Scaling of the Ferroelectric Field Effect Transistor and Programming Concepts for Non-volatile Memory Applications

vorgelegt von

Diplom-Ingenieur

Michael Fitsilis

aus Karditsa, Griechenland

Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online verfügbar.

Kurzfassung

Die zunehmende Bedeutung von nicht-flüchtigen Speichern für die Wissensgesellschaft im Informationszeitalter steht außer Frage. Bisher waren Speichersysteme unterteilt in schnelle jedoch flüchtige Arbeitsspeicher und langsame aber nicht-flüchtige Massenspeicher. Das langfristige Ziel ist aber der universelle Speicher, der hohe Zugriffsgeschwindigkeiten mit Nichtflüchtigkeit kombiniert. Ein Kandidat mit diesen Eigenschaften ist der ferroelektrische Feldeffekttransistor (FeFET), der Gegenstand der vorliegenden Arbeit ist.

Einleitend werden verschiedene nicht-flüchtige Speicherkonzepte vorgestellt. Die verschiedene Alternativen werden miteinander verglichen und es wird gezeigt, dass der FeFET eine Anzahl von Vorteilen gegenüber anderen Speicherbauelementen hat.

Anschließend werden die Grundregeln des Betriebes des FeFET, basierend auf dem MOSFET und dem ferroelektrischen Kondensator beschrieben. Mit einem Transistormodell und einem mathematischen Algorithmus für die Berechnung der ferroelektrischen Polarisation wird das Modell des FeFET hergeleitet.

Es wird auf die verschiedenen Herausforderungen für den FeFET eingegangen. Diese umfassen das Depolarisierungsfeld und den Leckstrom, die zur Verringerung der Polarisation und infolgedessen zu kurzen Datenspeicherzeiten führen. Im Falle des Leckstroms werden, basierend auf Ladungstransportmechanismen, Simulationen präsentiert mit Hilfe deren die Grenzen der Datenspeicherzeit geschätzt werden.

Die Möglichkeit der Miniaturisierung wird im Zuge der Bauelementskalierung für Speichersysteme ebenfalls untersucht und mit der für MOSFETs typische Skalierung verglichen. Es werden zwei Skalierungsregeln vorgestellt: „*constant gate stack scaling*" und „*variable gate stack scaling*". Die erste Regel kann, im Gegensatz zur zweiten, auch auf kleinste Dimensionen angewendet werden. Als Alternative zu der physikalischen Skalierung (Verkleinerung der Dimensionen) werden Mehrebenenzellen („Multi Level Cells" oder MLC) diskutiert.

Zunächst werden zwei Programmierkonzepte mit FeFETs vorgestellt: Das „*negative gate erase*" und das im Rahmen dieser Arbeit vorgeschlagene „*positive voltage erase*" Konzept. Sie werden auf Leistungsfähigkeit und Komplexität miteinander verglichen. Das „*positive voltage erase*" Konzept beseitigt die Notwendigkeit einer separaten Löschoperation und vereinfacht das Design eines Speicherchips nachhaltig.

Abschließend wird ein 1-Kbit Speicherchip vorgestellt, das auf dem „*positive voltage erase*" Konzept basiert. Das Design und die Simulation wurden auf dem sogenannten „schematic level" durchgeführt. Der Chip besteht aus einem Speicherarray und einer üblichen Peripherieelektronik (Decoder, Spannungstreiber, Leseverstärker).

Für die Simulationen wurden ein Schaltungssimulator und ein Bauelementsimulator eingesetzt.

Abstract

The importance of non-volatile memory for storage of digital information is without question. Research over the years has led to many different types of memory, each tailored to a specific need. Always, however, the search has continued for a universal type that combines high speed operation with non-volatility. One memory device with these properties is the ferroelectric Field Effect Transistor (FeFET), which is the object of study in this thesis.

First, a short introduction to non-volatile memories is given. Then a comparison of the various alternatives is made which shows that the FeFET has a number of advantages compared to other non-volatile memory devices.

