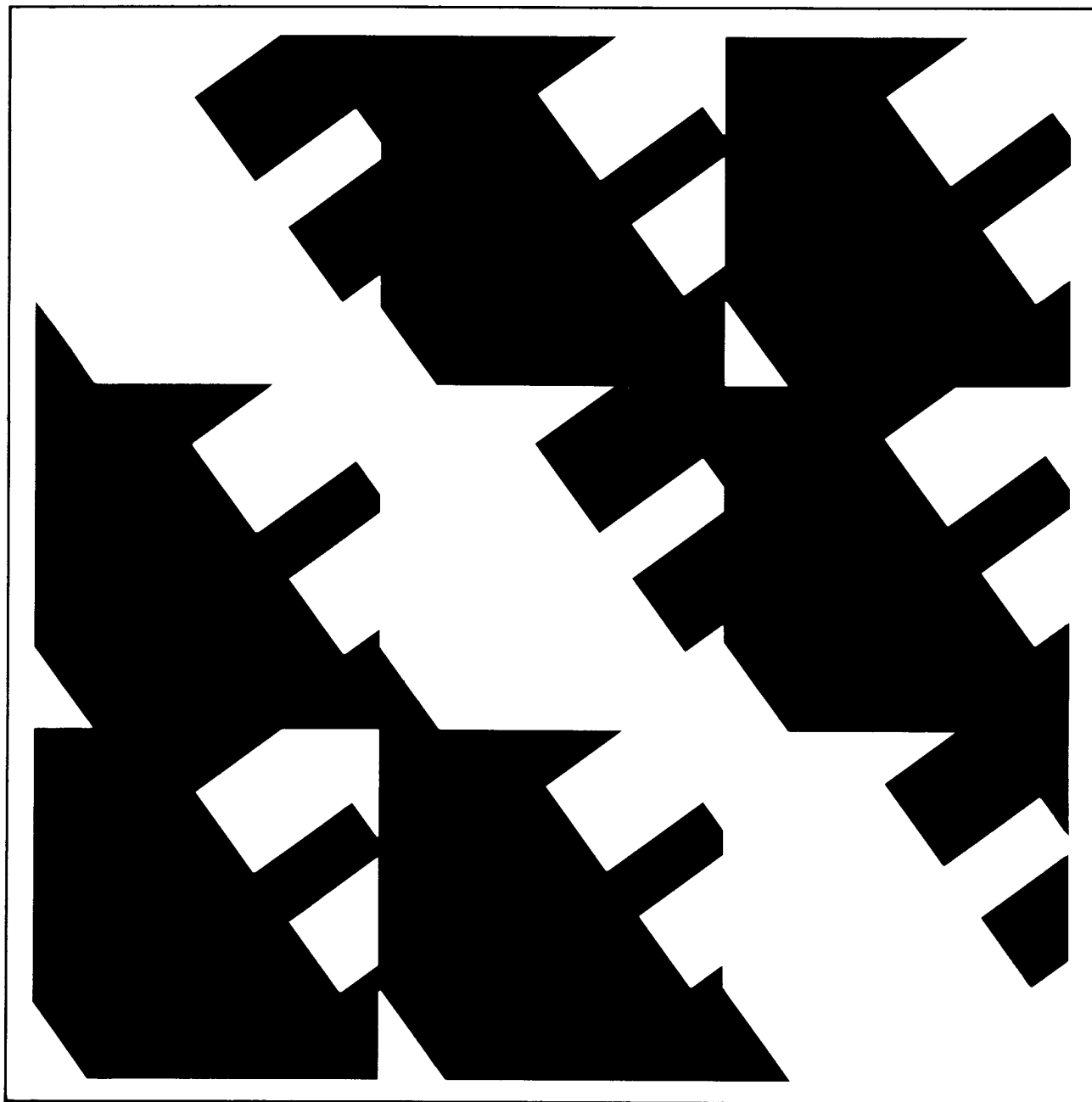


IEEE Standard Definitions and Characterization of Metal Nitride Oxide Semiconductor Arrays



ANSI/IEEE Std 641-1987



Published by The Institute of Electrical and Electronics Engineers, Inc. 345 East 47th Street, New York, NY 10017, USA

October 7, 1988

SH12195

An American National Standard

IEEE Standard Definitions and Characterization of Metal Nitride Oxide Semiconductor Arrays

1. Introduction

1.1 Comparison of MOS and MNOS Transistors.

A MNOS transistor is a type of IGFET similar to the MOS transistor. A basic IGFET or FET is a four-terminal device, namely gate, source, drain, and body in which the current flow from source to drain is modulated by a voltage on the gate with respect to the body. Detailed description of the current-voltage and capacitance-voltage of an IGFET is presented in [14]¹. In the MOSFET the gate insulator is a single layer of silicon dioxide whereas in the MNOS FET the gate insulator consists of a silicon nitride layer in addition to the

¹The numbers in brackets correspond to those of the references listed in Section 3.

