A Survey of Circuit Innovations in Ferroelectric Random-Access Memories

Ali Sheikholeslami, MEMBER, IEEE, AND P. Glenn Gulak, SENIOR MEMBER, IEEE

This paper surveys circuit innovations in ferroelectric memories at three circuit levels: memory cell, sensing, and architecture. A ferroelectric memory cell consists of at least one ferroelectric capacitor, where binary data are stored, and one or two transistors that either allow access to the capacitor or amplify its content for a read operation. Once a cell is accessed for a read operation, its data are presented in the form of an analog signal to a sense amplifier, where it is compared against a reference voltage to determine its logic level.

The circuit techniques used to generate the reference voltage must be robust to semiconductor processing variations across the chip and the device imperfections of ferroelectric capacitors. We review six methods of generating a reference voltage, two being presented for the first time in this paper. These methods are discussed and evaluated in terms of their accuracy, area overhead, and sensing complexity.

Ferroelectric memories share architectural features such as addressing schemes and input/output circuitry with other types of random-access memories such as dynamic random-access memories. However, they have distinct features with respect to accessing the stored data, sensing, and overall circuit topology. We review nine different architectures for ferroelectric memories and discuss them in terms of speed, density, and power consumption.

Keywords—*Ferroelectric memory, memory circuit design, nonvolatile memory.*

I. Introduction

For the last three decades, floating-gate memories have been the dominant class of nonvolatile memories in applications ranging from personal computers to consumer electronics. In recent years, however, ferroelectric memories have received more research attention as evidenced in recent inventions in this field. In the past three years, for example, there have been more than 320 patents granted by the U.S. patent office. More than 120 of these inventions have been granted during the past year alone. This increased level of activity is being driven by two motives: superior features of ferroelectric memories such as short programming time

and low power consumption and the emergence of new applications such as contactless smart cards and digital cameras.

Table 1 compares ferroelectric memories with electrically erasable and programmable read-only memories (EEPROM's) and Flash memories, two types of floating-gate memories, in terms of density, read-access time, write-access time, and the energy consumed in a 32-bit read/write. Enjoying a mature process technology, EEPROM's and Flash memories [1], [2] are superior to ferroelectric memories in terms of density. Also, they require less power compared to ferroelectric memories for read operations, a factor that will keep them popular in applications that demand numerous memory reads but only occasional memory writes. An example of such applications is an identity card where an identity code is programmed into the memory once but read many times afterwards.

Ferroelectric memories, on the other hand, are superior to EEPROM's and Flash memories in terms of write-access time and overall power consumption, and hence target applications where a nonvolatile memory is required with such features. Two examples of such applications are contactless smart cards and digital cameras. Contactless smart cards require nonvolatile memories with low power consumption, as they use only electromagnetic coupling to power up the electronic chips on the card. Digital cameras require both low power consumption and fast frequent writes in order to store and restore an entire image into the memory in less than 0.1 s.

Another advantage of ferroelectric memories over EEPROM's and Flash memories is that they can be easily embedded as part of a larger integrated circuit to provide system-on-a-chip solutions to various applications [5], [6]. Future personal wireless connectivity applications that are battery driven, such as third-generation cellular phones and personal digital assistants, will demand large amounts (multiple megabytes) of nonvolatile storage to retain accessed Internet Web pages, containing compressed video, voice, and data. The density and energy efficiency of writing data to memory would seem to indicate that ferroelectric memory will play a major role in these types of consumer products.

As shown in Fig. 1, a ferroelectric memory technology consists of a CMOS technology with added layers on top for ferroelectric capacitors. Therefore, by masking parts of the

Table 1
Nonvolatile Memories and Their Typical Features

Nonvolatile Memory	Area/Cell (normalized)	Read Access-Time	Write (prog.) Access-Time	Energy* per 32b Write	Energy* per 32b Read
EEPROM	2	50ns	10μs	1μJ	150pJ
Flash Memory	1	50ns	100ns	2μJ	150pJ
Ferroelectric Memory	5 (†)	100ns	100ns	1nJ	1nJ

* Data given for the basic memory cell and may be different when designed into an integrated circuit.

† This number is based on nonstacked ferroelectric capacitor technology. More advanced fabrication technologies (i.e., with stacked capacitor [3]) offer comparable area/cell to that of EEPROM.
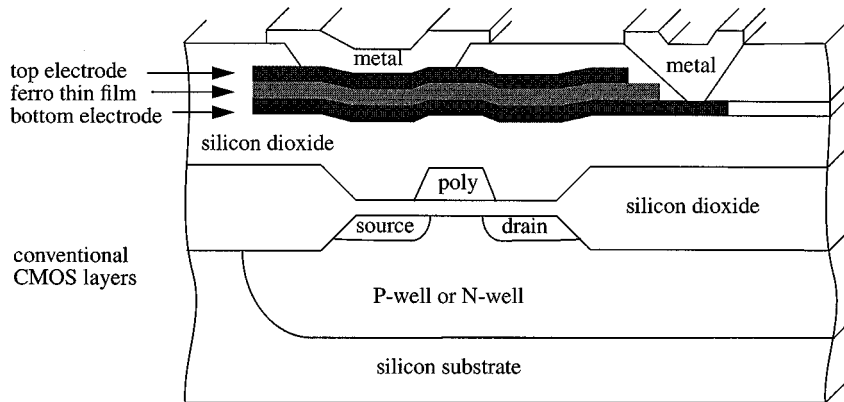

Fig. 1. Ferroelectric capacitor layers (two electrodes and a thin film of ferroelectric material) on top of a conventional CMOS process.

design that are not using ferroelectric capacitors, CMOS digital and analog circuits can be integrated together with ferroelectric memories, all in the same chip. Fig. 2 illustrates the cross section of a ferroelectric memory technology [3], [4] that allows the ferroelectric capacitors to sit directly on top of the transistors by means of stacked vias, hence reducing cell area.

Research on ferroelectric memories is proceeding on three fronts: material processing [7]–[14], modeling [15]–[17], and circuit design. On the material front, we offer a brief review of the historical evolution of ferroelectric materials for *ferroelectric* memories and compare them with *ferromagnetic* memories. This is presented in Section II of this paper, along with a brief review of the fundamental characteristics of ferroelectric materials and their usage as a medium for nonvolatile storage of binary data.

Accurate modeling of ferroelectric capacitors is essential to circuit simulation and design of ferroelectric memories. A circuit-based model, for example, can be integrated into circuit simulation tools such as HSPICE [18] to simulate various sections of a ferroelectric memory. We refer the interested readers to [15] for a recent survey of behavioral modeling of ferroelectric capacitors and to [16] for a recent modeling effort by authors of this paper.

Our main focus in this paper is on innovative circuit techniques. We survey research efforts on this front as it relates to circuit innovations for higher performance
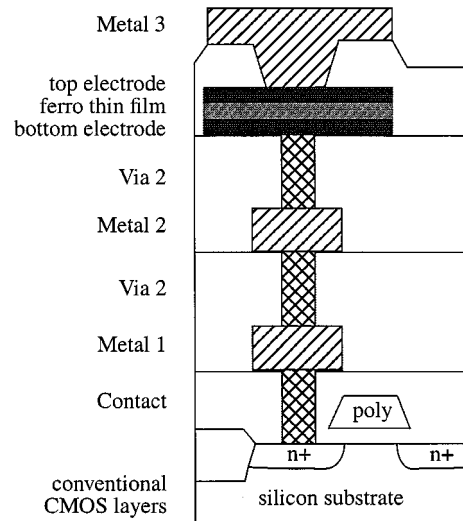

Fig. 2. Cross section of a ferroelectric memory technology [3] that uses three metal layers and allows stacked vias to minimize the memory cell area.

ferroelectric memories. Section III of this paper presents an overview of the basic circuit operations of ferroelectric memories and their terminology. Section IV of this paper presents circuit innovations that relate to reference voltage generation techniques. Six different methods are reviewed and compared in terms of their sensing speed, accuracy, area
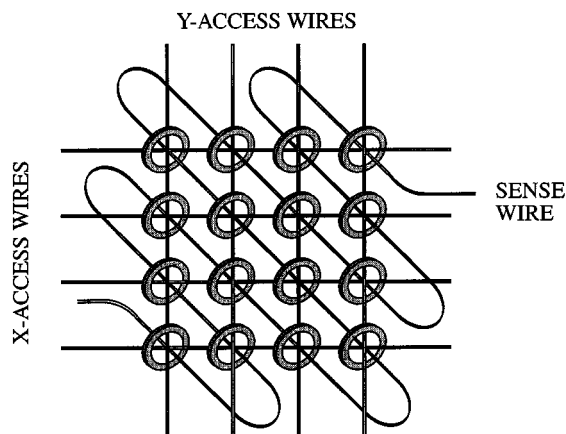
Fig. 3. Two-dimensional array of ferromagnetic cores. Each core is accessed by a simultaneous current pulse on an x-access and y-access wire [19].

overhead, and sensing complexity. Two of these methods are presented for the first time in this paper. Section V reviews circuit innovations at the architectural level and highlights their advantages and disadvantages over conventional architectures. Section VI offers a brief overview of technology trends. Last, Section VII presents our conclusions on the circuit design aspects of ferroelectric memories.

II. BACKGROUND

The underlying principles of operation for ferroelectric capacitors and ferromagnetic cores are similar. In this section, we present a brief review of the principles of operation of ferromagnetic memories and their counterparts in ferroelectric memories.

A. Ferromagnetic Cores

Prior to the 1950's, ferromagnetic memories (also known as core memories) were the only type of random-access, nonvolatile memories [19]. A core memory, as shown in Fig. 3, consists of a regular array of tiny magnetic cores that can be magnetized in one of two opposite directions, hence storing binary data in the form of a magnetic field. A write access into a core consists of sending simultaneous current pulses through the core via its x-access and y-access wires. Depending on the directions of the current pulses, a core is magnetized in a "0" or a "1" direction. The basic assumption here is that only the core that receives two simultaneous current pulses is affected. All the remaining cores, including those that receive one current pulse or none, retain their original magnetization.

A read access consists of a write access followed by sensing. We write a "0" to the core in order to discover the original data content of the core. If the original content of the core is a "1," writing a "0" would mean changing the magnetic direction of the core. This induces a large current spike on the sense wire. On the other hand, there will be no current spike on the sensing wire if the original content of the core was also a "0." Therefore, by sensing the presence of a current spike on the sensing wire, the original data of the accessed core are determined.
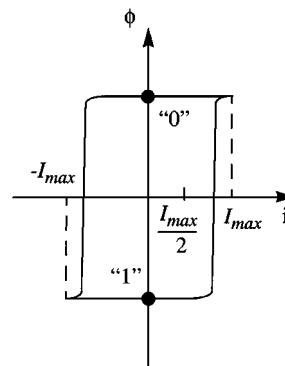

Fig. 4. Hysteresis loop characteristic of a ferromagnetic core of Fig. 3.

The read operation as explained above is destructive since a "0" is written to any core that is accessed for a read. The original data, however, are saved at the sense amplifier and can be restored back into the accessed core. In other words, a read access is only complete after the second write that restores the original data.

The success of the core memory was due to its simple architecture and a square-like hysteresis loop characteristic of the core, as shown in Fig. 4. The architecture allows storing a memory bit at every intersection of the access wires, resulting in a scalable, relatively dense array of cells. This architecture, with a few modifications that make it more suitable for higher capacities, has been adopted by many generations since the core memory. In fact, most semiconductor memories today, including dynamic random-access memories (DRAM's), EEPROM's, and ferroelectric random-access memories (FRAM's), use a very similar row–column architecture.

A square hysteresis loop, as shown in Fig. 4, displays the magnetic flux ϕ of a core in terms of the electric current i passing through it. The two points corresponding to a zero current are marked to represent the two nonvolatile magnetic states of the core. The rapid transition from one nonvolatile state to the other, defined by the hysteresis loop, is called magnetic switching and occurs over a small range of applied current. This characteristic is a key to accessing the core without disturbing its neighbors. A core will not switch states unless the current passing through it is larger than half of the maximum current. Each access wire of Fig. 3 carries only half of the maximum current. Therefore, only the core at the intersection of two wires carrying simultaneous current pulses could switch its state. All other cores remain unchanged since they receive no current or at most half of the maximum current, which is not enough for switching. The designer should of course ensure that no more than a single horizontal wire and a single vertical wire are activated at the same time.

Ferromagnetic cores were too bulky and soon became too expensive compared to much smaller, less power consuming, semiconductor memories. Ferroelectric memories were a suitable substitute for ferromagnetic memories as the hysteresis loop characteristics of ferroelectric capacitors and those of the ferromagnetic cores were similar. The term
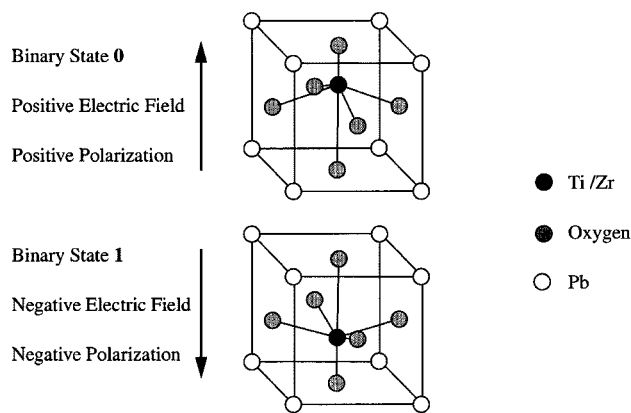
Fig. 5. Two stable states in a ferroelectric material known as PZT: the orientation of the spontaneous polarization is reversed by applying a proper electric field.

Fig. 6. Hysteresis loop characteristic of a ferroelectric capacitor. Remanent charge (Q_r), saturation charge (Q_s), and coercive voltage (V_C) are the three important parameters that characterize the loop. The + and − signs beside the capacitor symbol represent the applied voltage polarity.

*ferro*electric was adopted to convey this similarity despite the lack of *iron* in ferroelectric materials. However, the technological difficulties in reliably cointegrating ferroelectric thin films with silicon substrates have kept the ferroelectric memory densities a few orders of magnitude behind the densities of DRAM's and EEPROM's. In fact the commercially available 256-kbit FRAM from Ramtron [20] is three orders of magnitude behind the 256-Mbit DRAM from Micron [21]. Nevertheless, many believe that ferroelectric memories will be the dominant memory technology in the future [22]. The circuit innovations that are discussed in this paper unveil a continuous endeavor among those in the field toward this future.

B. Ferroelectric Capacitors

A ferroelectric capacitor is physically distinguished from a regular capacitor by substituting the dielectric with a ferroelectric material [23]. In a regular dielectric, upon the application of an electric field, positive and negative charges will be displaced from their original position—a concept that is characterized by polarization. This polarization, or displacement, will vanish, however, when the electric field returns back to zero. In a ferroelectric material, on the other hand, there is a spontaneous polarization—a displacement that is inherent to the crystal structure of the material and does not disappear in the absence of electric field. In addition, as illustrated in Fig. 5, the direction of this polarization can be reversed or reoriented by applying an appropriate electric field.

One family of well-established ferroelectric materials that have been used for over a decade are known as perovskites, having the formula ABO_3. A widely used material in this family is lead zirconate titanate (PZT) with the formula $Pb(Zr_xTi_{1-x})O_3$. Fig. 5 illustrates a unit cell of this material. The central atom in this unit cell is either titanium (Ti) or zirconium (Zi), depending on the contribution of each atom to the material formula. An appropriate electric field is capable of displacing the central atom from its previous stable position and, thereby, changing the polarization state of the unit cell.

