CHALMERS
UNIVERSITY OF TECHNOLOGY

A content-addressable memory cell with MNOS transistors

Citation for the original published paper (version of record):
Jeppson, K., Peterson, G., Carlstedt, G. (1973). A content-addressable memory cell with MNOS transistors. IEEE Journal of Solid-State Circuits, 8(5): 338 - 343.
http://dx.doi.org/10.1109/JSSC.1973.1050414

A Content-Addressable Memory Cell with MNOS Transistors

GUNNAR CARLSTEDT, GÖRAN P. PETERSSON, AND KJELL O. JEPPSON

Abstract—This paper describes a new associative memory cell in which MNOS transistors are used as storage elements. The memory can perform functions as a read-only memory (ROM) and at the same time as a read–write memory. The cell can be read as a random-access memory (RAM) or as a content-addressable memory (CAM). As an CAM certain bits can be masked out, i.e., not compared with the stored bits. The comparison can also be controlled from the memory by the stored words. Since the word length or combinations of normal words can be stored in one word of the memory, fewer memory cells are needed than in an ordinary memory. Searches for groups of words (prime implicands) can be performed. Memory cells with an area of 5000 –m have been built to demonstrate the feasibility of the MNOS-CAM.

I. Introduction

ASSOCIATIVE memories of content-addressable memories (CAM) have been used for a long time and have been manufactured in many technologies [1]–[5]. The memory cells have been expensive because of the complexity of the storage cell. Mundy *et al.* [6] have published a report on a simple associative memory cell with five MOS transistors and four connections. The cell area was 12 500 μm^2. In this paper we will describe a memory with a cell area of 5000 μm^2 using MNOS transistors as storage elements.

The switching behavior of the MNOS transistor has been analyzed by several authors [7]–[9]. Carlstedt and Svensson [10] have shown reading, writing, and storage characteristics of memory cells built by MNOS transistors.

In the first part of this paper the logical behavior of the memory cell will be analyzed in detail. In the next part certain access circuits will be described, and finally measurements on a memory cell with an area of 5000 μm^2 are described.

The memory cell described is a p-channel device.

II. The Memory Cell

In Fig. 1 the memory cell is shown. The memory cell consists of two MNOS transistors, T_{1nm} and T_{2nm}. All cells are arranged in a square pattern. In the n direction all gates and drains are connected to the G_m and D_m lines, respectively. In the m direction, all sources are connected. One word is stored in the memory cells with

Manuscript received March 19, 1973; revised June 19, 1973.
The authors are with the Research Laboratory of Electronics III, Chalmers University of Technology, Göteborg, Sweden.

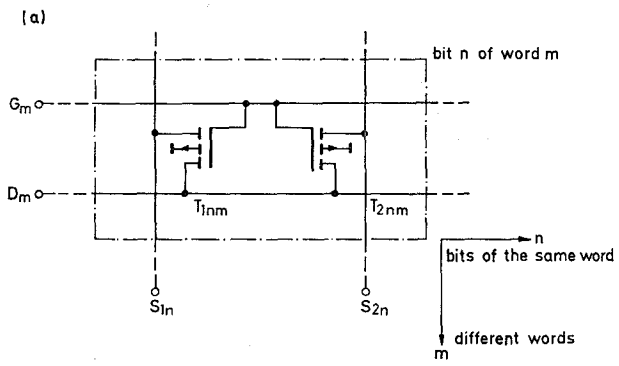


Fig. 1. (a) Memory cell. (b) Organization of a 4 word 4-b memory. Words are stored horizontally.

the same m coordinate. If the threshold voltage of an MNOS transistor has such a value that the transistor conducts during a read operation, the transistor contains a logical one, if not, it contains a logical zero. The contents stored in the memory transistors are described by the variables T_{1nm} and T_{2nm} for transistors T_{1nm} and T_{2nm}, respectively. A logical zero or one can be stored in both T_{1nm} and T_{2nm}. Therefore the memory cell has four states. The stored word is named X_m and its content is

$$X_m = X_{0m}X_{1m} \cdots X_{nm} \cdots X_{(N-1)m}$$

if the word is N bits long.

The relations between the four states of the memory cell and the stored word are as follows.