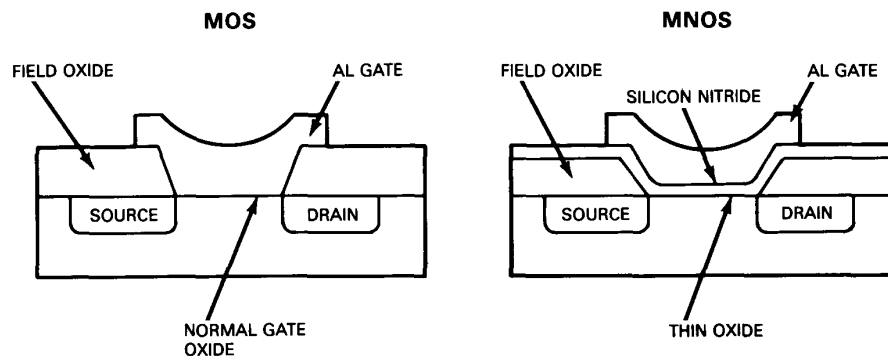
silicon dioxide. Figure 1 compares the structural differences between MNOS and MOS transistors.

The MNOS FET and MOSFET exhibit similar current-voltage characteristics, namely

- (1) A region where the drain current I_d varies quadratically with the drain-to-source voltage V_{ds} at constant gate-to-source voltage V_{gs}
- (2) A saturation region where I_d is very nearly independent of V_{ds} .

The gate-to-source voltage V_{gs} at which a defined minimum I_d is observed is defined as a threshold voltage. This threshold voltage is fixed for a given MOSFET. The threshold voltage at the onset of strong inversion is the sum of a number of terms including contributions from the gate metal-semiconductor work function difference ϕ_{ms} , the surface potential ϕ_s , the charge in the

Fig 1
MOS and MNOS Transistors



depletion region Q_b , and the net effective charge at the insulator-semiconductor interface Q_i , in the gate area. The threshold voltage may be written

$$V_{gs(th)} = \phi_{ms} + \phi_s - (Q_b/C_{gs}) - Q_i/C_{gs}$$

where

$$C_{gs} = \text{gate capacitance}$$

However, for a MNOS FET, when the silicon dioxide layer is thin (1 nm–4 nm), the threshold voltage of the FET can be altered by applying a voltage pulse on the gate with respect to the source, drain, and body. This threshold modification is attributed to the modification of charges trapped in the trapping sites of the silicon nitride gate dielectric at and near the silicon nitride silicon dioxide interface. Therefore, the threshold voltage may be written

$$V_{gs(th)} = \phi_{ms} + \phi_s - (Q_b/C_{gs}) - (Q_i/C_{gs}) - (qt/C_{gs})$$

where

- qt = net effective contribution of the trapped charges in the gate dielectric
- v = indicates time dependence of the MNOS threshold voltage
- q = indicates the time dependence of the trapped charges in the gate dielectric

The amplitude and duration of the voltage pulse required to alter the threshold voltage depend on the physical dimensions of the MNOS FET. When the gate of the MNOS FET with a thin oxide layer is subjected to repeated alternate positive and negative pulses, the device can be set to a more-positive and more-negative threshold voltage, respectively, thus providing a binary state necessary for a digital memory element. The decay rate of the threshold voltage from the high or low state is set by the device process and geometrical and electrical parameters. The decay of the threshold voltage can be made to extend over years even after all source of power to the device is removed. Thus a MNOS FET can be used as a binary, nonvolatile semiconductor memory element. The pulse amplitude and duration required to write and erase a MNOS FET memory element varies from ± 20 v to ± 30 v and 1 μ s to 1 ms for a large variety of these devices.