Another more recently developed family of ferroelectric materials is known as layered perovskites (or $Y - 1$ family),
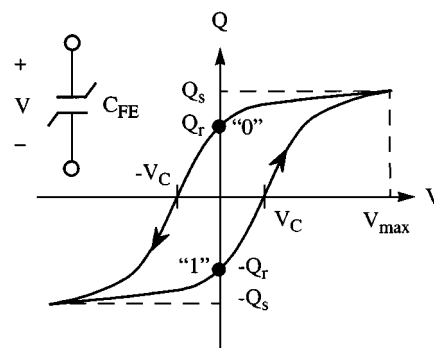
where oxide layers interleave perovskite layers in a lattice structure. A well-studied member of this family with bismuth oxide layers is SBT, with the formula $SbBi_2Ta_2O_9$. Compared to PZT, SBT promises better endurance characteristics such as fatigue (a gradual degradation of capacitor charge with repeated cycling of the capacitor [9]) and imprint (the tendency of a ferroelectric capacitor to prefer one state over the other if it stays in that state for a long period of time [11]), as well as low voltage of operation to below 1 V.

Although the polarization of each individual unit cell is tiny, the net polarization of several domains—each consisting of a number of aligned unit cells—can be large enough for detection using standard sense amplifier designs. The gross effect of polarization is a nonzero charge per unit area of the ferroelectric capacitor that exists at 0 V and does not disappear over time. The polarization charge, which is simply referred to as the capacitor charge in this paper, responds to the voltage across the capacitor in the same way as the magnetic flux of a ferromagnetic core responds to the current through the core. In this sense, ferroelectric capacitors are duals of ferromagnetic cores.

A hysteresis loop for a ferroelectric capacitor, as shown in Fig. 6, displays the total charge on the capacitor as a function of the applied voltage. When the voltage across the capacitor is 0 V, the capacitor assumes one of the two stable states: "0" or "1." The total charge stored on the capacitor is Q_r for a "0" or $-Q_r$ for a "1." A "0" can be switched to a "1" by applying a negative voltage pulse across the capacitor. By doing so, the total charge on the capacitor is reduced by $2Q_r$, a change of charge that can be sensed by the sense circuitry as explained later in this paper. Similarly, a "1" can be switched back to a "0" by applying a positive voltage pulse across the capacitor, hence restoring the capacitor charge to $+Q_r$. These characteristics are all very similar to those of a magnetic core except for the following: the hysteresis loop of a ferroelectric capacitor does not have sharp transitions around its coercive points: $-V_c$ and $+V_c$. This reflects a partial switching of electric domains in a ferroelectric capacitor, and further implies that even a voltage half of $V_{\max}$ can disturb the state of the capacitor. As a result, it is impossible to
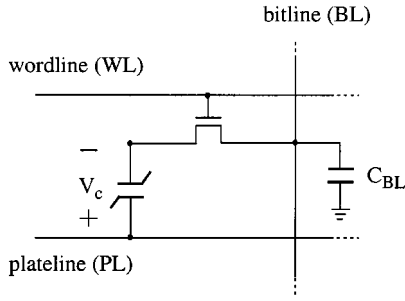
Fig. 7. Ferroelectric 1T-1C memory cell. "C_{BL}" represents the total parasitic capacitance of the bitline.
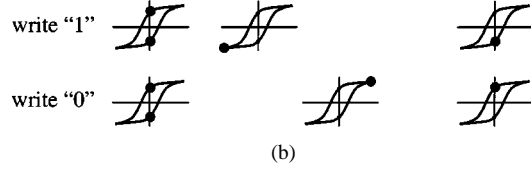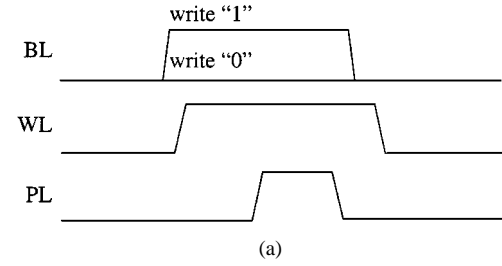

Fig. 8. (a) Timing diagram for a write operation of the memory cell shown in Fig. 7. (b) The state sequence for the memory cell capacitor as a "1" or a "0" being written into the cell. The initial state of the capacitor, in both cases, does not affect the subsequent states of the capacitor.

access a ferroelectric capacitor in a cross-point array without disturbing the capacitors on the same row or column.

One approach to remedy this situation would be to modify the ferroelectric material processing in order to create a square-like hysteresis loop. In this case, ferroelectric capacitors can form a ferroelectric memory array in the same way as magnetic cores form a core memory. This is a technical development that we may expect in near future.

Another approach is to modify a ferroelectric memory cell by including a transistor in series with the ferroelectric capacitor as shown in Fig. 7. The transistor, called the access transistor, controls the access to the capacitor and eliminates the need for a square-like hysteresis loop. When the access transistor is off, the FE capacitor remains disconnected from bitline (BL) and hence cannot be disturbed. When the access transistor is ON, the FE capacitor is connected to the bitline and can be written to or read by the plateline (PL). In other words, the presence of an access transistor in series with the ferroelectric capacitor compensates for the softness of its hysteresis loop characteristics and blocks unwanted disturb signals from neighboring memory cells.

Including a transistor in series with a ferroelectric capacitor is just one example of a powerful concept: overcoming a device imperfection by a circuit technique. In the rest of this paper, we explain in detail other circuit techniques, such as sensing schemes and reference voltage generation, that are invented to circumvent various ferroelectric device and process shortcomings.

III. BASIC MEMORY-CELL OPERATION

A ferroelectric memory cell, known as 1T-1C (one transistor, one capacitor) cell, is shown in Fig. 7. The cell consists of a single ferroelectric capacitor that is connected to a PL at one end and, via an access transistor, to a BL at the other end. The cell is accessed by raising the wordline (WL) and hence turning ON the access transistor. The access is one of two types: a write access or a read access. We explain each access separately in the following.

The timing diagram for a write operation is shown in Fig. 8(a). To write a "1" into the memory cell, the BL is raised to V_{DD}. Then the WL is raised to $V_{DD} + V_T$ (known as boosted V_{DD} [24]), where V_T is the threshold voltage of the access transistor. This allows a full V_{DD} to appear across

the ferroelectric capacitor ($-V_{DD}$ according to the voltage convention adopted in Fig. 7). At this time, the state of the ferroelectric capacitor is independent of the initial state of the FE capacitor, as shown in Fig. 8(b). Next, the PL is pulsed, that is, pulled up to V_{DD} and subsequently pulled back down to ground. Note that the WL stays activated until the PL is pulled down completely and the BL is driven back to zero. The final state of the capacitor is a negative charge state S_1 (defined as digital "1" in this paper). Finally, deactivating the WL leaves this state undisturbed until the next access.

To write a "0" into the cell, the BL is driven to 0 V prior to activating the WL. The rest of the operation is similar to that of writing a "1" as shown in Fig. 8.

The timing diagram for a read access is shown in Fig. 9. A read access begins by precharging the BL to 0 V, followed by activating the WL (Δt_0). This establishes a capacitor divider consisting of C_{FE} and C_{BL} between the PL and the ground. During Δt_1, the PL is raised to V_{DD}. This voltage is divided between C_{FE} and C_{BL}, the parasitic capacitance of the bitline, according to their relative capacitance. Depending on the data stored, the capacitance of the FE capacitor can be approximated by C_0 or C_1, as shown in Fig. 11. Therefore, the voltage developed on the bitline (V_x) can be one of the two values V_0 or V_1

$$V_x = \begin{cases} V_0 = \dfrac{C_0}{C_0 + C_{BL}} V_{DD} & \text{if the stored data is a 0} \\ V_1 = \dfrac{C_1}{C_1 + C_{BL}} V_{DD} & \text{if the stored data is a 1.} \end{cases}$$
(1)

At this point, the sense amplifier is activated to drive the BL to full V_{DD} if the voltage developed on the BL is V_1, or to 0 V if the voltage on the BL is V_0. The WL is kept activated until the sensed voltage on the BL restores the original data back into the memory cell and the BL is precharged back to 0 V.
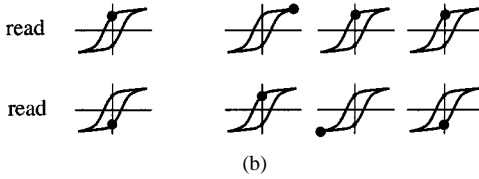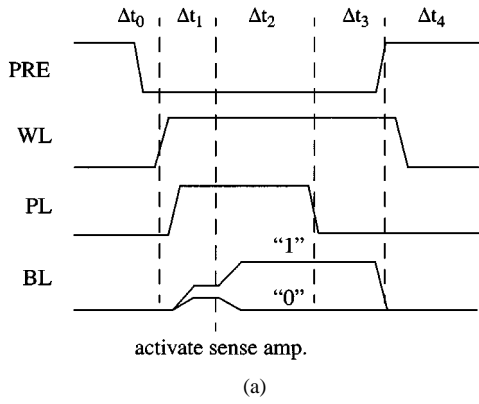
Fig. 9. (a) Timing diagram for a read operation of the memory cell shown in Fig. 7. (b) The state sequence for an initially "0" and an initially "1" memory cell capacitor as it undergoes the read operation.
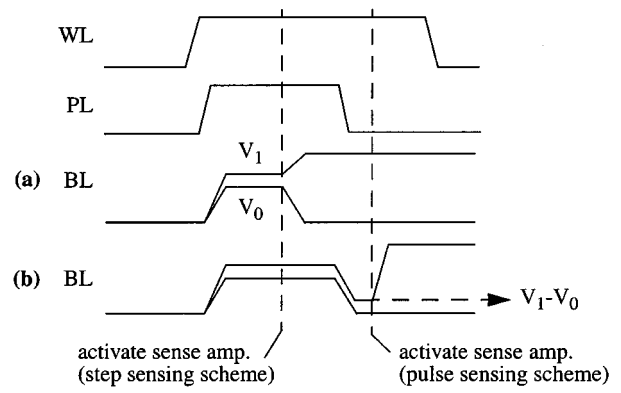

Fig. 10. Timing diagram for a read operation based on (a) step-sensing scheme and (b) pulse-sensing scheme.

Fig. 11. Hysteresis loop is approximated by two linear capacitors C_0 and C_1.

A. Sensing Schemes

The read access as presented above is known in the literature as the *step-sensing approach*, since a *step* voltage (the rising edge of a pulse) is applied to the PL prior to *sensing*. An alternative is the *pulse-sensing* approach in which a full *pulse* is applied to the PL prior to activating the sense amplifiers (refer to Fig. 10). The charge transferred to the BL in a pulse sensing scheme is either zero for a stored "0" or $2Q_r$ for a stored "1." Equivalently, the voltage developed on the BL is either 0 V for a stored "0" or $V_1 - V_0$ [refer to (1)] for a stored "1."

In both step- and pulse-sensing schemes, the voltage difference on the BL that is developed by a stored "1" and a stored "0" is equal to $V_1 - V_0$. The common-mode voltage, however, is equal to $(V_1 + V_0)/2$ in the step-sensing approach, as compared to $(V_1 - V_0)/2$ in the pulse-sensing approach. Therefore, the step-sensing approach provides a higher common-mode voltage on the BL that simplifies the sense amplifier design when a bias voltage is required. Another advantage of the step-sensing approach is that it provides a faster read access, as the sensing does not wait for the PL to be pulled low.

Both step- and pulse-sensing approaches restore "1," but only the step-sensing approach fully reinforces a "0." To substantiate this point, note that during a read operation in a step-sensing approach, an FE capacitor storing a "0" experiences a voltage sequence [25] of 0 V, $V_{DD} - V_0$, V_{DD}, and 0 V (a full-V_{DD} excursion). In a pulse-sensing approach, the corresponding voltage sequence is 0 V, $V_{DD} - V_0$, and 0 V (a $V_{DD} - V_0$ excursion). None of the two voltage sequences upsets the original data (i.e. "0") of the FE capacitor. However, the latter provides a weak reinforcement of "0" by applying a voltage less than V_{DD} across the capacitor. This seems to
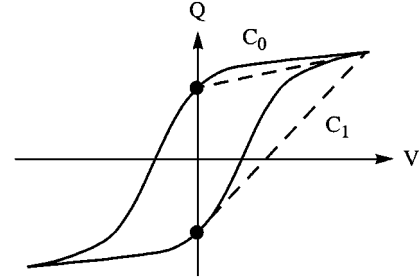
deteriorate the capacitor's long-term retention performance [26]. To remedy this situation, a second pulse must be applied to the PL to fully restore the "0" into the capacitor. This implies that the cycle time for the pulse-sensing approach can be twice as large as that of the step-sensing approach.

The pulse-sensing approach applies both the leading edge and the trailing edge of the voltage pulse to the FE capacitor prior to sensing. The trailing edge eliminates the non-switching part of polarization that was introduced on the BL by the rising edge and therefore bypasses the effect of non-switching part of polarization and its process variations altogether. This seems to be the only advantage of the pulse-sensing approach over the step-sensing approach.

So far, we have assumed that the sense amplifier can discriminate between a "0" and "1" voltage signal on the BL. This is only possible if a reference voltage, midway between a "0" and a "1" signal, is provided for the sense amplifier. In the next section, we explain the challenges in generating a reference voltage and review their proposed solutions.

IV. REFERENCE VOLTAGE GENERATION

The voltage that appears on the bitline, V_x, is V_0 or V_1, as defined by (1). A sense amplifier determines if V_x is equal to V_0 or V_1 by comparing it against a reference voltage (V_{REF}) that is ideally halfway between V_0 and V_1. The difference between V_x and $V_{REF}(V_x - V_{REF})$ is amplified to V_{DD} or V_{SS} depending on whether $V_x > V_{REF}$ or $V_x < V_{REF}$, respectively.

There are a few circuit challenges in generating an accurate reference voltage. One challenge is that only approxi-
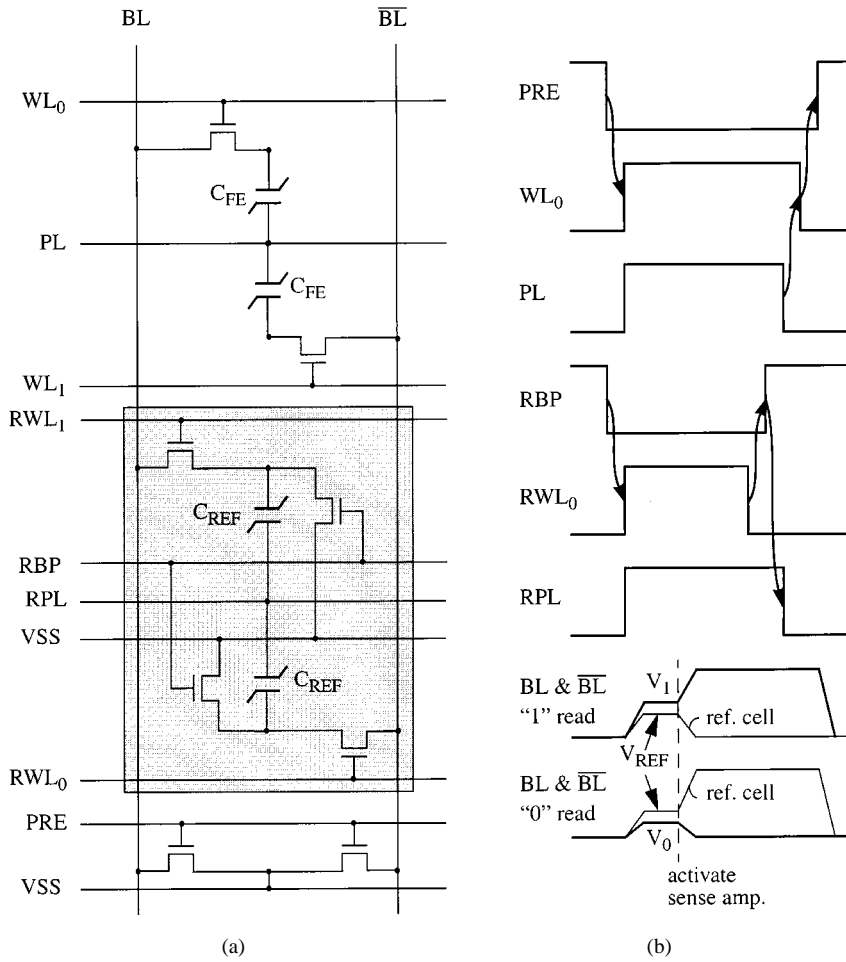
Fig. 12. (a) A pair of reference cells (shaded area) for a column of memory cells and (b) timing diagram of a read operation [27]. WL_0 and RWL_0 are raised simultaneously to access the memory cell on the left and the reference cell on the right of the column. WL_1 and RWL_1 are kept at 0 V (not shown) throughout the operation [27].

mate values of V_0 and V_1 are known in advance. The parameters that determine V_0 and V_1, as shown in (1), are process dependent and time dependent. C_{BL}, for example, depends on the diffusion capacitance of the access transistors connected to the bitline. C_0 and C_1 are cell dependent, can vary across the memory array, and degrade over time due to fatigue—a gradual degradation of capacitor charge with repeated cycling of the capacitor [9]. In other words, the cells that are accessed more often degrade faster than the less-accessed cells.

