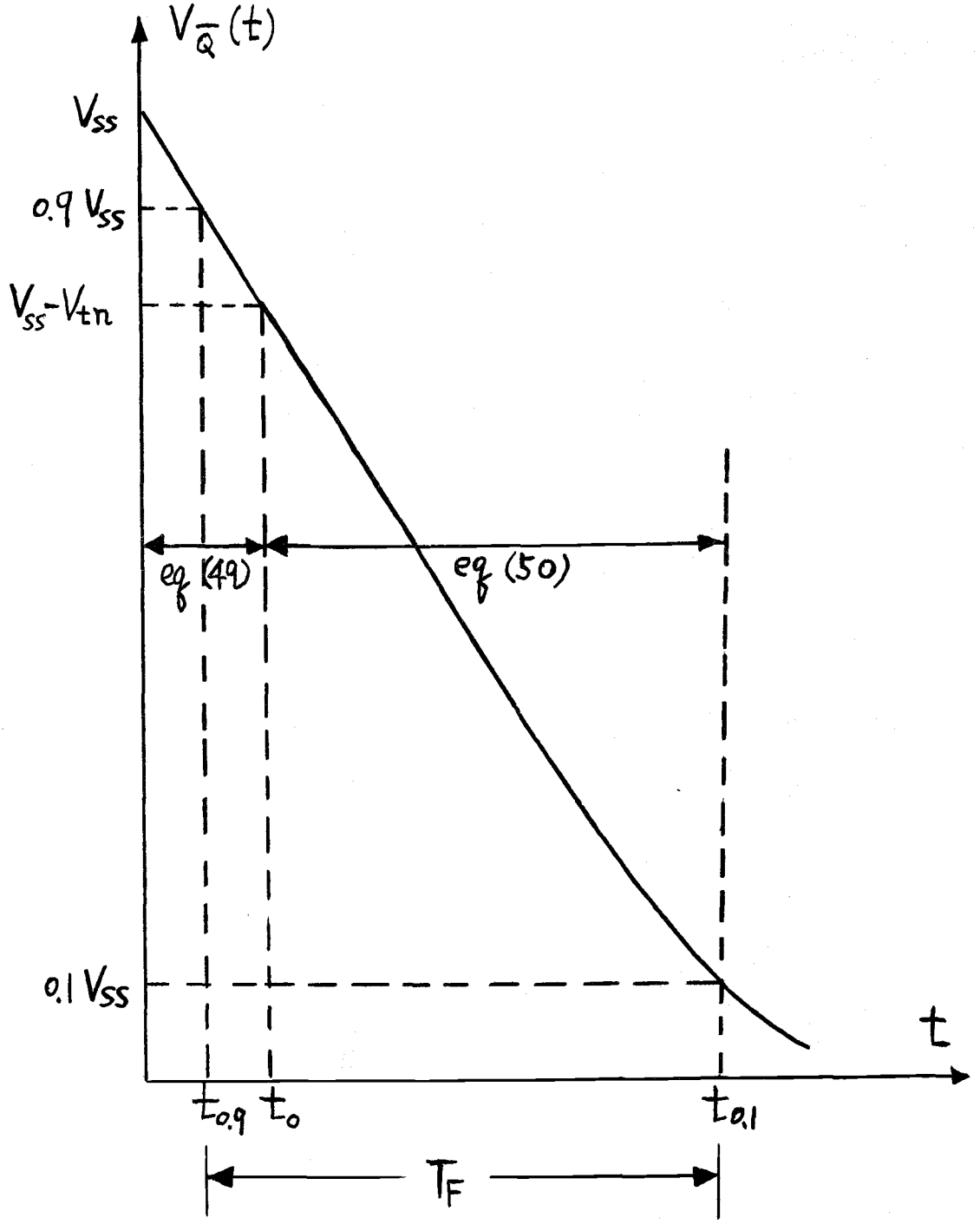


Figure 29. Step response of cell.

$V_{\bar{Q}}$ the response, which is the case of writing "1" into the cell. Because writing of "0" is the dual case, V_Q will have the same fall time as $V_{\bar{Q}}$ before.

2. Cell Pulsed Response

The cell in Figure 22 can be represented by the following block diagram consisting of two inverters.

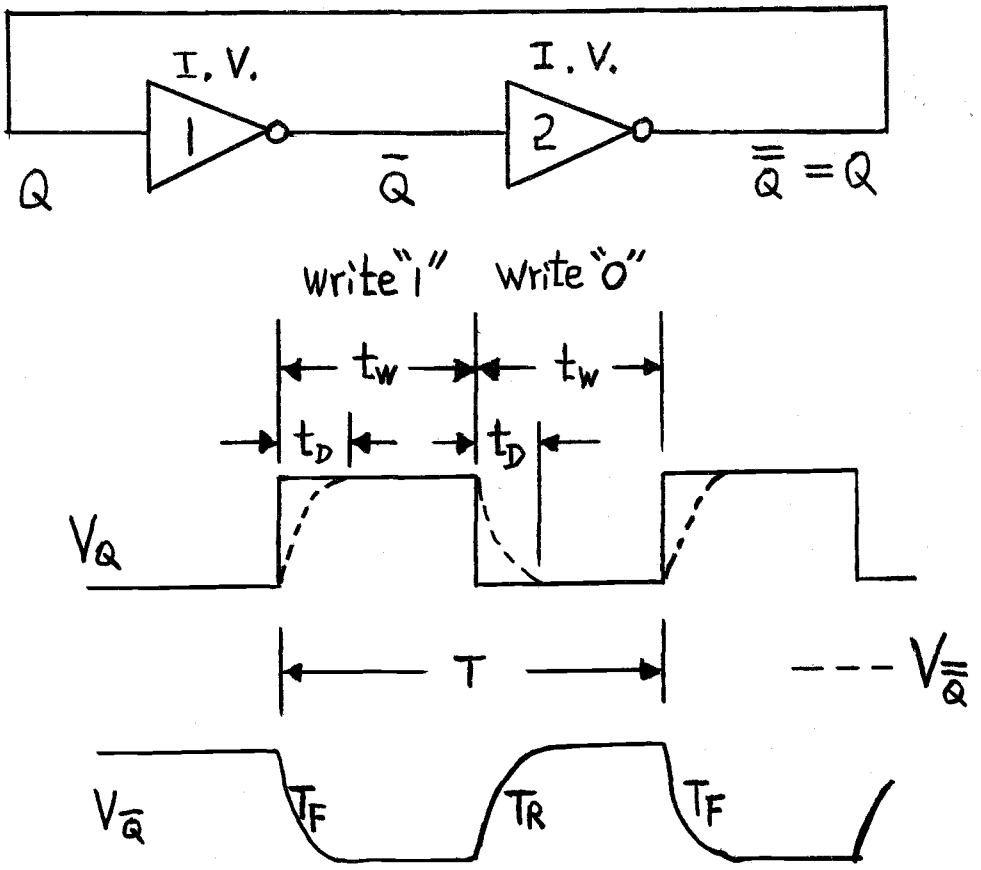


Figure 30. Cell pulsed response.

The first inverter consists of Q_1 and Q_3. The second inverter consists of Q_2 and Q_4 shown in Figure 22.

