H. S. Lee

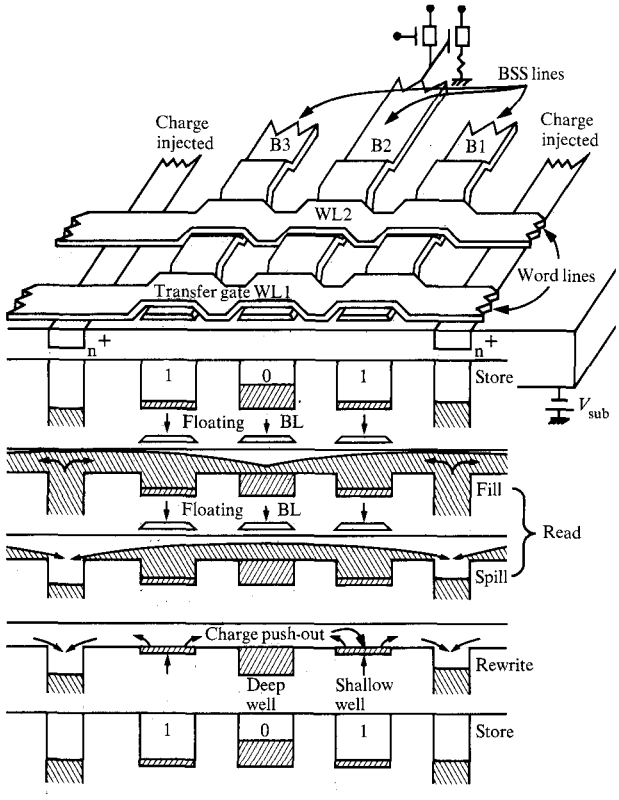
Analysis of the Merged Charge Memory (MCM) Cell

Abstract: This paper describes a new MOS dynamic RAM (Random-Access Memory) cell which utilizes a merged surface charge transistor structure. The merged charge memory (MCM) cell uses a polysilicon electrode as both a bit sense line and common plate for a column of storage capacitors. The MCM structure is self-aligned, contactless and free of closely spaced p-n junctions. Its spatial density approaches the conceptual limit of the intersection formed by two orthogonal lines or $4W^2$ where W is the minimum geometry feature. The cell area utilization efficiency is improved because of this simplicity. Preliminary experimental results and ASTAP simulations based on the charge control equivalent circuit for a dynamic potential well are described. Implications for chip design constraints are discussed, and the advantages and limitations of MCM are highlighted where appropriate.

Introduction

"What is beyond the one-device cell?" This is the question confronting us now. If we follow the evolutionary path of the RAM storage cell, i.e., reducing the number of elements in the existing cell rather than resorting to smaller lithographic dimensions and voltage scaling of the existing cell, we must come to an inevitable choice of merging two of the three basic elements found in any one-device cell, namely, bit-sense line, word line, and storage electrode. Merging the word line and the storage electrode has already been demonstrated by A. F. Tasch et al. as the "charge coupled RAM cell" [1], but merging the *bit line* and the storage electrode results in the MCM cell, as demonstrated in this paper.

Figure 1 The MCM array. Charge is stored at the silicon surface under the intersection of the bit sense/storage line and the transfer gate word line. The surface potential wells vary, and the charge flows bidirectionally at different times in the read/rewrite cycle.



The application of surface charge transistor structures to the single-transistor random access memory cell has been demonstrated by Engeler et al. [2]. To further simplify such a structure, one approach is to use a single electrode bit storage/sense (BSS) line that performs the functions of both the storage capacitor electrode and the bit sense line of the single-transistor cell. Parallel to every N BSS lines is a diffusion stripe, which is introduced to serve as a *"fill and spill"* [3] charge injector to both the right and the left N storage cells, whenever one of the transfer gate word lines perpendicular to the BSS lines is turned on.

The salient features of the random access memory cell described in this paper are: more storage area for a given cell area, i.e., better cell area utilization efficiency, no contact holes, and no closely spaced diffused p-n junctions in the entire memory array. Thus, it is potentially suitable for a very high density memory.

In the following section, the principles of operation, a computer simulation, and preliminary experimental results on an 8×8 array of these cells are given. Some chip design considerations are discussed, with advantages and limitations pointed out as appropriate.

Memory array and description of operation

In the merged RAM [4] array, charge is stored as minority carriers in a potential well under the cross-over point of the word line and the BSS line as shown in Fig. 1. For clarity only three rows of charge storing potential wells between the diffusion injectors are shown. The charge storage potential wells on a particular BSS line are defined by regions of thin oxide; they are separated from the adjacent storage potential wells along the BSS line by intervening areas of thick oxide. During a write operation, the BSS line is either kept at the high voltage V_{BH} to create a deep potential well, or pulsed halfway down to $V_{BH}/2$ to create a shallow potential well, depending on whether a ZERO or ONE is to be written. The charge injected from the N^+ diffusion stripes is gated and set [3] by the word line voltage ($V_{WL} \approx V_{BH}/2$) into a succession of pulsed storage capacitors of deep and shallow potential wells, resulting in an array of pulsed charge puddles as shown in Fig. 1.

The signal ONE or ZERO is stored as a near-empty or a half-filled level, respectively, in the potential well at the storage node (the BSS line is kept at the high voltage V_{BH}). The information in a particular word is read by first floating the BSS lines and pulsing the diffusion stripes to ground. The word line is then pulsed on from a negative offset voltage to $V_{BH}/2$. As the injected charges from the diffusion stripes fill up the potential wells, output signals proportional to the net charge poured into the wells are induced in the corresponding BSS lines. The potential wells that are initially only slightly filled receive more electrons than the wells already half filled, and thereby cause a greater drop in the corresponding BSS line voltages. Thus the difference between the ONE and ZERO signal comes primarily from the difference in the initial stored charge levels or the associated surface potential differences. The storage potential wells are now all filled and indistinguishable; thus the readout is destructive.

To expedite the read operation and subsequent rewriting of the individual bits back into the storage potential wells, the *charge-equilibration method* [3] is used. During the read/rewrite time with the word line voltage V_{WL} on, all the potential wells under any combination of ONE and ZERO bit patterns must be *filled and set* [3] in the shortest possible time, irrespective of the distances from the charge injecting diffusion stripes. This is accomplished by first pulsing the diffusion stripes to ground ($V_s = 0$), (with the substrate voltage negative, $V_{sub} = -3$), to maximize the overdrive ($V_{WL} - V_s - V_{th}$) and reduce the time for charge injection. The threshold voltage is V_{th}. After all the potential wells have been overfilled, drainage of the excess charges in the entire channel is expedited by the N^+ diffusion stripes at both ends of the channel, which are pulsed back to reverse biased charge sinks. During the drainage time the charges remaining in the potential wells under all the BSS lines approach equilibrium with the barriers under the word line gate.

Since readout is destructive, a read/rewrite operation may be carried out by a gated flip-flop for the sense/refresh circuit. A word is rewritten by clamping each BSS line at the high voltage (V_{BH}) for a deep potential well (ZERO) or pulsed down to $V_{BH}/2$ for a shallow potential well (ONE), while the word line remains on. The *pushing out* [5] of charge associated with rewriting ONE is clearly shown in Fig. 1. Since a sense latch can amplify the voltage difference by pulling the BSS line which reads ONE toward $V_{BH}/2$, a read/rewrite operation automatically results in pushing charge out from a potential well storing ONE. After the excess charge has been drained from the channel, the word line voltage is turned off. The charge puddles are now isolated from each other. Finally, all the shallow wells are pulsed back (in case of a sense latch, this can be done by restoring BSS lines to V_{BH}) to the same bit voltage V_{BH} as those of deep wells for information storage as shown in Fig. 1. It must be emphasized that the drainage of the excess charges in the channel is necessary to avoid interference between the wells, the details of which are explained next.

• *Word line off voltage and interference*

During the storage time, each bit location receives disturbing pulses from write operations at bit locations sharing the same BSS line and the other word lines. An "off" word line must provide an adequate potential barrier margin to prevent charge from spilling out of a disturbed ZERO well (half filled when the BSS voltage equals V_{BH} but filled when the BSS voltage is $\approx V_{BH}/2$) to adjacent bits along the "off" word line. The criterion is shown in the following equation for the case when the thin oxide thicknesses under the word line and the BSS line are equal. The surface potential ψ_s of the pulsed well with a given magnitude of electron density q_n is given by the depletion approximation as [6, 7]

$$\psi_s = V_g - V_{FB} \frac{q_n + q_B}{C_{ox}} \approx \psi_{s0} - \frac{q_n}{C_{ox}}, \tag{1}$$

where

V_g = voltage applied to electrode,
V_{FB} = flatband voltage,
C_{ox} = thin oxide capacitance per unit area under bit line or word line in the channel,
$q_B(\psi_s)$ = bulk depletion charge magnitude per unit area, and
ψ_{s0} = the surface potential with no free electronic charge.

