DESIGN AND CHARACTERIZATION OF MIS DEVICES

BY

KJELL O. JEPPSON

TECHNICAL REPORT NO 69

1976

SCHOOL OF ELECTRICAL ENGINEERING

CHALMERS UNIVERSITY OF TECHNOLOGY

GÖTEBORG, SWEDEN

DESIGN AND CHARACTERIZATION OF MIS DEVICES

Akademisk avhandling
som för avläggande av teknisk doktorsexamen
vid Chalmers tekniska högskola
försvaras vid offentlig disputation
i E-sektionens sammanträdesrum,
Hörsalsvägen 11, 4tr, Chalmersområdet, Göteborg
måndagen den 17 januari 1977, kl. 13.15

av
Kjell O Jeppson

School of Electrical Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden

Technical Report No 69

DESIGN AND CHARACTERIZATION OF MIS DEVICES

by

Kjell Jeppson

Submitted to the School of Electrical Engineering,
Chalmers University of Technology, in partial
fulfilment of the requirements for the degree of
Doctor of Technology

Göteborg, December 1976

Design and characterization of MIS devices

by

Kjell Jeppson

This thesis comprises the following papers[1]:

A. G. Carlstedt, G. P. Petersson and K. O. Jeppson, "A content-addressable memory cell with MNOS transistors", in *IEEE Journal of Solid-State Circuits*, vol. 8, no. 5, pp. 338-343, Oct. 1973, doi: 10.1109/JSSC.1973.1050414

B. K. O. Jeppson and C. M. Svensson[2], "Negative Bias Stress of MOS devices at high electric fields and degradation of MNOS memory devices", Research Report No 38, submitted to Journal of Applied Physics. Presented at 1st Non-Volatile Semiconductor Memory Workshop, Vail, Colorado, August 17-19, 1976. Published in *Journal of Applied Physics*, 48, pp. 2004-2014 (1977); https://doi.org/10.1063/1.323909

C. Kjell O. Jeppson, Christer M. Svensson, "Unintentional writing of a FAMOS memory device during reading", *Solid-State Electronics*, 19, 6, pp. 455-457 (1976) doi: 10.1016/0038-1101(76)90006-X

D. K. O. Jeppson and J. L. Gates, "The Effects of Impurity Redistribution on the Subthreshold Leakage Current in CMOS N-channel Transistors", *Solid-State Electronics*, 19, 1, pp. 83-85 (1976) doi: 10.1016/0038-1101(76)90137-4

E. K. O. Jeppson[3], "Influence of the Channel Width on the Threshold Voltage Modulation in MOSFET´s", *Electronics Letters*, 11, 14, p. 297-299 (10 July 1975). doi: 10.1049/el:19750225

[1] This reference list has been updated adding full paper titles and Digital Object Identifiers (DOIs) March 28, 2023

[2] As of today (March 2023), this paper has been continuously cited every year since publication. In total, 914 citations.

[3] As of today (March 2023), this paper (which was the first paper to point out the influence of the channel width on the MOSFET threshold voltage) has been cited 57 times including in the fundamental MOS Transistor textbook by Tsividis and McAndrew.

Design and characterization of MIS-devices

Abstract[4]

This work is a part of the research performed at the Research Laboratory of Electronics (Elektronfysik III), concerning [metal-insulator-semiconductor] MIS field-effect devices. It deals with the properties of different memory devices, such as the [metal-nitride-oxide-semiconductor] MNOS and the [floating-gate avalanche-injection metal-oxide-semiconductor] FAMOS memory transistors, where the MIS structure is utilized for information storage. Paper A describes a new associative memory cell in which MNOS transistors are used as storage elements. Paper B describes the Negative Bias Stress of MOS devices at high electric fields with respect to the degradation observed in MNOS memory devices repeatedly operated at high write/erase gate voltages. Paper C deals with the FAMOS memory device and how the information may be unintentionally changed after a large number of read cycles. Paper D is concerned with some critical problems during fabrication of low threshold voltage CMOS circuits for digital watch applications. Paper E shows the influence of a narrow channel width on the threshold voltage in MOS transistors when modulated by the substrate-source voltage.