Then the principles of operation of the FeFET are described based on the operation of the MOSFET and the ferroelectric capacitor. Using a transistor model and a mathematical algorithm for calculating the ferroelectric polarization, the FeFET model is derived.

Further, the various challenges that the FeFET faces are elaborated. These include the depolarization field and the leakage current that leads to the reduction of the remnant polarization and as a result, to short data retention times. For the case of the leakage current, simulations are presented based on current transport mechanisms to estimate the boundaries of data retention time for the device.

The miniaturization of the FeFET and comparison with the scaling of the MOSFET is considered next. Two scaling approaches are suggested, *variable* and *constant gate stack scaling*, of which the latter is applicable to even smaller dimensions than the former. As an alternative to physical miniaturization (dimension shrinking), multilevel cells (MLC) are discussed.

Two programming concepts with FeFETs are then investigated. One uses *negative gate erase* and the other a *positive voltage erase* method. They are compared in terms of efficiency and ease of realization. The *positive voltage erase* concept does away with the need for a separate erase operation and simplifies the memory chip design.

Finally, a 1-Kbit chip based on the *positive voltage erase* concept is introduced. The design and simulation were performed in schematic level. The memory design includes the FeFET matrix and peripheral electronics (decoders, voltage drivers, sense amplifiers).

For the simulations a circuit simulator and a device simulator were deployed.

Contents

1 INTRODUCTION

1.1 Non-volatile memories

Non-volatile solid state memories are of great importance in information technology. Compared to magnetic and optical mass storage devices they are more robust and have faster access times. This is mainly due to the fact that they don't consist of any moving parts. Under this category fall EEPROM (Electrically Erasable Programmable Read Only Memory), Flash (based on the floating gate transistor just like EEPROM), MRAM (Magnetic Random Access Memory) and FRAM (Ferroelectric Random Access Memory). With Flash and EEPROM electric charges are stored in a floating electrode isolated by an oxide. In MRAM a soft-magnetic layer and in FRAM a ferroelectric one is polarized in one of two directions. In FLASH and MRAM the read out mechanism is resistive, that is the memory element is switched between a low and a high resistive state. The resistivity is sensed at read out. For this reason the read out operation is non-destructive, in contrast to FRAM. The difference is that in FRAM at read out the ferroelectric is polarized to saturation polarization and the displacement current is detected. That is, the property that holds the information, the polarization of the ferroelectric, is modified and has to be refreshed afterwards.

Another type of non-volatile memory is the ferroelectric field effect transistor (FeFET) that offers a number of advantages compared to the other alternatives. It combines the fast operation of FRAM with the simple single transistor cell structure of Flash and additionally has a non-destructive readout so that information does not need to be rewritten (refreshed) after readout. Because it improves on the current technologies, it presents itself as the ideal candidate for a future generation universal type of memory.

A summary of the properties of the various non-volatile memory types is given in Table 1.1.

	Flash Memory NAND	FRAM	FeFET based Memory	MRAM
Write Access	μs	ns	ns	ns
Read Access	ns	ns	ns	ns
Erase necessary	yes	no	no	no
Destructive readout	no	yes	no	no
Data retention	10 years	10 years	max. 17 days[**]	10 years
Write endurance	$>10^5$ rewrites	unlimited rewrites[*]	---	$>10^{15}$ rewrites[***]
Basic Cell	1T	1T-1C	1T	1T-1MTJ

[*][66], [**][47], [***][67]

Table 1.1 Properties of the various non-volatile memories.