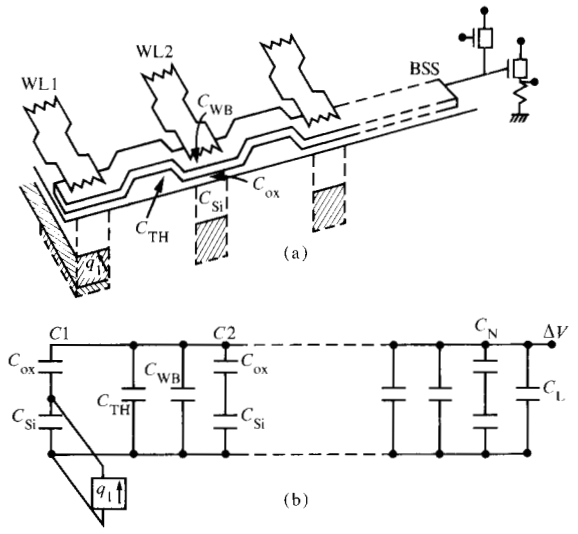


Figure 2 (a) Schematic cross-sectional view of the MCM along a bit storage/sense line. (b) A simple equivalent circuit of the floating bit storage/sense line.

Due to gating and setting by the word line voltage $V_{WL(on)}$ ($\approx V_{BH}/2$) and allowing sufficient time for drainage of excess carriers, the surface potentials underneath the BSS line and the word line must be equal. Thus from (1) the charge retained in the charge puddle at the end of write is

$$\left.\frac{q_n}{C_{ox}}\right|_{DEEP} = V_{BH} - V_{WL(on)} + \left.\frac{q_n}{C_{ox}}\right|_{WL(on)}, \tag{2}$$

and

$$\left.\frac{q_n}{C_{ox}}\right|_{SHALLOW} = V_{BL} - V_{WL(on)} + \left.\frac{q_n}{C_{ox}}\right|_{WL(on)}. \tag{3}$$

Here V_{BH}, V_{BL} ($\approx \frac{1}{2}V_{BH}$) are the bit line voltages for writing a deep and shallow well respectively, and $q_n|_{WL(on)}$ is the small residual charge in the channel at the end of incomplete write spill. Now, the condition for avoiding spillover by a pulse disturbance during store is

$$\psi_s(q_n = 0)|_{WL(off)} < \psi_s(q_n \text{ DEEP})|_{V_{BL}}. \tag{4}$$

From (1), (2), and (4), the required off voltage of the word line is

$$V_{WL(off)} < V_{BL} + V_{WL(on)} - V_{BH} - V_{th} < -V_{th}. \tag{5}$$

Typically $V_{WL(off)}$ is the negative of the threshold voltage.

Unlike the method just described, which avoids the charge spillover by providing additional potential barrier height through use of a negative off word line voltage, a combination of p^+ surface implantation and medium oxide thickness under the word line (which provides low surface potential, hence a potential barrier when word line is off at ground) can also be used together with a zero off word line voltage.

• *Induced bit line voltage and bit line capacitance*

A cross-sectional view of a BSS line in Fig. 2(a) identifies the several capacitances involved. A corresponding equivalent circuit for an N-bit, BSS line shown in Fig. 2(b) can be used to calculate the output voltage when a net charge q_1 is injected into a selected potential well. [An arbitrary bit pattern is stored in the rest of the potential wells ($j = 2$ to N) along the same BSS line.]

The analysis of this circuit [6, 8] yields the output voltage ΔV_{Bit} as

$$\Delta V_{Bit} = \frac{\dfrac{q_1 C_1}{C_{Si} C_{ox}}}{\dfrac{C_1}{C_{ox}} + N\left(\dfrac{C_{TH}A_{TH}}{C_{ox}A_{st}} + \dfrac{C_{WB}A_{WB}}{C_{ox}A_{st}}\right) + \dfrac{C_L}{C_{ox}} + \sum_{j=2}^{N} \dfrac{C_j(q_j)}{C_{ox}}}, \tag{6}$$

where

$$\frac{C_j(q_j)}{C_{ox}} = \left\{1 + \frac{2(V_g - V_{FB})C_{ox}^2}{qN_A\varepsilon_{Si}\varepsilon_0}\left[1 - \frac{q_j(V_g - V_{th})}{q_0(V_g - V_{FB})}\right]\right\}^{-\frac{1}{2}}, \tag{7}$$

$$q_0 = C_{ox}(V_g - V_{th}), \tag{8}$$

C_1 is the series capacitance of the oxide capacitance C_{ox} and the depletion capacitance C_{Si} of the selected potential well, and q_j is the electron charge density (for a near-empty or half-filled well) in the unselected jth storage well along the same BSS line. In order to increase the signal output, it is obvious from Eqs. (6) and (7) that C_{ox} should be maximized, while the depletion capacitance C_{Si}, the parasitic thick oxide capacitance $C_{TH}A_{TH}$, the parasitic capacitance between the bit/storage line and the word line $C_{WB}A_{WB}$, and the input capacitance C_L of the sense amplifier should be minimized.

Since the signal ONE or ZERO is stored as a near-empty or half-filled well (BSS lines at the high voltage V_{BH}), the variations of C_{Si} from cell to cell can be made negligible if lower substrate doping and thinner gate oxide is used. Thus the output signal variations, for sensing the same bit but with different bit patterns stored in the rest of the potential wells along the same BSS line, can be held within a few percent. One disadvantage of the merged BSS line structure is the fact that the bit sense line capacitance and the storage area capacitance are interdependent, resulting in an inflexible capacitive transfer ratio.

If charge storage area has to be maintained as small as possible, a very sensitive sense amplifier that can be laid out within the area limited by the bit pitch will be needed. To solve such a problem one can use a twin cell approach in which two adjacent potential wells per cell are used to store the information in a complementary form and the sensing is differential.

• *Charge transfer sensing and discussion*

The discussion in the previous section is based on voltage sensing of the floating bit line. However, with the use of the charge transfer preamplifier [9], it is possible to obtain a charge output comparable to the net incremental charge poured into the storage well during reading (i.e., the output signal can be made proportional to the storage area and can be made less sensitive to the BSS line capacitance, when a sufficient preamplification time period is allowed [10]). When combined with the better cell area utilization of the MCM structure (because of the structural simplicity, a large percentage of the cell area can be allocated to the storage area), it is then feasible to obtain signal outputs comparable to those of a one-device cell of the same cell area.

For a given set of layout ground rules it is obvious that the MCM cell area can be smaller than a one-device cell area. On the other hand, for a one-device cell to achieve the equivalent minimum cell area, it requires a more aggressive technology and a voltage scaling to alleviate limitations of dimensional scaling, such as "short-channel" [11] and "hot-electron" [12] effects.

Equivalent circuit and simulation

To simulate read/rewrite transients of a series of charged potential wells operating with arbitrary bit patterns, it is imperative to use equivalent circuits and a computer-aided circuit simulation program such as ASTAP.

The development of an equivalent circuit for a dynamic potential well together with the use of ASTAP simulation [13] enables one to analyze and design an MCM array with full knowledge of time-dependent voltage and current waveforms as well as charge transient waveforms. Concurrently, one can handle support circuit transients and parasitic capacitances associated with integrated circuit layouts. The remainder of this section highlights the results of such simulations, emphasizing behavior which is unique to MCM. Because dynamic operation of the potential well is inherent to MCM, an equivalent circuit model for the potential well itself is first described. This model is then incorporated into ASTAP to simulate an eight-bit section of the MCM.

To simulate dynamic conditions of a potential well underneath either the BSS line or transfer word line, the charge control model [7, 14, 15] is used. By using the gradual channel approximation, first the current across the potential well, the net mobile charge, and the net immobile depletion charge in the potential well are derived as functions of the terminal voltages. Subsequently, all the intrinsic inter-terminal nonlinear capacitances (which account for the displacement currents between the channel and gate as well as the channel and substrate) are evaluated as functions of the terminal voltages and are readily applicable to ASTAP (see the Appendix and Fig. 3). A nonlinear resistive switch is assumed for the inter-electrode region along the word line so as to decouple the depletion capacitance from those of the adjacent potential wells when their gate voltages are off. The value of the nonlinear resistor is assumed to vary, from a very small value to a very large value, as the gate voltage varies from above to below the threshold voltage during fall time.