T_{1nm}	T_{2nm}	
0	0	There are both a logical zero and one stored in the memory cell (X_{nm} = 1 or 0, i.e., do not care).

0	1	There is a logical zero stored in the memory cell ($X_{nm} = 0$).
1	0	There is a logical one stored in the memory cell ($X_{nm} = 1$).
1	1	The word does not use this bit (bit X_{nm} not used in X_m).

The memory can be controlled by the connections D_m, G_m, S_{1n}, and S_{2n}.

III. Writing Process

Writing is accomplished with a three-phase write cycle as described by Carlstedt and Svensson [10]. The two logical states correspond to different threshold voltages of the MNOS transistor and are defined by the polarity and amplitude of the write voltage across the insulator. The write process for a single MNOS transistor is shown in Fig. 2. During the first step, the threshold voltage is moved in the negative direction. During the next step the transistor assumes a well defined more positive threshold voltage, due to the opposite polarity of the second write pulse. During the third step the threshold voltage can selectively be changed to a more negative value. We shall now analyze in detail the behavior of the memory cell during the write cycle.

During the first phase of the write cycle, all gates G_m of the memory cells in the addressed word m have a large negative potential. The source and drain connections S and D are grounded. Because of the large negative potential all transistors are heavily conductive and therefore the write voltage across the insulator of the MNOS transistor equals the applied gate voltage. All threshold voltages are shifted to a more negative value.

During the second phase a large positive voltage is applied to the gates of the MNOS transistors in the addressed word. No transistor is conductive because of the large positive gate potential. The voltage across the insulator of the MNOS transistor therefore equals the voltage between the gate and the substrate. The threshold voltage now changes in the positive direction to the well defined more positive value.

During the third phase all gates of the addressed word have a large negative voltage applied. It is desired that some of the transistors retain the threshold voltage established in phase 2. The common drain connection is negatively biased (normally −15 V). If the threshold voltage of the transistor is to be changed, its source connection is grounded, if not, the source is biased to the negative drain potential. Since the write voltage is negative, there is a conductive channel between the source and the drain. If the source is biased, the effective write voltage across the insulator is reduced to a value at which the threshold voltage cannot be changed.

The transistors that are shifted during this third phase are heavily conducting. Large currents flow through the drain. The power dissipation in these transistors can be high enough to cause damage. To eliminate this, rectifying or resistive elements must be introduced

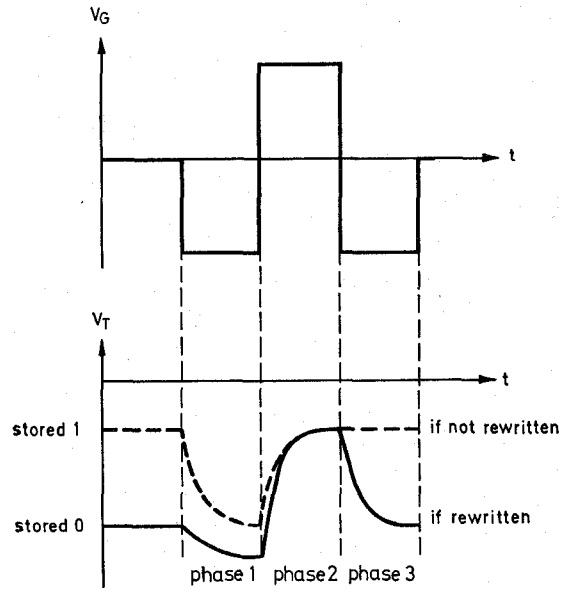


Fig. 2. Change of the threshold voltage during the three phase write process in a memory transistor. During the first phase the threshold voltage is not very well defined because of leakage and different logical states. During the second phase a well defined threshold voltage is established. In the third phase the information is written in the transistor.

in the drain connection of the MNOS transistors. The rectifier must be directed so that the transistor is able to conduct when the source has a higher potential than the drain. The D_m line can now be floating or grounded during writing. No current flows through the rectifiers. The rectifying element can be an MOS transistor or a Schottky diode. A memory cell modified with MOS transistors as rectifiers is shown in Fig. 6(a) and will be evaluated in a following section.