If pulsed voltage is applied to node Q, as writing is in progress, the write-pulse width t_w must be longer than

the pair delay time t_D of inverter 1 and inverter 2 in cascade to ensure proper switching of the cell. The pair delay time was given by Burns (1) as:

$$t_D \approx 0.9\, T_N \left[\frac{1}{(1-\alpha_n)^2} + \frac{1}{\beta(1-\alpha_p)^2} \right] \tag{53}$$

Definitions of T_N, α_n, α_p, β are given in equations (51), (38), (39), and (37) respectively. Equation (53) is restricted to $\beta<1$; for $\beta>1$ case change T_N to T_p, where

$$T_p = \frac{C}{K_p V_{ss}} \tag{54}$$

The maximum switching rate of the cell is:

$$\frac{1}{T} = \frac{1}{2t_D} \tag{55}$$

When the cell is in operation, after the write time, t_w, a finite amount of time must be allowed to read; this read time, t_r, is determined by the external read/write circuitry and the computing system requirement. Equation (55) gives an upper limit of how "often" write-in is permitted in a cell. However, in actual operation when read time, t_r, and clearance time for changing address is allowed, the write rate is much less than that shown in equation (55).

D. CELL POWER DISSIPATION

The cell in Figure 22 is redrawn below with nodal capacitors C_Q and $C_{\bar{Q}}$ shown:

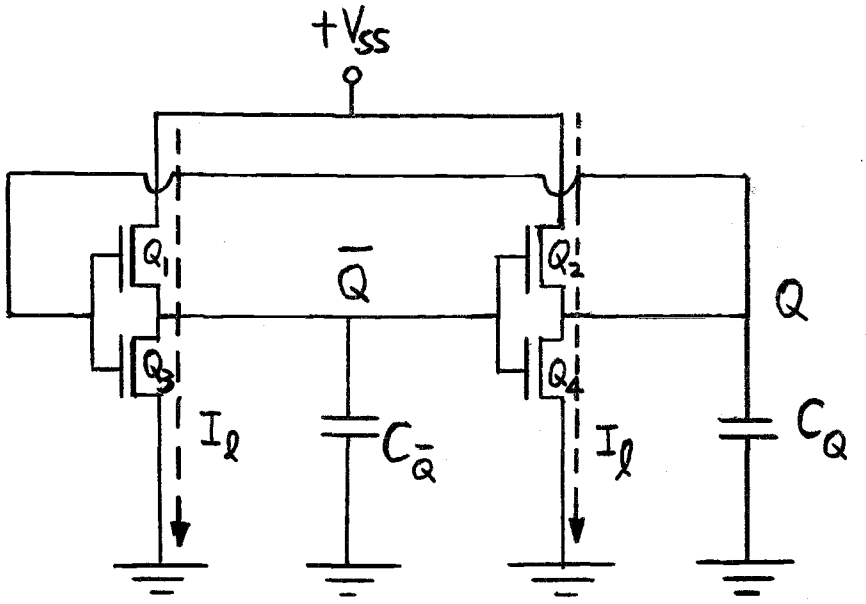


Figure 31. CMOS storage cell with nodal capacitors

There are two kinds of power dissipation within the cell. The stand-by d.c. power dissipation is due to leakage current flow (I_l) from + V_{ss} to ground. In either stable state "1" or "0", between + V_{ss} and ground there is always an off device (refer to Table 4) which permits only leakage current flow. The amount of this leakage current depends on the reverse biased drain channel junction of the device. The d.c. power dissipation for the cell is then:

$$P_{d.c.} = 2\ V_{ss}\ I_l \tag{56}$$

For V_{ss} = 20 volts, $I_l = 10^{-7}$ ampere, P_{dc} is only four micro-watt per cell at stand-by, which is quite low.

The dynamic a.c. power dissipation for any inverter has been studied by Burns (1) as:

$$P_{\text{inverter a.c.}} \approx C V_{ss}^2 f \tag{57}$$

where C is the output capacitor. Since the cell in Figure 31 can be considered as two inverters in cascade, its dynamic a.c. power dissipation is then twice that of an inverter:

$$P_{\text{a.c.}} = 2 C V_{ss}^2 f \tag{58}$$

where $C = C_Q = C_{\bar{Q}}$, f is the operating frequency. For $C = 15 \times 10^{-12}$ farad, V_{ss} = 20 volts, P_{ac} is 12 milli-watt at one mega-hertz.

IV. DN-MOS IN CMOS RAM CELL

A. MONOLITHIC STRUCTURAL CROSS-SECTION

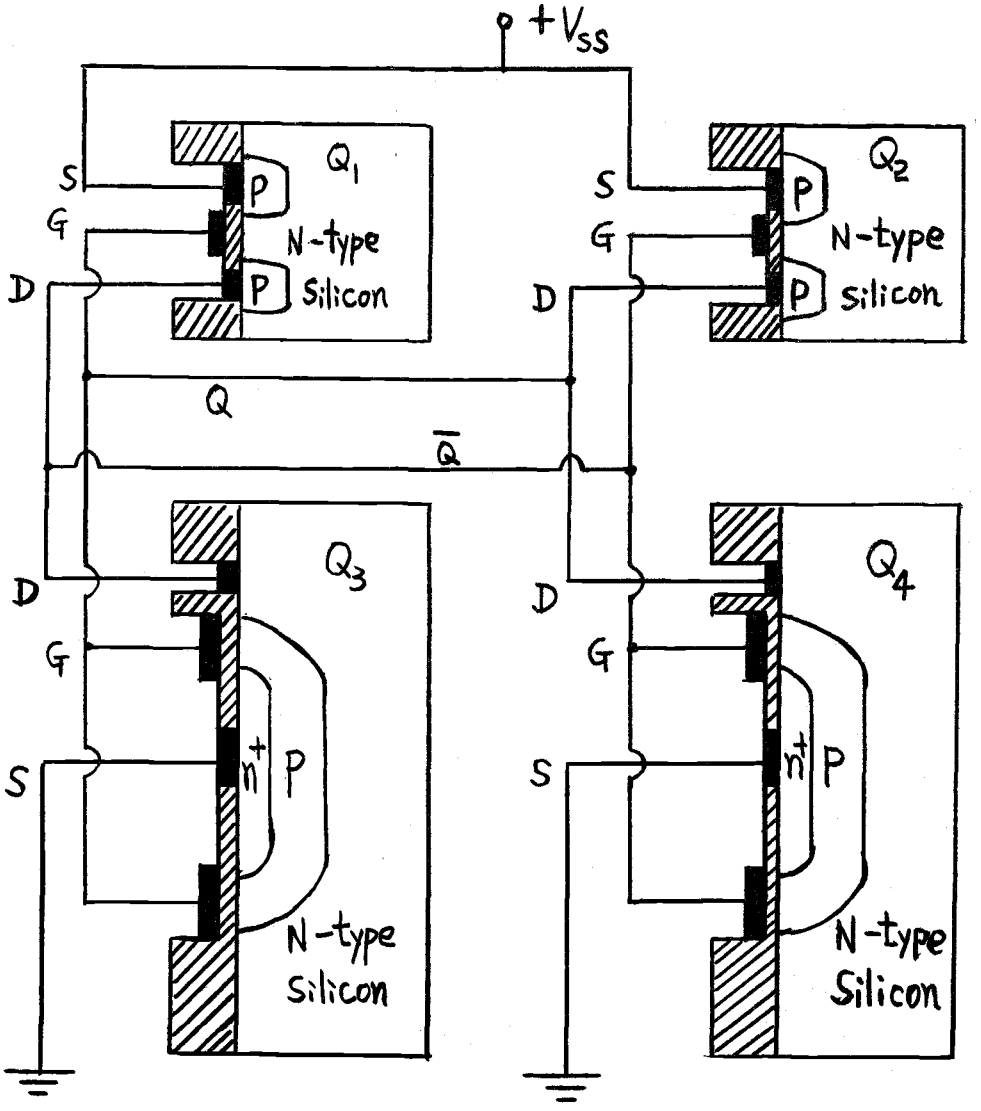


Figure 32. Monolithic storage cell. (Isolation not shown)

Figure 32 shows two DN-MOS transistors, Q_3 and Q_4, being used in the implementation of the storage cell shown in Figure 22. The drain of Q_3 and Q_4 are of different voltage when the cell is in operation, therefore isolation diffusion is needed around the DN-MOS drain region. The isolating p-region is diffused through the n-layer until it reaches the p-type substrate.