1.2 Ferroelectricity – FRAM – FeFET

In ferroelectric based memories information is stored via the polarization state. Polarization in one of two directions is interpreted as "0" or "1". Figure 1.1 shows one of the two stable states in a ferroelectric crystal. In FRAM the ferroelectric is used as a capacitor in the configuration shown in Fig. 1.2. The FeFET is derived from the 1T-1C (one transistor - one capacitor) FRAM basic cell by integrating the ferroelectric in the gate stack of the select transistor below the gate electrode, thus resulting in a smaller basic cell.
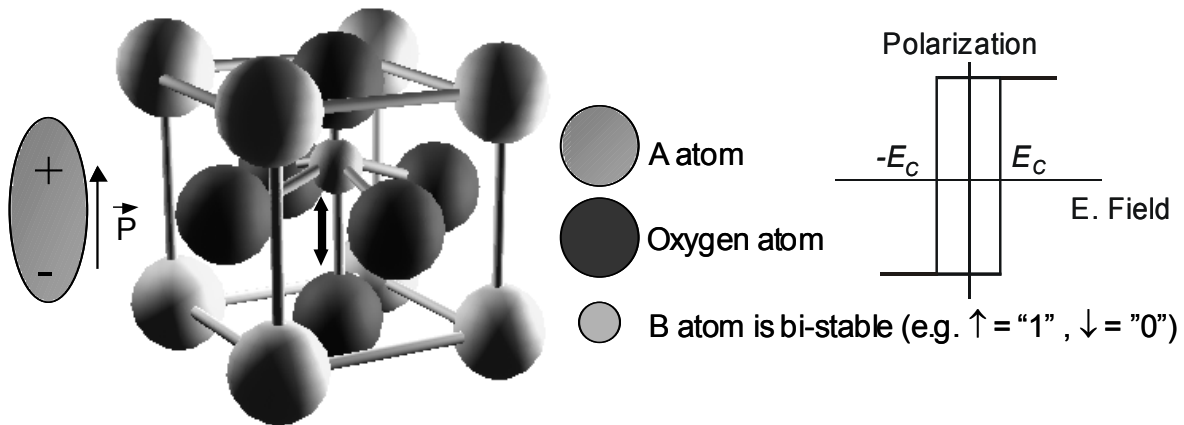
Fig. 1.1 A bi-stable ferroelectric crystal of perovskite type ABO₃ (left). The positive ion in the center has two stable states and in relation to the negatively charged oxygen ions gives the crystal two polarization states. Switching between the two states is possible by applying an electric field in the direction of the polarization, greater than the coercive field strength E_C as shown in the hysteresis plot of a single ferroelectric dipole (right) [77].
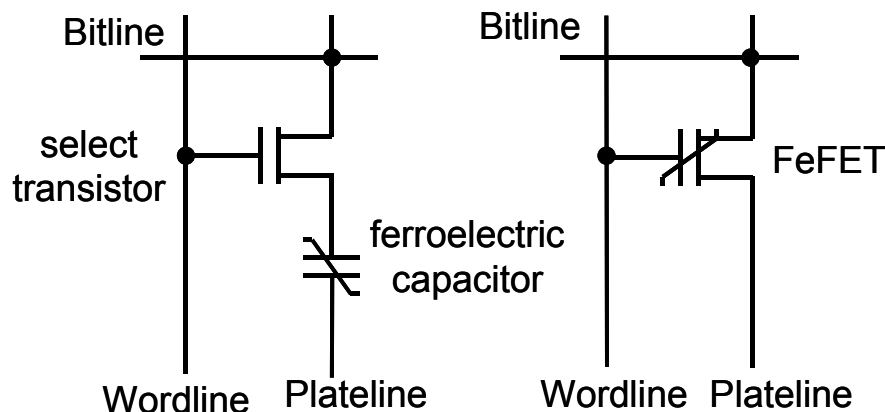
Fig. 1.2 FRAM (left) and FeFET (right) basic cells.

Because of the smaller cell, a higher integration density is possible. Moreover, as will be shown further on, this new structure enables a non-destructive read out. The 1-T (single transistor) structure of the FeFET is also found in the floating gate transistor (Flash) that currently dominates the market for non-volatile memories, because of the very high memory

capacities possible. Because of the similarities between the two devices, a comparison is made in section 2.4.