The major current-voltage characteristics of a MNOS transistor are defined by two equations given in Table I. These equations relate the current I_d flowing between drain and source of the transistor, to the voltage between drain and source

V_{ds} and the voltage applied to the gate V_{gs} . The dependence of I_d on both voltages V_{ds} and V_{gs} only occurs when V_{ds} is smaller than V_{gs} . The equation governing this situation describes the nonsaturation region (or triode region) of the MNOS. When V_{ds} is larger than V_{gs} , the drain current I_d becomes independent of V_{ds} and is governed only by the value of V_{gs} . This condition describes the saturation region of a MNOS. It occurs most commonly when drain and gate leads of the device are tied together.

Table 1
Equations for Simplified MNOS
Drain Current I_d

Nonsaturation Region: $ V_{ds} < V_{gs} - v_{gs(th)}$
$I_d = k' (W/L) [2 (V_{gs} - v_{gs(th)}) V_{ds} - V_{ds}^2]$
Saturation Region: $V_{ds} \geq V_{gs} - v_{gs(th)}$
$I_d = k' (W/L) (V_{gs} - v_{gs(th)})^2$
where
$k' = \mu C_{gs}/2$

The detailed electrical characteristics of a MOSFET and MNOS FET have been reported in IEEE Std 581-1978 [1]. This standard is restricted to the use of a MNOS FET nonvolatile memory element in an IC memory array.

1.2 Description of this Standard. This standard has ten sections.

Section 1 is an introduction to the MNOS device and memory array.

Section 2 defines symbols and definitions used in the remaining sections.

Section 3 provides the reader with references that contain added detail on specific concepts.

Section 4 defines the MNOS arrays and their functional operations. This section serves as the compendium of MNOS arrays that are currently used.

Section 5 defines MNOS array retention. It summarizes the various factors affecting memory retention. Techniques of measurement and data format, and presentation of the array retention properties are defined.

Section 6 defines and describes the MNOS array endurance property. Factors affecting MNOS endurance are summarized and techniques used to

verify array endurance are presented. The relationship between endurance and retention is discussed and a method of presenting array endurance characteristics is described.

Section 7 summarizes reliability considerations for MNOS arrays and identifies failure rates for both noncatastrophic retention failures and end-of-life catastrophic endurance limit failures. Some applicable reliability test measurement techniques are also discussed in this section.

Section 8 defines the testing methodology necessary to establish the unique properties of the MNOS array for both the consumer and the producer. Definition of array parameter failures, minimum identity of test measurements used, and data format requirements are described.

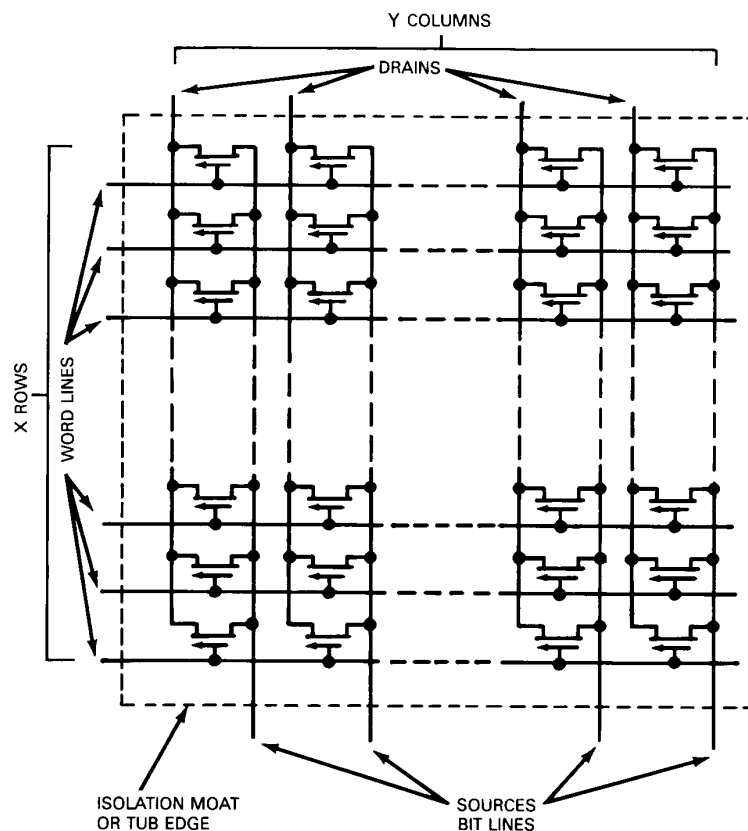
Section 9 deals with radiation effects on MNOS arrays. Measurement techniques and tests for defining hardness are discussed.

Section 10 describes another nonvolatile memory technology, called floating-gate.