IV. Reading Process

The reading process has a single phase cycle, and it is accomplished by applying a negative voltage to the drain D_m of the addressed word m, and by applying a read voltage on the gates. The read voltage has a value chosen to lie between the two threshold voltages. All source connections are grounded. A current in the source connection shows that a logical one is stored in the addressed transistor.

V. Interrogation Process

During the interrogation process all words in the memory perform the same logical function. The memory is searched by the comparand

$$Y = y_0 y_1 \cdots y_n \cdots y_{N-1}.$$

This is accomplished by applying a read voltage to all the gates of the memory. The logical variables S_{1n} and S_{2n} define the voltages on the source connections of a memory cell. S is logical one when the source connection is negative (normally −15 V) and low when it is grounded. A current will flow through the negative biased drain connection D_m if the function

$$I_m = \sum_{n=0}^{N-1} (T_{1nm}S_{1n}^* + T_{2nm}S_{2n}^*)$$

is true. S_{1n}^* means the inverse of S_{1n}. This current can be used to detect different modes depending on the organization. We shall now describe the different possible cases.

$S_{1n} = 1$, $S_{2n} = 1$: The bit n in the memory cannot give any current through the connection D_m. In an associative memory this combination can be used to mask out a certain bit in the comparand. The masked out bit is not compared with the corresponding bit in the stored words during the interrogation process.

$S_{1n} = Y_n$, $S_{2n} = Y_n^*$: If $T_{1nm} = T_{2nm}^* = X_{nm}$ an exclusive or logical function is performed between the variable X_{nm} and Y_n. If one or more bits, X_{nm} in a certain word is not equal to the corresponding bit Y_n of the comparand, a current will flow through the connection D_m. A mismatch is found. With this combination of variables the content is compared.

If $T_{1nm} = T_{2nm} = 0$ no mismatch can be found. Therefore this combination of T can be used to mark that the value of the bit is of no importance, i.e., both a logical zero and one are stored. This feature may be used when functions are stored in the memory, as in a read-only memory (ROM). Therefore not all states of the function have to be stored separately, i.e., a considerably lower memory capacity is needed to store a certain function.

If $T_{1nm} = T_{2nm} = 1$ a mismatch is found since either S_{1n} or S_{2n} is a logical zero. This feature can be used to control word length, when the stored word must have a length that is shorter (or equal) to the length of the comparand Y.

$S_{1n} = 0$, $S_{2n} = 0$: This case indicated that the bit in the comparand has the value of both a logical zero and one. There is no interrogation current *if and only if* the value of the stored bit also is both a logical zero and one (see $T_{1nm} = T_{2nm} = 0$). This mode can therefore be used when a search for prime implicands is performed.

VI. Access Circuits for the Associative Memory

The memory described has the connections S, G, and D that are connected to circuitry controlling the memory.

In Fig. 3 a circuit for the S_{1n} connection is shown. During the write process the WRITE input is a logical one and transistor T_4 is conducting. During the two first phases of the write cycle the transistor T_{6n} is conducting, thus grounding the S_{1n} connection. When writing is to be inhibited during the third phase, the DATA_{1n} input is zero. This means that T_{3n} is nonconducting and the potential of S_{1n} approaches V_C. The content stored in the memory transistor T_{1nm} is the inverse of DATA_{1n}.

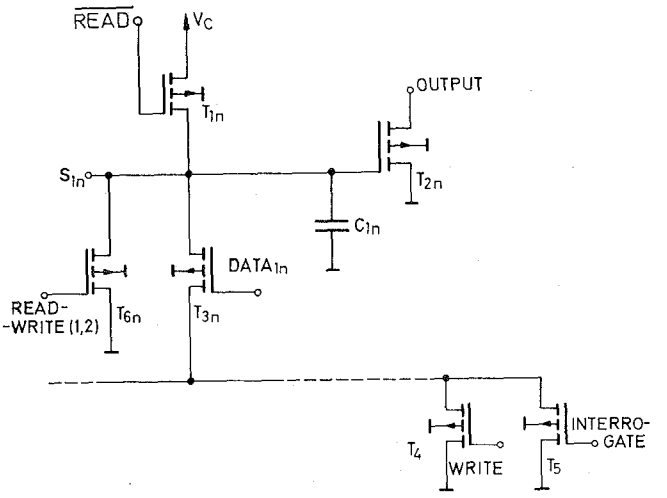


Fig. 3. An access circuit for the S connection. The circuit gives the voltages needed for writing, reading, and searching.