[4] This abstract has been edited adding full explanations of acronyms MIS, MNOS and FAMOS.

Chapter 1. Introduction

The development of the metal-oxide-semiconductor (MOS) technology during the last decade has resulted in the evolution of large scale integrated (LSI) circuits. Among other things, this has given the semiconductor memory a dominant role in the memory market. The advantages of semiconductor memories are their low price, their high performance, e.g., in speed, and their high packing density.

However, one drawback of all conventional random access semiconductor memories is that they are volatile, i.e., they lose their information upon power failure. The existing technology of nonvolatile semiconductor memories only provides for so called Read Mostly Memories (RMM) [a type of memory that can be read fast but written to only slowly]. In these memories it is easy to read information but either impossible or difficult to write new information. The simplest memory of this type is the Read Only Memory (ROM), where a permanent information pattern is stored in the memory during fabrication. There are also Programmable Read Only Memories (PROMs) where the information is written during a special programming step (e.g. Fusible-Link Memories). However, once written, the information cannot be erased. A more advanced PROM is the floating-gate avalanche-injection MOS (FAMOS) memory, which will be discussed later, where information may be erased by exposure to [a flash of] UV-light. [One of] the most promising nonvolatile semiconductor memory is the metal-nitride-oxide semiconductor (MNOS) memory which will be described in the following section. The MNOS memory is usually characterized as an Electrically Alterable Read Only Memory (EAROM).

For nonvolatile storage of information that needs to be continuously updated, usually magnetic memories, such as discs, tapes and tape cassettes are used. Because of the size and complexity of the magnetic memories, many small semiconductor systems lack a nonvolatile store. As a result of the search for a nonvolatile random access semiconductor memory cell, extensive research has been devoted to the MNOS memory transistor. This transistor was first suggested as a memory device in 1967 [1, 2]. The properties of the MNOS device have been characterized by many workers and many theories for the operation of MNOS memory transistors have been proposed as recently reviewed [3]. Also, the Research Laboratory of Electronics at Chalmers University of Technology has had a research program on MNOS devices involving most of their properties.

Chapter 2. The nonvolatile MNOS memory device

The MNOS transistor is like an ordinary MOS transistor, but its insulator consists of a double layer of silicon dioxide [SiO_2] and silicon nitride [Si_3N_4]. The different thicknesses of these layers (10 - 30 Å and 400 - 1000 Å, respectively) and their different electrical properties lead to the ability of storing charge at the interface [4]

Information may be stored as either positive or negative charge at the interface. The two logical states of the memory transistor are thus defined as a high and a low threshold voltage. The information may be retrieved by detecting in effect the threshold voltage as current or no current flow through the transistor at a predetermined gate read voltage.

The positive or negative charge is injected to the interface by tunneling of holes or electrons from the silicon through the oxide and into the nitride, where the charge is trapped close to the interface. The writing characteristics have been studied by Lundstrom and Svensson [5, 6], and by many others.

The MNOS memory transistor has very good storage properties and a possible memory retention time on the order of 1 - 10 years. The retentivity is limited by the slow leakage of the stored charge through the oxide layer. The discharge mechanisms have been studied by Lundkvist et al. [7, 8]. They proposed a model for the discharge behavior according to which the [memory] retention time may be evaluated from short-time measurements of units under stress.

The use of the MNOS-memory transistor in memory arrays has been described by Carlstedt and Svensson [9]. Many applications of the MNOS transistor in memory systems have been suggested. It may be used for nonvolatile storage in numerical control systems, point-of-sale terminals, in preset radio and TV tuning, and for telephone auto-dialing systems.

Paper A

In paper A, we show that the MNOS transistor is ideally suited for application in Content Addressable Memories (CAMs). In the content addressable mode of memory operation, the address is not needed for storing and retrieving information. Instead, information is simultaneously sensed from all specified memory locations and compared with the search information. Since all associated information can be retrieved, the stored information is assigned priority in a multiple output situation. A computer memory that can be operated in the content-addressable mode is called content-addressable memory or associative memory. The content-addressable read is called a match or search operation.