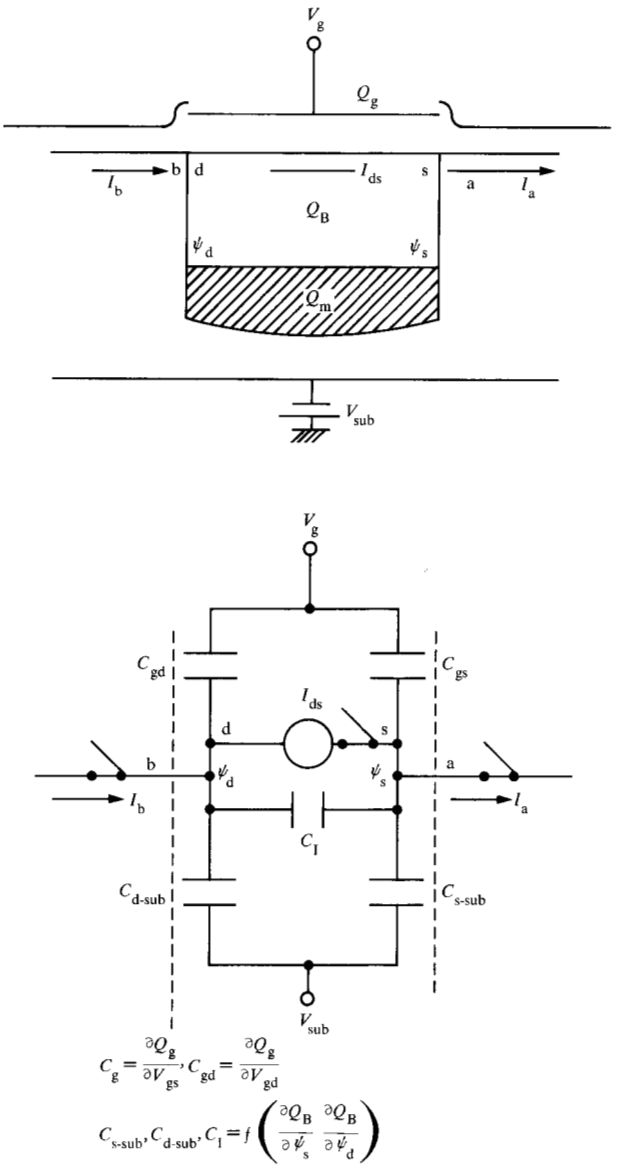


Figure 3 Charge control circuit model of a potential well.

In the MCM RAM the potential wells under the BSS lines are at all times storing some charge, either the residual charge (because of incomplete charge transfer during the "spill" phase of the write process) or half-filled charge. Thus, they possess a non-zero channel conductance (except during the power-up period) and a

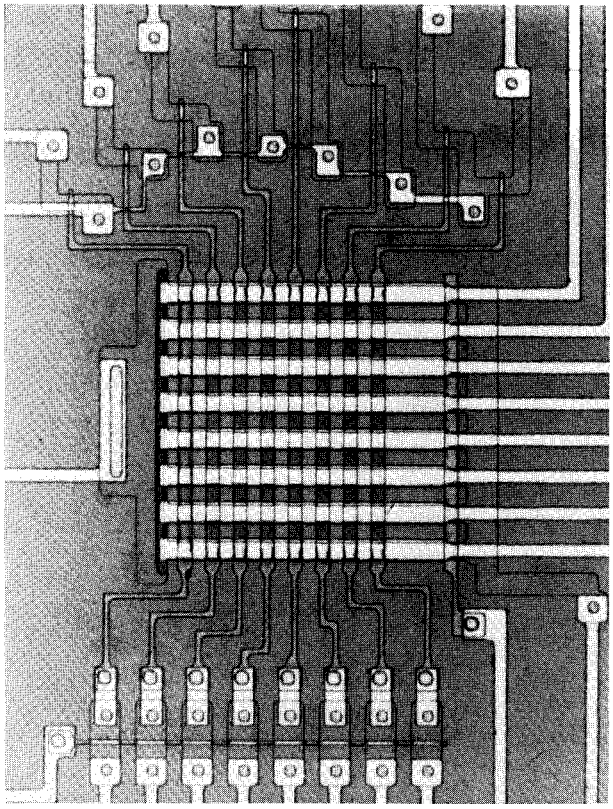


Figure 4 Photomicrograph of an 8 × 8 MCM test structure.

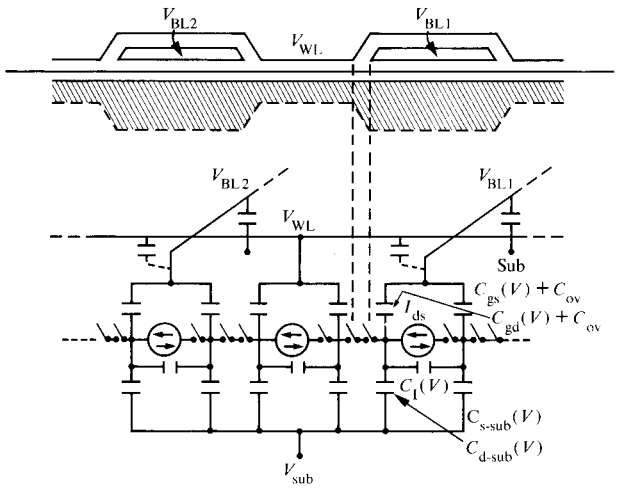


Figure 5 Equivalent circuit model for the MCM potential array used in ASTAP simulation.

negligible inversion layer time constant. Therefore, the MCM channel is readily represented in equivalent circuit format.

To validate the MCM concept and the circuit simulations, we processed a simple 8 × 8 array of MCM storage cells. A photomicrograph of the test structure is shown in Fig. 4. Each poly BSS line is connected to a charge sensing circuit which consists of a reset and a source follower FET.

As shown in Figs. 1 and 5, the equivalent circuit model is used to simulate both the dynamic behavior of the potential wells under BSS lines and the potential wells underneath the word line. The circuit model also includes the right and left n^+-p junction used as pulsed charge injectors, and eight charge sensing circuits connected to the eight BSS lines. All the circuit parasitic capacitances, such as the capacitance between the word line and the BSS line, the series capacitance of the storage thin oxide capacitance and the depletion capacitance of the unselected potential well, and a lumped capacitance of the bit line (eight bits long) are included. The capacitances associated with the layout of the sense circuit and the nine-pF scope-probe capacitance were also included in the ASTAP simulation. The read/rewrite cycle is explained in relationship to the timing diagram shown in Fig. 6, which shows the word line voltage, charge injector voltage, reset gate voltage, and examples of BSS line voltages (bit line 1 and bit line 6) for read/rewrite ONE and ZERO, respectively. Reading is performed by first turning off the reset gate voltage, floating the BSS line, and pulsing the charge injecting diffused lines to ground. Because of the substrate bias, the injecting diodes are still reverse biased so that charges can flow only under the selected word line. After turning on the selected word line and inundating all the potential wells, irrespective of the well position and bit pattern, the injecting diodes are strongly reverse biased. This allows the excess charge to flow out from the entire channel until the surface potential underneath the bit storage lines approaches those under the word line. The corresponding signals for the stored ONE and ZERO appearing on the bit lines during the fill and spill process are also shown in Fig. 6.

As mentioned, readout is destructive because the potential wells are all half filled (ZERO) and indistinguishable after a read operation. The rewrite operation is essentially a "selective spill" process. As we see in Fig. 6, the reset is first turned on and the bit lines are either held at 8.5 V or pulsed down to 4.5 V, depending on whether we are writing a ZERO or a ONE. Selected shallow potential wells ($V_g = 4.5$) now push out excess charges (similar to the CCD push-out effect) and these charges are drained off by the strongly reverse biased diodes at the ends. The word line is then turned off, which isolates the potential wells, and the shallow potential wells are pulsed back to the V_{BH} store condition.

Thus we have simulated the charge transfer characteristics of an eight-bit experimental test array with spatially and temporally varying bit patterns.

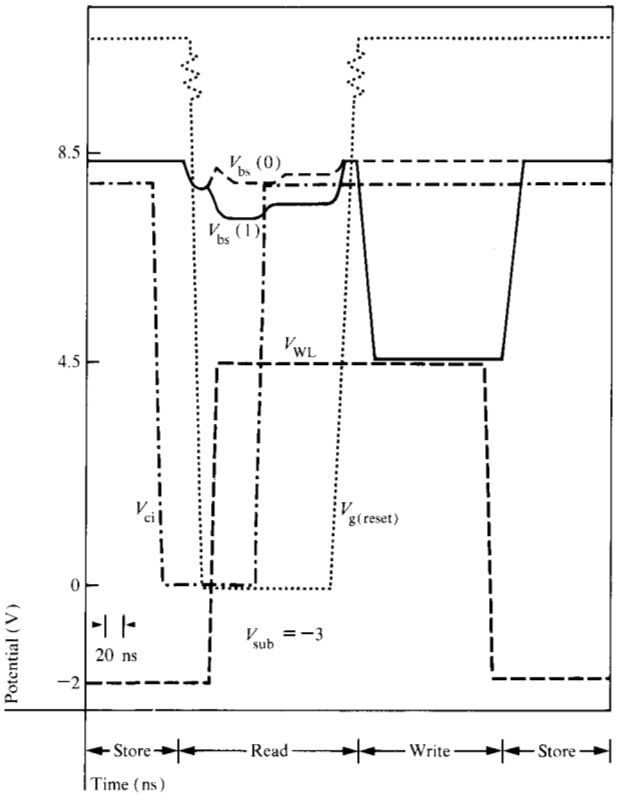


Figure 6 Read/rewrite timing diagram for operating the MCM test structure. The BSS line voltage waveforms show read ZERO, write ZERO, and read ONE, write ONE operation.