Usually the breakdown voltage of n^+-p junction of the DN-MOS transistor is not high enough to stand the + V_{ss} supply. If low threshold voltage of both p-channel and n-channel MOS transistor can be fabricated through ion-implantation process (4), then the cell can operate on low V_{ss}; n^+-region can be used as drain.

When the cell operates on low V_{ss}, the d.c. and dynamic a.c. power dissipation can be saved drastically. This can be seen from equations (56) and (58).

When n^+-region of the DN-MOS transistor is used as drain, the nodal capacitance between drain and ground will be the reverse biased n^+-p junction capacitor instead of the larger reverse biased p-n layer junction capacitor (see Figure 32). The fall time T_F and pair delay time t_D both become shorter (see equations 51, 52, 53), resulting in faster switching cell; the a.c. dynamic power is saved too (see equation 58).

transistors.

Another advantage is that there is no special process involved in obtaining the goal of faster switching and/or smaller size cell. Standard process sequence arrangement makes this method attractive to industrial practice.

C. HYBRID REALIZATION

A hybrid cell is built using devices as stated in section II C 2. Figure 33 below shows the hybrid cell.

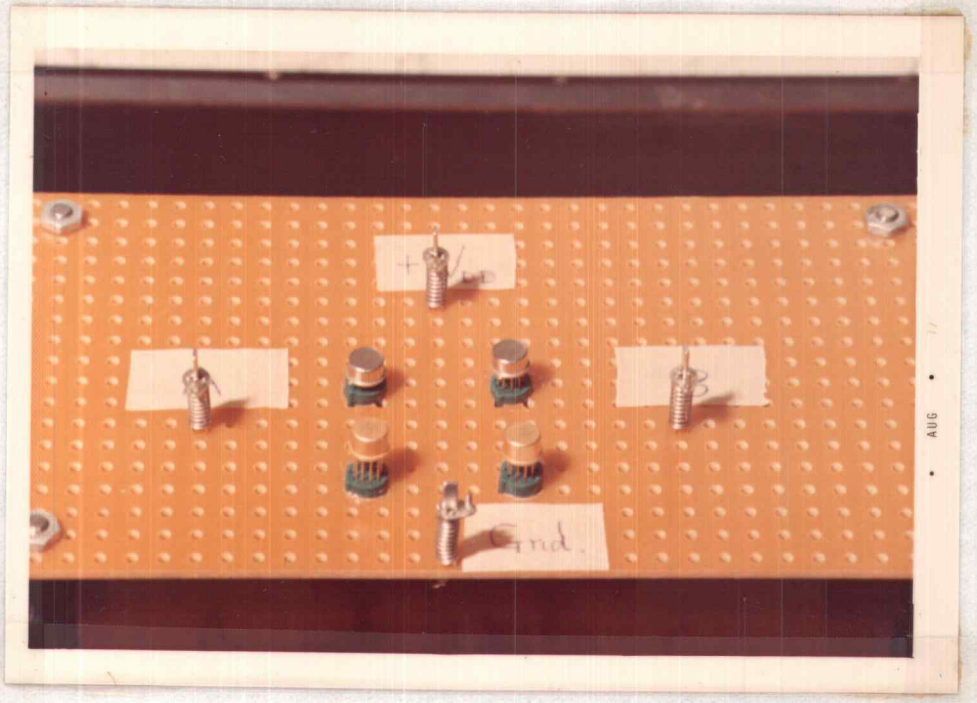


Figure 33. Hybrid cell.

In each TO-5 capsule shown in Figure 33 there is a die like that shown in Figure 8.

The cell switching response is tested using the schematic shown in Figure 34.

B. ADVANTAGES OF DN-MOS TRANSISTOR IN CMOS RAM CELL

Rearranging equation (52) with the aid of equations (51) and (35), the fall time T_F can be expressed in terms of biasing condition and device physical parameters as:

$$T_F = \left(\frac{C}{V_{ss}}\right)\left(\frac{2\,t_{ox}}{\mu_e \epsilon_{ox}}\right)\left(\frac{L_n}{W_n}\right)\left[\frac{\alpha_n - 0.1}{(1-\alpha_n)^2} + \frac{\tanh^{-1}\left(\frac{0.9-\alpha_n}{1-\alpha_n}\right)}{1-\alpha_n}\right] \qquad (59)$$

One advantage can be seen clearly from the fact that equation (59) shows that T_F is directly proportional to L_n -- the n-type channel MOS transistor channel length. Using the DN-MOS transistor, L_n can be made close to 1 μM (10^{-6} meter) as compared to the regular photolithographic limit of 5 μM. As a result, T_F is made five times shorter by introducing the DN-MOS transistor in the CMOS RAM cell, assuming the other parameters in equation (59) remain unchanged.

T_F is a determining factor of how long it takes to write "1" or "0" into the memory cell. A faster cell needs larger W_n/L_n ratio for the n-type MOS transistors in that cell, which usually means larger silicon area on the chip. The area saving nature of incorporating DN-MOS transistors in CMOS RAM cell can be explained from the example that in order to have the same W_n/L_n ratio for the n-type MOS transistor, using the DN-MOS transistor needs only (1/5) W_n as compared to use of the conventional n-type MOS