Content-addressable memories have been manufactured in many memory technologies such as cryogenic and magnetic technologies and more recently using bipolar and MOS-integrated circuits. A major difficulty has been high cost because of storage cell complexity. Very often associative search was best performed by storing information in low-cost random-access memories and then searching the memory word-by-word by special algorithms looking for associated information. However, such a search is very time consuming. From this we may infer that a CAM is an efficient hardware substitute for a random-access memory (RAM) and a searching algorithm.

With this is mind, we proposed an associative memory cell using MNOS-transistors. Since the properties of an MNOS transistor are very well suited for associative memory application, this

memory cell is very simple. First of all, the memory is nonvolatile, and the readout is nondestructive. Secondly, the MNOS transistor fits the associative memory organization in a natural way. Our memory cell consists of two MNOS memory transistors facing each other like the two halves in a flip-flop memory cell. To reduce current flow during the write cycle, two rectifying or resistive elements must be added in the drain connection of the MNOS-transistors. In our memory cell MOS transistors were used as rectifying diodes. The cell area was 5000 μm^2 but may be further reduced, e.g., by using Schottky diodes instead of· MOS transistors. The memory cell is accessed by four logic signal lines. The MNOS transistor gates are connected in a word line and all drains are connected through the rectifying elements to a common, word sense or match line. There are two bit-lines connected to each memory cell through the MNOS-transistor sources. These two bit-lines contain the desired information (comparand).

The memory cell can store four states of information. Because of the many available combinations, it is possible to control the masking of certain bits not only from the comparand but also from the cell itself.

In paper A the principle of the memory cell is described as well as the layout and fabrication of the cell. The write, read and search operations are discussed, and some access circuits are suggested that minimize the number of connections to the memory chip. Measurements were made on a memory cell to demonstrate the feasibility of the MNOS-CAM. Our conclusion (in 1972) was that, with our small-area cell and with the suggested access circuits, a 1-kbit memory was quite feasible.

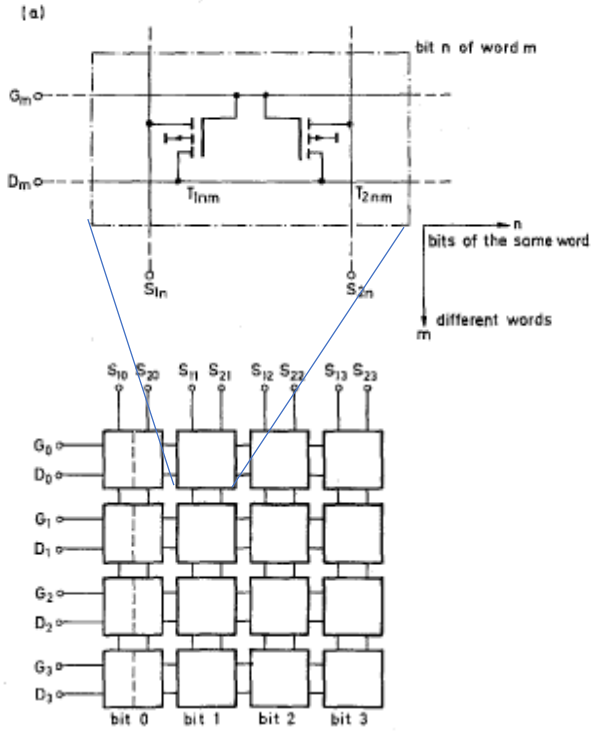

Fig. 1. The CAM-cell (top) and the content-addressable memory array (bottom). From [A].