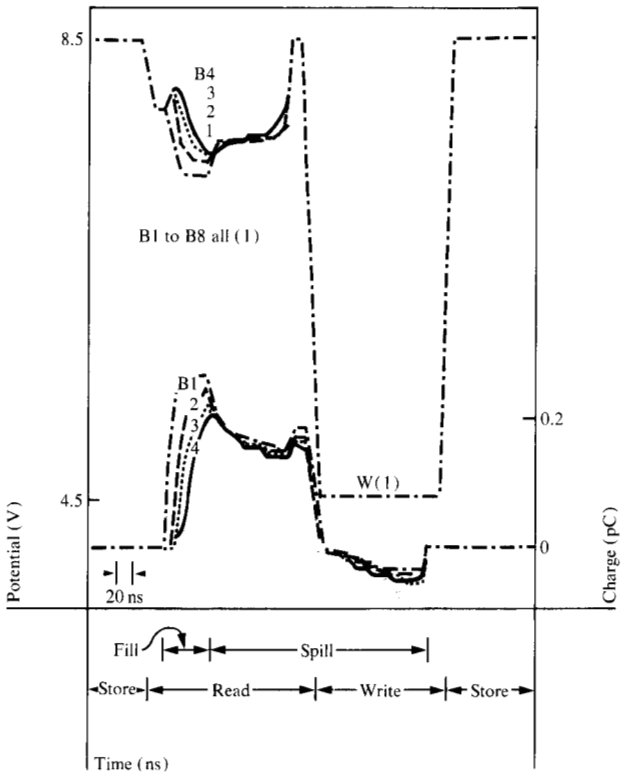


Figure 7 Simulation of the read/rewrite string of ONEs and the corresponding transient waveforms of the incremental charges ΔQ_i in the potential wells.

• *Charge transfer "read fill" time*

Since the ONE signal is stored as a near-empty potential well, reading a ONE requires more charge to be injected from the charge injecting diffusions than does reading a ZERO. Thus the longest charge transfer time is associated with reading a series of stored ONEs. Figure 7 shows an example of simulating the read/rewrite process of eight ONEs. Since electrons are injected from diffusions at both ends of the bit string, only bits 1 to 4 are shown; bits 8 to 5 are mirror images. The transient waveforms of the BSS lines during the read/rewrite process as well as the incremental charges ΔQ_i if the ith bit during the time the word line is on are also shown in Fig. 7, where

$$\Delta Q_i(t) = \int_{\text{WL(on)}}^{t} [I_{i\text{L}}(t) - I_{i\text{R}}(t)]\, dt. \tag{9}$$

Here $I_{i\text{L}}(t)$ = current flowing out of the ith BSS well at the left edge and $I_{i\text{R}}(t)$ = current flowing into the ith BSS well at the right edge.

The overetched experimental device dimensions (all in μm) which were used in the simulation include a BSS channel length of 6.35, a channel width of 9, a transfer word line channel length of 8.38, and an eight-bit array channel length of 127. The time allocated for read fill is the time between the turn-on of the word line and the time at which the diffusion stripes are pulsed back to high draining voltage; it is 50 ns. The allocated "read spill" time (for draining excess charges) is 85 ns, while the rewrite time (selected spill) is 125 ns. The time at which each BSS potential well receives a charge signal Q_{sig} equal to

$$Q_{\text{sig}} \approx A_{\text{st}} C_{\text{ox}} (V_{\text{BH}} - V_{\text{WL}}) \approx 0.16 \text{ pC} \approx 10^6 \text{ electrons} \tag{10}$$

is plotted against the square of the effective channel lengths from diffusion to BSS lines 1 through 4 (8.38, 23, 37.8, and 52.6 μm, respectively) as shown in Fig. 8. The boundary condition at the injecting diffusion is similar to that of IGFETs so that the time delay is expected to be proportional to $L^2/\mu(V_{\text{WL}} - V_{\text{s}} - V_{\text{th}})$ [16]. The linear relationship shown in Fig. 8 confirms this L^2 dependence.

Since the allocated read fill time (50 ns) is longer than the receiving time for the bit 4 charge signal, excess charges are introduced along the entire channel. However, they are drained in the subsequent spill time. This is clearly shown in the incremental charge and floating bit line voltage waveforms in Fig. 7.

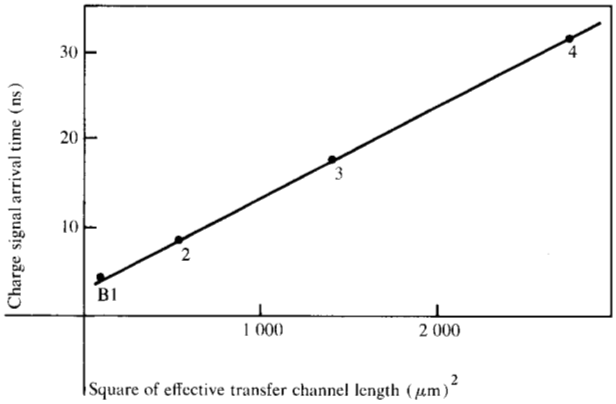


Figure 8 Charge signal arrival time vs L^2 for bits 1 to 4. The pattern is all ONEs.

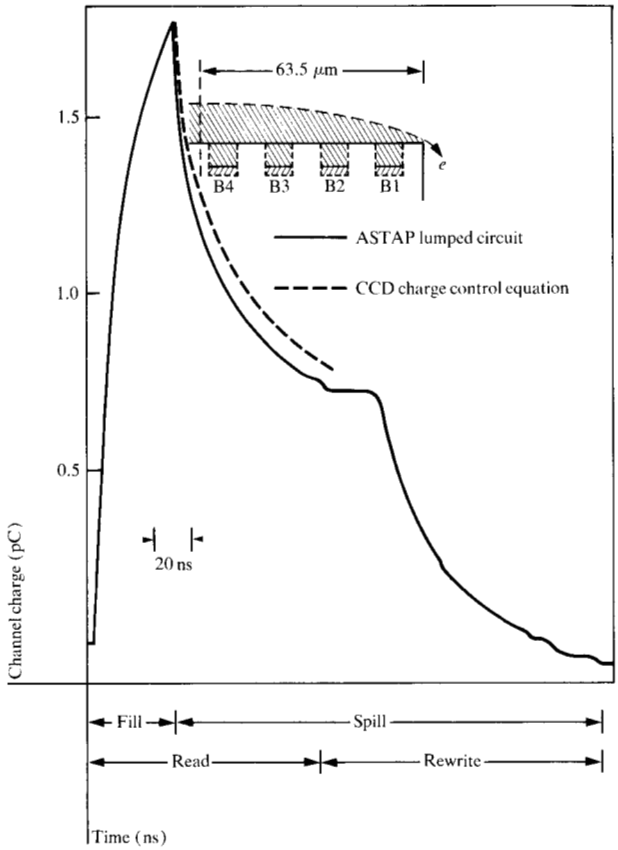


Figure 9 Incremental channel charge transient. The inset shows the boundary condition of the spilling of excess charges.

• *Charge transfer read spill time*

During the read spill time the excess charge Q_{ex} in the channel is drained at both ends. In the case of all ONEs the excess charge distribution is symmetric with respect to the center line between bit 4 and bit 5. Thus, the boundary condition becomes similar to that of CCD, and the CCD charge transfer equation (6) can be applied to estimate the read spill time. The equation is

$$\frac{Q_{(t)}}{Q_{ex}} = \frac{e^{-(t/\tau_p)}}{1 + K(1 - e^{-t/\tau_p})}, \tag{11}$$

where

$$\tau_p = \frac{4}{\pi^2}\frac{L^2}{D},$$

$$K = \frac{Q_{ex}}{WL}\frac{\mu}{2D\,C_{ox}},$$

D is the diffusion constant, μ is the mobility, and L is half of the total channel length $(127\ \omega)/2$.

$$Q_{ex} = \int_{WL(on)}^{t_{EF}} [I_{4L}(t) - I_{Cl}(t)]\,dt - 4Q_{sig}, \tag{12}$$

where t_{EF} is the end of read fill time, $I_{Cl}(t)$ is the current flowing into the channel at the diffusion, and $I_{4L}(t)$ is the current flowing at the left edge of the fourth BSS well. The value of the integration is obtained from the ASTAP simulation.