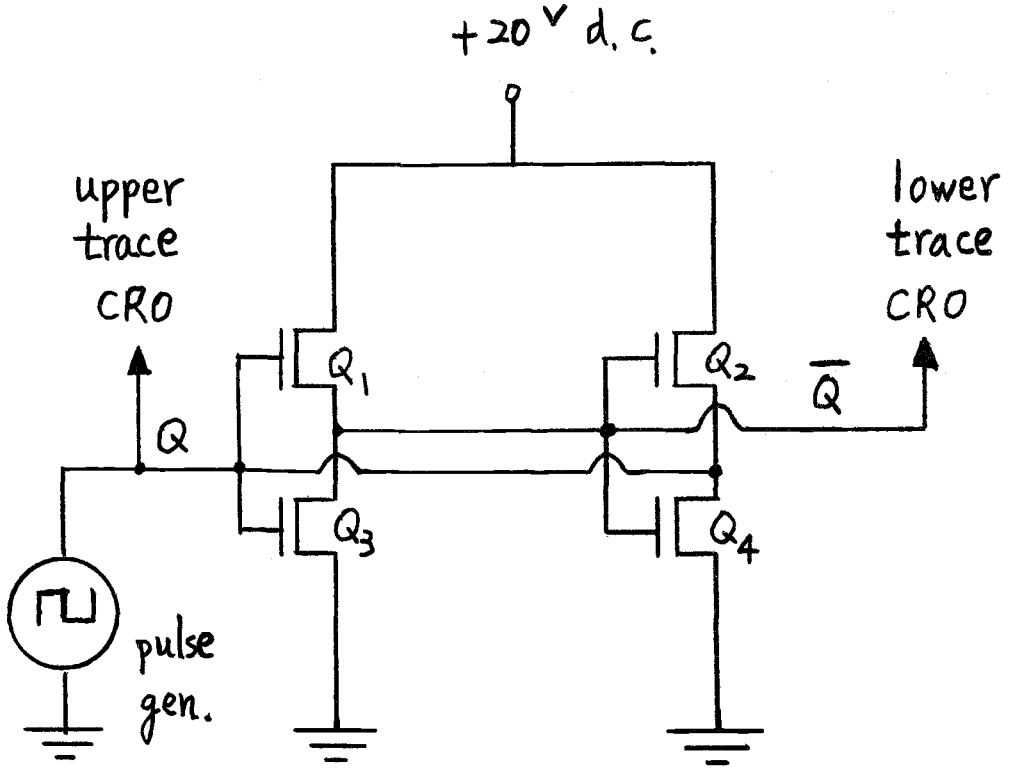


Figure 34. Cell switching test schematic.

Rutherford B 7 type pulse generator is used as source V_Q. Tektronix type 516 dual trace oscilloscope is used to trace both V_Q and $V_{\overline{Q}}$. Results are shown in Figures 35 and 36.

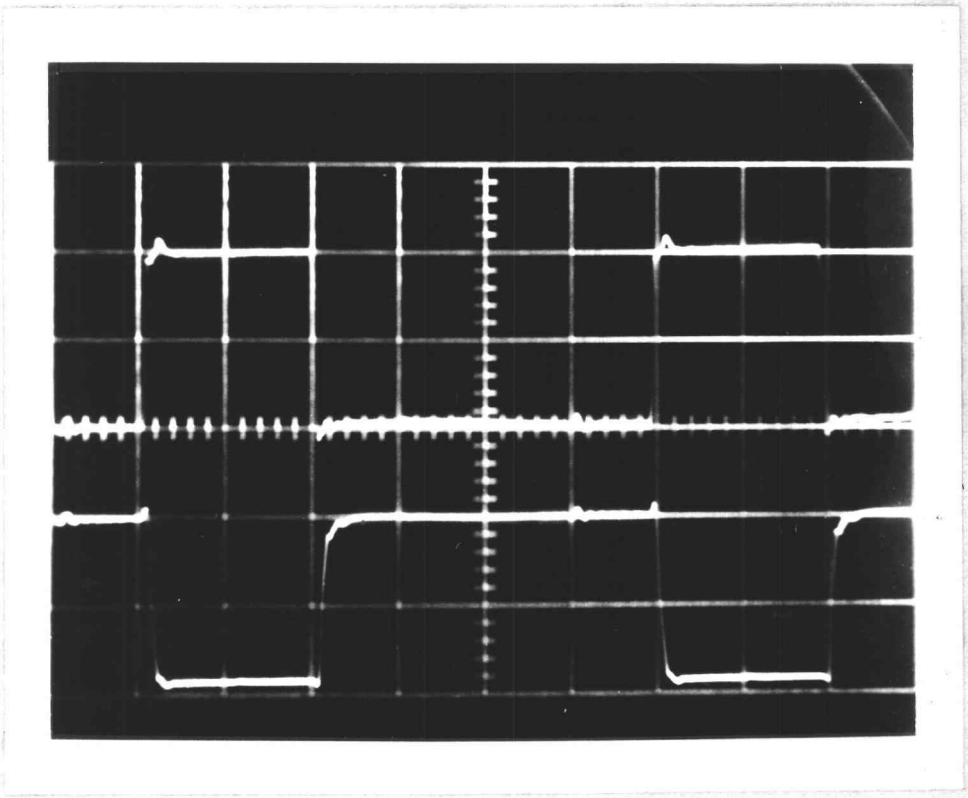


Figure 35. Cell switching.

Upper trace V_Q, lower trace $V_{\bar{Q}}$

Vertical: ten volts per major division

Horizontal: 500 ns per major division

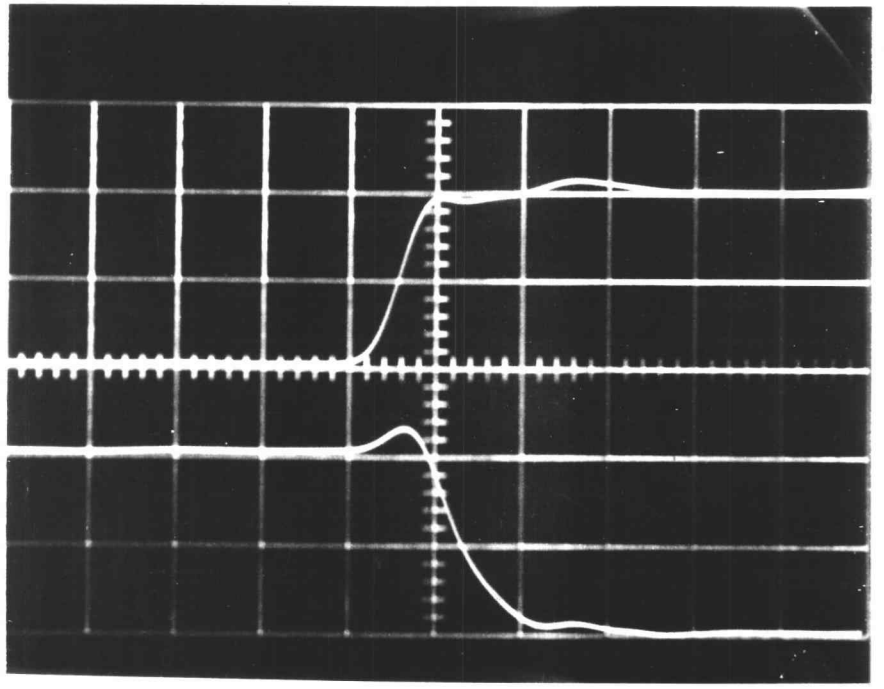


Figure 36. Enlarged cell switching.

Upper trace V_Q, lower trace $V_{\bar{Q}}$

Vertical: ten volts per major division

Horizontal: 40 ns per major division

As indicated in equation (59), the fall time T_F is strongly dependent on the value of C, the nodal capacitance at node $\bar{Q}$. The actual intrinsic C of the cell is of the same order as the stray capacitance in the setup shown in Figure 34 (10 to 20 x 10^{-12} farad), so the fall time indicated in Figure 36 is not that of the cell alone.* In order to clarify equation (59), a large capacitor of 1,000 x 10^{-12} farad is deliberately used to simulate C, so that stray capacitance can be neglected. This leads to the following T_F testing schematic:

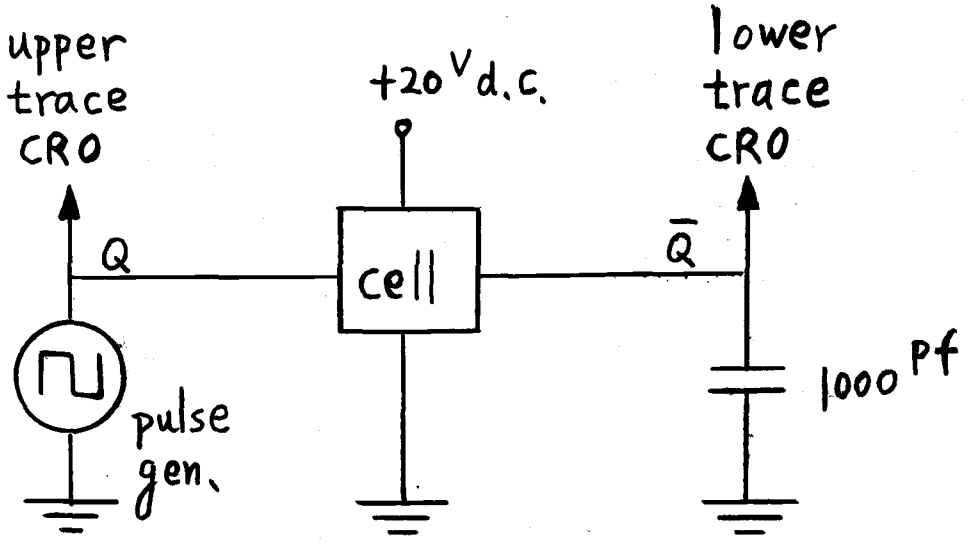


Figure 37. T_F testing circuit.