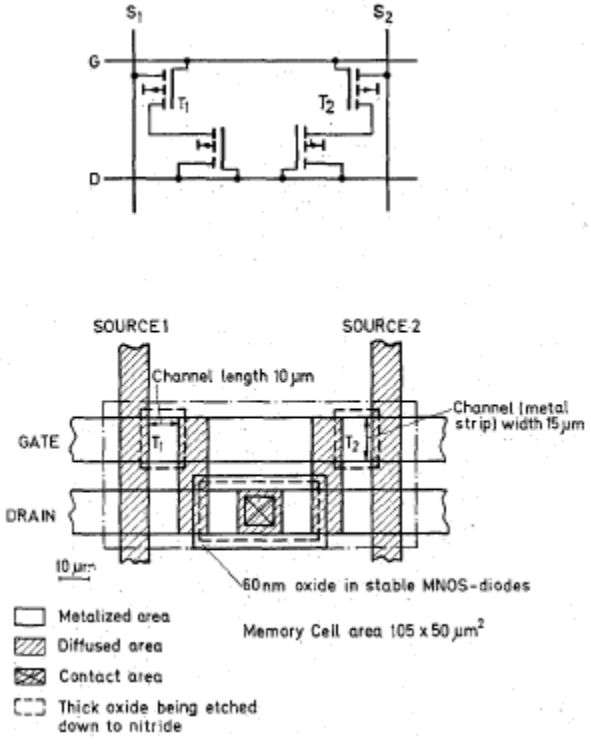

Fig. 2. CAM-cell circuit schematic (top) and layout (bottom). From [A].

Note: These figures from paper A were not shown in the original version of the thesis.

Paper B

Paper A described one of several applications of the MNOS transistor in nonvolatile memory systems. For several years it was hoped that MNOS memories would be used as conventional LSI RAMs. However, due to some performance imperfections they are now finding their market as an EAROM (Electrically Alterable Read Only Memory}. First of all, the write/erase cycle is considerably longer than the read cycle (10-1000 times). However, the most serious shortcoming for RAM applications of the MNOS device is the wear-out after extensive write/erase cycling. This means that the use of the MNOS transistor as a read/write memory is limited to a certain number of write/erase cycles (typically 10^6 cycles). The degradation effect shows up as a reduced retention time accompanied by a decrease in the size of the memory window. Also, the threshold window moves towards more negative voltages. It is the high oxide field (700-1600 MV/m) during the write/erase cycle that introduces changes in device properties. In 1972, the negative shift of the threshold window was shown to be caused by a strong increase in the density of the Si-SiO_2 interface traps. This effect was proposed to be caused by the same effect that is observed in regular MOS devices subjected to negative bias stress (NBS).

Since 1972, not many papers have been published and little progress has been made in the understanding of the MNOS degradation effects. Paper B is an attempt to clarify the situation by proposing a model for surface trap generation due to negative bias stress. The formation of surface traps is a very important and interesting phenomenon much more complicated than that of charge just being trapped at the interface. We believe that new surface traps may only be formed if the atomic structure at the interface is changed. In this way surface traps are formed at the silicon-silicon dioxide interface where the atomic structure of the oxide is changed. These changes should be basically the same in both MOS and MNOS structures. We nave therefore chosen to study NBS of MOS devices because these measurements are easier to evaluate. NBS measurements on MNOS devices are more complex to evaluate mainly because of difficulties in accurately determining the electric field in the oxide. This field is affected by the stored charge which is continuously increased during NBS.

Therefore, this paper gives a detailed study of the increase of the density of surface traps in MOS structures after NBS at temperatures (25 – 125°C) and fields (400 – 700 MV/m) comparable to those used in MNOS devices. These measurements extend earlier measurements on the NBS effect in MOS-capacitors which were performed at high temperatures and low fields. The NBS behavior of MOS capacitors was characterized through studies of the time dependence of the surface trap formation at different stress voltages and temperatures. Two mechanisms were observed to create new surface traps during NBS. The first mechanism appears at moderate fields and is believed to be [reaction-] diffusion controlled. The time dependence of this type of process was shown to agree with experiments. The other mechanism, which is dominant at very high fields, is a hole tunneling process.