The dependence of the incremental channel charge on read spill time is calculated by using Eq. (11) and is compared with the ASTAP simulation. Figure 9 shows a reasonably close agreement between the two methods, and supports the usefulness of the "lumped charge potential well" circuit model for ASTAP simulation. At the end of the allocated read spill time, some excess charges are still present. However, they are further drained in the subsequent rewrite transient.

• *Surface potential and BSS line voltage waveforms*

In the lumped-charge circuit model shown in Fig. 3, the surface potentials at the edges of a potential well are described by the voltages across $C_{s\text{-}sub}$ and $C_{d\text{-}sub}$. For potential wells under the BSS line, these voltages are equal, due to the presence of stored charges.

Figure 10 depicts the simulated equivalent surface potential at the right edge of the potential wells for bits 1 to 4. The difference in the surface potentials during "store" time for ONE and ZERO is approximately 3.6 V, which is less than the maximum possible difference of 4 V$(V_{BH} - V_{WL})$. Also, a slight position dependence of the surface potential is seen; for instance, the fourth well has slightly lower potential than the first well even though they are both storing a ONE. Similarly, the third and second wells are storing a ZERO, but the third well has lower potential than the second well. These results are consistent with the fact that an incomplete spill in the previous cycle results in more residual charge left for potential wells farther from the charge-injecting diffusions.

In the read fill operation all the wells are inundated, and the surface potentials under the wells approach the same value as that of the injecting diffusion. Since the diffusion is pulsed down to zero and the substrate is at −3 V, the surface potential with respect to the substrate approaches 3 V. In the subsequent read spill and rewrite operations the surface potentials of the BSS wells are eventually set by the surface potential under the word line. After rewrite, the bit lines are all pulsed to 8.5 V and return to the original store condition, concluding a read/rewrite cycle.

The BSS line voltage waveform is shown in Fig. 11. At the beginning of a read fill operation the reset gate turns off, causing an initial drop on all the floating bit lines. As soon as the word line turns on, electrons are injected into the channel. Since bits 1 and 4 are storing ONE, and are nearly empty, voltages on the BSS lines associated with them decrease the most. The change in the floating BSS voltage is proportional to the net electron charge under the gate. The longer time delay of the bit 4 response is, as described before, due to the longer transit time.

Bits 2 and 3 are storing ZERO and their potential wells are half filled. As the word line turns on, their BSS voltages increase first, due to the capacitive coupling. The corresponding incremental charge transients $Q_2(t)$ and $Q_3(t)$ show some charge slopping [17]. Subsequently, the excess charge is introduced, and the BSS voltages decrease to some value above the BSS floating voltage at the initial droop. At the end of the read fill operation the difference in the BSS floating voltages for reading ONE and ZERO is fully developed, as shown in Fig. 11.

During the read spill time, all the floating BSS line voltages increase due to the excess charge transferring out of the entire channel and spilling into the strongly reverse biased end diffusions. The increase of the BSS line voltages is the same, since the net excess charges under the BSS lines are equal, and the increment of the floating BSS line voltage is proportional to the net excess charge transferred out of the well.

During rewrite, the reset gates are turned on and the BSS lines, which are now tied to the bit drivers, assume the writing voltages as shown in Fig. 11.

• *Source-follower output waveforms*

Figure 12 shows the simulated source-follower output waveforms of bits 1 to 4. They substantiate the experimental results which follow. The outputs are taken from a one-KΩ resistor in parallel with a nine-pF probe capacitance. Figure 13 shows the simulation of a differential output from the twin cell operation. Since the bit pattern used for the single cell is 10011001, the bit pattern for the twin cell operation becomes 0′1′0′1′. These waveforms again resemble the experimental results that follow.

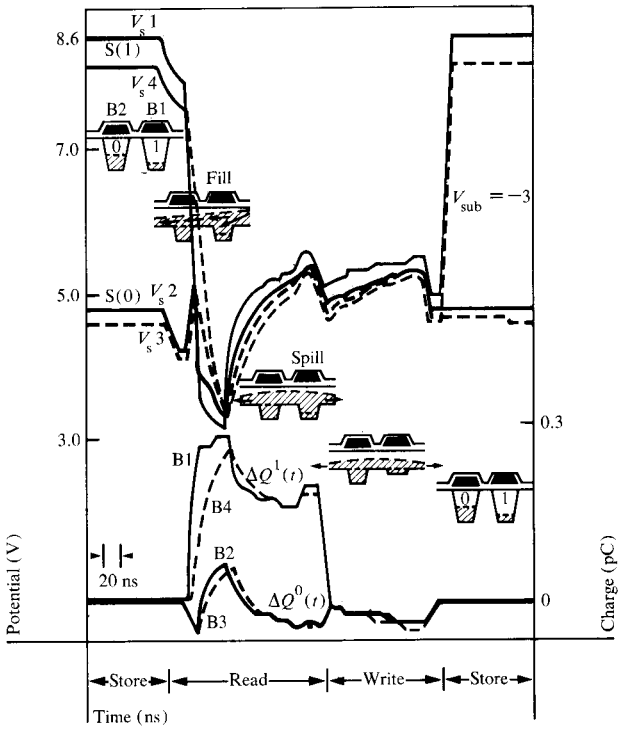


Figure 10 Transient surface potential voltages for bits 1 to 4. Spatial bit pattern for eight bits is 10011001. The corresponding incremental charge transients are also shown.

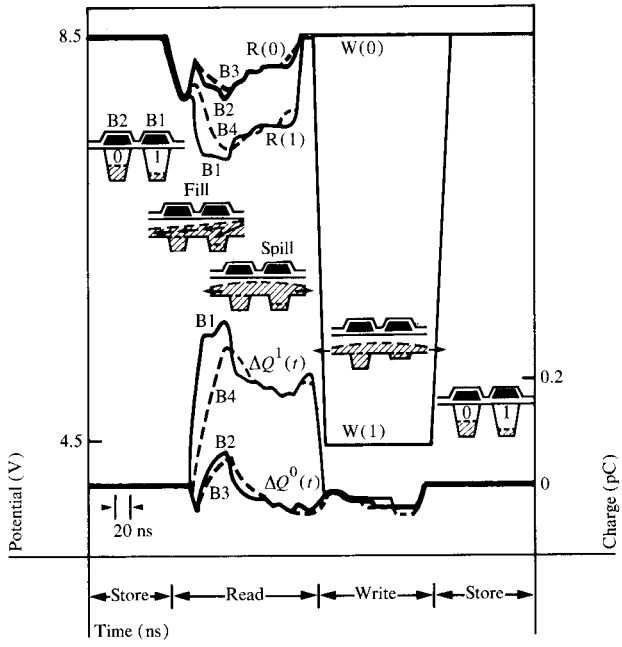


Figure 11 BSS line voltages and incremental charges for bits 1 to 4.

Experimental results

A photomicrograph of an 8 × 8 array of the MCM test structure is shown in Fig. 4. Each poly BSS line is con-