*Tektronix 516 oscilloscope has 23 x 10^{-9} second rise time which contributes part of the fall time also.

Figure 38 below gives the fall time result, which is 3.6 micro-second. Notice that at this time base the pulse generator is very close to an ideal step generator, which is what equation (59) is based upon. This is another reason for using the large output capacitor in the measurement of T_F.

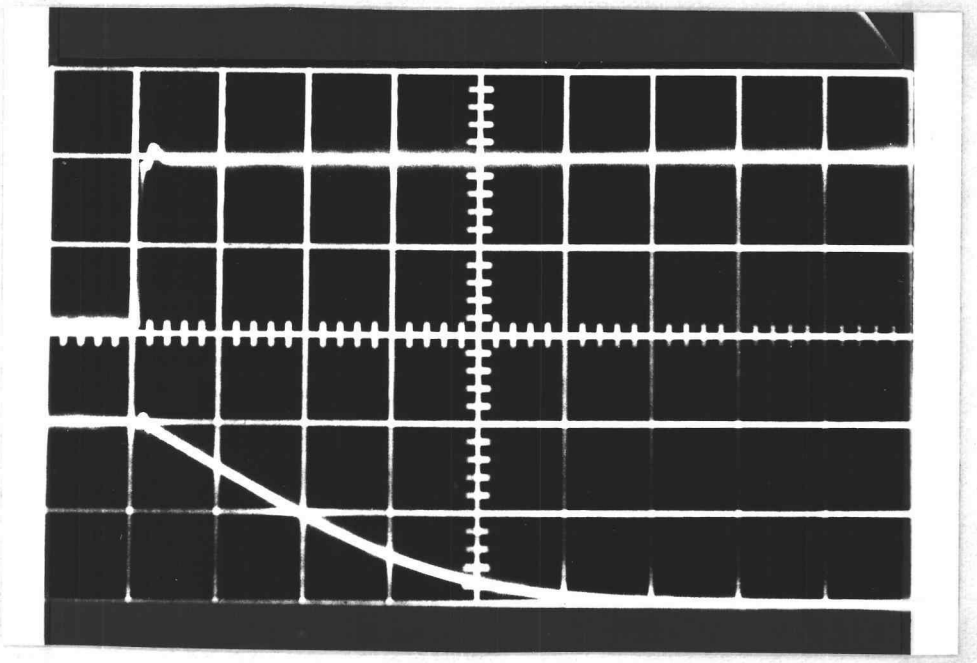


Figure 38. Fall time response.

Vertical: ten volts per major division

Horizontal: one micro-second per major division

T_F = 3.6 micro-second

The calculated fall time T_F is:

$$T_F = \frac{(10^{-9})}{(2.97\times10^{-5})20}\left[\frac{0.2-0.1}{(0.8)^2} + \frac{\tanh^{-1}\frac{0.9-0.2}{1-0.2}}{1-0.2}\right]$$

$$= 3.11 \times 10^{-6} \text{ sec} \qquad (60)$$

where equations (52) and (51) were used and:

$C = 10^{-9}$ farad

$K_N = 2.97 \times 10^{-5}$ amp/v^2

$V_{ss} = 20$ volts

$\alpha_n = \frac{4}{20} = 0.2$

The calculated value is quite close to the measured value of 3.6 micro-second.

The lower trace of Figure 38 is exactly what has been predicted in Figure 29: a linear decay first, followed by a monotonic decay.

The difference between calculated and measured value of fall time is due to the fact that K_N has been assumed constant, which is true for first order approximation of MOS equations, but the actual device behaves differently. As source drain voltage varies, K_N varies also, which is not accounted for during the previous calculation. Equa-(59) gives a very good estimation of what fall time should be, but not the exact value.

D. MONOLITHIC REALIZATION

A set of masks for monolithic CMOS RAM cell was fabricated. The composite of this set of masks is shown in Figure 39:

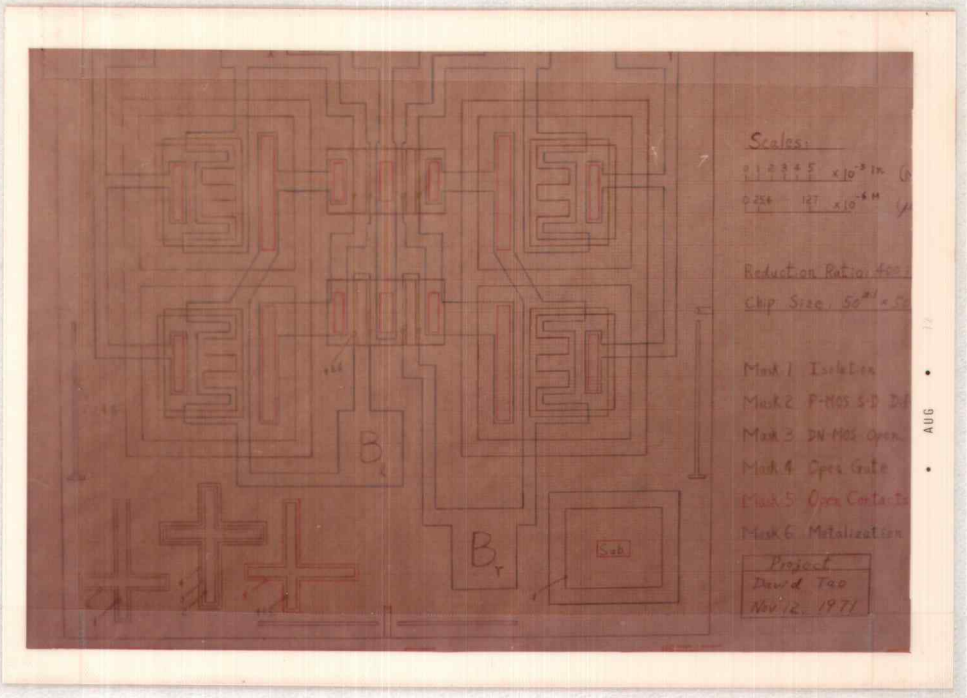


Figure 39. CMOS RAM cell composite.

The fabrication sequence is essentially the same as that stated in section II C 1 for discrete devices except that a long isolation diffusion of 48 hours drive-in time (see section II C 3c) is performed before p-MOS source and drain diffusion to isolate drain regions of two different DN-MOS transistors in the CMOS RAM cell. From Figure 17 it is clear that 48 hours drive-in will diffuse through an n-type layer of 8×10^{-6} meter thickness.