1.3 MNOS Array Organization. The undecoded MNOS array is an array of MNOS FET nonvolatile memory elements arranged in columns of source and drain diffusions and rows of gate electrodes (see Fig 2). Except for test purposes, the undecoded array is not used in practice. The rows of gate electrodes are decoded for selection of memory access and similarly the source and drain columns are organized for sensing and bit column selection. In addition, isolation of the MNOS array is needed so that the erase operation can be implemented using the opposite voltage polarity as the read and write operations of the MNOS array.

As a result of continuing design and application innovations, array organizations, organization name designations, and acronyms vary with time

Fig 2
Undecoded MNOS Array X Rows by Y Columns



and usage. In the following paragraphs describing the several commonly available MNOS arrays, names and acronyms are for reference only, and represent common usage to date.

Different array organizations can be achieved depending on the decoding of the gate electrodes. The EAROM is the simplest organization conceptually. An erase operation in which every MNOS element in the array is set to the high-conductance state is first completed. This erase operation is the block-erase operation. Then, individual gate electrode lines (or word lines) are decoded for the write operation. The EAROM is also commonly known as the EEPROM, a designation derived originally from the floating-gate programmable read-only memory technology. In the EEPROM designation, in-situ alterability of stored data is emphasized. This is different from the EPROM erase operation based on ultraviolet irradiation. In the case of the WAROM, voltages to the MNOS array substrate, the individual gate electrode lines, and the diffusion columns are manipulated so that only MNOS transistors sharing a common gate electrode line are erased and written per memory operation.

In a typical MNOS memory organization, data entry (and hence read out) are completed in parallel for each word line. In the BORAM, data is input and output serially. Within this device, there is a serial-to-parallel converter circuit so that in the MNOS array, data entry is efficiently completed in the parallel mode. With typical BORAM organizations, selection of the word line is by way of the decoding of input address bits.

To circumvent the lengthy erase and write times necessary for all MNOS memory operations, a *shadow memory* organization is sometimes used. In this realization the MNOS array is the back-up memory array into which the MOS RAM inputs data by way of a bulk write operation during power down of the IC memory device. With the shadow MNOS memory organization, the memory normally operates at typical MOS RAM speed for read and write, yet achieves data storage non-volatility with MNOS backup storage.

Very often MNOS arrays are identified as either single-element-per-bit storage or two-element-per-bit storage. Single-element-per-bit storage requires a sensing circuit that must differentiate between the low- and high-conductance MNOS thresholds for the two logic states. In the case of the two-element-per-bit storage, the sense amplifier simply compares the threshold states of the two MNOS transistors and sets or resets a read-out latch depending on the comparison of the two thresh-

old voltages. The two-elements-per-bit storage provides more than twice the memory window in addition to added tolerance to process variations from wafer lot to wafer lot. The single-element-per-bit storage has the obvious advantage of increased storage density.

1.4 Restrictions. For illustration purposes this standard assumes p-channel MNOS FET memory elements and fixed-threshold p-channel IGFET for peripheral circuits. The extension of this standard to n-channel technology is straightforward. Even though the memory effect of MNOS FET is of an analog nature, this standard is concerned only with its use in digital applications.

Only those memories that store information as a result of charge tunneled into traps in the dielectric are considered in this standard. Semiconductor nonvolatile memory elements such as the floating-gate avalanche-injection MOS [6] and variations will be addressed by a separate IEEE standard. Section 10 contains a brief summary of floating-gate memory technology. However, many structures representing other variations of device programming but used in memory arrays similar to the arrays used for the MNOS FET can be tested by applying this standard with slight modifications.

The tri-gate [5] (protected source-drain) and the single-gate MNOS FET nonvolatile memory element differ slightly in write-erase characteristics, retention, and endurance. These differences can lead to modifications of the design of the peripheral circuit in a MNOS memory array. Because of the predominant use of the tri-gate structure at this time, this standard focuses on this structure. However, in most cases the discussion is also applicable to the single-gate structure. A more detailed discussion of these two types of MNOS FET is presented in Section 4 of this standard. Thereafter the tendibility of this standard to the single-gate structure is quite clear.

2. Definitions and Symbols

This section describes definitions and symbols used throughout this standard.

2.1 Definitions

address. A character or group of characters that identifies a register, a particular part of storage, or some other data source or destination (see JESD 100-82 [3]).

bit. A unit of information in the binary numeration system. Also a unit of storage capacity of a memory.

bit-retention time. The retention time for one address location.