Another circuit challenge is the design of a reference scheme that is robust to various ferroelectric capacitor imperfections such as *relaxation* [10] and *imprint* [11]. Relaxation refers to a partial loss of remanent charge in a microsecond time regime if the capacitor is left unaccessed following a sequence of continuous cycling. Transient dynamics of this phenomenon aside, from a circuit point of view, relaxation means a smaller V_1 and a larger V_0 than those predicted by (1). Imprint refers to the tendency of a ferroelectric capacitor to prefer one state over the other if it stays in that state for a long period of time. In a memory circuit, imprint manifests itself as a voltage offset in both V_0 and V_1, and hence as a voltage offset in V_{REF}.

This all implies that no fixed value of V_{REF} can be used across the chip, but rather a variable reference voltage is required to accurately track the process variation and the ferroelectric material degradation. In this section, we discuss circuit ideas for reference voltage generation that are suggested for the 1T-1C architecture. The circuit ideas that resort to architectural changes for reference voltage generation will be discussed in Section V.

A. One Oversized Reference Capacitor Per Column (1C'/BL)

A conventional approach in generating a reference voltage for a column of memory cells is shown in Fig. 12(a). The shaded area on the lower part of the figure shows the reference voltage circuit, which consists of two cells (one reference cell per bitline) with their dedicated reference wordlines (RWL_0 and RWL_1) running through the array. RWL_0 controls access to a row of reference cells that are connected to the $\overline{BL}$ and is activated when any even-numbered wordline is activated. Similarly, RWL_1 controls access to a row of reference cells that are connected to the BL and is activated

when any odd-numbered wordline is activated. For example, RWL_1 and WL_1 are activated simultaneously as they target opposite bitlines (BL and $\overline{BL}$).

A reference capacitor C_{REF} always stores a "0," but it is sized larger than C_{FE} such that its projected signal on the BL, V_{REF}, is midway between V_0 and V_1 [24], [27]. The timing diagram used by Sumi *et al.* [27] for a read operation is shown in Fig. 12(b). The bitlines, as well as the storage nodes of the reference cells, are precharged to 0 V prior to a read operation. WL_1 and RWL_1 are activated together, followed by a simultaneous step voltage on the PL and the RPL. The $\overline{BL}$ is raised to V_{REF}, while the BL is raised to either V_0 or V_1, depending on the stored data's being a "0" or a "1," respectively. Since V_{REF} is halfway between V_0 and V_1, activating the sense amplifier sends the bitline of higher voltage (BL or $\overline{BL}$) to V_{DD} and the bitline of lower voltage to 0 V.

Note the timing of RBP relative to those of RWL_1 and RPL. The storage nodes of the reference cell are pulled down to the ground prior to pulling down the RPL back to 0 V. This is to guarantee that C_{REF} experiences only a positive (or 0 V) voltage across it, and hence operating in a nonswitching mode independent of the sensed data. C_{REF} is therefore subjected to no cycling at all (not fatigued) and maintains its signal level much longer than the fatigue-prone life of C_{FE}.

Adding a reset transistor to each reference cell, which was first introduced by Sumi *et al.* [27], has an additional benefit: assume that the BL is driven to 0 V while $\overline{BL}$ is driven to V_{DD} by the sense amplifier. Just prior to deactivating the RWL_1, the voltage across C_{REF} is full V_{DD}, with the PL at V_{DD} and the BL at 0 V. By deactivating the RWL_1, one node of C_{REF} is still at V_{DD}, while the other node (the storage node) is floating. Pulling down the PL pulls down the storage node with itself to $-V_{DD}$ (in fact, to $-V_f \cong -0.7$ V due to the forward-biased diode present from source to substrate of the access transistor). Therefore, without resetting the storage node to 0 V, there will be a voltage of V_f built up across C_{REF} that could contribute to a next immediate read (before the storage node leaks its charge to the substrate). By including a reset transistor in the reference cell, the storage node can be grounded by raising RBP prior to pulling down the PL, hence avoiding the problem at the expense of negligible increase in the array area.

The accuracy of V_{REF} in this scheme is strongly correlated with the accuracy of the size of C_{REF}. An accurate size for the reference capacitor is hard to estimate theoretically and requires some empirical measurements. One method to alleviate this design issue is to make the reference capacitor adjustable (or programmable). In other words, include incremental reference capacitors to be added in parallel to an original approximately sized capacitor. An advantage of doing so is the ability to compensate for a possible charge degradation over time by adding incremental reference capacitors to the original capacitor.

Another method, proposed by Miyakawa *et al.* [30], is to leave the reference capacitor alone but make the voltage level on the RPL adjustable using five address bits (32 voltage levels). The bits are selected to produce the desired reference voltage level on the bitline for each individual chip. These bits are then written permanently into five flip-flops by laser-cutting selected fuses in the flip-flop circuits. The bits and,

(a)

precharge
for the next operation
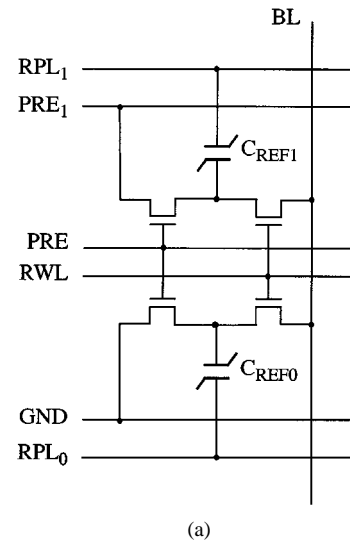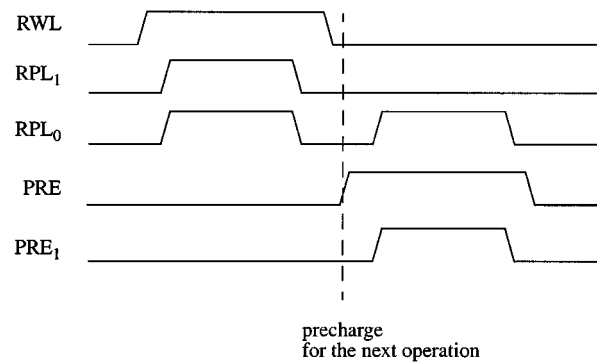
(b)

Fig. 13. (a) Reference cell for a 2 × 0.5C/BL reference scheme and (b) timing diagram for a read operation of the reference cell [32].

hence, the adjusted voltage level for the RPL, will be restored during the chip power-up.

In a slightly different variation, Jung *et al.* [29] suggest generating an adjustable reference voltage using a transistor and adjustable resistors only. This adjustable voltage is then stored on a gate-oxide capacitor and subsequently shared with the bitline capacitance to produce the final reference voltage for sensing.

B. Two Half-Sized Reference Capacitors Per Column (2 × 0.5C/BL)

Ideally, a reference cell should generate a voltage on the bitline that is midway between V_0 and V_1, that is, $(V_0 + V_1)/2$. A straightforward approach to generating such a voltage is proposed by Lowrey *et al.* [32], as shown in Fig. 13. A reference cell consists of two capacitors C_{REF0} and C_{REF1}, each with two separately controlled access transistors. The capacitors are half the size of a memory cell capacitor C_{FE}, with C_{REF0} always storing a "0" and C_{REF1} always storing a "1." Therefore, C_{REF0} is expected to generate $V_0/2$ and C_{REF1} to generate $V_1/2$ on the bitline, a total of $(V_0 + V_1)/2$ if they are accessed simultaneously by raising the RWL, RPL_0, and RPL_1.

The actual reference voltage generated by this scheme can be approximated by the following formula:

$$V_{\text{REF}} = V_{\text{DD}} \left(\frac{C_0/2 + C_1/2}{C_0/2 + C_1/2 + C_{\text{BL}}} \right) \qquad (2)$$

where C_0 and C_1 are the approximate ferroelectric capacitance as defined in Fig. 11. This voltage should be compared with an ideal reference voltage $(V_0 + V_1)/2$

$$(V_0 + V_1)/2 = V_{\text{DD}} \left(\frac{C_0/2}{C_0 + C_{\text{BL}}} + \frac{C_1/2}{C_1 + C_{\text{BL}}} \right). \qquad (3)$$

It can easily be shown that V_{REF} in (2) is always larger than the ideal reference voltage in (3), independent of the values of C_0, C_1, and C_{BL}. In other words, V_{REF} is closer to V_1 than it is to V_0, and, hence, the noise margin for a "1" is smaller than its optimum value. Depending on the C_0/C_{BL} and C_1/C_{BL}, this noise margin can be small enough to force the designer to choose a larger than optimum C_{FE}.

Another drawback of this reference scheme is that it fatigues the reference cells faster than the regular memory cells by accessing the reference row each time a memory row is accessed. This implies that the generated V_{REF} [(2)] reduces at a rate faster than that of an ideal reference voltage [(3)]. Moreover, writing a "0" into C_{REF0} and a "1" into C_{REF1} can cause imprint in these capacitors. Fortunately, this can be resolved by shuffling the data between C_{REF0} and C_{REF1} every time they are accessed.

It is possible to eliminate the overestimation of V_{REF} by modifying this scheme, as presented in the next two sections. Different fatigue rates for the reference cells and the memory cells, however, are common among all reference schemes that share a row of reference cells with an array (several rows) of memory cells.

C. One Half-Sized Reference Cell Per Half-Column (0.5C/0.5BL)

The output voltage of a capacitor divider circuit is a function of the input voltage and the capacitance ratio. Therefore, to ensure that V_{REF0} is equal to V_0, $C_{\text{REF0}}/C_{\text{BL}}$ must be equal to C_0/C_{BL}. Similarly, to ensure that V_{REF1} is equal to V_1, $C_{\text{REF1}}/C_{\text{BL}}$ must be equal to $C_1/C_{\overline{\text{BL}}}$, assuming that C_{REF0} and C_{REF1} dump their charge on BL while C_{FE} dumps its charge on $\overline{\text{BL}}$. With C_{REF} being half the size of C_{FE}, we propose splitting the reference cell and the bitline into two parts, as shown in Fig. 14(a). The top part includes C_{REF0} and $C_{\text{BL}}/2$, and the bottom part includes C_{REF1} and $C_{\text{BL}}/2$. Since both the reference capacitor and C_{BL} are divided by two, their ratio stays the same as the original ratio. The voltages generated on the BL_TOP and BL_BOT are V_0 and V_1, respectively. When the two parts are shorted together by raising the BLS, an average of $(V_0 + V_1)/2$ will appear on the bitline. A timing diagram that performs this reference voltage generation is shown in Fig. 14(b).

We have assumed, in the last two sections, that a full-size capacitor scales its charge to exactly half for a half-sized capacitor. This assumption is not quite valid considering the capacitor perimeter scales differently than its area and, hence,
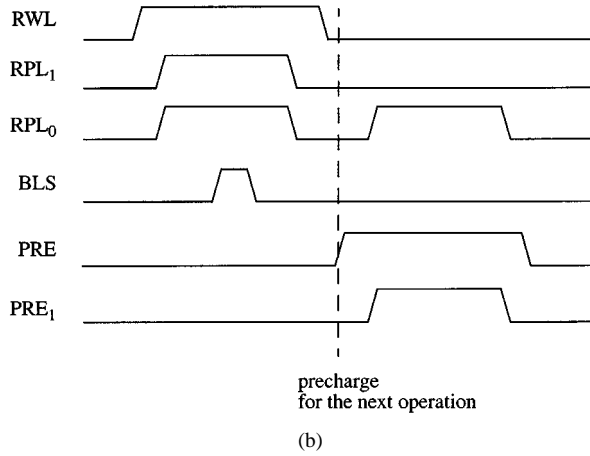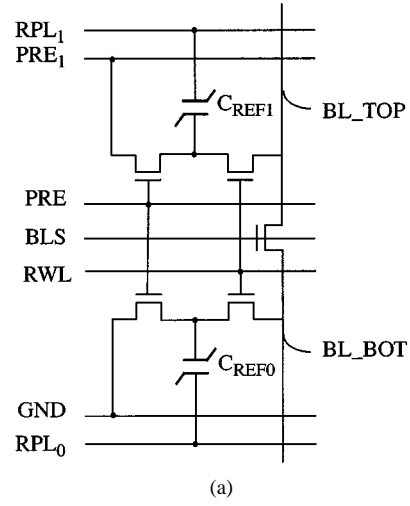

Fig. 14. (a) Reference cell for a 0.5C/0.5BL reference scheme and (b) timing diagram for a read operation of the reference cell.

the peripheral parasitic capacitance of a half-area capacitor is not necessarily equal to half of the peripheral capacitance of a full-sized capacitor. Moreover, both the 2×0.5C/BL and the 0.5C/0.5BL schemes are only practical when half-sized ferroelectric capacitors are allowed in the VLSI process. If the capacitors used in the memory cells are minimum-sized capacitors, then a half-sized capacitor is not allowed and this scheme cannot be used. In the next section, we look at another approach in reference voltage generation that eliminates this concern.

D. Two Full-Sized Reference Capacitors Per Two Columns (2C/2BL)

In the previous section, we divided a reference cell between the two halves of a bitline and, hence, used reference capacitors that were half the cell capacitors. Another approach, proposed by Wilson and Meadows [33], shares a reference cell using a full-sized ferroelectric capacitor between two adjacent columns. A block diagram of a memory architecture using this approach is shown in Fig. 15, where each memory cell is indicated by a circle and each reference cell (shared by adjacent bitlines) is indicated by a rectangle.
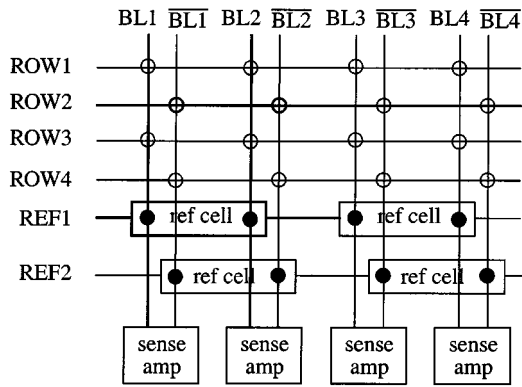
Fig. 15. Block diagram of an architecture using the 2C/2BL reference scheme. The block in thick lines is further explored in Fig. 16.

The blocks drawn with thick lines in this figure are further expanded in Fig. 16(a) to show the circuit details of this scheme.

The reference cell connected to the BL_1 has a ferroelectric capacitor storing a "1" (C_{REF1}), and the reference cell connected to the BL_2 has a ferroelectric capacitor storing a "0" (C_{REF0}). Both C_{REF0} and C_{REF1} are identical in size to a cell capacitor, C_{FE}. Each reference cell has an additional access transistor (controlled by the RP) that is used to restore its original data into the cell after a read operation is complete.

Fig. 16(b) shows a timing diagram of a read operation. By raising the RWL and the RPL, C_{REF1} generates V_1 on BL_1 while C_{REF0} generates V_0 on BL_2. The accessed memory cells generate their own data (V_x) on $\overline{BL}_1$ and $\overline{BL}_2$. Next, BL_1 and BL_2 are shorted together by raising EQ momentarily to share their charge. This results in a common reference voltage on both BL_1 and BL_2 that is equal to ($V_0 + V_1$)/2. At this time, the sense amplifiers are activated to discriminate between V_x and V_{REF}.