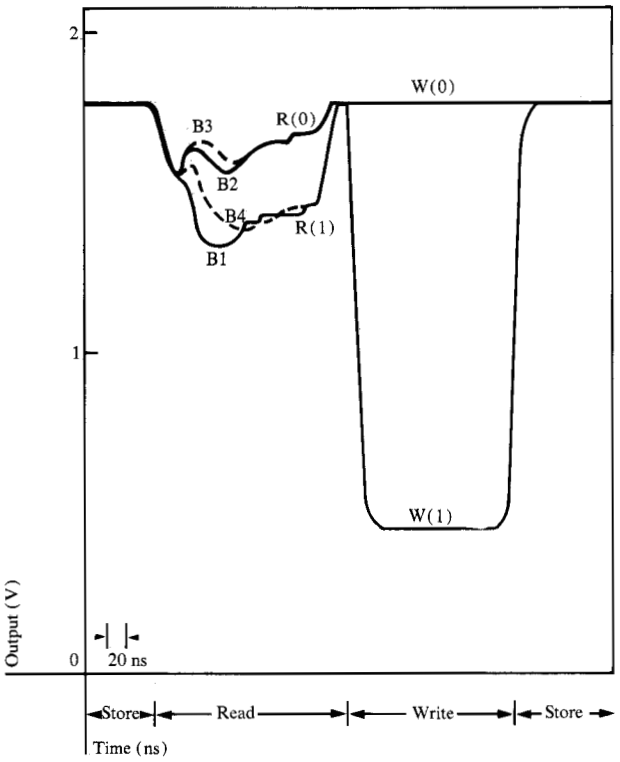
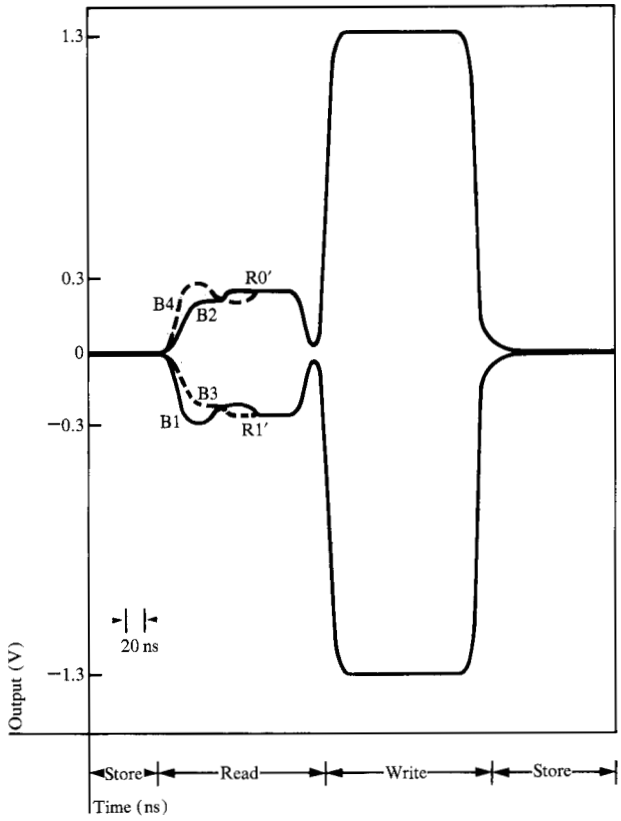


Figure 12 Source follower outputs from bits 1 to 4.

Figure 13 Source follower outputs for twin cell operation.



nected to a source follower sensing circuit at the lower end, and to a reset FET at the top. All eight BSS lines share a common reset gate, which can either float all BSS lines simultaneously for the read operation, or connect them to individual bit drivers for writing. The eight word lines are individually accessible. They gate and set the charges from the charge injecting diffusions into the MCM array. A control gate is made for the left charge-injecting diffusion for testing double-ended vs single-ended charge injections. All the BSS lines are designed to have the same parasitic capacitances. The area of the cell is 20 μm × 15 μm or 300 μm^2, the storage capacitance area is 65 μm^2, and the corresponding capacitance is approximately 0.05 pF. The layout of the cell and the process employed were by no means optimized. An n-channel self-aligned silicon gate process using a 2 Ω-cm p-type substrate and a gate oxide of 50 nm (500 Å) was available and was therefore used.

Characteristic waveforms for MCM operation are shown in Fig. 14. The top trace shows the reset gate pulse; the second trace shows the pulse applied to the diffusion stripes; the third shows the pulse applied to a bit line in the time sequence of write ONE, write ZERO, and write ONE; the fourth shows the word line voltage waveform. The bottom trace shows the output from the corresponding source follower, which monitors the voltage waveform on the bit line. In time sequence it shows the write ONE/read ONE/write ZERO/read ZERO and write ONE/read ONE transient. An enlarged output waveform is shown in Fig. 15(a). From the measured gain of the source follower, the bit line voltage difference for ONE and ZERO is estimated at 600 mV. Figure 15(b) shows the output waveform for the read ONE/rewrite ONE/read ONE/rewrite ONE operation. Notice the close similarity to the ASTAP simulation given in Fig. 12.

By introducing the disturb pulses (pulse down to 4.5 V) to simulate writing on adjacent word lines during the store condition and subsequently reading out the disturbed cell, it was confirmed that the negative offset (−2 V) for the word line off voltage was very effective in avoiding charge spillover to adjacent bits along the word line. (See the analysis in the section entitled "Word line off voltage and interference.")

Tests were made for comparing dc charge-injection (fill only) vs pulsed charge injection (fill and spill) from the diffusion stripes. With the disturb pulse operating during store, and operating with time-wise as well as space-wise bit pattern variations, the difference in bit bit line voltage between a reading for ONE and that for ZERO was measured as a function of the variation in injected diffusion voltage. For fill and spill injection, the bit line voltage differences were consistent for diffusion voltage variations from 3 V (threshold cutoff) down to 0 V. Under dc injection the ONE and ZERO difference

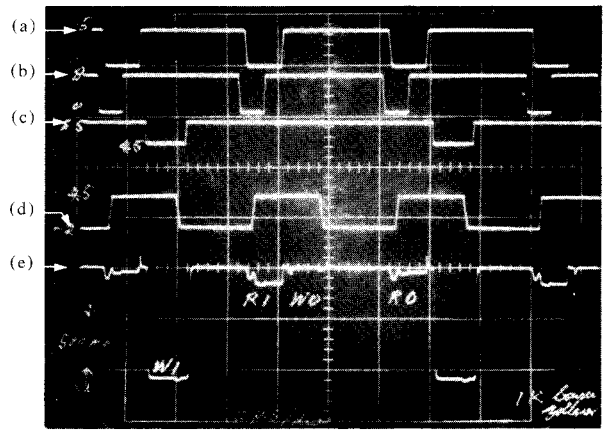
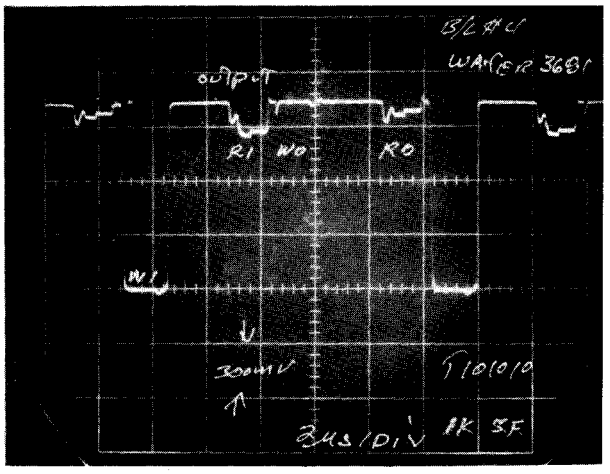
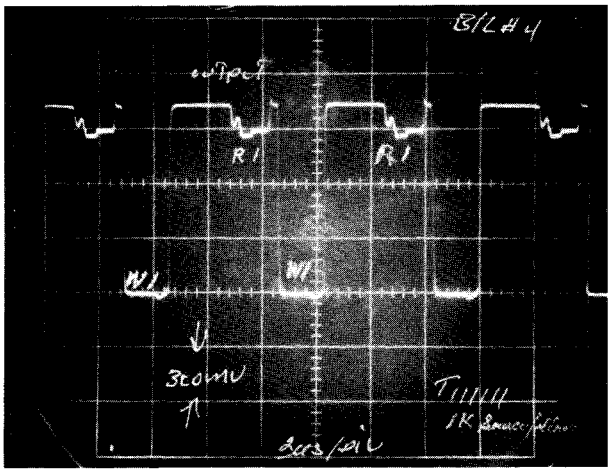


Figure 14 Trace (a)—reset gate voltage (0 to 15 V); trace (b)—charge injecting diffusion (0 to 8 V); trace (c)—bit line voltage (4.5 to 8.5 V); trace (d)—word line voltage (−2 to 4.5 V); and trace (e)—the source follower output (bottom trace 500 mV/div). Horizontal scale 2 μs/div. Write ONE read ONE, write ZERO read ZERO timing.



(a)



(b)

Figure 15 (a) Source follower output for write ONE read ONE, write ZERO read ZERO timing, 300 mV/div., 2 μs/div. (b) Same as (a) above except write ONE read ONE write ONE read ONE timing.

existed near the cutoff but quickly degraded as the injected diffusion voltage decreased toward 0 V. In fact, near 0 dc injection voltage, the bit line voltages for reading ONE and ZERO were either equal or changed into reading the opposite. This failure in reading the signal for the dc injection is expected since the excess charge in the channel can readily spill into the adjacent wells if a disturb occurs during the store condition.

In twin cell operation, a bit is stored as complementary conditions on two adjacent patented wells. In Fig. 16, the second and third traces are bit line voltages on bits 5 and 6. Notice the disturb pulse applied immediately after the write ZERO operation. The output shows the read 0′/write 1′/disturb/read 1′/write 0′/disturb/read 0′ time sequence. The twin MCM operation is superior to single MCM in providing twice the signal level and better common noise rejection. In addition, it does not require reference voltage generation for sensing.