More specifically, we assume in our model that the silicon-silicon dioxide (Si/SiO_2) interface contains a large number of defects which normally are electrically inactive. These defects are believed to be “dangling” silicon bonds that become inactive by binding to hydrogen. During stress, these defects become electrically active through a chemical reaction in which the weakly bonded hydrogen reacts with the SiO_2. This reaction leaves a trivalent silicon atom, with one

"dangling" bond, acting as a surface trap. The hydrogen forms a silanol group (Si-OH) in the oxide, leaving another charged trivalent silicon atom in the oxide. The reaction thus formed a surface trap and a surface charge at the defect site. Our measurements indicate that this process is diffusion controlled. This means that a stable surface trap is only formed if the silanol group at the interface diffuses away into the bulk oxide. The rate at which a new surface trap is formed, is then controlled by the diffusion rate of the silanol group.

At very high fields (>6.3 MV/cm) the formation of surface traps increases very rapidly with the applied field and becomes independent of temperature. In this second mechanism, we suggest that holes are injected and trapped in the oxide and that these holes then create surface traps. However, this second mechanism requires further clarification regarding the atomic model.

Possible electrochemical reactions which could occur at the Si-SiO_2 interface were discussed with respect to the results of the measurements. Finally, the NBS behavior of MOS capacitors was compared with the degradation behavior of MNOS capacitors, and it was found that the MNOS devices followed the [reaction-] diffusion-controlled mechanism discussed above.

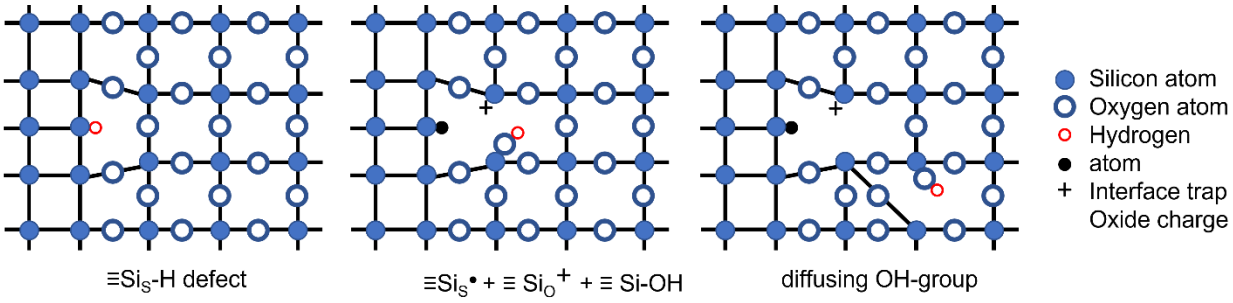

Fig. 3. Schematic two-dimensional representation of the Si-SiO2 interface showing a) the Si-H defect, b) electrical activation of the ≡Si-H defect during negative-bias stress forming an interface trap (●) , an oxide charge (+), and a hydroxyl (OH) group that c) may diffuse in the oxide and be the reaction-diffusion-limited factor. (≡Si_S indicates a silicon atom bonded to three other silicon atoms while ≡Si_O indicates a silicon atom bonded to three oxygen atoms.)[5]

[5] This redrawn version of figure 8 in paper B was not in the original thesis (only in paper B) but has been added to this digital version.

Chapter 3. The non-volatile FAMOS-memory device

Paper C

In paper C the read-disturb, or unintentional writing, of a FAMOS memory device was studied. The FAMOS (floating-gate avalanche-injection MOS) transistor has a polycrystalline silicon gate that is electrically floating in the silicon dioxide. The transistor is normally off, i.e., storing a logical one. Information {logical zero) is written into the FAMOS memory cell by injection of hot electrons to the floating gate. These hot electrons are generated by avalanche breakdown in the drain-substrate junction and acquire large enough energy to surmount the Si-SiO_2 barrier. The negative charge on the gate gives it a negative potential and turns the device on. This device is useful as a PROM but not as a read write memory (RWM) since the information can only be erased through illumination with [a flash of] UV-light or X-rays [10].