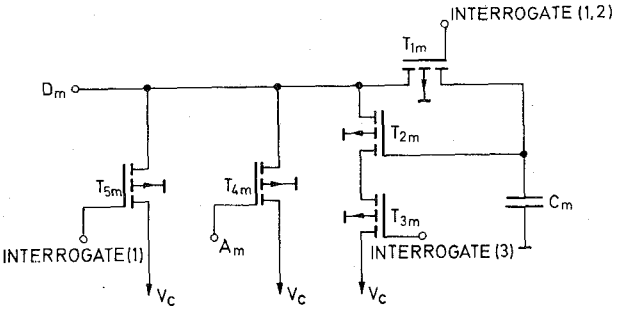


Fig. 4. Access circuit for the D connection. The circuit gives the voltages needed during the write and read cycles. The matched words can also automatically be read after interrogation.

During the read process transistors T_4 and T_5 are nonconducting. Transistor T_{2n} then acts as a driver for the output. Transistor T_{6n} conducts during a short time in the beginning of the read cycle and discharges the capacitance C_{1n}. If a logical one is stored, the transistor T_{2n} is conducting because C_{1n} is recharged by a current on the S_{1n} line.

During the interrogation process T_5 is conducting and the comparand Y_n (and its complement Y_n^*) are connected to DATA_{1n} and DATA_{2n}. S_{1n} varies as the inverse of DATA_{1n}.

The circuit shown in Fig. 4 can be used on the interrogation connections D of the memory. This circuit supplies the different voltage levels to the drain connection. When writing, the potential of the drain connection is not defined. During reading the input A_m is logical one in the addressed words and the transistor T_{4m} is then conducting. The drain potential of the addressed word is then near V_C.

The interrogation process is divided into three phases. During the first phase T_{1m} and T_{5m} are conducting. The capacitor C_m is charged. During the next phase only T_{1m} is conducting. The capacitor is discharged if any mismatch of the word is present. In the last phase, the word is read out from the memory. This is accomplished by making transistor T_{3m} conducting. Only in the matched words the drain potential is negative (towards V_C). In these words the information is being read out through the source circuits described previously.

VII. Layout and Fabrication

A simple memory circuit was fabricated to demonstrate the feasibility of the MNOS–CAM. The storage array consists of eight programmable MNOS transistors and eight stable MNOS diodes organized as two words with two bits per word.

The two stable MNOS diodes are used as rectifying elements in the drain connections. The choice of stable MNOS diodes for this purpose was made to ease manufacturing as it proved to be compatible with the fabrication process of the MNOS transistors. The disadvantage of using MNOS diodes is the high voltage drop across the diode due to the threshold voltage.

Fig. 5 shows the layout of the entire circuit and in Fig. 6 a single memory cell (one bit) is shown. The nominal channel width and length of the diodes and transistors are 15 μm and 10 μm, respectively. The dimensions are the minimum values that could be handled with our equipment and processes allowing a reasonable yield. We believe that the dimensions could be reduced by a factor of 2.

One memory cell occupies an area of 105 $\times$ 50 μm^2 (8.4 mil^2). A 1024-b memory (16 b, 64 words) would thus cover an effective memory area of 1.7 $\times$ 3.2 mm^2. The smallest spacings between the different areas on a photomask are 10 μm. The smallest alignment tolerance for two consecutive photomasks were chosen to be 2.5 μm.

The CAM memory was fabricated from 1–2 $\Omega\cdot$cm, (100), n-type silicon. The p diffusion was made from boron nitride wafers and a sheet resistance of about 30 Ω/square was obtained. A 60-nm oxide layer was grown and etched off, leaving oxide islands over the MNOS diode areas. The oxidation also acted as a drive-in step for the diffusion, resulting in a diffusion depth of about 2.5 μm. A double-layer of 2-nm thin oxide and 80-nm silicon nitride was fabricated forming the gate insulator in the memory transistor. The silane–ammonia method was used to deposit the silicon nitride. The wafer was then covered with a 350-nm layer of silicon dioxide under the metal strips. Finally the metallization pattern was defined in a double metal layer consisting of about 80-nm chromium and 300-nm gold.