BORAM. block-organized random-access memory. Also known as block-oriented random-access memory. A memory arrangement that permits random access to blocks of memory cells that are read using serial transmission methods.

NOTE: A block is a singly addressed large number of memory cells (greater than needed for a single computer word operation).

boots trap. A circuit design technique in which a junction point (node) is capacitively driven to a voltage of greater magnitude than that available from the device power supply.

column. A group of memory cells having a common sense amplifier that detects the state of the cell being addressed.

dose rate. The time rate of deposition of radiation expressed in grays in silicon per second (see IEEE Std 581-1978 [1]).

NOTE: The integration of dose rate over a specified period of time represents the total dose.

dynamic radiation test. A test of the instantaneous effects of radiation obtained by monitoring electrical and memory properties of interest continuously, during and immediately after exposure to a transient radiation pulse (see IEEE Std 518-1978 [1]).

EAROM. electrically alterable read-only memory. A synonym for an electrically erasable programmable read-only memory (see JESD 100-82 [3]).

EEPROM. electrically erasable programmable read-only memory. A reprogrammable read-only memory in which cells may be erased electrically and in which each cell may be reprogrammed electrically (see JESD 100-82 [3]).

endurance. The minimum number of data alteration cycles possible without catastrophic failure or degradation beyond the specified performance characteristics of any memory cell within an array.

EPROM. erasable programmable read-only memory. Same as EAROM, except erasure is implemented by exposure to ultraviolet light.

FET. field-effect transistor. A transistor in which the conduction is due entirely to the flow of majority carriers through a conduction channel controlled by an electric field arising from a voltage applied between the gate and source electrodes (see JESD 77-81 [2]).

first-bit access time (of a BORAM). The time interval between the application of addressing and enabling signals and the availability at an output of the first bit from a block of data.

fully decoded EAROM organization. An EAROM organization in which rows and columns of memory cells are addressable through on-chip decoding circuitry.

functional dynamic tests. Operation of the MNOS array at nominal speed to determine functional performance.

grays in silicon. A unit of absorbed dose as measured by its ionizing effect in silicon; 1 gray = 1 joule of energy deposited in a kilogram of irradiated silicon.

NOTE (1): This number can be translated into the density of electron-hole pairs in silicon by the equation $n_{eh} \sim 4 \times 10^{-5} \times \text{grays}$

NOTE (2): One gray in silicon = 100 rd in silicon. However, the unit rads is not an SI unit and therefore is deprecated. 1 rd = 0.01 Gy (See ANSI/IEEE Std 260-1978).

IGFET. insulated-gate field-effect transistor. A field-effect transistor having one or more gate electrodes that are electrically insulated from the channel (see JESD 77-81 [2]).

incremental total dose test. A test of the permanent changes induced by radiation obtained by a comparison of characteristics before and after exposure to successively higher increments of integrated total dose levels.

in-situ total dose test. A test of the instantaneous effects of an environment containing radiation obtained by monitoring the device characteristics during exposure to a relatively low, constant dose rate.

internal reference voltage. An internally developed bias voltage used as an applied gate voltage in the read mode.

memory capacity. The maximum number of bits that a memory is capable of storing.

memory cell. The smallest subdivision of a memory into which a unit of data has been or can be entered, in which it is or can be stored, and from which it can be retrieved (see JESD 100-82 [3]).

memory organization. An arrangement of memory cells, either by geometrical arrangement in rows and columns, or by the organization of the data to be stored.

memory window. The difference in threshold voltage between the low- and high-conductance logic states of a memory cell.

MNOS transistor. The acronym derives from the layer sequence in the gate region of the IGFET, namely, metal nitride oxide semiconductor. Usually it has a variable threshold voltage, but some devices with this layer sequence have fixed threshold voltages (see IEEE Std 581-1978 [1]).

MNOS memory section. The portion of an integrated circuit memory built using a MNOS technology.

MOS transistor. A type of IGFET, referring specifically to the layer sequence in the gate region of the IGFET, namely, metal oxide semiconductor (see IEEE Std 581-1978 [1]).

nonvolatile memory. A memory in which the data content is retained when power is no longer supplied to it (see JESD 100-82 [3]).

pattern insensitivity. The ability to write, store, and read streams of data without regard to the sequence of the data and the pattern of distribution of ones and zeros across the memory locations.

power-on retention time. The retention time with the memory biased in the *on* condition in a non-operating (that is, unlocked, deselected) mode.

radiation-induced data-loss characteristics. Characteristics that define stored data integrity as a function of total dose or dose rate after a given test pattern has been written into the device.