However, it has been reported that the floating gate may be charged by channel current-induced breakdown at drain voltages considerably lower than the gate-enhanced drain-to-substrate breakdown. Since the only difference between reading and writing of a FAMOS transistor is the magnitude of the negative voltage applied to the drain, the transistor could be unintentionally written during repeated read cycles.

The temperature dependence of this failure process was studied by us, and it was shown that the worst case for unintentional charging of a FAMOS device occurs at low temperatures. This temperature dependence is qualitatively explained by a model including the temperature dependence of pre-avalanche carrier multiplication and hot-electron injection.

In this study, measurements were done on discrete FAMOS transistors available as test transistors in the FAMOS memory NS 5202. The drain voltages. were comparable to the actual read voltage in a FAMOS memory circuit. Measurements were done at three different temperatures -40°, 25° and 70° C.

Spurious charging of a FAMOS device has been reported a problem and necessitated endurance tests in practical applications. A worst-case test of the (storing a “1”) memory retention time of such a circuit, should then be performed at low temperatures, as shown in paper C and not at high temperatures.

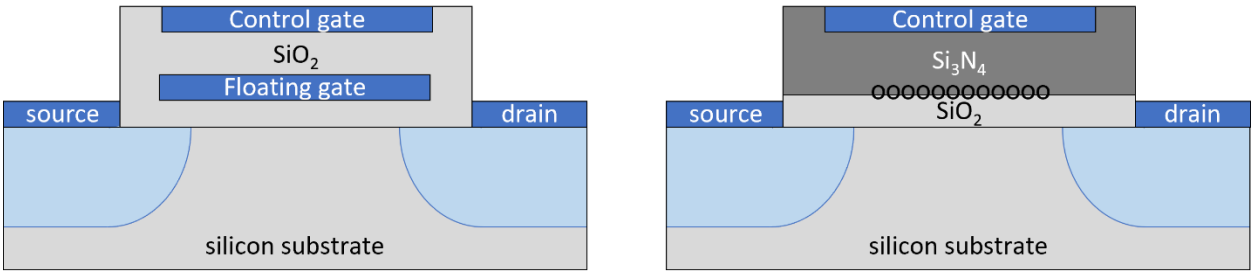

Fig. 4. Floating-gate avalanche-injection metal-oxide-semiconductor (FAMOS) transistor (left) and metal-nitride-oxide-semiconductor (MNOS) transistor (right).

Note: These figures showing cross-sections of the FAMOS and MNOS field-effect transistors were not in the original version of the thesis but have been added to the digital version.

Chapter 4. MOS-technology

Paper D

This paper deals with the problem of designing a bulk CMOS/FET process for fabrication of a digital watch circuit. The major concern in such a process is to obtain a well-defined, low threshold voltage giving the device a low, subthreshold-leakage current. As the name indicates, this leakage current flows for gate voltages less than the conventional, linearly extrapolated, threshold voltage. In a circuit with a low threshold voltage of 0.55 V and a battery voltage of only 1.1 V, there are very small margins and an excess, subthreshold-leakage current leads to excess power consumption.

Careful calculations show that some subthreshold leakage current is inherent to the MOSFET device. However, we found that in a CMOS circuit the actual leakage current in the n-channel device may be much larger than the inherent current. The major reason for this is the complexity of the n-channel device, which is formed in a p-type tub in the n-type substrate. The region of the substrate where the n-channel device is formed, therefore contains both boron and phosphorus impurity ions. This paper shows that such a substrate is very sensitive to impurity redistribution during thermal oxidation processes. The surface impurity concentration may change considerably and cause a large increase in the subthreshold leakage current. Th.is may be due to a leakage current flowing because the transistor is in weak inversion and/or because a leakage current flows under the field oxide around the edges of the channel region.