The FeFET is basically a MOSFET (Metal Oxide Semiconductor FET) with a ferroelectric layer inserted between the gate electrode and the silicon surface. This can be polarized by applying a voltage to the gate. For technological reasons, another layer has to be added between the silicon and the ferroelectric. As will be explained in section 2.6.2, this layer usually cannot be avoided. The integration of the ferroelectric in the MOSFET structure has proven to be a challenge with no effective solution to the present day. A short timeline of the FeFET research is given in the next section.

1.3 Timeline - Status

The FeFET has undergone a lot of research since its conception in the late 50s to this day. It has been constantly attempted to grow different ferroelectric materials on Si, either directly, resulting in the MFS (Metal Ferroelectric Semiconductor) FET, or indirectly using a dielectric buffer, thus resulting in the MFIS (Metal Ferroelectric Insulator Semiconductor) FET, or even including a floating electrode yielding the MFMIS (Metal Ferroelectric Metal Insulator Semiconductor) FET. Soon, however, it became evident that manufacturing was not trivial and the best devices showed retention times much less than the *10 years* set as a requirement by the industry. In [47] a retention time of $1.5 \cdot 10^6 s$ (*17 days*) is reported for an MFIS (Metal Ferroelectric Insulator Semiconductor) structure with BLT ($Bi_{3.45}La_{0.75}Ti_3O_{12}$) on Al_2O_3 on Si, which is the highest value that has been reported so far.

The FeFET research can be summed up in chronological order as follows ([9], [74]) :

- First patents filed on the FeFET [1]   (1957)
- Implementation with TGS (triglycine sulfate) on CdS [2]   (1963)
- First implementation on Si using $BiTiO_3$ [3]   (1974)
- Growth of oxygen free ferroelectric $BaMgF_4$ on Si to avoid
  the SiO_2 interface layer (MFS) [46], [80]   (1991)
- FeFET using $LiNbO_3$ on Si [48]   (1991)
- Working FeFET devices using $BaMgF_4$ [49]   (1993)
- MFMIS structure is proposed to achieve better interfaces [50]   (1995)
- MFMIS FETs with $PbZi_{1-x}Ti_xO_3$ (PZT) are reported [51], [78]   (1995)
- Y_2O_3, CeO_2, $SrTiO_3$, MgO, SiN are proposed as interface
  oxides [52-54], [86-87]   (1995-2000)
- $SrTiO_3$ is successfully grown on Si [55] and used as a gate
  dielectric in [85]   (1998-2000)
- MFMIS FETs with $SrBi_2Ta_2O_9$ (SBT) and high-k dielectrics
  are reported [56-59]   (1999-2000)
- A 1T-2C (1 transistor - 2 capacitors) FeFET cell structure is
  proposed [60], [61], [79]   (2000)
- Growth of ferroelectric $BaTiO_3$ on Si with a low interface states
  density [62]   (2000)
- Fatigue free $(Bi,La)_4Ti_3O_{12}$ (BLT) [63] is used in MFIS FETs [64]   (2001)
- Growth of $Pb_3Ge_5O_{11}$ (PGO) on $(Zr,Hf)O_2$ on Si (MFIS) [65]   (2002)
- Growth of Y_2MnO_3 (low-P_R ferroelectric) indirectly on Si (MFIS) [45]   (2003)
- MFIS FET with BLT on Al_2O_3 showing a retention time of
  $1.5 \cdot 10^6 s$ (*17 days*) [47]   (2004)

Still, almost 50 years after the conception of the FeFET, the ideal device remains elusive [5], since solutions to known problems, most notably retention loss, are yet to be found. It will take advances in process technology to perfect the gate stack, achieving low interface states densities, perhaps through the use of new materials.