VIII. Measurements

A few CAM memories were bonded in 16-pin TO-5 cans. The MNOS transistor gates were pulsed with

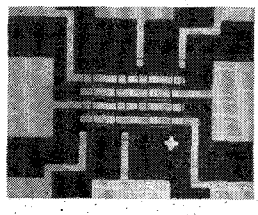


Fig. 5. Layout of the fabricated 2 $\times$ 2 memory. The active memory area is 210 $\times$ 100 μm^2.

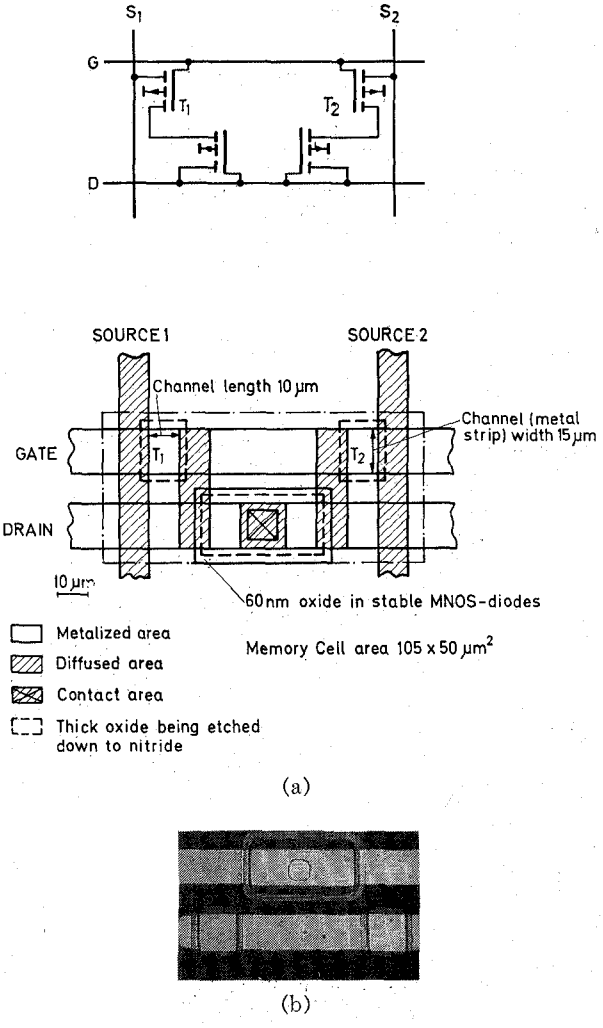
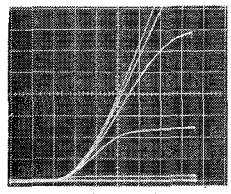


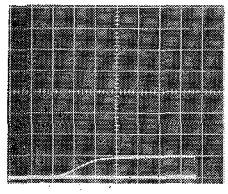
Fig. 6. Memory cell configuration. (a) The circuit diagram upper and the layout on the silicon wafer lower. (b) Photomicrograph of a fabricated memory cell.

1-ms write pulses of +35 V and −40 V resulting in threshold voltages about −3 V and −10 V. In Fig. 7 the characteristics of half a memory cell (one MNOS transistor in series with a stable MNOS diode) is shown for both a logical one and zero. The areas outside the active elements covered by a thin oxide-nitride-thick oxide layer showed a threshold voltage of −64 V. Thus it prevented parasitic transistor action in the circuit with the chosen write pulses.

After recycling the memory transistors stored logical



(a)



(b)

Fig. 7. The characteristics for one half memory cell when the cell stores either (a) logical one or (b) logical zero. Horizontal scale shows the voltage between drain and source connections with the scale 2 V/div. On the vertical scale the current through the transistors are shown with the scale 20 μA/div. The gate voltage has been varied in 6 steps of 2 V.