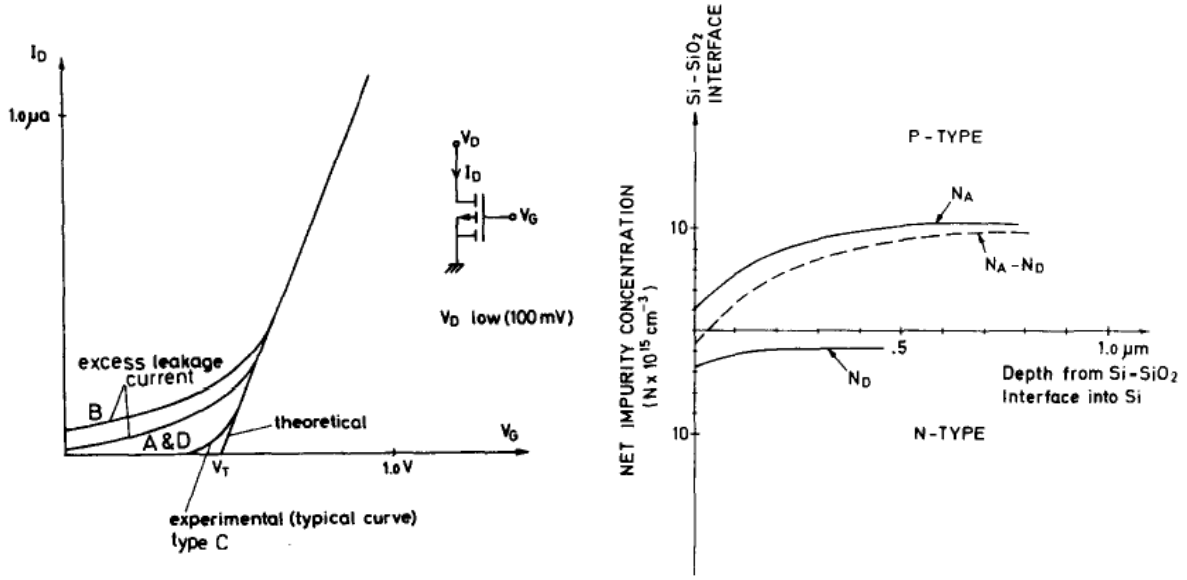

Fig. 5. Drain current vs. gate voltage showing excess subthreshold leakage (A, B, D) (left), and the effects of impurity redistribution showing that n-type layer can be formed at the Si/SiO2 interface leading to excess subthreshold leakage (right). These figures from paper D have been added to the digital version of the thesis.

Paper E

In paper D we studied the doping profile beneath the channel region and its influence on the channel leakage current. The doping profile was determined from high-frequency CV-measurements on MOS-capacitors since there is no method of determining the doping profile of the transistor itself. This led to the studies in paper E of the threshold voltage modulation, due to the substrate bias, which method may be used to determine the average substrate doping under the gate of an MOS tran-sistor. However, this method is also limited by the channel length and channel width which must be much larger than the depth of the depletion layer under the gate.

If the area of the depletion region is the same as the channel dimensions, the depletion charge that must be compensated from the gate in order to maintain constant current is determined only by the impunity concentration. However, in an actual transistor, the effective area of the depletion region is some fraction of the depletion depth narrower than the channel length and some fraction wider than the channel width. Several authors have presented theories of the influence of a short channel length on the threshold-voltage modulation. In paper E, the influence of the channel width on the threshold voltage modulation is pointed out. The main purpose of this paper is to show the importance of the channel width in narrow channel transistors. The result is an analytical expression of the threshold voltage as a function of substrate bias as well as channel geometry. The theory agrees well with experimental data.

In a paper published later [11] Kroell and Ackermann numerically solved Poisson´s equation to show the influence of the channel length and width on the threshold voltage. The general behavior of their solutions supports the theory presented in paper E.