1.4 Objectives

The target of this thesis is to use a model of the FeFET in combination with a simulation program to predict the scaling of the device and derive efficient scaling rules that can yield acceptable performance for the scaled devices. The miniaturization potential of the FeFET is investigated, because no memory technology is interesting unless it has long term prospects and can be competitive price/performance wise.
Further, different FeFET programming concepts are examined and a new programming concept is proposed that can simplify the FeFET based memory design. This concept is applied to design and simulate a memory chip.
Finally, calculations of the retention time are performed with consideration of the leakage current in the gate stack, and suggestions are made to reduce the depolarization field.

1.5 Simulation Tools

Two types of simulation programs are used in this thesis. One is a circuit simulator that uses a FeFET model to perform single device simulations up to complex circuit simulations with thousands of devices. The other is a device/process simulator (this type of simulation is known as TCAD or technology CAD) and is only suitable for single device simulations. The latter is physically more accurate, but also more complex.

1.5.1 Circuit Simulation

The circuit simulator is the one mostly used in this thesis. It solves the voltage and current equations in a circuit by using mathematical models for the different devices [71]. For the MOSFET the BSIM3v3 (Berkeley Short-channel IGFET[1] Model) model is used, that is also the basis for the FeFET model. Compared to a device simulation program, it is less precise, but combines very fast operation with acceptable accuracy. The BSIM3v3 model is a mathematical, semi-empirical MOSFET model, which solves the semiconductor equations analytically, by applying several simplifications (e.g. geometric and doping profile, depletion region approximation, charge sheet approximation). These simplifications make it ideal for use in a circuit simulator, for circuits with many thousands of transistors. Moreover, it is scalable from big structures down to channel lengths of about *0.15 µm* (Version 3v3). For smaller sizes the BSIM4 model was developed, that addresses the MOSFET physical effects into the sub-*100 nm* regime. Except in chapter 6, in most parts of this thesis the BSIM3v3 model is used. BSIM3v3 is the standard transistor model for deep sub-micron device simulation supported by the Electronic Industries Alliance (EIA). It was developed in the Berkeley University for the simulation of sub-micron technology, but can also be used for bigger structures.
The number of parameters, including the scaling parameters (binning parameters), is *~300*, although the latter are not used most of the time (default to zero). Many parameters are correlated and many have a use only in special applications (e.g. big structures, noise behavior, high frequencies, temperature dependence). For a simulation with the BSIM3v3

---

[1] Insulated Gate Field Effect Transistor

model a parameter extraction is first needed. There are special programs that are used for parameter extraction after a device is electrically characterized. The measurements needed for a full parameter extraction include S-Parameter, noise and temperature dependence measurements. However, as pointed out before, not all parameters are necessary for all types of simulation. In our case the important ones are: N_{CH}, N_{SUB} (channel and substrate doping concentrations respectively), X_T (channel doping depth) and d_{Ox} (oxide thickness). The effect of these parameters on the device functionality is examined in section 5.8. As an alternative to the (N_{CH}, N_{SUB}) pair of parameters, one can use (k_1, k_2) or (γ_1, γ_2). For a detailed description of the BSIM3v3 model see [6], [75]. These parameters are either known or can be extracted from measurement curves after appropriate fitting. Every chip manufacturer ('chip foundry') that makes CMOS circuits on order, supplies the BSIM parameters (sometimes also called SPICE Parameters) of their MOS transistors, so that designers can simulate their circuits before they move to chip layout. The MOSFET parameters used in chapter 9 for the chip simulation are those of a SIEMENS *0.5 μm* CMOS process.

1.5.2 Device-Process Simulation

A device simulator is basically a Partial Differential Equation (PDE) solver. It calculates and solves numerically the Poisson, the carrier transport and carrier continuity equations [38]. The Poisson equation relates the electrostatic potential to the charge concentration. The continuity equations describe the variation of the charge concentration in dependence of the transport mechanisms. The equations are solved numerically at the grid points of the device structure. A denser grid yields a more precise solution. The application of suitable initial conditions is very important for a fast convergence.
The simulations performed were 2-D (two-dimensional) with the 3rd dimension assumed *1μm*. 3-D simulations are also possible but are much more time and resource consuming.
A process simulator allows the simulation of the complete CMOS technology processing starting from a semi-conducting substrate to a fully functional (virtual) electronic device. It can be used alongside a device simulator, to describe structures that are later to be electrically characterized with the device simulator. It was used to extract the geometry of the MOSFET of Fig. 7.12 before performing a device simulation. One benefit of using this program is that it is possible, for example, to calculate the doping profile by specifying the parameters of the ion implantation procedure.