Discussion

The MCM structure is self-aligned, contactless, and free of closely spaced p-n junctions, thereby suggesting higher yields. Its spatial packing density approaches the conceptual limit of the intersection formed by two orthogonal lines. However, there are constraints: reduced charge storage capacity per unit area, inflexible storage capacitance C_{st} to BSS line capacitance C_{bs} ratio, and the extreme pitch constraints on the sense amplifiers. For comparable operating voltages and gate oxide thickness, the MCM charge storage capacity per unit area is approximately half that of the one-device cell. This is because MCM operates on a half-filled well concept whereas the one-device cell operates with a near-full storage well. The poor C_{st}/C_{bs} ratio, i.e., a rapid increase in the cell area for an increase in the C_{st}/C_{bs} ratio, is the price paid for the merging, although this can be alleviated by some innovations in the layout and by using light substrate doping and reducing parasitic capacitances. The merging of the bit line and the storage line into a BSS line also imposes an extreme pitch constraint on the sense amplifiers. At the time of writing, it seems there is no sense amplifier which is both small enough to fit into the pitch, and sensitive enough to detect worst case signals for a very long BSS line. However, the bit pitch constraint may be solved by having two BSS lines sharing a sense amplifier in a new chip organization or by

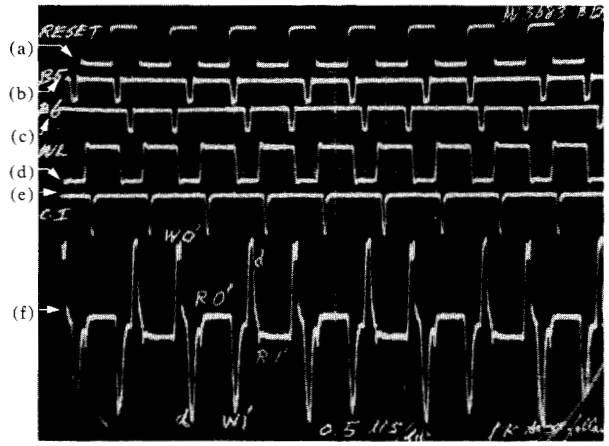


Figure 16 Twin cell operation. Trace (a)—reset gate voltage (0 to 15 V); trace (b)—BSS line 5 voltage (4.5 to 8.5 V); trace (c)—BSS line 6 voltage (4.5 to 8.5 V); trace (d)—disturb pulse follow after write 0 (half-filled well) in BSS line 5 and 6, word line voltage (−2 to 4.5 V); trace (e)—charge injecting diffusion (0 to 8 V); and trace (f)—source follower output. Time sequence write 0, disturb, read 0′, write 1′, disturb, and read 1′. Horizontal scale: 0.5 μs/div.

twin MCM operation as mentioned before. With the advent of the charge transfer preamplifier [9], it seems that some of the constraints mentioned above can be alleviated as mentioned early in the text.

Summary

A theoretical analysis, and equivalent circuit for a bilateral dynamic potential well and the associated ASTAP circuit simulation, as well as the experimental results for validating the MCM concept, have been presented. The MCM cell provides a spatial packing which approaches the conceptual limit of the intersection formed by two orthogonal lines. The constraints associated with such a cell have been described and the possible improvements discussed.

Acknowledgments

The author is indebted to W. D. Pricer for his original conception of MCM. Also, the author is grateful to B. Agusta and J. A. Hiltebeitel for their encouragement; to J. B. Dinklage and J. M. Wursthorn for process support; to L. G. Heller and R. R. Troutman for their comments; and to M. L. Plantier for his help in testing and instrumentation.

Appendix

• *Lumped charge equivalent circuit of a bidirectional potential well*

Since the response time for electron redistribution in a potential well is just the dielectric relaxation time of the inversion layer and is much shorter than the time constants of external circuits, the current flow, the charge storage, and the surface potential can be approximated by a charge control or lumped-charge model [7]. In fact, analytical solutions of charge-transfer in charge transfer devices (CTDs) based on the charge control concept have been shown to compare well with numerical solutions [7, 14, 15, 18]. Under dynamic conditions, in which charge is physically transferred from one region to another, one can approximate the detailed character of the transfer by a four-terminal equivalent circuit represented by voltage-dependent transfer conductance, as shown in Fig. 3. Any waveforms can be applied to the V_g terminal and substrate bias terminal V_{sub}. The virtual drain and source terminals d and s are at the edges but within the potential well.

The voltage-dependent distributed gate-to-channel capacitance is represented by two lumped capacitances as

$$C_{gs} = \frac{\partial Q_g}{\partial V_{gs}}, \quad C_{gd} = \frac{\partial Q_g}{\partial V_{gd}},$$

where Q_g is the total mobile charge (electrons) Q_m plus total depletion charge Q_B in the potential well, and V_{gs}, V_{gd} are the differences between the gate voltage and the instant surface potentials at the s and d points within the potential well.

The effect of the total immobile depletion charge Q_B within the potential well can be accounted for by first calculating

$$\frac{\partial Q_B}{\partial \psi_{ss}}, \quad \frac{\partial Q_B}{\partial \psi_{sd}},$$

and then decomposing the results into the voltage-dependent capacitances $C_{d\text{-}sub}$, $C_{s\text{-}sub}$ and C_I as shown in Fig. 3. Here ψ_{ss} and ψ_{sd} are the surface potentials at the s and d points, respectively. Since the capacitances associated with the mobile and the depletion charges are lumped at the edges, the gradual channel approximation may be assumed to be valid for the current I_{ds} across the potential well. The current I_b flowing into point b and the current I_a flowing out at point a of the potential well are, in general, different. The net mobile charge Q_m in the potential well at any instant can then be calculated by integrating the difference between I_b and I_a with respect to time.

We will now assume a simple equivalent circuit to represent a gapless overlapping inter-electrode region. The circuit consists of a) the overlapping capacitances C_{ov} at the edges, which are included into the C_{gd}, C_{gs} lumped capacitances, and b) a nonlinear resistor which varies from a very small value to a very large value as the gate voltage varies from above to below the threshold voltage. This nonlinear resistive switch is needed to decouple the depletion capacitance from those of the ad-

jacent potential wells when their gate voltages turn off. Since the charge transfer time is basically limited by the transfer of charges across the potential well, the error in approximating the self-induced field in the inter-electrode region by the simple equivalent circuit mentioned above is expected to be small. By following the basic assumption of the charge-control or lumped-charge model [7] I, Q_m and Q_b are derived in the following as functions of terminal voltages only. Substitution of the surface potential equation (1) into the electron current expression gives

$$I_{ds} = -W\mu C_{ox}\left(\frac{kT}{q} + \psi_{s0} - \psi_s\right)\frac{d\psi_s}{dx}. \tag{A1}$$

By using the gradual channel approximation and integrating Eq. (A1) over the length L and expressing I_{ds} in terms of the surface potentials at the edges ψ_{ss} and ψ_{sd}, we get

$$I_{ds} = \frac{W\mu C_{ox}}{L2}\left(\psi_{sd} - \psi_{ss}\right)\left(\psi_{sd} - \psi_{ss} - 2\psi_{s0} - \frac{2kT}{q}\right), \tag{A2}$$

and

$$\frac{d\psi_s}{dx} = \frac{(\psi_{sd} - \psi_{ss})\left(\psi_{sd} + \psi_{ss} - 2\psi_{s0} - 2\frac{kT}{q}\right)}{2L\left(\psi_s - \psi_{s0} - \frac{kT}{q}\right)}. \tag{A3}$$

The total mobile charge in the potential well is then

$$Q_m = W\int_0^L q_m dx = WC_{ox}\left[\psi_{s0}L - \int_{\psi_{ss}}^{\psi_{sd}} \frac{\psi_s}{\frac{d\psi_s}{dx}} d\psi_s\right]. \tag{A4}$$

After integrating and substituting the relations

$$V_{gs} = V_g - \psi_{ss}, \tag{A5a}$$

$$V_{gd} = V_g - \psi_{sd}, \tag{A5b}$$

and

$$\psi_{s0} \approx V_g - V_{th}, \tag{A5c}$$

we get

$$\begin{aligned} Q_m = {} & \tfrac{2}{3} WLC_{ox}\Big\{(V_{gs} - V_{th})^3 - (V_{gd} - V_{th})^3 \\ & + \tfrac{3}{2}\frac{kT}{q}\left[(V_{gs} - V_{th})^2 - (V_{gd} - V_{th})^2\right]\Big\} \\ & \div \Big\{(V_{gs} - V_{th})^2 - (V_{gd} - V_{th})^2 \\ & + \frac{2kT}{q}\left[(V_{gs} - V_{th}) - (V_{gd} - V_{th})\right]\Big\}. \end{aligned}$$

For reasons of simplicity, C_{gs} and C_{gd} are derived from Q_m instead of Q_g.