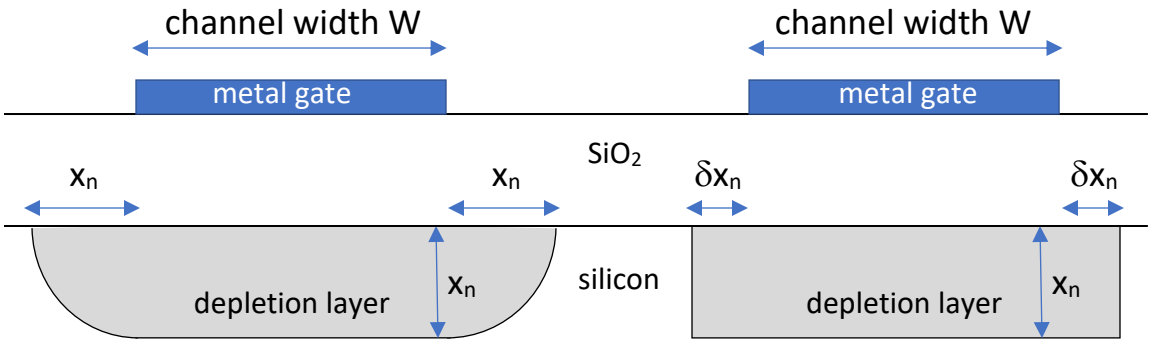

Fig. 4. Simple cross-sectional model views of narrow-channel MOSFETs where the left-hand first-order model is approximated by the square-box model to the right for estimating the extra depletion charge outside the channel region. These redrawn versions of figure 1 in paper E have been added to the digital version of the thesis.

Acknowledgement

I wish to acknowledge Professor Torkel Wallmark for his advice and support during the course of this work. The many discussions with the members of the group at the Research Laboratory of Electronics have also been stimulating. In particular, I am grateful to Dr Christer Svensson for his encouragement and many invaluable discussions. Thanks are also due to Monica Myregård and Karin Widmark for typing the manuscripts. I also wish to acknowledge the Microelectronics Division of Rockwell International where the work in paper D was performed.

References

1. H. A. R. Wegener, A. J. Lincoln, H. C. Pao, M. R. O´Connell and R. E. Oleksiak, Int. Electron Devices Meeting (IEDM), IEDM Technical Digest (Washington, DC 1967 Dec. 18-20 IEEE).
2. J. T. Wallmark and J. H. Scott, Jr., *RCA Rev*., 30, 335 (1969)
3. J. J. Chang, "Nonvolatile semiconductor memory devices," in *Proceedings of the IEEE*, vol. 64, no. 7, pp. 1039-1059, July 1976, doi: 10.1109/PROC.1976.10272.
4. P. Balk, Solid State Devices, ESSDERC 1973, (The Institute of Physics) (1974).
5. K. I. Lundstrom and C. M. Svensson, "Properties of MNOS structures," in *IEEE Transactions on Electron Devices*, vol. 19, no. 6, pp. 826-836, June 1972, doi: 10.1109/T-ED.1972.17500. https://doi.org/10.1109/T-ED.1972.17500.
6 . C. M. Svensson and K. I. Lundstrom, " Trap-assisted charge injection in MNOS structures", *J. Appl. Phys.*, 44, 4657 (1973) doi: 10.1063/1.1662016
7. L. Lundkvist, I. Lundstrom and C. Svensson, "Discharge of MNOS structures", *Solid-State Electronics*, 16, 7, pp. 811-823 (1973) doi: 10.1016/0038-1101(73)90178-0
8. L. Lundkvist, C. Svensson and B. Hansson, "Discharge of MNOS structures at elevated temperatures", *Solid-State Electronics*, 19, pp. 221-227 (1976) doi: 10.1016/0038-1101(76)90166-0
9. L. G. Carlstedt and C. M. Svensson, "MNOS memory transistors in simple memory arrays", in *IEEE J. Solid-State Circuits*, SC-7, 5, pp. 382-388 (1972) DOI: 10.1109/JSSC.1972.1052897
10. D. Frohman-Bentchkowsky, "A fully decoded 2048-bit electrically programmable FAMOS read-only memory," in *IEEE Journal of Solid-State Circuits*, vol. 6, no. 5, pp. 301-306, Oct. 1971, doi: 10.1109/JSSC.1971.1050191.
11. K. E. Kroell and G. K. Ackermann, "Threshold voltage of narrow channel field effect transistors", *Solid-State Electronics*, 19, pp. 77-81 (1976) doi: 10.1016/0038-1101(76)90136-2