Example: Boron implanting in vertical direction on a p-type substrate with a uniform concentration of *10^{14}* using an implantation dose of *10^{14} cm^{-2}*, a voltage acceleration of *10 kV* and selecting a gaussian profile, yields a doping depth of *~130 nm* as shown in Fig. 1.3.
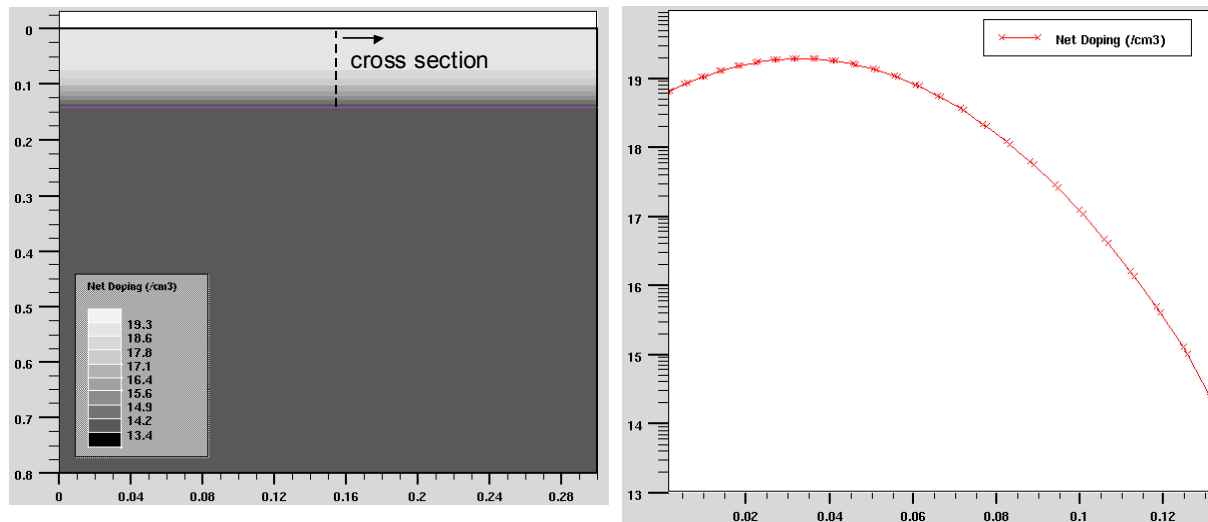
Fig. 1.3 (left) Simulation result of the ion implantation with ATHENA: gaussian distribution with a doping depth of ~130 nm and (right) doping profile (dimensions are in μm).

2 FeFET : Principles of operation

In order to present the principles of operation of the FeFET, first the MOS capacitor will be described and the basic equations will be given. The structure will then be expanded to include a ferroelectric layer in the gate stack, by introducing the effective gate voltage and the stack capacitance, yielding the MFIS capacitor, upon which the FeFET is based. In a similar way, the MOSFET model BSIM3v3 will be used as a basis for the FeFET model in the next chapter. A deeper mathematical analysis is given in [6], [7] (for the MOSFET) and [8], [24] (for the FeFET).