While planning the mask composite, a distance of greater than 25 μM (25×10^{-6} meter) between isolation edge and DN-MOS opening is allowed, because the lateral diffusion and depletion region may extend 13 μM from isolation edge and 7 μM from DN-MOS opening edge (see Figure 19); 25 μM will guarantee there is no accidental punch through.

The finished chip photomicrograph is shown in Figure 40:

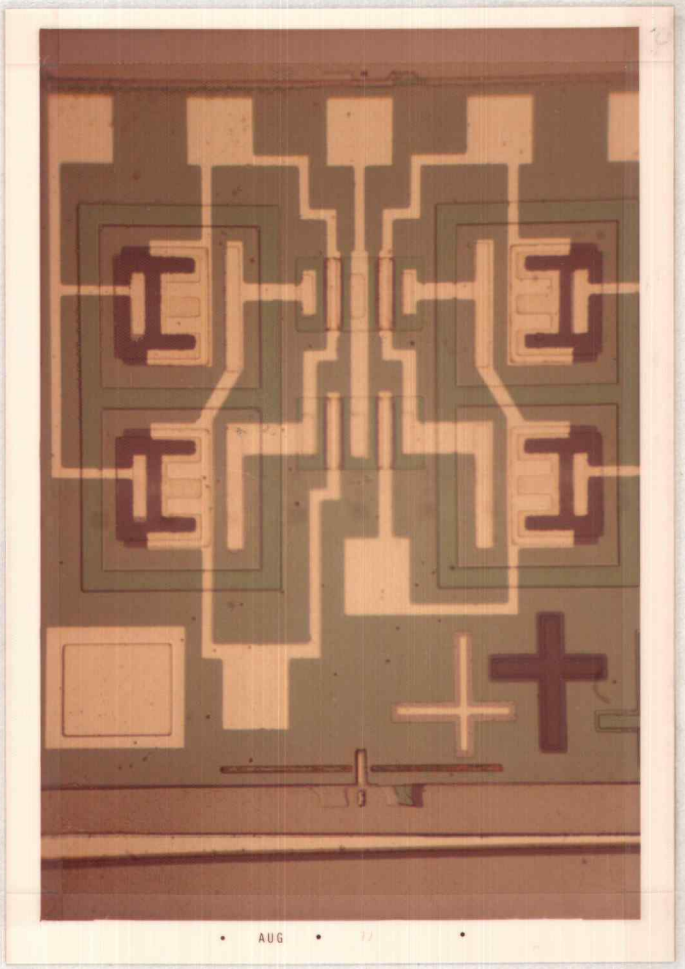


Figure 40. Monolithic CMOS RAM cell photomicrograph.

The chip is size 50 mil x 50 mil (1 mil = 10^{-3} inch). There are two cells per chip; the cell area is about 20 mil x 20 mil (or 400 mil^2 per cell). Most of the areas in Figure 40 are occupied by the isolation and the required spacing around the isolation region. Isolation comes with the monolithic complementary structure, while there is no isolation needed in the other mono-channel structure. The complementary structure has a very low stand-by d.c. power, but its monolithic realization utilizes more silicon area than mono-channel structure. This can be thought of as a trade off between density of function and power dissipation.

The yield of the monolithic CMOS RAM cell was not high; however, the good ones have about the same switching characteristics as its hybrid counterpart.

V. CONCLUSION

The purpose of this study was to investigate the use of the DN-MOS transistors in the CMOS RAM cell.

In the pursuit of this investigative goal, discrete DN-MOS transistors and conventional p-MOS transistors were fabricated on the same chip in order to pin down a proper fabrication sequence. The effect of the fabricating sequence on device threshold voltage and channel length were studied. The CMOS RAM cell operating characteristics as related to the characteristics of devices in the cell were investigated.

Advantages are obtained by using DN-MOS transistors in a CMOS RAM cell. They are faster switching speed for the same silicon area and no special process is needed in realization. A hybrid CMOS RAM cell was built first to justify some of the theoretical predictions; it was found that the fall time measured is close to that predicted. Finally, a monolithic CMOS RAM cell with DN-MOS transistors in it was realized. The switching characteristic of this cell is the same as the previous hybrid one.

For future investigation, it is suggested that ion-implantation process be used to control sheet resistance R_s, and as a result low reproducible threshold voltage devices can be obtained (see section II C 3c about R_s effect on V_{th}). The cell can then operate on low d.c.

voltage supply; both d.c. stand-by and a.c. dynamic power dissipations are saved appreciably. A high switching speed, high density, low power dissipation CMOS RAM cell with DN-MOS transistors is thus realized.

BIBLIOGRAPHY

1. Burns, J. R. Dec. 1964. Switching response of complementary-symmetry MOS transistor logic circuits. R.C.A. Review. p. 627-661.

2. Chou, P. C. 1972. Doctoral dissertation. Oregon State University. 221 p.

3. Cobbold, Richard S. C. 1970. Theory and application of field-effect transistors. Wiley-Interscience. 534 p.

4. Coppen, P. J. et al. Feb. 1972. A complementary MOS 1.2 volt watch circuit using ion-implantation. Solid State Electronics. p. 165-175.

5. Crawford, Robert H. 1967. MOSFET in circuit design. McGraw-Hill. 136 p.

6. Deal, B. E., M. Sklar, A. S. Grove and E. H. Snow. March 1967. Characteristics of the surface-state charge (Q_{ss}) of thermally oxidized silicon. Journal of Electrochem Society. p. 266-274.

7. Deal, B. E. et al. 1966. Barrier energies in metal-silicon dioxide-silicon structures. J. Phys. Chem. Solids. p. 1873-1879.

8. Dwight, Herbert Bristol. 1961. Tables of integrals and other mathematical data. Macmillan Co., 4th ed. 336 p.

9. Feller, A. 1969. Computer analysis and simulation of MOS circuits. I.E.E.E. International Solid-State Circuit Conference, Digest of Technical Papers. p. 134-135.

10. Friedrich, Joseph H. Sept. 1968. A coincident-select MOS storage array. I.E.E.E. Journal of Solid State Circuits. p. 280-285.

11. Frohman-Bentchkowsky, D. and A. S. Grove. Jan. 1969. Conductance of MOS transistors in saturation. I.E.E.E. Transaction on Electron Devices. p. 108-113.

12. Grove, A. S. 1967. Physics and technology of semiconductor devices. John Wiley and Sons. 366 p.

13. Kennedy, D. P. and R. R. O'Brien. May 1965. Analysis of the impurity atom distribution near the diffusion mask for a planar p-n junction. I.B.M. Journal of Research and Development. p. 179-186.

14. Lawrence, H. and R. M. Warner, Jr. Diffused junction depletion layer calculations. Bell System Technical Publications Monograph 3517. 51 p.

15. Lin, Hung Chang et al. Nov. 1969. Complementary MOS-bipolar transistor structure. I.E.E.E. Transaction on Electron Devices. p. 945-951.

16. Lindmayer, Joseph and Charles Y. Wrigley. 1965. Fundamentals of semiconductor devices. D. Van Nostrand Co. 486 p.

17. MOS random-access 256-bit memory. April 1969. Texas Instrument Application Report, Bulletin CA-127. 8 p.

18. Peirce, B. O. 1929. A short table of integrals. Boston, Ginn & Co. 3rd ed. p. 116-120.

19. Sigg, Hans J. et al. Jan. 1972. D-MOS transistor for microwave applications. I.E.E.E. Transaction on Electron Devices. p. 45-53.