At the end of a read operation, a "0" and a "1" must be restored in C_{REF0} and C_{REF1}, respectively, for the next read. This can be achieved by keeping the RPL high and pulsing the RP and the RS as shown in Fig. 16(b).

Similar to the previous approach, the 2C/2BL approach keeps C_{REF}/C_{BL} and C_{FE}/C_{BL} the same. Also, by keeping the same size C_{FE} and C_{REF} for both the memory cell and the reference cell, the voltage on a reference bitline better mimics that of the memory cell. As a result, the generated reference voltage is closer to the ideal value.

E. Adding Reference Cells to Rows (2C/WL)

All the reference schemes discussed so far include a row of reference cells in the array that is accessed by a separate wordline (RWL) and plateline (RPL). A common drawback of these schemes is that a memory cell and a reference cell are accessed at different rates and, hence, fatigued at different rates. For example, sequentially accessing n rows of an array fatigues the reference cells n times faster than each individual memory cell.

A circuit technique that does not suffer from this drawback is proposed by Papaliolios [34] and Wood [35]. As shown in Fig. 17, one reference cell is assigned to each row of the array,

(a)

begin a read   activate SA's   reset ref. capacitors
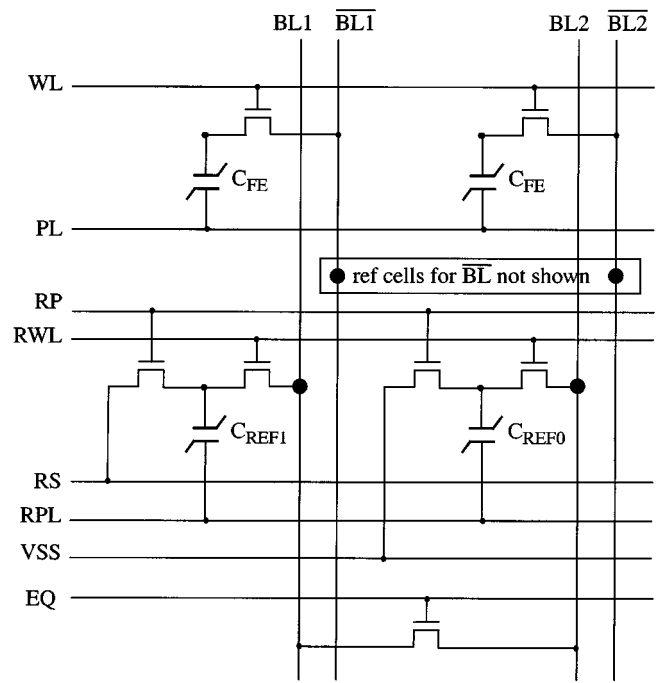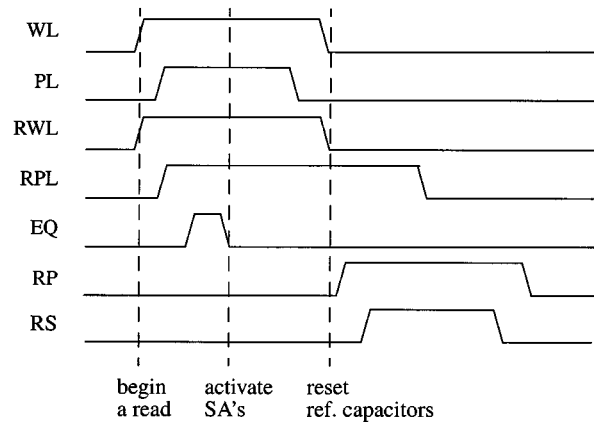
(b)

Fig. 16. (a) The circuit diagram and (b) its corresponding timing diagram for a 2C/2BL reference scheme. The timing diagram is a modified version of what is proposed in [33].

instead of each column of the array. As a result, a reference cell and a memory cell are accessed by the same WL and fatigued at the same rate.

A reference voltage V_{REF} is generated by shorting the RBL and $\overline{RBL}$ that are holding V_0 and V_1, respectively. There is a new design challenge to address in this architecture since there is only one reference bitline that must feed many sense amplifiers. Each sense amplifier adds its own capacitive load to the RBL and hence creates a capacitive imbalance between the RBL and the bitlines. Wood [35] proposes adding extra capacitance C_{ext} to each bitline to counterbalance the extra capacitance of the RBL.

The circuit diagram of the sense amplifier for this design is shown in Fig. 18. The sense amplifier buffers the BL and the RBL by its input transistors, converting V_x and V_{REF} into I_x and I_{REF}, respectively. I_x and I_{REF} feed the precharged-low
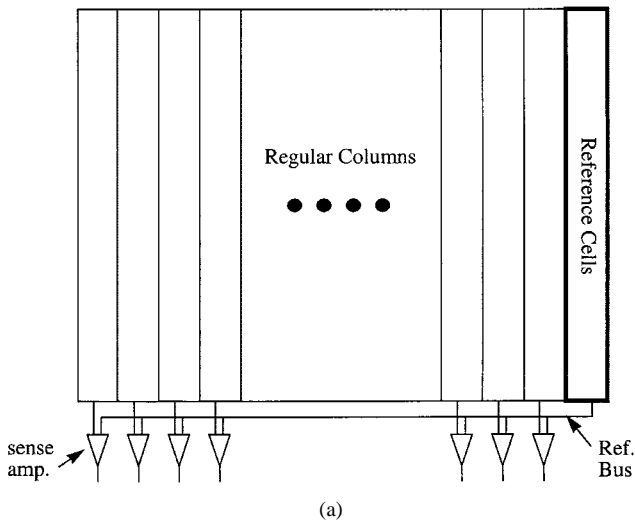
(a)
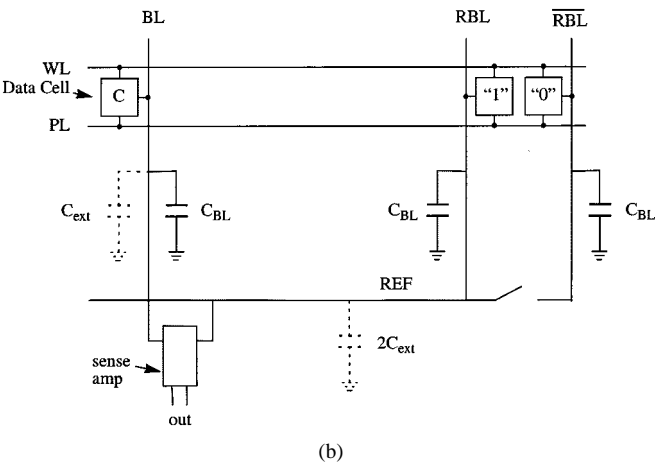
(b)

Fig. 17. (a) The block diagram of a 2C/WL reference scheme and (b) its corresponding circuit diagram [35].

nodes of a cross-coupled NMOS pair and compete in pulling up their corresponding nodes. The current with the higher magnitude will pull up its corresponding node faster and turns on the NMOS transistor of the opposite node. At this point, the pullup transistor controlled by SAP is turned on, activating the cross-coupled PMOS pair, to pull up the node with higher voltage to a full V_{DD}. This ends the sensing part of a read. A separate write-back circuit (not shown in Fig. 18) restores the read data into the cell via the BL.

Wood's proposal for balancing the bitline capacitance is difficult to achieve since only a simulated value of C_{ext} is known in advance. Also, adding extra capacitance to the bitline reduces the voltage available for sensing. Furthermore, due to nonlinearity of the voltage-to-current conversion in the sense amplifier, I_{REF} is not exactly the average of I_0 and I_1, assuming V_{REF} is the average of V_0 and V_1. In other words, assuming that the voltage-to-current conversion is expressed by $i = f(v), f((V_0 + V_1)/2) \neq 1/2(f(V_0) + f(V_1))$ unless f is a linear function. This translates to unequal noise margin for the "1" and the "0" signals.

A circuit technique that resolves these issues is shown in Fig. 19. All the bitlines in the array, including the RBL and the $\overline{RBL}$, are buffered by PMOS transistors at one end. The
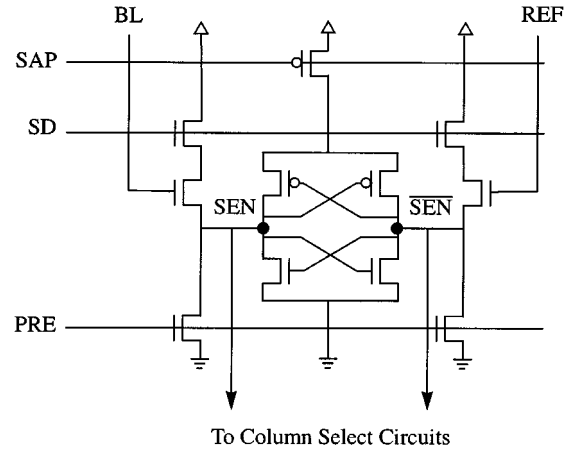

Fig. 18. Circuit diagram of a current-steering sense amplifier employed in the 2C/WL architecture [35].

PMOS transistors introduce little capacitive load to the bitlines and no capacitive imbalance. The nonlinear voltage-to-current conversion is bypassed in this design by first converting V_0 to I_0, V_1 to I_1, and then summing the two currents to generate $2I_{REF}$. The sum operation is simply achieved by tying the drains of the PMOS transistors together. Both $2I_{REF}$ and I_x are then mirrored via a set of NMOS transistors to generate I_{REF} and I_x, respectively, sinking charge from the opposite nodes of a cross-coupled PMOS pair. I_{REF} and I_x compete with each other in keeping their respective nodes down despite the source current supplied by the pullup transistor controlled by SAP. The current with the lower magnitude allows its node to be pulled up and turns ON the NMOS transistor pulling down the opposite node. Similarly, the current with the higher magnitude keeps down its corresponding node and keeps the NMOS transistor of the opposite side OFF. This process ends with one node pulled down to ground and the other node pulled up to V_{DD}.

There are two points to consider in the design of the sense amplifier of Fig. 19: the matching accuracy of the current mirrors and the sensing speed. The simple current mirrors used in this design can be replaced by more complicated current mirrors (such as the cascode current mirror [36]) to achieve better current matching. Obviously, the price for better matching is the added area of additional transistors (one additional transistor for the cascode current mirror) in each branch. The sensing speed of this design is limited by the capacitive load at the common node of the current mirror. This is based on the assumption that there is only one reference column for the entire array. If there are more than one reference column, they will divide the capacitive load and, therefore, provide higher sensing speed and better matching accuracy.

All the reference schemes discussed so far add only a fraction to the total array area. Obviously, the smaller the area overhead, the more attractive the reference scheme to high-density ferroelectric memories. However, the area overhead is not the primary design target for a low-density ferroelectric memory; rather, a higher accuracy is. Including one reference cell for each memory cell in the array consumes almost twice as much area as any of the reference schemes discussed in this paper. The reward, however, is a robustness that is hard
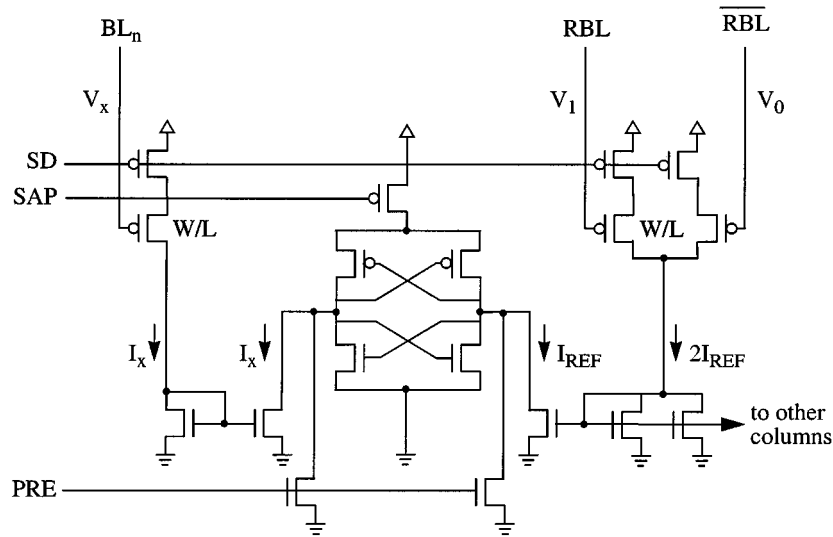
Fig. 19.   Circuit diagram of an improved 2C/WL sensing scheme.

to match by any other scheme. We present this in the following section.

F. A Self-Referenced, Fully Differential Architecture (2T-2C)

A fully differential cell [37] consists of two transistors and two capacitors (2T-2C), as shown in Fig. 20. A 2T-2C cell can be viewed as two adjacent 1T-1C cells sharing the same WL and PL but storing opposite data. C_{FE} dumps its charge on the BL while $C_{\overline{FE}}$ dumps its charge on the $\overline{BL}$. Since $C_{\overline{FE}}$ always stores a data value opposite to C_{FE}, the voltage difference between BL and $\overline{BL}$ is one of $V_1 - V_0$ or $V_0 - V_1$ depending on whether the stored data in C_{FE} is a "1" or a "0," respectively. This is twice the voltage level that is available for sensing in a 1T-1C architecture. The price paid for doubling the signal is doubling the cell area. Although this can be afforded for lower density memories (less than 256 kb), there is a drive toward using the 1T-1C architecture for higher density memories (beyond 1 Mb).

Similar to using a reference cell per row scheme, C_{FE} and $C_{\overline{FE}}$ fatigue at the same rate since they are accessed simultaneously. Also, the physical proximity of the two capacitors results in better matching characteristics compared to previous schemes. This makes the 2T-2C cell one of the most robust cells in ferroelectric memories.

G. Summary

Table 2 compares the six reference cells discussed in this section in terms of speed, density, accuracy in reference generation, sensing circuit complexity, and fatigue immunity. Among those presented, the fully differential (2T-2C) cell is the most robust in terms of sensing and the noise margin, but it is only suitable for low densities. Among the reference schemes using the 1T-1C cell, the 2C/2BL and 2C/WL schemes have superior sensing complexity and fatigue immunity, respectively.
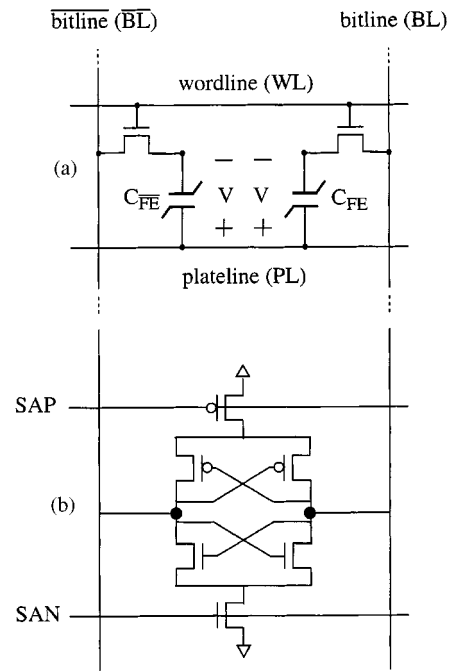

Fig. 20.   Circuit diagram of (a) a fully differential memory cell and (b) a sense amplifier.

V. FERROELECTRIC MEMORY ARCHITECTURE

Ferroelectric memories have borrowed many circuit techniques from DRAM's due to similarities of their cells and DRAM's mature architecture [38]. A *folded-bitline architecture*, for example, that was first introduced to replace an older *open-bitline architecture* in DRAM [39], [40] is now well adopted in FRAM [41]. The bitlines are *folded* to lie on the same side of a sense amplifier, as shown in Fig. 21, instead of lying *open* on opposite sides of the sense amplifier, to reduce chances of any bitline mismatch that could occur due to process variations.