2.1 n-MOS Capacitor

In an n-MOS[2] (Metal Oxide Semiconductor) capacitor, applying a voltage to the gate yields the following results (The structure is shown in Fig. 2.2. It is assumed that no charges are present in the oxide $Q'_{Ox} = 0$.) [72]:

a) If the gate voltage is equal to the flatband voltage $V_{GB} = V_{FB}$, the capacitor is in flatband condition. The energy bands in the semiconductor are flat (not bent). There is no charge at the gate electrode or the semiconductor. The flatband voltage V_{FB} is given by Eq. 2.1 where Φ_{MS} is the workfunction difference (in Volt) between the gate electrode and the semiconductor given by Eq. 2.2.

$$V_{FB} = \Phi_{MS} - \frac{Q'_{Ox}}{C'_{Ox}}, \quad \text{where} \tag{2.1}$$

$$\Phi_{MS} = \Phi_M - \Phi_S = \Phi_M - \chi - \frac{E_g}{2q} - \phi_t \cdot \ln\left(\frac{N_{SUB}}{n_i}\right). \tag{2.2}$$

E_g is the semiconductor energy gap ($E_g = E_c - E_v$), χ the electron affinity ($q \cdot \chi = E_{vacuum} - E_c$), q the electron charge, ϕ_t the thermal potential, N_{SUB} the substrate and n_i the intrinsic doping concentration, the latter being $\approx 10^{10}$ cm^{-3} at room temperature ($T = 300$ K). C'_{Ox}[3] is the oxide capacitance per area.

b) If $V_{GB} < V_{FB}$, the capacitor is in accumulation and positive charges accumulate in the semiconductor near the interface to the oxide. The energy bands are bent, but this band-bending occurs at the interface at a small depth. The voltage drop is only across the oxide.

c) If $V_{FB} < V_{GB} < V_{TH}$, the semiconductor is in depletion. A negatively charged zone (space charge zone) builds in the semiconductor that grows with higher voltage. The negative charge comes from the mobile acceptor ions being repelled by the positive gate voltage towards the substrate leaving behind immobile negative ions.

---

[2] substrate is p-type

[3] $C'_{Ox} = C_{Ox}/(L \cdot W)$. In some equations the capacitance C_{Ox} [F] is used. Similarly Q'_{Ox} is the oxide charge per area.

d) If $V_{GB} > V_{TH}$, the semiconductor is in inversion and a conducting channel of electrons builds at the top from the inversion of the p-type semiconductor into n-type. The threshold voltage V_{TH} is the onset of inversion and is calculated by Eq. 2.5.

Figure 2.1 shows the energy bands of the MOS capacitor for the different operation regimes.
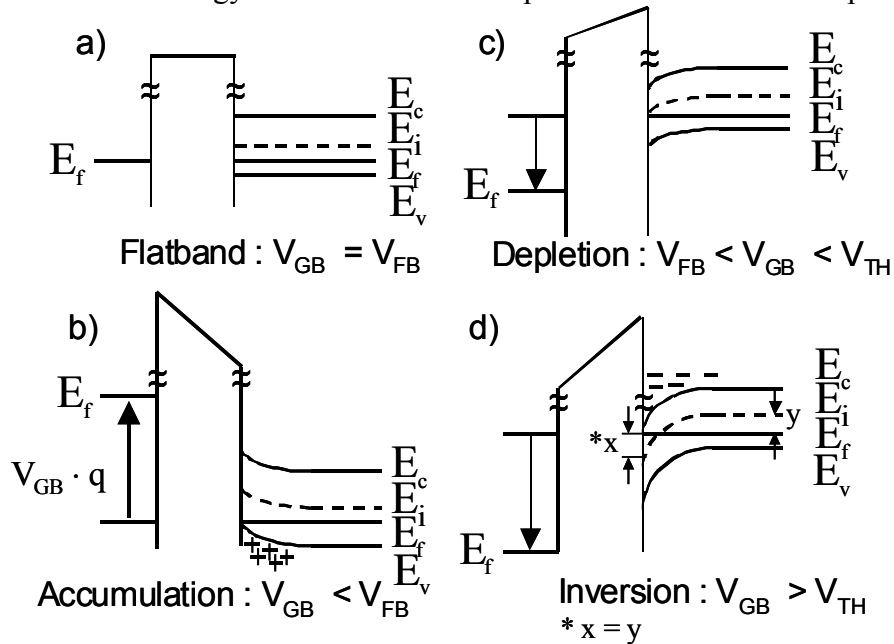
Fig. 2.1 The operation regimes in an n-MOS capacitor.