RAM. random-access memory. A memory that permits access to any of its address locations in

any desired sequence with similar access time to each location (see JESD 100-82 [3]).

NOTE: The term RAM, as commonly used, denotes a read/write memory.

read disturb cycles. The number of consecutive read cycles that occur before a memory state becomes indistinguishable, due solely to reading.

retention failure. The inability to correctly sense the state of a memory cell within the limits of device specifications dependent on the time period of data storage.

retention longevity. The time elapsed between the instant of writing a data pattern into a memory and the time when the read failure rate exceeds some predetermined value.

NOTE: This definition allows for *soft* errors and fits reliability prediction and specification methods.

retention time. The time interval between the instant of writing a memory pattern into a memory and the first retention failure.

retention time at maximum read rate. The retention time using the maximum specified read rate on a single address.

row. A group of memory cells having a common internal address line.

single-transistor cell. A memory cell consisting of one transistor.

sense amplifier voltage window. The minimum difference in threshold voltage required by a sense amplifier to differentiate between the low- and high-conductance MNOS thresholds for the two logic states.

standby power. The power consumption while the chip is not performing any read or write operation.

static radiation test. A test of the permanent changes induced by radiation obtained by a comparison of characteristics before and after exposure (see IEEE Std 581-1978 [1]).

two-transistor cell. A memory cell consisting of two MNOS transistors.

unpowered retention time. The retention time with all terminals of the memory at ground vol-

tage, except for an occasional read cycle to test the device state.

WAROM. word-alterable read-only memory. A memory that permits erasing and writing to memory cells constituting a computer word without disturbing any other word.

worst-case retention failure. The change of state of any memory cell from the last-written state, even if the state later returns to that last-written state.

worst-case retention time. The time interval between the instant of writing a memory pattern into a memory and the first worst-case retention failure.

2.2 Letter Symbols for Quantities

C_{gs} = gate capacitance

k' = gain factor

I_d = drain current

L = length of the channel in the direction of current flow

Q_b = charge density in depletion region

Q_i = charge density in insulator

qt = net effective trapped charges in the gate dielectric

V_{ds} = drain-to-source voltage

V_{gs} = gate-to-source voltage

$V_{gs(th)}$ = threshold voltage (fixed)

$v_{gs(th)}$ = MNOS transistor threshold voltage

W = width of channel

μ = carrier mobility

ϕ_{ms} = contact potential between gate metal and silicon

ϕ_s = surface potential

2.3 Letter Symbols for Units of Measurement

μs = microsecond

ms = millisecond

nm = nanometer

3. References

- [1] IEEE Std 581-1978, IEEE Standard Definitions, Symbols, and Characterization of Metal-Nitride-Oxide Field-Effect Transistors.²
- [2] JESD 77-81, Letter Symbols, Abbreviations, Terms, and Definitions for Discrete Semiconductor and Optoelectronic Devices.³
- [3] JESD 100-82, Terms, Definitions, and Letter Symbols for Microcomputers and Memory Integrated Circuits.
- [4] BRUCKER, G. J. and RICHARDSON, D. Radiation and Semiconductors: The Basic Parameters. *Electronic Engineering*, Aug 1980, pp 73-75.
- [5] CRICCHI, J. R., BLAHA, F. C., and FITZPATRICK, M. D. The Drain-Source Protected MNOS Memory Device and Memory Endurance. Technical Digest of the International Electron Devices Meeting. Washington, DC, 1973, pp 126-127.
- [6] FROHMAN-BENTCHKOWSKY, D. FAMOS—A New Semiconductor Charge Storage Device. *Solid-State Electron*, 17, 1974, pp 517-529.
- [7] GRANT, E. I. and LEAVENWORTH, R. S. *Statistical Quality Control*, 4th ed. New York: McGraw-Hill Book Co, 1972.
- [8] HSIA, Y. MNOS LSI Memory Device Data Retention Measurements and Projections. *IEEE Transactions on Electron Devices*, vol ED-24, no 5, May 1977, pp 568-577.
- [9] JEPPSON, K. O. and SVENSSON, C. M. Retention Testing of MNOS LSI Memories. *IEEE Journal of Solid-State Circuits*, vol SC-14, no 4, Aug 1979, pp 723-729.
- [10] RAYMOND, J. P. Radiation Testing of Microcircuits, *Semiconductor International*, Sept 1980, pp 115-124.
- [11] RELIABILITY REPORT. Reliability of Military Electronic Equipment. *Report by Advisory Group on Reliability of Electronic Equipment*, Office of the Assistant Secretary of Defense (Research and Engineering), Superintendent of Documents, Government Printing Office, Washington, DC, 1957.