Table 2
Comparison of Various Reference Generation Schemes

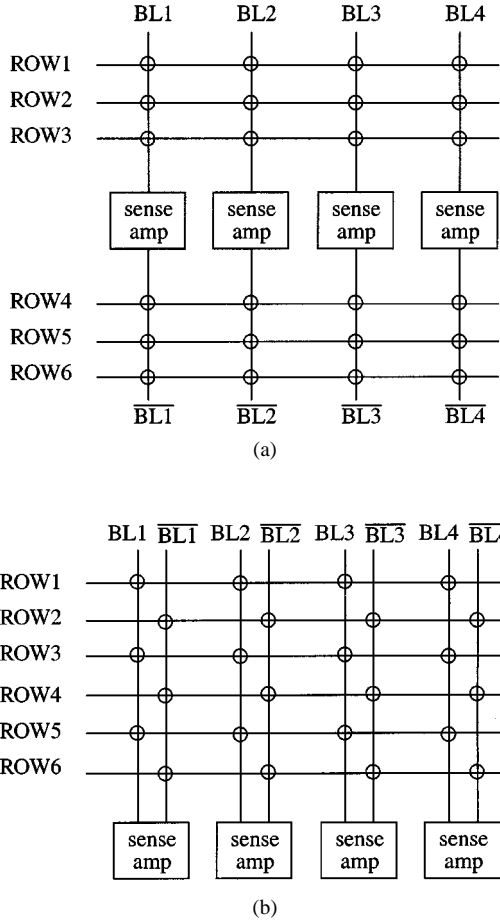| Basic Cell	Reference Scheme	speed	density	accuracy	sensing complexity	fatigue immunity	publication
1T-1C	1C′/BL	++	+	+	+++	+	[24][27]
	2x0.5C/BL	+	+	++	+++	+	[32]
	0.5C/0.5BL	+	+	+++	+++	+	this paper
	2C/2BL	++	+	+++	+++	+	[33]
	2C/WL	++	+	+++	+	+++	[34][35]
2T-2C	1C/Cell	+++	-	+++	+++	+++	[37]

Fig. 21. Block diagram of a ferroelectric memory with (a) an open-bitline architecture and (b) a folded-bitline architecture.

On the other hand, the requirement for pulsing the plateline in an FRAM has called for original circuit techniques that were not required in a DRAM. There are various memory architectures that have been developed for an FRAM with moving plateline. We discuss these architectures in Section V-A to V-D.

A plateline is slow to move due to its relatively high capacitance. This has motivated the researchers to innovate a constant-plateline architecture in which a data access is accomplished with the PL tied to $V_{DD}/2$, similar to a DRAM architecture. There is an obvious speed advantage to this ar-

chitecture, but it has its own disadvantages. First, it requires a regular refresh to all the memory cells, as we will see later in this paper. Second, it reduces the voltage range across the ferroelectric capacitor in the memory cell compared to that of the moving-plateline architecture. Referring to Fig. 22, if the PL is fixed to a constant voltage (such as $V_{DD}/2$), then the maximum range of voltage across the FE capacitor is V_{DD}, which is achieved by moving the BL from "0" to V_{DD}. On the other hand, if the PL is movable from "0" to V_{DD}, the voltage range across the FE capacitor is from $-V_{DD}$ to $+V_{DD}$ (a total of $2V_{DD}$). In this case, the effective voltage range of the hysteresis loop is $2V_{DD}$, which is twice the voltage range in the constant-plateline architecture. This extra width can ease the circuit design substantially when the capacitor's coercive voltage is a large fraction of V_{DD}.

Section V-E describes in detail a nondriven-plateline (constant-plateline) architecture. Section V-F describes a technique that only moves the plateline when it does not contribute to the memory access time. A dual-mode architecture is presented in Section V-G, followed by three architectures in Section V-H to J that describe architectures with nonconventional ferroelectric memory cells.

A. Wordline-Parallel Plateline (WL//PL)

Fig. 23 shows a simplified block diagram of a WL//PL architecture. As its name suggests, the PL is run parallel to the WL in this architecture. When a WL and PL pair is activated, an entire row that shares the same WL and PL is accessed at once. It is impossible, in this architecture, to access a single cell without accessing an entire row. This is in fact common in almost every RAM since the adjacent cells in a row store the adjacent bits of a byte, which are accessed simultaneously. Sometimes, the PL in this architecture is shared between two adjacent rows to reduce the array area by eliminating a metal line [24]. In this case, the unaccessed cells connected to an activated PL can be disturbed. This is due to the voltage that develops across the FE capacitors of the nonselected cells with the active PL. Ideally, one expects this voltage to be zero because the storage nodes of the cells should be floating. However, the parasitic capacitance of a storage node forms a capacitor divider with the FE capacitor itself and produces a nonzero voltage across the FE capacitor. For a stored "0" data, the disturb voltage is in the direction
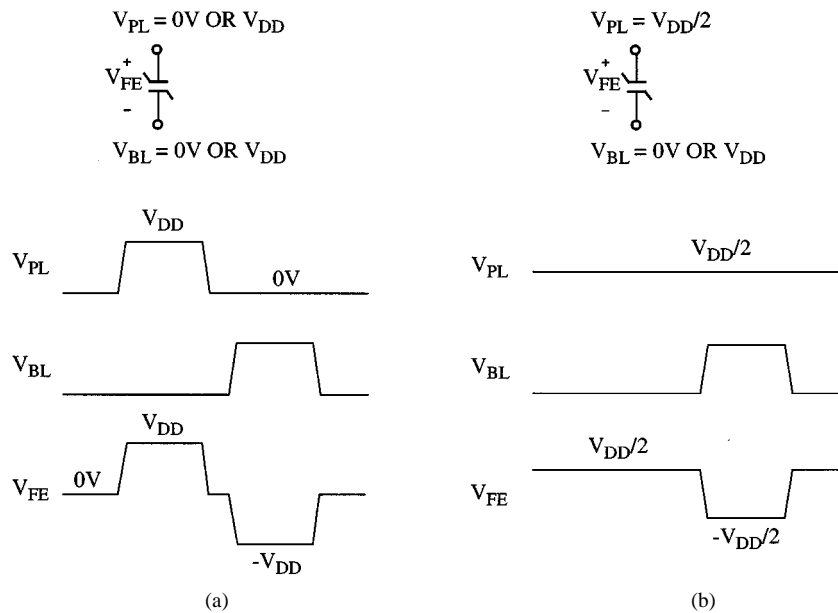
Fig. 22. (a) Moving-PL architecture: the voltage range of V_{FE} is $2V_{DD}$. (b) Nondriven-PL architecture: the voltage range of V_{FE} is V_{DD}.

that reinforces the "0." For a stored "1" data, however, the disturb voltage is in the direction of flipping the data. If this voltage is small enough (much less than the coercive voltage of the FE capacitor), it can be ignored. Otherwise, a data "1" can be flipped by a sequence of small voltage disturbances.

B. Bitline-Parallel Plateline (BL//PL)

Fig. 24 shows an array architecture in which the PL is run parallel to the BL [42], hence the name BL//PL for the architecture. Unlike the previous architecture, only a single memory cell can be selected by a simultaneous activation of a WL and a PL. This is the memory cell that is located at the intersection of the WL and the PL. It is possible to select more than one memory cell in a row by activating their corresponding platelines.

This architecture absorbs the function of a y-decoder in the selection of the platelines. In fact, the activation of the sense amplifiers is controlled by the same signal as the PL. Therefore, only one sense amplifier is activated if only one memory cell needs to be accessed. This reduces the power consumption significantly. On the other hand, if an entire row needs to be accessed, then all the platelines are selected simultaneously, hence increasing the dynamic power consumption due to charging and discharging the platelines.

The main disadvantage of this architecture is that activating a PL could disturb all the cells in the corresponding column [42]. This is very similar to the situation discussed for the WL//PL architecture with PL shared between two adjacent rows.

C. Segmented Plateline (Segmented PL)

A WL//PL architecture is power consuming and relatively slow because the PL is activated in its full length to access all the cells in the row at once. Also, a BL//PL architecture
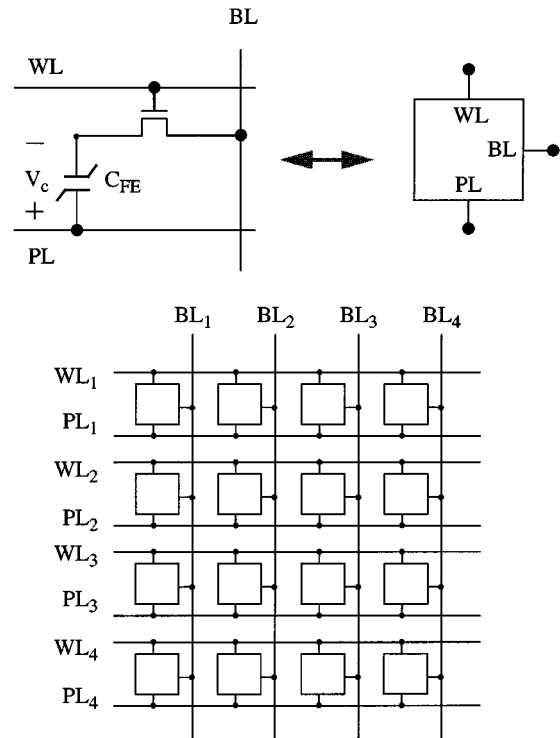

Fig. 23. Block diagram of a ferroelectric memory with WL//PL architecture.

could be power consuming if multiple platelines are activated to access multiple memory cells in the selected row. For larger arrays, the PL can be segmented into local platelines (LPL's) that run parallel to the WL and controlled by a global plateline (GPL) that runs parallel to the BL [27], [43]. As shown in Fig. 25, a GPL is ANDed with the WL to generate the signal for the LPL. Since the LPL is only connected to a few memory cells (eight in this example), it can respond much faster than a PL in the WL//PL architecture.
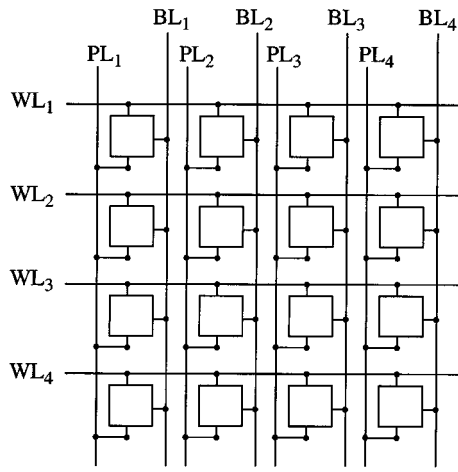
Fig. 24. Block diagram of a ferroelectric memory with BL//PL architecture.

Fig. 25. Block diagram of a ferroelectric memory with segmented-PL architecture [43].

Also, since the GPL is gated by the WL, there is no disturbance to the nonselected cells in the column, as it was in the BL//PL architecture.

Among the three architectures discussed so far, the segmented-PL architecture seems to be the most feasible architecture for a large-density ferroelectric memory. A compromise between speed and power consumption can be made by choosing the number of LPL's per GPL.

D. Merged Wordline/Plateline (ML) Architecture

A WL and its neighboring PL in a WL//PL architecture can be merged to form a single merged line[1] (ML) in an architecture proposed by Kang *et al.* [44]. Fig. 26 shows the circuit diagram of two 1T-1C memory cells or a single 2T-2C cell connected to two ML's (ML_1 and ML_2) and two bitlines (BL_n and BL_{n+1}). In the following, we describe the read/write operation of this cell, considering the 1T-1C option only.

A read/write operation is a multiphase operation since a merged line plays the double role of both a wordline and a plateline. We explain the write operation through the timing diagram of Fig. 27(a), in which we intend to write a "0" into C_1 and a "1" into C_2. The write operation consists of four distinct phases. During Δt_0, the BL_n is set to 0 V while BL_{n+1} is set to V_{DD}. During Δt_1, both ML_1 and ML_2 are raised to V_{DD}, forcing a "0" into C_1. Next, during Δt_2, ML_1 is pulled down to ground, leaving C_1 unchanged but forcing a "1" into C_2. During Δt_3, ML_1 is raised back to V_{DD} while ML_2 is pulled down. This is to write a "1" into C_1 if BL_n were at V_{DD} but, in this example, does not change the state of C_1. Finally, both the ML_1 and ML_2 are pulled down to ground ending the write operation.

A read operation begins by precharging the bitlines and followed by simultaneously raising the ML_1 and the ML_2, as shown in Fig. 27(b). The bitlines are sensed during Δt_1, and the sensed data are written back into the cell during Δt_2 and Δt_3, similar to a write operation.

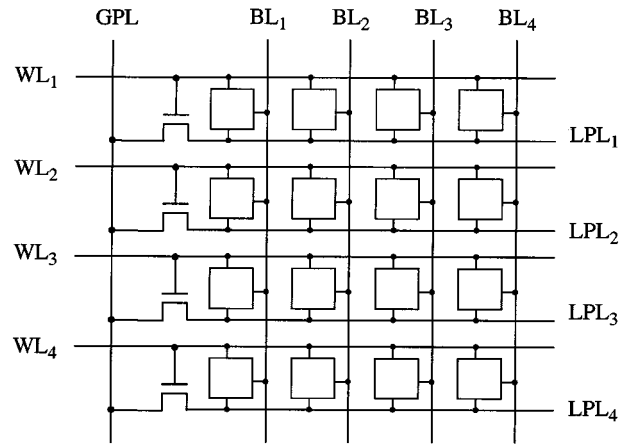[1]Kang *et al.* use the term *split wordline* in their paper [44] to refer to a merged wordline.
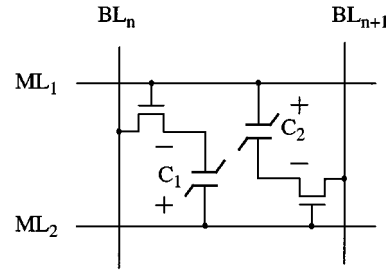

Fig. 26. Circuit diagram of a pair of ferroelectric memory cells for a merged-line architecture [44].

Compared to a WL//PL architecture, the ML architecture enjoys a shorter read access time, as the PL capacitance is now divided equally between two merged lines. This allows the merged line to respond twice as fast, assuming that the original wordline capacitance is negligible compared to the original plateline capacitance. The read/write cycle time, however, remains the same in both architectures. This is due to the fact that four transitions are required for a full read/write operation, instead of two in a WL//PL architecture. Finally, the ML architecture can achieve higher density compared to the WL//PL architecture due to reduced number of access wires, that is, using a single merged line instead of a wordline and a plateline. This higher density comes at the expense of complicated processing steps [44] such as stacking the bottom electrode of the FE capacitor right on top of the transistor gate and providing a side contact (plug) to the top electrode from the transistor source/drain area.

E. Nondriven Plateline Architecture

All the architectures discussed so far in this paper are plateline-driven architectures. That is, they require the PL to move from ground to V_{DD} in order to access a memory cell. In this section, we discuss an architecture proposed by Koike *et al.* [45] that keeps the PL at a constant voltage throughout the entire read/write operations, hence the name *Nondriven Plateline* (NDP) architecture. By not moving the PL, this architecture reduces the read/write access time substantially, as shown in Fig. 28.
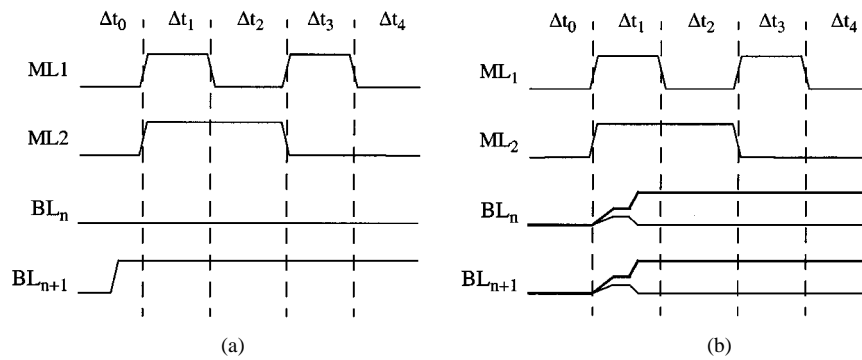
Fig. 27. Timing diagram for (a) a write operation and (b) a read operation of the memory cell shown in Fig. 26.
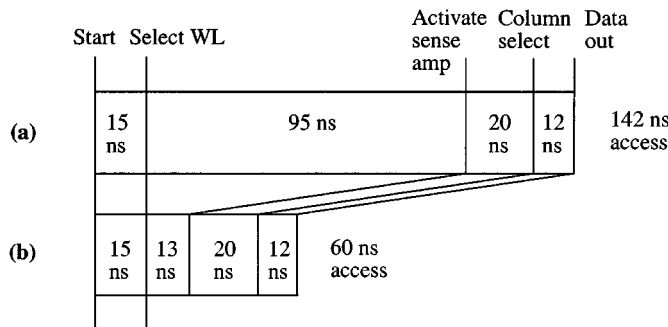

Fig. 28. Typical access time comparison of (a) a conventional driven-plateline architecture and (b) the NDP architecture [45].