20. Sze, S. M. and G. Gibbons. Sept. 1966. Effect of junction curvature on breakdown voltage in semiconductors. Solid State Electronics. p. 831-845.

21. Wallmark, J. Torkel and Harwick Johnson. (Eds.) 1966. Field-effect transistors, physics, technology and applications. Prentice Hall. 376 p. (Chap. 12)

22. Warner, Raymond M., Jr. (Ed.) 1965. Integrated circuit design principles and fabrication. Motorola Semiconductor Product Division Staff. McGraw-Hill. 385 p.

APPENDICES

APPENDIX I

DESCRIPTION OF PROCESS STEPS

a. FIELD OXIDATION

Five ohm-cm n-type silicon of (100)* orientation is used as the starting material. Five ohm-cm corresponds to 10^{15} per cm^3 doping level. The choice of this doping level and (100) orientation on threshold voltage will be discussed in a later section. Silicon wafer is precleaned through the following sequence.

Trichloro-ethylene (TCE) rinse, acetone rinse, deionized (DI) water rinse, nitrogen (N_2) blow dry, 30 seconds buffered hydrofluoric acid (Buff. HF)** etch, DI rinse, N_2 dry.

The purpose of this sequence is to get rid of any grease or silicon dioxide that might be on the starting wafer surface. The oxidation cycle is three hours in wet oxygen atmosphere,*** followed by one hour dry oxygen atmosphere all in 1100°C oxidation furnace. Field oxide thickness at this stage is about 12,000 Å (1.2×10^{-6} meter), which is pink in color under microscope. The oxide is thick enough so that after successive cleaning etches in steps b and c it will still be able to protect field silicon from diffusion and yield a high field

*Miller index for crystal orientation

**Buff. HF : 1 part (48%) HF to 4 parts saturated NH_4F aqueous solution by volume at room temperature

***Wet oxygen : oxygen passing through hot deionized water at 95°C

threshold voltage.*

High field threshold voltage is needed to prevent accidental onset between any two diffused pockets, as shown in Figure A below. If field oxide is ten times thicker than gate oxide, it takes approximately 10 V_{th} to turn on the unwanted MOS structure in Figure A. Such high voltage usually is prohibited from the chip.

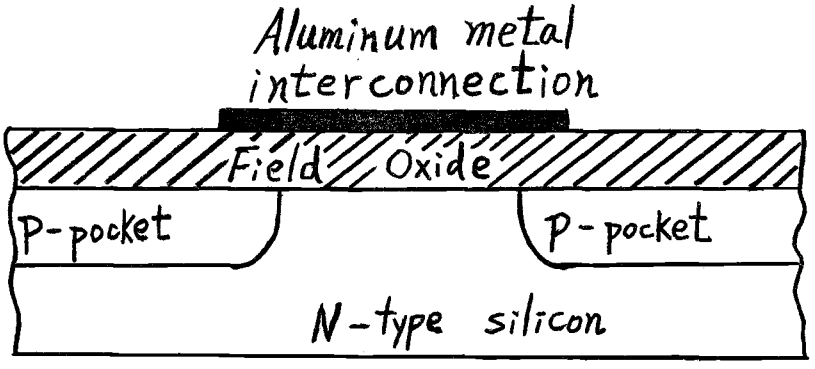


Figure A. Accidental MOS structure between two diffused pockets.

*Refer to equations (4) and (5). V_{th} approximately proportional to t_{ox} for large t_{ox}

b. P-MOS SOURCE AND DRAIN DIFFUSION

Standard photo-resist technique is used to open source and drain windows for the p-MOS transistor. Etching time inside buffered HF is ten minutes, followed by DI rinse, photo-resist removal using acetone, DI rinse again, then nitrogen blow dry.

Predeposition is 20 minutes inside 950°C furnace with boron nitride solid wafer as impurity source and nitrogen atmosphere. It was found that the residue on the surface of the wafer can be cleaned by boiling in DI water for ten minutes, followed by 30 seconds etch in buffered HF. The sheet resistance after predeposition is measured to be 150 ohms per square. This value is obtained from the diffused resistor pattern which has 20 squares; a resistance of 3,000 ohms gives the mentioned sheet resistance.

Drive-in diffusion is done in 1100°C furnace; the first three hours in wet oxygen atmosphere, then one hour in dry oxygen atmosphere. Total drive-in time at this stage is four hours. During this drive-in period some oxide is grown on top of the diffused p-MOS source and drain area, which will prevent the same area from being diffused during step (c), which contains the DN-MOS p and n^+ diffusion. The first three hours in wet oxygen produces the major portion of the needed protection oxide thickness.

In the following steps the diffused region experiences

the additional heat cycle as: 20 minutes at 950°C, five hours at 1100°C, one hour at 1000°C, 15 minutes at 1100°C, 20 minutes at 400°C. Together with the four hours at 1100°C in this step, the calculated total $D_p t$ is 0.95 $(\mu M)^2$. For practical calculation purpose later, we use ten hours at 1100°C as effective $D_p t$ product.

c. DN-MOS OPENING, P-DIFFUSION FOLLOWED BY N^+-DIFFUSION

Opening is done by standard photo-resist technique and ten minutes etch in buffered HF.

P-diffusion predeposition is the same as in step (b). Boron nitride is the source, post cleaning is done through ten minutes boiling in DI water followed by 30 seconds buffered HF etch. Resulting sheet resistance is 150 ohms per square.

P-diffusion drive-in is five hours at 1100°C with nitrogen atmosphere. Post cleaning is a 20 seconds etch in buffered HF.

N^+-diffusion is done through the same window opening as p-diffusion. It is a one step diffusion, which has constant surface concentration all the time, the impurity distribution is complementary error function. A commercial grade of phosphorus film by Emulsitone Co. (P.O. Box 99, Livingston, N.J.) is used as source. The film is diluted with DI water one to one by volume at room temperature, so

that when it is applied on the wafer it will spin evenly on the regular photo-resist spinner. Diffusion time is one hour at 1000°C with nitrogen atmosphere. Residue is cleaned for ten minutes in boiling DI water and 20 seconds buffered HF etch.

By the same token as in step (b), the drive-in time for p-region is six hours for practical calculation purpose, and the diffusion time for N^+-region is one hour.

d. GATE OPENING FOLLOWED BY GATE OXIDATION

Standard photo-resist technique and eight minutes buffered HF etch opens the gate area. The gate area at this stage may have some contamination ions. A five minutes immersion in hot (80°C) diluted sulfuric acid (reagent grade, one to one by volume aqueous solution) followed by 20 seconds etch in fresh clean buffered HF is used to remove some of the contamination. It is believed that hot sulfuric acid immersion will grow a thin layer of oxide which contains the unwanted ions, then the contaminated oxide is etched away from the gate area.