By neglecting the kT/q terms, and differentiating with respect to V_{gs} and V_{ds}, the lumped capacitances are for $V_{gs} - V_{th} > V_{ds} \geq 0$.

$$C_{gs} = \tfrac{2}{3} WLC_{ox}\left\{1 - \left[\frac{V_{ds} - (V_{gs} - V_{th})}{V_{ds} - 2(V_{gs} - V_{th})}\right]^2\right\}, \tag{A7a}$$

and

$$C_{gd} = \tfrac{2}{3} WLC_{ox}\left\{1 - \left[\frac{V_{gs} - V_{th}}{V_{ds} - 2(V_{gs} - V_{th})}\right]^2\right\}, \tag{A7b}$$

where V_{th} is derived at the s point within the potential well for $V_{gd} - V_{th} > V_{sd} \geq 0$.

$$C_{gs} = \tfrac{2}{3} WLC_{ox}\left\{1 - \left[\frac{(V_{gd} - V_{th})}{V_{sd} - 2(V_{gd} - V_{th})}\right]^2\right\}, \tag{A8a}$$

and

$$C_{gd} = \tfrac{2}{3} WLC_{ox}\left\{1 - \left[\frac{V_{sd} - (V_{gd} - V_{th})}{V_{sd} - 2(V_{gd} - V_{th})}\right]^2\right\}, \tag{A8b}$$

where V_{th} is defined at the d point within the potential well.

To extend the validity range beyond the pinch-off region and to reduce the number of solution passes required by the ASTAP program, a smooth varying function which fits nicely to Eqs. (A7) and (A8) is found. For $V_{ds} > 0$ and $V_{gs} - V_{th} > 0$, then

$$C_{gs} = \tfrac{1}{3}\left[1 + \tanh\left(0.55 + \frac{V_{ds}}{V_{gs} - V_{th}}\right)\right] + C_{ov}, \tag{A9a}$$

and

$$C_{gd} = \tfrac{1}{3}\left[1 + \tanh\left(0.55 - \frac{V_{ds}}{V_{gs} - V_{th}}\right)\right] + C_{ov}. \tag{A9b}$$

For $V_{sd} \geq 0$ and $V_{gd} - V_{th} > 0$,

$$C_{gs} = \tfrac{1}{3}\left[1 + \tanh\left(0.55 - \frac{V_{sd}}{V_{gd} - V_{th}}\right)\right] + C_{ov}, \tag{A10a}$$

and

$$C_{gd} = \tfrac{1}{3}\left[1 + \tanh\left(0.55 + \frac{V_{sd}}{V_{gd} - V_{th}}\right)\right] + C_{ov}, \tag{A10b}$$

where the parasitic overlapping capacitance C_{ov} between the gate electrodes is included.

The total immobile depletion charge in the potential well according to the depletion approximation is given by

$$Q_B = W\int_0^L q_B dx = W\sqrt{2K_{Si}\varepsilon_0 q N_A}\int_{\psi_{ss}}^{\psi_{sd}} \frac{\sqrt{\psi_s}}{(d\psi_s/dx)}\, d\psi_s. \tag{A11}$$

After integrating and eliminating a common factor $(\sqrt{\psi_{sd}} - \sqrt{\psi_{ss}})$, we obtain

$$Q_B = 4WL\sqrt{2K_{Si}\varepsilon_0 qN_A}$$
$$\times\left[\tfrac{1}{5}(\psi_{sd}^2+\psi_{ss}^2+\sqrt{\psi_{ss}\psi_{sd}}\,(\psi_{ss}+\psi_{sd}+\sqrt{\psi_{ss}\psi_{sd}})\right.$$
$$\left.-\tfrac{1}{3}\left(\psi_{s0}+\frac{kT}{q}\right)(\psi_{ss}+\psi_{sd}+\sqrt{\psi_{ss}\psi_{sd}})\right]$$
$$\div\left[\left(\sqrt{\psi_{sd}}+\sqrt{\psi_{ss}}\right)\left(\psi_{sd}+\psi_{ss}-2\psi_{s0}-\frac{2kT}{q}\right)\right]. \quad \text{(A12)}$$

By incorporating into C_I, we obtain $C_{d\text{-sub}}$ and $C_{s\text{-sub}}$ as follows:

If $\psi_{sd}=\psi_{ss}=V$,

$$C_{d\text{-sub}}+C_{s\text{-sub}}=\frac{\partial Q_B}{\partial\psi_{ds}}=\frac{\partial Q_B}{\partial\psi_{ss}}. \quad \text{(A13)}$$

If $\psi_{ss}=0$,

$$C_I+C_{d\text{-sub}}=\frac{\partial Q_B}{\partial\psi_{ds}}. \quad \text{(A14)}$$

Hence,

$$C_{d\text{-sub}}=C_{s\text{-sub}}=\frac{WL}{4}\left(K_{Si}\varepsilon_0\Big/\sqrt{\frac{2K_{Si}\varepsilon_0 V}{qN_A}}\right), \quad \text{(A15)}$$

and

$$C_I=\frac{WL}{4}\left(K_{Si}\varepsilon_0\Big/\sqrt{\frac{2K_{Si}\varepsilon_0 V}{qN_A}}\right)$$
$$\times\left(\tfrac{16}{15}\times\frac{3V^2=13\psi_{s0}V+10\psi_{s0}^2}{V^2-4\psi_{s0}V+4\psi_{s0}^2}-1\right). \quad \text{(A16)}$$

Examination of Eq. (A16) shows that, as V tends to the ψ_{s0}, the equivalent of the pinch-off voltage, C_I tends to zero, which reflects the validity limit of the gradual channel approximation. For ASTAP applications, a small nonzero value can be assigned to C_I so that it does not approach zero nor become exactly zero in value.

References

1. A. F. Tasch, Jr., R. C. Frye, and H. S. Fu, "The Charge Coupled RAM Cell Concept," *IEEE J. Solid-State Circuits* **SC-11,** 58 (1976).
2. W. E. Engeler, J. J. Tiemann, and R. D. Baertsch, "A Surface-Charge Random-Access Memory System," *IEEE J. Solid-State Circuits* **SC-7,** 330 (1972).
3. M. F. Tompsett, "Surface Potential Equilibration Method of Setting Charge in Charge-Coupled Devices," *IEEE Trans. Electron Devices* **ED-22,** 305 (1975).
4. H. S. Lee and W. D. Pricer, "Merged Charge Memory (MCM), A New Random Access Cell," IEDM Technical Digest, 1976, p. 15.
5. A. M. Mohsen, T. C. McGill, M. Anthony, and C. A. Mead, "Push Clocks: A New Approach to Charge-Coupled Device Checking," *Appl. Phys. Lett.* **22,** 172 (1973).
6. W. S. Boyle and G. E. Smith, "Charge-Coupled Semiconductor Devices," *Bell Syst. Tech. J.* **49,** 598 (1970).
7. H. S. Lee and L. G. Heller, "Charge-Control Method of Charge Coupled Device Transfer Analysis," *IEEE Trans. Electron Devices* **ED-19,** 1270 (1972).
8. J. J. Tiemann et al., "Intra Charge-Transfer Structures for Signal Processing," *IEEE Trans. Electron Devices* **ED-21,** 300 (1974).
9. L. G. Heller, D. P. Spampinato, and Y. L. Yao, "High Sensitivity Charge-transfer Sense Amplifier," *IEEE J. Solid-State Circuits* **SC-11,** 596 (1976).
10. Private discussion with L. G. Heller, IBM System Products Division, Essex Junction, VT, October 1976.
11. H. S. Lee, "An Analysis of the Threshold Voltage for Short-channel IGFETs," *Solid-State Electron.* **16,** 1407 (1973).
12. R. R. Troutman, "Hot Electron Emission from a Semiconductor Under Low-Level Avalanche Multiplication," *IEDM Tech. Digest,* 578 (1976).
13. *ASTAP Program No. 5796-PBH,* ASTAP General Information Manual No. GH20-1271-0, IBM Corporation, White Plains, NY, 1973.
14. C. N. Berglund and K. K. Thornber, "Incomplete Transfer in Charge-Transfer Devices," *IEEE J. Solid-State Circuits* **SC-8,** 108 (1973).
15. A. M. Mohsen et al., "Charge Transfer in Overlapping Gate Charge-Coupled Devices," *IEEE J. Solid-State Circuits* **SC-8,** No. 3 (1973).
16. J. R. Burns, "Large-Signal Transit-Time Effects in the MOS Transistor," *RCA Rev.* **30,** 15 (1969).
17. R. A. Haken, "Supply Charge Isolation—A Simple Surface Potential Equilibration Charge-Injection Technique for Charge-Coupled Devices," *IEEE J. Solid-State Circuits* **SC-11,** 189 (1976).
18. C. H. Sequin and M. F. Tompselt, *Charge Transfer Devices,* Academic Press Inc., New York, 1975.

Received December 10, 1976; revised April 7, 1977

The author is located at the IBM System Products Division laboratory, Essex Junction, Vermont 05452.