We describe the function of this architecture by traversing the timing diagram of a read operation, shown in Fig. 29. Prior to a read operation, all the storage nodes are held at $V_{DD}/2$ against the leakage current of the PN junctions of the access transistors. This voltage is maintained by regularly refreshing all the storage nodes to $V_{DD}/2$. With the PL also fixed at $V_{DD}/2$, an FE capacitor experiences 0 V, or a small positive voltage that is controlled by the voltage compensation cycle.

The read operation begins by precharging all the bitlines to ground. Then, the WL (WL$_1$ in Fig. 29) is activated, initiating a charge sharing between the storage nodes in the selected row and their corresponding bitlines. This pulls down all the storage nodes close to ground and forces a voltage slightly less than $+V_{DD}/2$ across all the FE capacitors in the selected row. This voltage, instead of V_{DD} in previous architectures, is used to switch the capacitors storing a "1." Koike *et al.* [45] claim that $V_{DD}/2$ is enough to switch an SrBi$_2$Ta$_2$O$_9$-based FE capacitor, although it might not be enough to switch a more conventional, PZT-based FE capacitor. Let us assume that the voltage that appears on the BL at this point is V_0 or V_1, depending on whether the stored data is a "1" or a "0," respectively. The sense amplifier is then activated driving the BL to V_{DD} if it senses a V_1, or to ground if it senses a V_0.

Next, BL$_1$ and BL$_2$ are shorted together while the WL is held high. This forces the storage nodes back to their original voltage $V_{DD}/2$. Finally, the WL is pulled low, ending the row access. As shown in Fig. 29(b), a cell storing a "0" experiences a sequence of 0 V, $+V_{DD}/2$, $+V_{DD}/2$, 0 V. This sequence keeps the state of the FE capacitor intact. On the
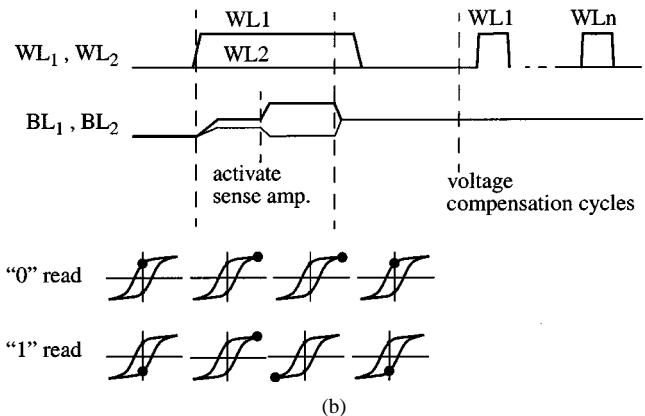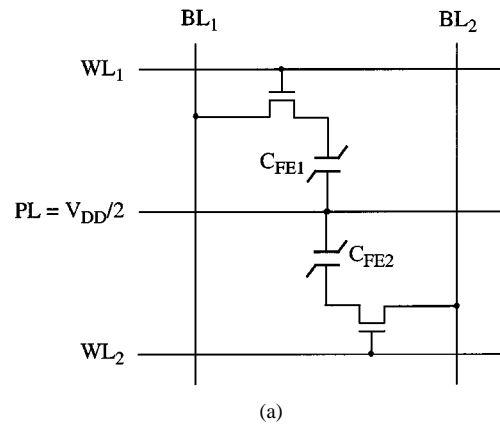

Fig. 29. (a) Circuit diagram of an NDP architecture and (b) its corresponding timing diagram for a read operation and voltage compensation cycles [45].

other hand, a cell storing a "1" experiences a sequence 0 V, $+V_{DD}/2$, $-V_{DD}/2$, 0 V. This sequence switches the state of the FE capacitor twice and, hence, restores the original data into the cell.

A voltage compensation cycle, as shown at the end of read operation in Fig. 29, is performed by sequentially raising the wordlines while keeping the bitlines all at $V_{DD}/2$.

A write access is similar to the read access except for the BL that is forced to V_{DD} or 0 V to write a "1" or "0" to a cell, respectively.

As mentioned earlier, this architecture reduces the access time by not moving the PL. Koike *et al.* [45], [46] have shown

that this architecture shaves 82 ns off an otherwise 142-ns access time in a 1-Mb FRAM prototype fabricated in a 1.0-μm CMOS technology.

F. Bitline-Driven Architecture

The nondriven-plateline architecture, explained in the previous section, limits the voltage available for switching an FE capacitor to $V_{DD}/2$. Also, by tying the PL to $V_{DD}/2$, it requires regular refreshing of the storage nodes to $V_{DD}/2$. Hirano *et al.* [28], [29] proposed a different read scheme in a similar architecture that ties the PL to ground, instead of $V_{DD}/2$, and precharges (drives) the BL to V_{DD} prior to activating the WL. This allows a full V_{DD} to develop across the FE capacitor for the read operation. Furthermore, by tying the PL to ground instead of $V_{DD}/2$, the need for refreshing is eliminated. After the read is complete, the PL is pulsed momentarily to V_{DD} to restore the read data.

A circuit diagram for this architecture and its corresponding timing diagram for a read/write operation is shown in Fig. 30. The circuit in the shaded area is used to drive both BL and $\overline{BL}$ to either V_{DD} or 0 V, depending on whether $\overline{DBLH}$ or BBLL is active, respectively. During Δt_0, all the bitlines are precharged to 0 V by an active DBLL. During Δt_1, all the bitlines are precharged to V_{DD}. This is just prior to activating a wordline. The wordline (WL$_0$) is activated during Δt_2 and stays active until the read/write is complete, whereas the PL is only pulsed during Δt_4, that is, after the read data are sensed by the sense amplifiers.

The bitline-driven architecture reduces the access time substantially by not moving the PL during the read, but does not reduce the read cycle time as it requires moving the PL for the restore operation. The segmented-PL architecture, explained earlier in this paper, can be combined with the bitline-driven architecture to further reduce the cycle time and the access time [28], [29].

G. Dual-Mode Ferroelectric Memories

We mentioned earlier in this paper that fatigue has always been a problem with ferroelectric capacitors. A dual-mode ferroelectric memory provides an easy solution to this problem by cycling ferroelectric capacitors only during a power-down and a power-up. During regular mode of operation, an FE capacitor is used as a nonswitching capacitor in its positive state and hence not cycled.

As an example of a dual-mode ferroelectric memory [48], consider a WL//PL architecture with all its platelines fixed to the ground. This implies that the voltage across a ferroelectric capacitor is always positive or zero, independent of the voltages of the bitline and the wordline. In this mode, the memory acts like a DRAM, with similar read/write/refresh operations. During a power-down, the data content of a row is sensed by the sense amplifiers, then written back to the FE capacitors by pulsing the PL. The ferroelectric memory is now in its nonvolatile mode, although not used until the next power-up. After the power-up, the content of a memory cell is sensed by a sense amplifier via pulsing the PL. Then the plateline is fixed to the ground, and the content of the sense
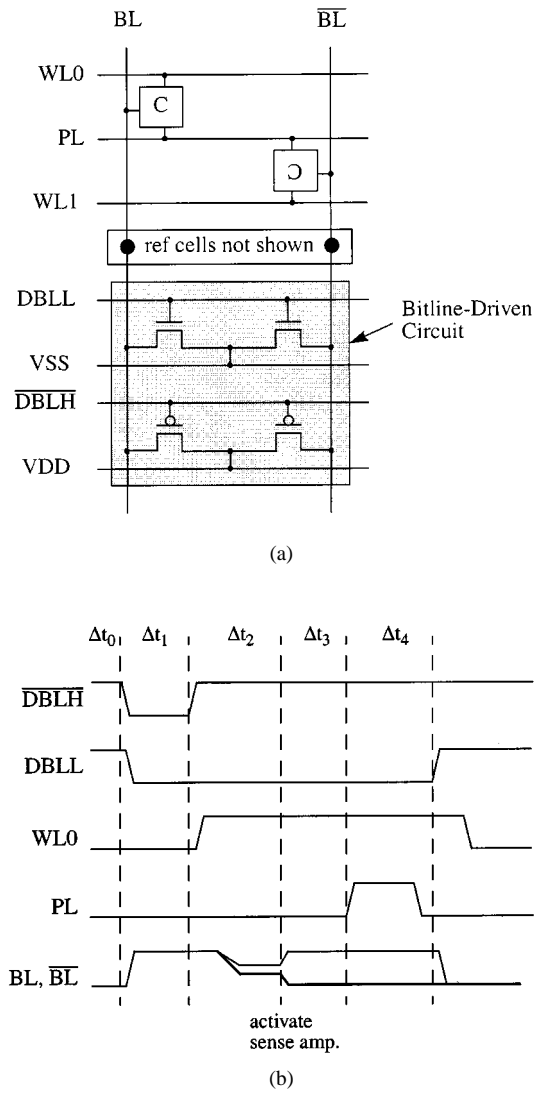

(a)

activate
sense amp.

(b)

Fig. 30. (a) Bitline-driven architecture and (b) its read/rewrite timing diagram [29].

amplifier is written back to the cell. At this point, the ferroelectric memory resumes operations in its volatile mode.

Shadow SRAM [49] is another example of a dual-mode architectures in which the ferroelectric capacitor is not used at all during regular operations. Each cell in a shadow SRAM consists of two halves as shown in Fig. 31. The top half of the circuit is a regular SRAM cell. The bottom half, shown shaded, is equivalent to a fully differential ferroelectric memory cell. During the regular operation, the memory is essentially an SRAM as the bottom half of the cell is isolated from the top half by a low voltage on STO. When a power shutdown is going to happen, STO is raised to V_{DD}, therefore connecting the two parts together. STO, in effect, acts like a wordline for the ferroelectric memory cell while the internal nodes of the SRAM cell act as its bitlines. As shown in the timing diagram of Fig. 31, the PL is pulsed during Δt_1 to force the internal nodes of the SRAM cell to write their stored data into the FE capacitors. The memory is ready for a power shutdown as soon as the STO line is pulled low, disconnecting the FE capacitors from the SRAM cell.
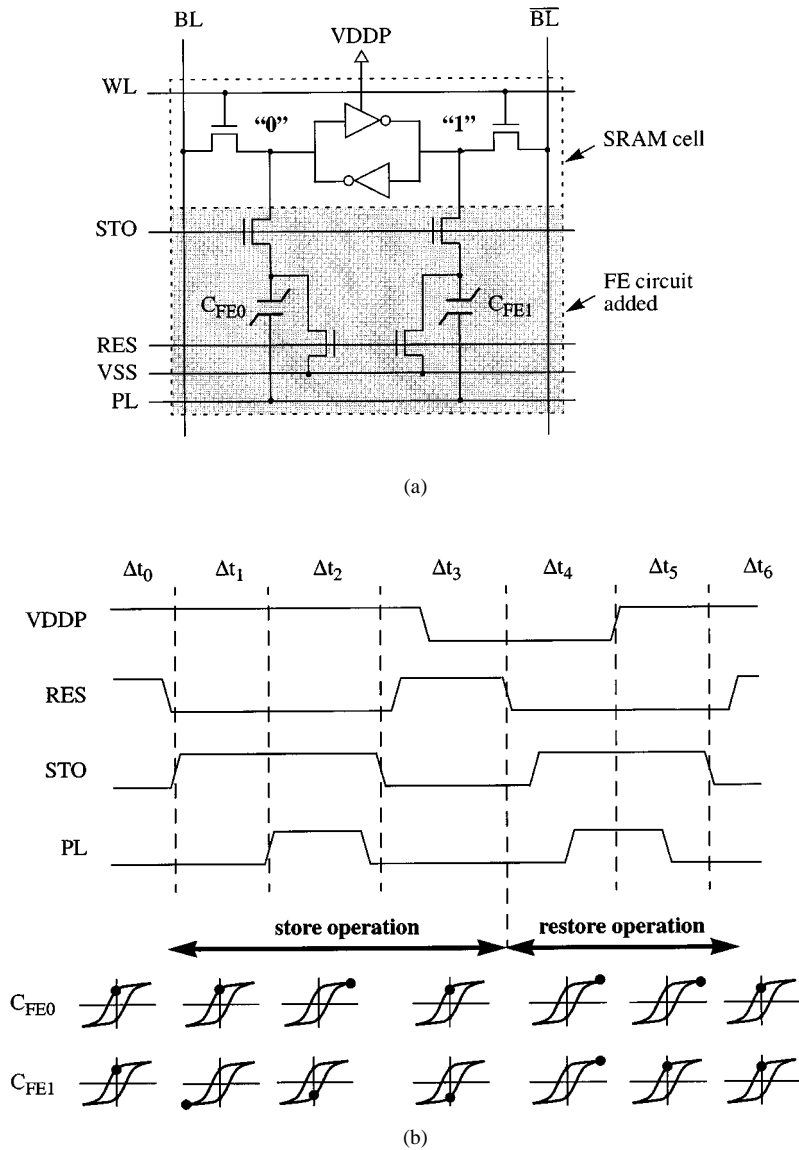
Fig. 31. Circuit diagram of a shadow SRAM and (b) its corresponding timing diagram for a store/restore operation [49].

A timing diagram for a restore operation follows that of a store operation (refer to Fig. 31). Prior to a power-up, STO is activated to connect the FE capacitors to storage nodes of the SRAM cell. Then the PL is pulsed to force the capacitors' stored charge onto the internal nodes of the SRAM cell. The FE capacitor with a stored "1" produces a higher voltage on its corresponding storage node, compared to the voltage produced by the other capacitor. This provides a proper initial condition on the storage nodes of the SRAM cell to move them to the right logic level during the power-up (Δt_5). At this point, STO is deactivated to disconnect the FE capacitors from the storage nodes of the SRAM cell. RES is activated during Δt_3 and beyond Δt_6 as a precautionary measure to maintain a reset voltage of 0 V across the FE capacitors.

All the architectures discussed so far in this section use the conventional 1T-1C memory cell. In the following, we present three architectures that employ nonconventional memory cells.

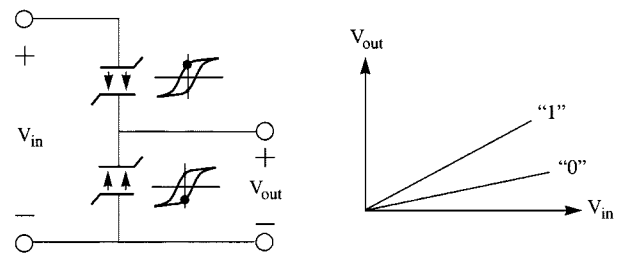Fig. 32. Circuit diagram of a transpolarizer and its input–output voltage charcteristic.