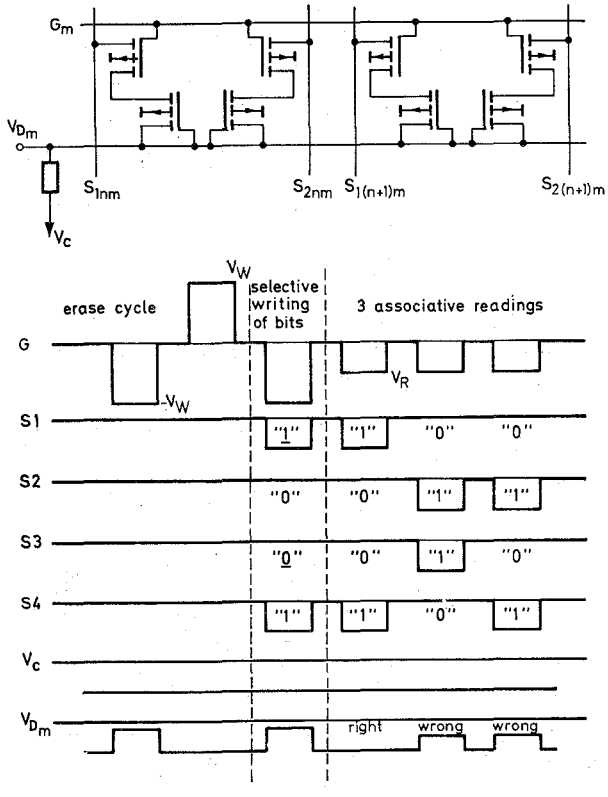


Fig. 8. Write and associative search cycle for a two bit word, V_W and V_R correspond to the write and read voltage, respectively.

ones, i.e., had low threshold voltages of about -3 V. Logical zeros were then selectively written with a negative pulse while the source terminals of the transistors that should remain in the logical one state were biased to -15 V. The conductive channels assumed the potentials of their sources since the drain connections are insulated from each other through the reverse biased MNOS diodes. The voltage drop across the gate dielectric will then be reduced to a lower value

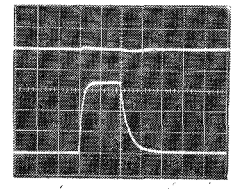


Fig. 9. Output signals on the drain connections measured with an oscilloscope during the interrogation process. The upper trace shows a word that matches the comparand and the lower trace shows a mismatch. On the vertical scale the output voltage is shown (2 V/div) and on the horizontal the time (scale 5 μs/div).

at which a change in threshold voltage cannot take place. After this three phase write cycle, the memory could be tested by reading the information in each word. The drain terminal of the word was negatively biased and a read voltage of -6 V was applied to the gate. A current pulse at the source output indicates the logical state.

The write and associative read for a two-bit word is shown in Fig. 8. In the two words of the memory the information **10** and **01** were written, i.e., in the first word the four transistors stored the information **1001** and in the second word **0110**. The word **10** was compared with the memory information by applying -15, 0, 0, -15 V at the source terminals. The drain terminals of the two words were biased with -15 V via two load resistors. The read voltage was 6-V, 10-μs pulses. In the first word the information is correct and no transistor is conducting. In the second word, the information is wrong in both of the cells and one transistor in each cell is conducting. This is because the source terminal of the transistor with low threshold voltage is put to ground and the transistor becomes conducting. Fig. 9 shows the output signal measured with an oscilloscope.

Conclusion

An associative memory can be built with memory cells consisting of four or less MNOS transistors. A memory cell has been built using only 5000-μm^2 silicon area. This small cell area and the access circuits described should make a 1-kb memory quite feasible. Because of the use of MNOS transistors, the memory cell is nonvolatile. The memory cell has four logical states. These states can be used to mask the cell, store a logical zero, store a logical one, or store both a logical one and zero. The memory cell can be accessed randomly or associatively.

The purpose of the fabricated cell was to demonstrate its logical functions rather than its speed. However, the cell has no intrinsic speed limitations compared to conventional MOS logic.