Gate oxidation is done in 15 minutes at 1100°C, first with five minutes in wet oxygen atmosphere then ten minutes in nitrogen atmosphere. Gate oxide thickness is 1200 Å (1.2×10^{-7} meter) which has blue color under microscope.

e. OPEN CONTACTS TO SOURCE AND DRAIN, METALIZATION AND EXCESS ALUMINUM REMOVAL

Standard photo-resist technique and eight minutes buffered HF etch open the source and drain area for metalization. Aluminum is evaporated in a vacuum evaporator to form the required contacts to source and drain, gate metal, and bonding pads. After evaporation the wafer is annealed for 20 minutes at 400°C to improve adherence between aluminum and silicon or silicon dioxide in a nitrogen atmosphere. Excess aluminum is removed by etch in hot (70°C) aluminum etch* for 40 seconds with photo-resist defining the pattern. During the etching process, bubbles keep on generating from the excess aluminum area; when the generation stops, it signals the completion of the removal process. Rinse the sample in running DI water. Photo-resist residue can be removed by using acetone with the help of a cotton swab.

*Aluminum etch: 80 parts H_3PO_4 (concentrated 85%), 16 parts DI water, 5 parts HNO_3 (concentrated 70%). All by volume at room temperature.

APPENDIX II

NUMERICAL DATA OF I_{ds} VERSUS V_{gs} FOR P-MOS TRANSISTOR AND DN-MOS TRANSISTOR

Table A. $\sqrt{|I_{ds}|}$ versus $|V_{gs}|$ for P-MOS Transistor

$\lvert -V_{gs} \rvert$ volt	$\lvert -I_{ds} \rvert$ amp	$\sqrt{\lvert I_{ds} \rvert}$ $(\text{amp})^{\frac{1}{2}}$
20	9.0×10^{-3}	9.49×10^{-2}
18	7.6×10^{-3}	8.70×10^{-2}
16	6.0×10^{-3}	7.75×10^{-2}
14	4.8×10^{-3}	6.92×10^{-2}
12	3.6×10^{-3}	6.00×10^{-2}
10	2.4×10^{-3}	4.90×10^{-2}
9	1.7×10^{-3}	4.12×10^{-2}
8	1.2×10^{-3}	3.46×10^{-2}
7	700×10^{-6}	2.64×10^{-2}
6	350×10^{-6}	1.87×10^{-2}
5	100×10^{-6}	1.00×10^{-2}

Table B. $\sqrt{I_{ds}}$ versus V_{gs} for DN-MOS Transistor

V_{gs} volt	I_{ds} amp	$\sqrt{I_{ds}}$ $(\text{amp})^{\frac{1}{2}}$
20	13.6×10^{-3}	11.65×10^{-2}
18	9.4×10^{-3}	9.70×10^{-2}
16	6.2×10^{-3}	7.86×10^{-2}
14	4.0×10^{-3}	6.31×10^{-2}
12	2.2×10^{-3}	4.69×10^{-2}
10	790×10^{-6}	2.81×10^{-2}
9	480×10^{-6}	2.19×10^{-2}
8	280×10^{-6}	1.67×10^{-2}
7	140×10^{-6}	1.18×10^{-2}
6	60×10^{-6}	7.74×10^{-3}
5	20×10^{-6}	4.46×10^{-3}

APPENDIX III

DN-MOS TRANSISTOR IMPURITY PROFILE

NUMERICAL VALUES

Table C. Numerical Values for the Plot of DN-MOS P-Region Impurity Profile

$$N_A(x) = 6.7 \times 10^{17} e^{\frac{-x^2}{2.16}}$$

$x^{\mu M}$	x^2	$x^2/2.16$	$e^{x^2/2.16}$	$N_A(x)cm^{-3}$
0	0.00	0.0000	1.00	6.70×10^{17}
0.2	0.04	0.0185	1.018	6.60
0.4	0.16	0.074	1.077	6.22
0.6	0.36	0.167	1.180	5.68
0.8	0.64	0.296	1.343	4.99
1.0	1.00	0.463	1.59	4.20
1.2	1.44	0.667	1.95	3.44
1.4	1.96	0.907	2.48	2.70
1.6	2.56	1.185	3.27	2.05
1.8	3.24	1.500	4.48	1.50
2.0	4.00	1.850	6.36	1.05
2.2	4.84	2.240	9.39	7.15×10^{16}
2.4	5.76	2.670	9.65	6.95
2.6	6.76	3.140	23.3	2.88
2.8	7.84	3.620	37.5	1.79
3.0	9.00	4.160	63.5	1.05
3.2	10.24	4.750	115.7	5.82×10^{15}
3.4	11.56	5.350	211.0	3.18
3.6	12.96	6.000	403.0	1.66
3.8	14.44	6.690	825.0	8.13×10^{14}
4.0	16.00	7.400	1656.0	4.05

Table D. Numerical Values for the Plot of DN-MOS n^+-Region Impurity Profile*

$$N_D(x) = 10^{21} \operatorname{erfc} \frac{x}{0.6}$$

$x^{\mu M}$	$\frac{x}{0.6}$	$\operatorname{erf} \frac{x}{0.6}$	$\operatorname{erfc} \frac{x}{0.6}$	$N_D(x)$ cm^{-3}
0.0	0.000	0.00000	1.00000	1.00×10^{21}
0.2	0.333	0.36231	0.63769	6.37×10^{20}
0.4	0.666	0.65374	0.34626	3.46×10^{20}
0.6	1.000	0.84270	0.15730	1.57×10^{19}
0.8	1.333	0.94059	0.05941	5.94×10^{19}
1.0	1.666	0.98153	0.01847	1.84×10^{18}
1.2	2.000	0.99532	0.00468	4.68×10^{17}
1.4	2.333	0.99903	0.00097	9.70×10^{17}
1.6	2.666	0.99984	0.00016	1.60×10^{16}
1.8	3.000	0.99998	0.00002	2.00×10^{16}

*Values of error function from Peirce's table (18).

Table E. DN-MOS Transistor Net Impurity Profile Numerical Values

$x^{\mu M}$	$N_D - N_A$		$(N_D - N_A)$	cm^{-3}
0.0	(10000-6.70)	$\times 10^{17}$	+ 9.99330	$\times 10^{20}$
0.2	(6370-6.60)	$\times 10^{17}$	+ 6.36340	$\times 10^{20}$
0.4	(3460-6.22)	$\times 10^{17}$	+ 3.45378	$\times 10^{20}$
0.6	(1570-5.68)	$\times 10^{17}$	+ 1.56432	$\times 10^{19}$
0.8	(594-4.99)	$\times 10^{17}$	+ 5.8901	$\times 10^{19}$
1.0	(184-4.20)	$\times 10^{17}$	+ 1.7980	$\times 10^{18}$
1.2	(46.8-3.44)	$\times 10^{17}$	+ 4.336	$\times 10^{17}$
1.4	(9.70-2.70)	$\times 10^{17}$	+ 7.00	$\times 10^{16}$
1.6	(1.60-2.05)	$\times 10^{17}$	- 3.5	$\times 10^{17}$
1.8	(2.00-15.0)	$\times 10^{16}$	- 1.3	$\times 10^{17}$
2.8	(1-17.9)	$\times 10^{15}$	- 1.69	$\times 10^{16}$
3.0	(1-10.5)	$\times 10^{15}$	- 9.5	$\times 10^{15}$
3.2	(1-5.82)	$\times 10^{15}$	- 4.82	$\times 10^{15}$
3.4	(1-3.18)	$\times 10^{15}$	- 2.18	$\times 10^{15}$
3.6	(1-1.66)	$\times 10^{15}$	- 6.6	$\times 10^{14}$
3.8	(1-0.813)	$\times 10^{15}$	+ 1.87	$\times 10^{14}$
4.0	(1-0.405)	$\times 10^{15}$	+ 5.94	$\times 10^{14}$

+ : net impurity is n-type donor
- : net impurity is p-type acceptor