H. Transpolarizer-Based Architectures (Transpolarizer-Based Memory Cell)

A transpolarizer [50]–[53] consists of two identical ferroelectric capacitors in series, as shown in Fig. 32. If opposite states are stored in its capacitors, a transpolarizer acts as a capacitor divider with a dividing ratio determined by the digital
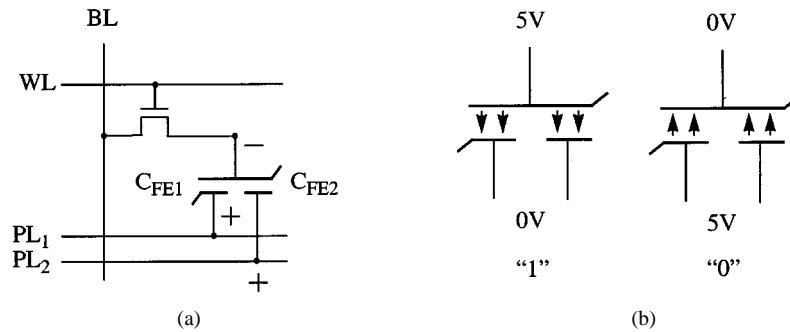
Fig. 33. (a) Circuit diagram of a transpolarizer-based cell and (b) applied voltage corresponding to writing a "1" or a "0" into a memory cell.

states of its capacitors. A "0" is stored in a transpolarizer by storing a positive state in its top capacitor and a negative state in its bottom capacitor. Similarly, a "1" is stored by reversing the two states. The capacitor divider ratio (V_{out}/V_{in}) is less than 0.5 for a stored "0" and greater than 0.5 for a stored "1." This is because an FE capacitor exhibits a higher capacitance in its negative state compared to its positive state when subjected to a positive voltage.

The main idea of using a transpolarizer to store binary data was first introduced in the 1950's [50]. Three decades later, Eaton et al. [51]–[53] proposed a transpolarizer-based memory cell that was suitable for implementation in VLSI. Only last year, however, a fully functional, 8-kbit test chip was reported by Tanabe et al. [54], [55] to successfully demonstrate this architecture. In the following, we describe the read and write operations in this architecture based on [54].

The memory cell, as shown in Fig. 33, consists of a transpolarizer and an access transistor. The cell, which is referred to as 1T-2C by Tanabe et al. [54], is connected to two plate-lines PL_1 and PL_2. To write a "0" or a "1" into the cell, the WL is activated, the BL is respectively grounded or raised to V_{DD}, and both PL_1 and PL_2 are raised simultaneously to V_{DD}. With the voltage convention adopted in Fig. 33, writing a "0" into the cell is equivalent to storing positive states in both capacitors. Similarly, writing a "1" into the cell is equivalent to storing negative states in both capacitors.

A read operation is performed using the timing diagram of Fig. 34. Prior to activating the WL, the bitlines are precharged to $V_{DD}/2$, the PL_1 is raised to V_{DD}, and the PL_2 is grounded. Therefore, at the beginning of Δt_2, the storage node of the cell (not shown in the figure) moves slightly above or below $V_{DD}/2$ depending on the stored data's being "1" or "0," respectively. Then, the WL is activated during Δt_2 to let the bitline voltage drop slightly below $V_{DD}/2$ for a stored "0," or rise slightly above $V_{DD}/2$ for a stored "1." This voltage change on the bitline is then sensed by the sense amplifier during Δt_3.

The readout scheme, as described above, is destructive and, hence, is followed by a write-back to the cell during Δt_4 and Δt_5. PL_2 is pulled up during Δt_4 and pulled down simultaneously with PL_1 while the BL is kept at its previous value by the sense amplifier. A write-back is completed during Δt_5
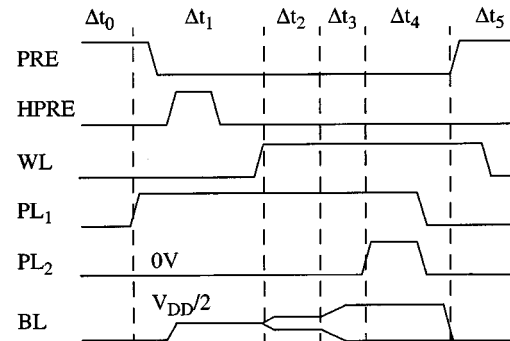

Fig. 34. Timing diagram for a read operation of a transpolarization-based cell shown in Fig. 33. PRE and HPRE are used to precharge the BL to 0 V and $V_{DD}/2$, respectively.

when the WL is deactivated after PRE forces both the BL and the storage node to the ground.

Two important points about this architecture are worth mentioning.

1) Unlike architectures using the 1T-1C cell, the transpolarizer-based architecture enjoys a much simpler, and more reliable, reference scheme: the reference voltage is simply $V_{DD}/2$, which is not affected by process variations or by ferroelectric capacitor degradations.

2) A transpolarizer-based (1T-2C) cell requires an additional ferroelectric capacitor and an extra plateline compared to the 1T-1C cell. However, as mentioned by Tanabe et al. [55], the 1T-2C cell requires smaller capacitance value compared to that of the 1T-1C cell to develop the same signal level on the BL.

These two points make the transpolarizer-based architecture suitable for high-density ferroelectric memories.

I. Cross-Point Array of Ferroelectric Gain Cells

Fig. 35 shows a memory architecture that eliminates the plateline altogether and achieves a high-density ferroelectric memory with nondestructive readout scheme [56]. The architecture consists of a crosspoint array of active ferroelectric memory cells, called gain cells. A gain cell, shown in Fig. 35, consists of an FE capacitor, a linear capacitor, and a single transistor. The linear capacitor adds almost no area cost to the cell, as it is only an extended part of the gate over the diffusion area. Also, interface-related problems are of less concern for the FE capacitors, as they are formed on a single
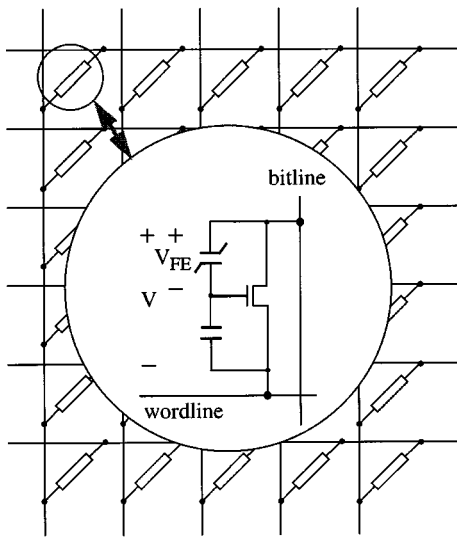
Fig. 35. Crosspoint array of gain cells [56]. No plateline is required in this architecture.

Fig. 36. Current–voltage characteristic of the gain cell shown in Fig. 35.

side of the FE material. The series combination of the FE capacitor and the linear capacitor forms a capacitor divider, similar to the transpolarizer, with its output amplified by the transistor; hence the name gain cell.

All wordlines and bitlines are at $V_{\rm DD}$ in a standby mode of operation, hence applying 0 V to all cells in the array. During the write operation [56], a sequence of multilevel signals is applied to the wordlines and the bitlines that guarantee a full writing voltage of $\pm V_{\rm DD}$ across the selected cells and only an insufficient $\pm V_{\rm DD}/3$ across the nonselected cells. The FE capacitors are assumed undisturbed under $\pm V_{\rm DD}/3$.

A read operation begins by precharging all the bitlines to $V_{\rm DD}$, followed by slightly lowering the voltage of the WL from $V_{\rm DD}$. The bitlines begin to discharge via the gain cells to the WL. The cells storing a "0" have a higher sink current compared to those storing a "1," as shown in the current–voltage characteristic of the cells in Fig. 36. Therefore, the bitlines connected to selected cells storing a "1" are pulled lower compared with those connected to cells storing a "0." This voltage slide on the bitlines are then sensed by the sense amplifiers, ending a nondestructive read operation.

J. Chain FRAM (NAND Architecture)

A chain FRAM (CFRAM) architecture [57], [58] is similar to a NAND Flash memory architecture [1], in which a number of memory cells are grouped in series to share a single contact to the bitline. In a CFRAM, a chain of memory cells shares a single contact to the bitline at one end and a single contact to the plateline at the other end. The combined effect of reducing the number of contacts on the bitline and the plateline is reduced access time and reduced chip area.

Fig. 37 shows the circuit diagram and a cross-section layout of two groups of cells, called *cell blocks* [57]. A cell block is terminated by a BL at one end and a PL at the other end. In a standby operation, all the wordlines are at $V_{\rm DD}$ to guarantee 0 V across all the FE capacitors in the block. In the active operation, a cell is accessed by grounding its corresponding WL and raising the Block-Select (BS)
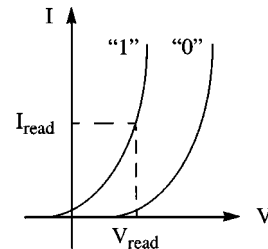
signal. All remaining wordlines remain high, allowing the BL voltage and the PL voltage to reach the selected cell via the chain of access transistors.

Takashima *et al.* [57] have shown that both the driven-plateline and the nondriven-plateline read schemes (refer to Section V-E) can be incorporated in this architecture. No refreshing is required in the latter case since all the FE capacitors are short-circuited during a standby operation via their parallel transistors.

For a fixed number of cells per bitline, the CFRAM architecture has less bitline capacitance compared to the 1T-1C architecture because only one cell in a cell block requires a direct contact to the bitline. The rest of the cells contribute less capacitance to the bitline as they share diffusion area with neighboring cells. This allows the designer to increase the number of cells per bitline by increasing the number of cells (N) per cell block. On the other hand, increasing N has an adverse effect on the readout delay of the cell as both the equivalent series resistance and the equivalent parasitic capacitance of the cell block increase with N. Also, increasing N beyond a certain point ($N = 32$ according to [57]) increases the total bitline capacitance as the parasitic capacitance of the transistors in series become the dominant part of the bitline capacitance. Therefore, a compromise must be made among the bitline capacitance, the readout delay, and the chip size. Takashima *et al.* [57] have shown that at 1024 cells per bitline, and 16 cells per cell block, the bitline capacitance of the CFRAM is comparable to that of the 1T-1C architecture while the total chip area is reduced to 63% of that of the 1T-1C architecture.

Takashima *et al.* [57] have pointed out that a small transient disturbance occurs across the nonselected FE capacitors during read/write operation, due to the ON resistance of the parallel transistors. It is claimed [57], [58], however, that the amplitude of the disturbance voltage is in the tolerable range if $N \geq 16$ and the data write is enforced over at least 8 ns in a 0.45-μm technology.

K. Summary

Table 3 summarizes the main features of ferroelectric memory architectures discussed in this section. Except for the SRAM-based and the DRAM-based architectures, all listed architectures offer nonvolatile memory operations.

The architectures using the 1T-1C memory cell require a reliable reference voltage that can be generated by one of the techniques listed earlier in Table 2. A transpolarizer-based architecture requires only a fixed reference voltage $V_{\rm DD}/2$
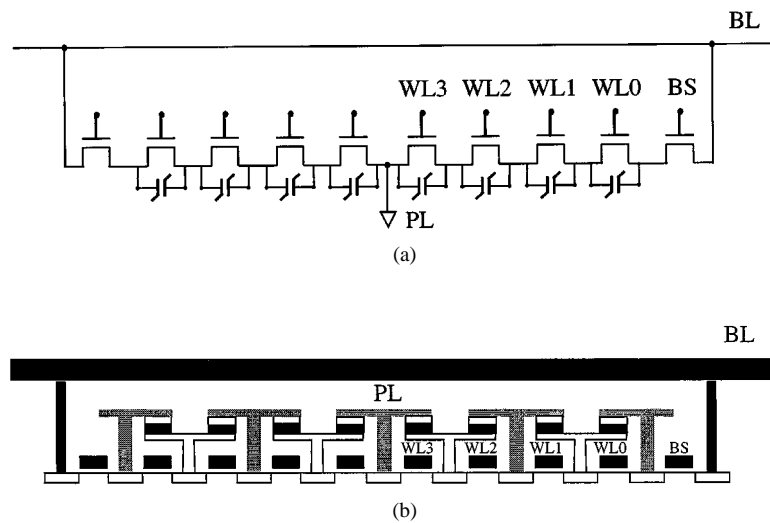
Fig. 37. (a) Circuit diagram of two cell blocks and (b) its corresponding cross-section layout in a chain FRAM architecture [57].

Table 3
Features of Various Ferroelectric Memory Architectures

Basic Cell	Architecture	Advantages	Disadvantages	Ref.
1T-1C	WL-Parallel PL	no disturb	—	—
	BL-Parallel PL	less power	disturb	[42]
	Segmented PL	less power, less disturb, high density	high area overhead (for low density)	[43]
	Merged WL-PL	high density	multi-phase control signals	[44]
	NonDriven PL	high speed	refreshing required, $V_{DD}/2$ for switching	[45][46]
	Bitline Driven	high speed	—	[28][29]
SRAM cell + FE capacitor	Dual-Mode: SRAM-based	less fatigue, higher speed	low density, more power	[52]
DRAM cell + FE capacitor	Dual-Mode: DRAM-based	less fatigue,	refreshing required, low density	[53]
1T-2C	Transpolarizer-based	fixed reference voltage, high density	—	[50][51] [54][55]
Gain Cell	Gain-cell Array	high density	multi-level control signals	[56]
Chain 1T-1C	Chain FRAM	high density	—	[57][58]

while achieving high density using an area/cell comparable to that of the 1T-1C cell. Two other architectures that promise high densities are the gain-cell array and chain FRAM.

VI. FUTURE TRENDS

Looking toward the future, we anticipate progress in three areas: density, access and cycle times, and use as an embedded memory in system-on-chip technology.

The density of commercial ferroelectric memories has improved dramatically over the past three years from 64 to 256 kb, with 1-Mb densities expected soon. This is entirely consistent with a Moore's-law improvement in density that has held true for several generations of DRAM, SRAM, and Flash memories. Ferroelectric process innovations such as stacked capacitors foreshadow further storage density improvements. If this trend continues, conservative estimates would indicate that commercial 64-Mb densities are on the horizon. Memory densities of 1 Gb appear to be technically feasible [59] using 0.18-μm technology.

Memory access and cycle times have not changed dramatically over the past few generations of commercial ferro-

electric memory IC's, even though new applications will undoubtedly demand improvements. Technology-independent architectural enhancements such as pipelined architectures conceivably could allow access times in the 5–10-ns regime. Whether commercial memories appear with this capability will depend on whether high-performance applications can be identified that require nonvolatile storage.

System-on-chip integration will require high-density embedded memories compatible with application-specific integrated circuit (ASIC) processes. Ferroelectric process technology can be adapted to minimally affect ASIC fabrication processes [3] at a very attractive overall die cost. An especially noteworthy feature is that ferroelectric memory technology can save several mask steps over EEPROM technologies and will eliminate the need for the associated high-voltage processes [8], a particularly troublesome requirement. Furthermore, ferroelectric memories offer simultaneously high-speed, low-voltage, and low-energy write capability, a feature set that in combination is truly unique amongst nonvolatile memory technologies. Collectively, these factors will serve as the enabling driver for many forthcoming applications, first in contactless smart cards and then in wireless Web-enabled phones and personal digital assistants.

VII. Conclusion

Ferroelectric memory research has developed in three principle directions: material processing and technology, capacitor modeling, and circuit design. In this paper, we focused on circuit innovations in reference voltage generation techniques and ferroelectric memory architectures.

Robust design of a reference voltage generator is essential to a successful design of any ferroelectric memory architecture using the 1T-1C cell, especially for densities of 1 Mb and beyond. The 1T-2C (transpolarizer-based) architecture, using a fixed reference voltage and cell size similar to the 1T-1C cell, is a suitable candidate for high-density ferroelectric memories. The 2T-2C architecture remains the preferred architecture for lower densities due to its robustness to process variations and its proven performance.

References

[1] B. Ricco, G. Torelli, M. Lanzoni, A. Manstretta, H. Maes, D. Montanari, and A. Modelli, "Nonvolatile multilevel memories for digital applications," *Proc. IEEE*, vol. 86, pp. 2399–2421, Dec. 1998.

[2] P. Pavan, R. Bez, P. Olivi, and E. Zanoni, "Flash memory cells—An overview," *Proc. IEEE*, vol. 85, pp. 1248–1271, Aug. 1997.

[3] K. Amanuma, T. Tatsumi, Y. Maejima, S. Takahashi, H. Hada, H. Okizaki, and T. Kunio, "Capacitor-on-Metal/Via-Stacked-Plug (CMVP) memory cell for 0.25 μm CMOS embedded FeRAM," in *Tech. Dig. IEEE Int. Electron Devices Meeting*, 1998, pp. 363–366.