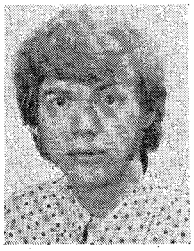
References

[1] W. L. McDermid and H. E. Peterson, "A magnetic associative memory," *IBM J. Res. Develop.*, vol. 5, pp. 59–63, Jan. 1961.
[2] R. W. Ahrons, "Superconductive associative memories," *RCA Rev.*, vol. 24, pp. 325–354, Sept. 1963.
[3] R. Igarashi and T. E. Yaita, "An integrated MOS transistor

associative system with 11 ns 100 ns cycletime," in *1967 Spring Joint Comput. Conf., AFIPS Conf. Proc.*, vol. 30. Washington, D.C.: Thompson, pp. 499–506.
[4] R. F. Herlein and A. V. Thompson, "An integrated associative memory element," in *ISSCC Dig. Tech. Papers*, Feb. 1969, pp. 42–43.
[5] J. T. Koo, "Integrated-circuit content-addressable memories," *IEEE J. Solid-State Circuits*, vol. SC-5, pp. 208–215, Oct. 1970.
[6] J. L. Mundy, J. F. Burgess, R. E. Joynson, and C. A. Neugebauer, "Low-cost associative memory," *IEEE J. Solid-State Circuits*, vol. SC-7, pp. 364–369, Oct. 1972.
[7] D. Frohman-Bentchkowsky, "The metal-nitride-oxide-silicon (MNOS) transistor characteristics and applications," *Proc. IEEE*, vol. 58, pp. 1207–1219, Aug. 1970.
[8] K. I. Lundström and C. M. Svensson, "Properties of MNOS structures," *IEEE Trans. Electron Devices*, vol. ED-19, pp. 826–836, June 1972.
[9] L. Lundkvist, I. Lundström, and C. Svensson, "Discharge of MNOS structures," to be published.
[10] L. G. Carlstedt and C. M. Svensson, "MNOS memory transistors in simple memory arrays," *IEEE J. Solid-State Circuits*, vol. SC-7, pp. 382–388, Oct. 1972.

Gunnar Carlstedt was born in Göteborg, Sweden, on March 5, 1946. He received both the M.Sc. degree in electrical engineering, and the Ph.D. degree from the Chalmers University of Technology, Göteborg, Sweden, in 1969 and 1973, respectively.

Since 1969 he has been a member of the Research and Teaching Staff of the Research Laboratory of Electronics at Chalmers University of Technology. He has been a consultant for the Swedish Computer manufacturer SAAB-Scania. His research has been concerned with the architecture of both small and large computers and memory technology.



Göran P. Petersson was born in Karlskrona, Sweden, on January 26, 1946. He received the M.Sc. degree in electrical engineering from the Chalmers University of Technology, Göteborg, Sweden, in 1969.

Since 1969 he has been a member of the Research and Teaching Staff of the Research Laboratory of Electronics at Chalmers University of Technology. His research has been concerned with MOS and MNOS physics and MNOS memory applications.

Kjell O. Jeppson was born in Göteborg, Sweden, on June 3, 1947. He received the M.Sc. degree in electrical engineering from the Chalmers University of Technology, Göteborg, Sweden, in 1970.

Since 1970 he has been a member of the Research and Teaching Staff of the Research Laboratory of Electronics at Chalmers University of Technology. His research has been concerned with MNOS physics and memory applications.

A Bipolar Four-Phase Dynamic Shift Register

D. KASPERKOVITZ

***Abstract*—A 512-b dynamic shift register is integrated on 6.4-mm² active chip area. The frequency range is from 100 Hz to 3 MHz. At 1 MHz the power dissipation is 20 mW. The performance of the shift register is insensitive to spread in process parameters because the information is regenerated in each cell. The comparison of the measured and the calculated working range shows the essential influence of all parasitic capacitances.**

Introduction

GENERALLY speaking, there are two different ways to increase the bit density of integrated parallel and serial memories. Firstly, technology can be improved resulting in smaller tolerances and therefore in smaller chip areas of integrated components and memory cells. Secondly, within a given technology with more or less fixed layout rules the circuit designer can try to find a circuit with the required electrical function that needs as few components per cell as possible. For dynamic shift registers fabricated with standard bipolar technology the "best" circuit with the least number of components per cell is already known for years. It is the bucket-brigade delay line [1], where each cell consists of only one transistor.

However, if bipolar bucket-brigade delay lines are used as long shift registers for digital information, one has to include discriminating and amplifying stages after 30 to 100 cells [2]. This is due to the systematic fault in charge-deficit transfer from one bucket to the following, caused partially by the base current of the bipolar transistor. An integrated discriminating and

Manuscript received March 22, 1973; revised June 15, 1973.
The author is with the Philips Research Laboratories, Eindhoven, The Netherlands.