²IEEE publications are available from the Service Center, The Institute of Electrical and Electronics Engineers, 445 Hoes Lane, PO Box 1331, Piscataway, NJ 08855-1331.

³JESD publications are available from Electronic Industries Association, 2001 Eye Street, NW, Washington, DC 20006.

- [12] SCHUERMEYER, F. L. Test Results on a MNOS Memory Array. *IEEE Transactions on Electron Devices*, vol ED-24, no 5, May 1977, pp 564-568.
- [13] SCHUERMEYER, F. L. and YOUNG, C. R. Endurance Studies on MNOS Devices. *Journal of Applied Physics*, vol 49, no 8, Aug 1978, pp 4556-4559.
- [14] SZE, S. M. *Physics of Semiconductor Devices*, ch 10. New York: Wiley Interscience, John Wiley & Sons, 1969.
- [15] WILLIAMS, R. A. and BEGUWALA, M. E. The Effect of Electrical Conduction of Si_3N_4 on the Discharge of MNOS Memory Transistors. *IEEE Transactions on Electron Devices*, vol ED-25, no 8, Aug 1978, pp 1019-1023.

4. Basic MNOS Cell Operation in an Integrated Circuit

4.1 Background. A MNOS cell grouped with other MNOS cells in a MNOS array and with other support MOS transistors on the same substrate form a complete MNOS memory integrated circuit. It is necessary to define the functions of these cells in various configurations. These functions are erasing, writing, and reading.

4.2 General Description. A representation of the physical structure of a MNOS device is shown in Fig 1. As in the case of a normal MOS transistor, this device has source and drain regions of diffused or implanted n- or p-type dopant for n- or p-channel devices respectively and a gate electrode of metal or polysilicon. A MNOS transistor differs from a MOS transistor in the structure of the gate dielectric that isolates the gate electrode from the silicon substrate. In a MNOS transistor, the dielectric is a two-layer structure of silicon nitride over silicon dioxide, whereas a MOS device utilizes only an oxide layer. In the case of a MNOS memory device, the oxide layer is made very thin, in the order of 2.5 nm.

As in a standard p-channel MOS transistor, the application of a negative voltage to the gate, equal to or greater than the threshold voltage $V_{\text{gs(th)}}$ will result in an inversion of the substrate surface under the gate forming a conducting channel that allows current to flow between drain and source when a potential difference exists between them. The property that makes a MNOS transistor both practical and very attractive as a memory ele-

ment is the possibility of changing the $V_{\text{gs(th)}}$ of the device. This is accomplished by the application of a (sufficiently) large positive voltage to the gate electrode. As a result, electrons tunnel from the silicon of the channel region into the gate dielectric, where they become trapped. These negative charges decrease the threshold voltage of the p-channel transistor. Conversely, when a large negative potential is applied to the gate electrode, holes are injected from the silicon to become trapped in the gate dielectric. This causes an increase in the threshold voltage of the p-channel transistor.

The polarity of the trapped charge is ascertained by the application of a voltage intermediate to the two threshold voltage extremes generated by the high positive or negative voltages. If the trapped charge is more positive, the transistor will not turn on, and no current can flow in its channel. If, on the other hand, the trapped charge is more negative, the applied intermediate voltage will turn on the transistor, and current can be made to flow in the channel.

Viewing the (p-channel) MNOS transistor now as a memory cell in a circuit, the application of a large positive voltage becomes the act of erasing, and the application of a large negative voltage becomes the act of writing. The intermediate potential used to determine the state of conduction is the read voltage of the memory cell. These quantities are discussed in greater detail in the following paragraphs.

4.3 Cell Functions

4.3.1 Erasing and Writing. As a result of the application of a positive gate tunneling potential with respect to the device substrate, a net negative charge is trapped in the nitride as shown in Fig 3(a). This causes a positive shift in threshold voltage and places a p-channel device in the high-conduction or erased state. The same operation would result in a positive $V_{\text{gs(th)}}$ shift for a n-channel MNOS transistor, which would, however, remain in the low-conduction or written state.

Conversely, a p-channel written state with a negative $V_{\text{gs(th)}}$ shift results from a negative gate tunneling potential (see Fig 3(b)). The same operation produces an erased n-channel transistor.