[4] J. Yamada, T. Miwa, H. Koike, H. Toyoshima, K. Amanuma, S. Kobayashi, T. Tatsumi, Y. Maejima, H. Hada, H. Mori, S. Takahashi, H. Takeuchi, and T. Kunio, "A 128 kb FeRAM macro for a contact/contactless smart card microcontroller," in *ISSCC Dig. Tech. Papers*, 2000, pp. 270–271.

[5] R. E. Jones Jr., "Ferroelectric nonvolatile memories for embedded applications," in *Proc. IEEE Custom Integrated Circuits Conf.*, 1998, pp. 431–438.

[6] T. Miwa, J. Yamada, Y. Okamoto, H. Koike, H. Toyoshima, H. Hada, Y. Hayashi, H. Okizaki, Y. Miyasaka, T. Kunio, H. Miyamoto, H. Gomi, and H. Kitajima, "An embedded FeRAM macro cell for a smart card microcontroller," in *Proc. IEEE Custom Integrated Circuits Conf.*, 1998, pp. 439–442.

[7] D. Jung, S. Lee, B. Koo, Y. Hwang, D. Shin, J. Lee, Y. Chun, S. Shin, M. Lee, H. Park, S. Lee, K. Kim, and J. Lee, "A highly reliable 1T/1C ferroelectric memory," in *Symp. VLSI Circuits Dig. Tech. Papers*, 1998, pp. 122–123.

[8] P. Zurcher, R. E. Jones, P. Y. Chu, D. J. Taylor, B. E. White Jr., S. Zafar, B. Jiang, Y. T. Lii, and S. Gillespie, "Ferroelectric nonvolatile memory technology: Applications and integration challenges," *IEEE Trans. Comp., Packag., Manufact. Technol. A*, vol. 20, pp. 175–181, June 1997.

[9] R. Moazzami, C. Hu, and W. H. Shepherd, "Endurance properties of ferroelectric PZT thin films," in *Tech. Dig. IEEE Int. Electron Devices Meeting*, Dec. 1990, pp. 181–184.

[10] R. Moazzami, N. Abt, Y. Nissan-Cohen, W. H. Shepherd, M. P. Brassington, and C. Hu, "Impact of polarization relaxation on ferroelectric memory performance," in *Tech. Dig. Symp. VLSI Technology*, May 1991, pp. 61–62.

[11] J. M. Benedetto, M. L. Roush, I. K. Lloyd, and R. Ramesh, "Imprint of ferroelectric PLZT thin-film capacitors with lanthanum strontium cobalt oxide electrodes," in *Proc. 9th Int. Symp. Appl. of Ferroelectrics*, 1994, pp. 66–69.

[12] R. Moazzami, P. D. Maniar, R. E. Jones Jr., A. C. Campbell, and C. J. Mogab, "Ferroelectric PZT thin films for low-voltage nonvolatile memory," in *Proc. IEEE Int. Nonvolatile Memory Technology Conf.*, 1993, pp. 44–47.

[13] E. Fuji, T. Otsuki, Y. Judai, Y. Shimada, M. Azuma, Y. Uemoto, Y. Nagano, T. Nasu, Y. Izutsu, A. Matsuda, K. Nakao, K. Tanaks, K. Hirano, T. Ito, T. Mikawa, T. Kutsunai, L. D. McMillan, and C. A. Paz de Araujo, "Highly-reliable ferroelectric memory technology with bismuth-layer structured thin films ($Y - 1$ family)," in *Tech. Dig. IEEE Int. Electron Devices Meeting*, 1997, pp. 597–600.

[14] P. Y. Chu, D. J. Taylor, P. Zurcher, B. E. White, S. Zafar, B. Jiang, B. Melnick, R. E. Jones Jr., and S. J. Gillespie, "Effects of film thickness and process parameters on the properties of $SrBi_2Ta_2O_9$ ferroelectric capacitors for nonvolatile memory applications," in *Proc. 10th Int. Symp. Appl. of Ferroelectrics*, 1996, pp. 329–332.

[15] A. Sheikholeslami and P. G. Gulak, "A survey of behavioral modeling of ferroelectric capacitors," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 44, pp. 917–924, July 1997.

[16] A. Sheikholeslami, P. G. Gulak, H. Takauchi, H. Tamura, H. Yoshioka, and T. Tamura, "A pulse-based, parallel-element macromodel for ferroelectric capacitors," IEEE Trans. Ultrason., Ferroelect., Freq. Contr., June 1999, to be published.

[17] B. Jiang, J. C. Lee, P. Zurcher, and R. E. Jones Jr., "Modeling ferroelectric capacitor switching using a parallel-elements model," *Integrat. Ferroelect.*, vol. 16, pp. 199–208, 1997.

[18] "HSPICE User's Manual," Meta-Software, Inc., vol. 2, 1996.

[19] J. A. Rajchman, "A survey of magnetic and other solid-state devices for the manipulation of information," *IRE Trans. Circuit Theory*, pp. 210–225, Sept. 1957.

[20] "RAMTRON's FM1808 Product Specifications," Ramtron Int. Corp., Colorado Springs, CO, 1999.

[21] "MICRON's MT48LC64M4A2 Data Sheet," Micron Tech. Inc., Boise, ID, 1999.

[22] E. M. Philofsky, "FRAM—The ultimate memory," in *Proc. 7th IEEE Int. Nonvolatile Memory Technology Conf.*, 1996, pp. 99–104.

[23] J. C. Burfoot and G. W. Taylor, *Polar Dielectrics and Their Applications*. Berkeley, CA: Univ. of California Press, 1979.

[24] W. Kraus, L. Lehman, D. Wilson, T. Yamazaki, C. Ohno, E. Nagai, H. Yamazaki, and H. Suzuki, "A 42.5 mm^2 1 Mb nonvolatile ferroelectric memory utilizing advanced architecture for enhanced reliability," in *Symp. VLSI Circuits Dig. Tech. Papers*, 1998, pp. 242–245.

[25] N. Abt, "Electrical measurement of ferroelectric capacitors for nonvolatile memory applications," *Proc. Materials Research Society Symp.*, vol. 200, pp. 303–312, 1990.

[26] Y. Chung, M. Choi, S. Oh, B. Jeon, and K. Suh, "A 3.3-V 4-Mb non-volatile ferroelectric RAM with a selectively-driven double-pulsed plate read/write-back scheme," in *Symp. VLSI Circuits Dig. Tech. Papers*, 1999, pp. 97–98.

[27] T. Sumi, N. Moriwaki, G. Nakane, T. Nakakuma, Y. Judai, Y. Uemoto, Y. Nagano, S. Hayashi, M. Azuma, E. Fujii, S. Katsu, T. Otsuki, L. McMillan, C. P. de Araujo, and G. Kano, "A 256 kb nonvolatile ferroelectric memory at 3 V and 100 ns," in *ISSCC Dig. Tech. Papers*, 1994, pp. 268–269.

[28] H. Hirano, T. Honda, N. Moriwaki, T. Nakakuma, A. Inoue, G. Nakane, S. Chaya, and T. Sumi, "2-V/100-ns nonvolatile ferroelectric memory architecture with bitline-driven read scheme & nonrelaxation reference cell," in *Symp. VLSI Circuits Dig. Tech. Papers*, 1996, pp. 48–49.

[29] ——, "2-V/100-ns nonvolatile ferroelectric memory architecture with bitline-driven read scheme and nonrelaxation reference cell," *IEEE J. Solid-State Circuits*, vol. 32, pp. 649–654, May 1997.

[30] T. Miyakawa, S. Tanaka, Y. Itoh, Y. Takeuchi, R. Ogiwara, A. M. Doumae, H. Takenaka, I. Kunishima, S. Shuto, O. Hidaka, S. Ohtsuki, and S. Tanaka, "A 0.5 μm 3 V 1T1C 1 Mbit FRAM with a variable reference bitline voltage scheme via a fatigue free reference capacitor," in *ISSCC Dig. Tech. Papers*, 1999, pp. 104–105.

[31] D. J. Jung, B. G. Jeon, H. H. Kim, Y. J. Song, B. J. Koo, S. Y. Lee, S. O. Park, Y. W. Park, and K. Kim, "Highly manufacturable 1T1C 4 Mb FRAM with novel sensing scheme," in *Tech. Dig. IEEE Int. Electron Devices Meeting*, Dec. 1999, pp. 279–282.

[32] T. A. Lowrey and W. L. Kinney, "Folded bit line ferroelectric memory device," U.S. Patent 5 541 872, July 30, 1996.

[33] D. R. Wilson and H. B. Meadows, "Voltage reference for a ferroelectric 1T/1C based memory," U.S. Patent 5 572 459, Nov. 5, 1996.

[34] A. G. Papaliolios, "Dynamic Adjusting Reference Voltage for Ferroelectric Circuits," U.S. Patent 5 218 566, June 8, 1993.

[35] S. W. Wood, "Ferroelectric memory design," M.A.Sc. thesis, Univ. of Toronto, Toronto, ON, Canada, 1992.

[36] D. Johns and K. Martin, *Analog Integrated Circuit Design*. New York: Wiley, 1997, pp. 137–138.

[37] J. T. Evans and R. Womack, "An experimental 512-bit nonvolatile memory with ferroelectric storage cell," *IEEE J. Solid-State Circuits*, vol. 23, pp. 1171–1175, Oct. 1988.

[38] W. Kinney, "Signal magnitudes in high density ferroelectric memories," *Integrat. Ferroelect.*, vol. 4, pp. 131–144, 1994.

[39] R. F. Harland, "MOS one-transistor cell RAM having divided and balanced bit lines coupled by regenerative flipflop sense amplifiers and balanced access circuitry," U.S. Patent 4 045 783, Aug. 30, 1977.

[40] R. C. Foss, "The design of MOS dynamic RAMs," in *ISSCC Dig. Tech. Papers*, 1979, pp. 140–141.

[41] W. L. Larson, "Non-volatile ferroelectric memory with folded bit lines and method of making the same," U.S. Patent 5 371 699, Dec. 6, 1994.

[42] R. Womack and D. Tolsch, "A 16 kb ferroelectric nonvolatile memory with a bit parallel architecture," in *ISSCC Dig. Tech. Papers*, 1989, pp. 242–243.

[43] R. E. Jones Jr., "Ferroelectric nonvolatile random access memory having drive line segments," U.S. Patent 5 373 463, Dec. 13, 1994.

[44] H. B. Kang, D. M. Kim, K. Y. Oh, J. S. Roh, J. J. Kim, J. H. Ahn, H. G. Lee, D. C. Kim, W. Jo, H. M. Lee, S. M. Cho, H. J. Nam, J. W. Lee, and C. S. Kim, "Multi-phase driven split word line ferroelectric memory without PL," in *ISSCC Dig. Tech. Papers*, 1999, pp. 108–109.

[45] H. Koike, T. Otsuki, T. Kimura, M. Fukuma, Y. Hayashi, Y. Maejima, K. Amanuma, M. Tanabe, T. Matsuki, S. Saito, T. Takeuchi, S. Kobayashi, T. Kunio, T. Hase, Y. Miyasaka, N. Shohota, and M. Takada, "A 60 ns 1 Mb nonvolatile ferroelectric memory with non-driven cell plate line write/read scheme," in *ISSCC Dig. Tech. Papers*, 1996, pp. 368–369.

[46] ——, "A 60 ns 1 Mb nonvolatile ferroelectric memory with a non-driven cell plate line write/read scheme," *IEEE J. Solid-State Circuits*, vol. 31, pp. 1625–1634, Nov. 1996.

[47] H. Fujisawa, T. Sakata, T. Sakiguchi, O. Nagashima, K. Kimura, and K. Kajigaya, "The charge-share modified (CSM) precharge-level architecture for high-speed and low-power ferroelectric memory," *IEEE J. Solid-State Circuits*, vol. 32, pp. 655–661, May 1997.

[48] H. Takata, T. Mnich, and D. Novosel, "Dual mode ferroelectric memory reference scheme," U.S. Patent 5 737 260, Apr. 7, 1998.

[49] K. Dimmler and S. Eaton, "Memory cell with volatile and non-volatile portions having ferroelectric capacitors," U.S. Patent 4 809 225, Feb. 28, 1989.

[50] C. F. Pulvari, "The transpolarizer: An electrostatically controlled circuit impedance with stored setting," in *Proc. IRE*, June 1959, pp. 1117–1123.

[51] S. S. Eaton Jr. and M. Parris, "One transistor memory cell with programmable capacitance divider," U.S. Patent 4 914 627, Apr. 3, 1990.

[52] ——, "SRAM with programmable capacitance divider," U.S. Patent 4 918 654, Apr. 17, 1990.

[53] ——, "DRAM with programmable capacitance divider," U.S. Patent 4 910 708, Mar. 20, 1990.

[54] N. Tanabe, S. Kobayashi, T. Miwa, K. Amanuma, H. Mori, N. Inoue, T. Takeuchi, S. Saitoh, Y. Hayashi, J. Yamada, H. Koike, H. Hada, and T. Kunio, "High tolerance operation of 1T/2C FeRAM's for the variation of cell capacitors characteristics," in *Dig. Tech. Papers Symp. VLSI Technology*, 1998, pp. 124–125.

[55] N. Tanabe, S. Kobayashi, H. Hada, and T. Kunio, "A high density 1T/2C cell with $V_{cc}/2$ reference level for high stable FeRAMs," in *Tech. Dig. IEEE Int. Electron Devices Meeting*, 1997, pp. 863–866.

[56] M. Aoki, H. Takauchi, and H. Tamura, "Novel gain cell with ferroelectric coplanar capacitor for high-density nonvolatile random-access memory," in *Tech. Dig. IEEE Int. Electron Devices Meeting*, 1997, pp. 942–944.

[57] D. Takashima and I. Kunishima, "High-density chain ferroelectric random access memory (chain FRAM)," *IEEE J. Solid-State Circuits*, vol. 33, pp. 787–792, May 1998.

[58] D. Takashima, S. Shuto, I. Kunishima, H. Takenaka, Y. Oowaki, and S. Tanaka, "A sub-40 ns random-access chain FRAM architecture with a 7 ns cell-plate-line drive," in *ISSCC Dig. Tech. Papers*, 1999, pp. 102–103.

[59] D. Takashima, Y. Oowaki, and I. Kunishima, "Gain cell block architecture for gigabit-scale chain ferroelectric RAM," in *Symp. VLSI Circuits Dig. Tech. Papers*, 1999, pp. 103–104.

Ali Sheikholeslami (S'98–M'99) is an Assistant Professor in the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada, and holds the L. Lau Junior Chair in Electrical and Computer Engineering. His research interests are in the areas of analog and digital integrated circuits, VLSI memory design (including SRAM, DRAM, and content-addressable memories), ferroelectric memories (circuit design and performance modeling), and multiple-valued memories. He has worked with industry on various memory design projects in the past few years. He worked on a modular SRAM project with Nortel, Ottawa, ON, in 1994, on an embedded DRAM project with Mosaid, Ottawa, in 1996, and on behavioral modeling of ferroelectric capacitors with Fujitsu, Kawasaki, Japan, in 1998. He is a coauthor for several journal and conference papers on ferroelectric memory design and modeling as well as multiple-valued memories. He received two U.S. patents in the area of content-addressable memories in 1998 and 1999.

P. Glenn Gulak (S'82–M'83–SM'96) received the Ph.D. degree from the University of Manitoba, Canada.

He is a Professor in the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada and holds the L. Lau Chair in Electrical and Computer Engineering. His research interests are in the areas of memory design, circuits, algorithms, and VLSI architectures for digital communications. He has supervised an active research group in the area of ferroelectric memories since 1990. From 1985 to 1988, he was a Research Associate in the Information Systems Laboratory and the Computer Systems Laboratory at Stanford University.

Dr. Gulak is a Registered Professional Engineer in the province of Ontario. He has received several teaching awards for undergraduate courses taught in both the Department of Computer Science and the Department of Electrical and Computer Engineering at the University of Toronto. He received a Natural Sciences and Engineering Research Council of Canada Postgraduate Scholarship. He has served on the ISSCC Signal Processing Technical Subcommittee since 1990 and currently serves as the Technical Program Chair for ISSCC.