The commonality in defining states is derived from a circuit view of the device, the written state being the low-conduction state in both p- and n-channel technologies and the erased state being the high-conduction state in both cases.

4.3.2 Reading. The difference between the threshold voltage of a device in the erased state

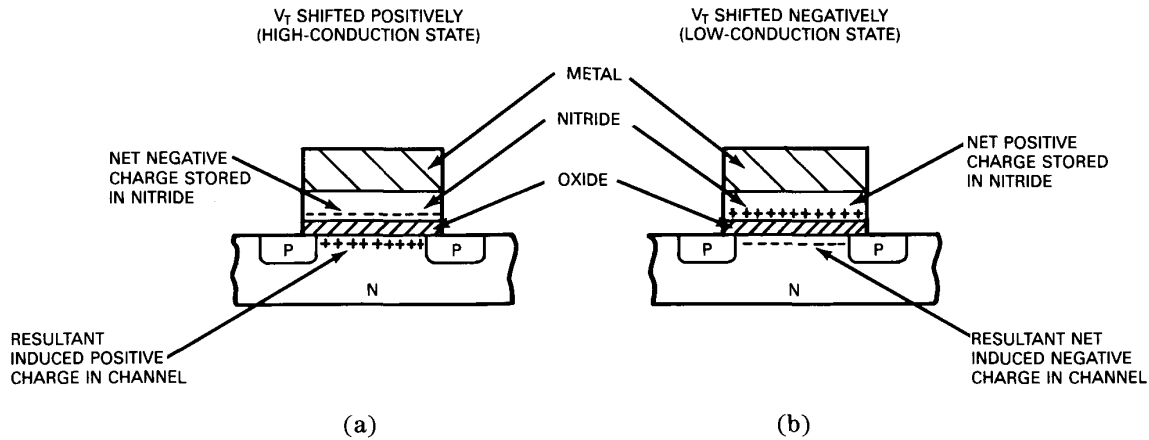


Fig 3
States of MNOS Transistor
(a) Erased Transistor (b) Written Transistor

$V_{(th)E}$ and that of one in the written state $V_{(th)W}$ can be several volts depending on fabrication variables and write/erase voltages and times. This fact provides for a convenient method of evaluating the condition of the cell. By the application of a gate reference or read voltage, V_{GR} , set somewhere between the two threshold voltages, a written p-channel device, whose $V_{gs(th)}$ will be more negative than the applied reference voltage, will not conduct, but an erased device will. Conduction or nonconduction may be interpreted by the sensing logic circuitry as either a high or low logic state.

To summarize, the following relationships are true for both p- and n-channel MNOS memory devices:

$$|V_{(th)E}| < |V_{GR}| < |V_{(th)W}|$$

$V_{(th)E}$ = erased-state threshold voltage
 $V_{(th)HC}$ = high-conduction threshold voltage
 $V_{(th)W}$ = written-state threshold voltage
 $V_{(th)LC}$ = low-conduction threshold voltage
 $V_{(th)E} = V_{(th)HC}$
 $V_{(th)W} = V_{(th)LC}$

4.3.3 Write Inhibit. To prevent an erased transistor from being written or a written transistor from being erased, gate and substrate voltages must be equalized. This is typically accomplished by having a drain bias voltage close to that of the gate voltage with the source open, so that during the write operation, the source voltage is pulled

towards the drain voltage, thus equalizing gate and substrate voltages.

4.3.4 Alternative Cell Structure. A disadvantage of the MNOS-gate structure described above results from continually stressing the gate dielectric by the repeated application of the high-voltage erase/write pulses, in particular in the areas where the dielectric overlaps the source and drain regions. Observed failure mechanisms include a complete dielectric breakdown that may occur in these areas.

Figure 4 shows a structure known as a tri-gate transistor that is used to overcome this disadvantage. In this structure the thin (2.5 nm) oxide region is reduced to a small area in the center of the channel, leaving the side portions of the device with a much thicker (45 nm) oxide that significantly improves the yield and characteristics of the MNOS memory element. The memory characteristics of the device are governed by the thin oxide, central region of the gate that behaves similarly to the flat-gate device previously described. One difference is that the high-conduction threshold will now be determined by the fixed $V_{gs(th)}$ of the side portions, which is close to that of a normal MOS transistor fabricated under the same conditions.

4.4 Cell Concepts and Structures. Three different cell concepts have been used in practical MNOS memory arrays.

- (1) Single-transistor cell
- (2) Two-transistor cell