The different regions of operation can be observed in a *C-V* (capacitance vs. voltage) measurement as shown in Fig. 2.2 (left). To obtain the low frequency (or quasi-static) curve, the capacitance must be measured while the structure is in equilibrium. For the high-frequency curve, the small signal capacitance is measured by superposition of a dc voltage source that is slowly varied, and a low amplitude, high frequency sinusoidal signal generator. Because of the applied high frequency signal, equilibrium is not reached past the threshold voltage, since the minority carriers in the inversion layer cannot generate rapidly enough.

Fig. 2.2 C-V curve of an n-MOS capacitor (left) and measurement setup (right). Measurement is usually performed with an LCR meter.

If the ac signal frequency is low enough (quasi-static measurement), then the capacitance increases beyond the threshold, because there is equilibrium at every voltage step of the dc voltage source and the minority carriers in the inversio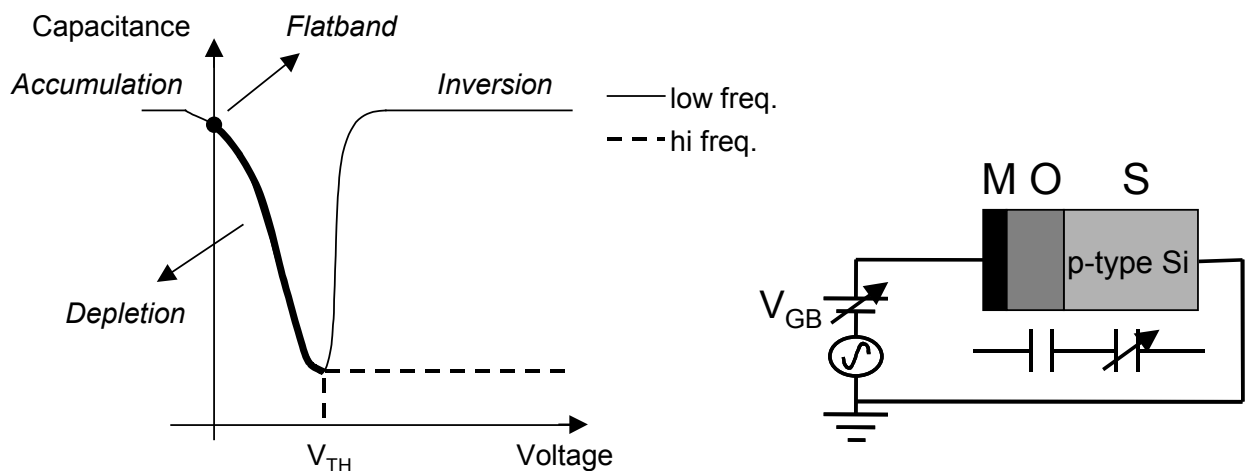n layer are able to follow the variation of the ac signal. In the high frequency measurement the inversion layer appears transparent, because the generation-recombination rates of the minority carriers cannot keep up with the rapidly changing ac signal, thus the capacitance remains at the minimum level (oxide capacitance in series with the maximum depletion region). In depletion the capacitance is lower, because the oxide capacitor is in series with a semiconductor capacitor of variable thickness. The depletion depth is maximal at threshold voltage and beyond, and depends on the doping concentration of the substrate. In accumulation and in inversion the gate voltage drops only across the gate oxide.

The total sum of voltages across the gate stack is given by [7]

$$V_{GB} = V_{FB} + V_{Si} + V_{Ox},$$

(2.3)

where V_{Si} is the voltage in the (depleted) semiconductor and V_{Ox} the voltage across the oxide. The maximum depletion depth is given by

$$X_D = \sqrt{\frac{2\varepsilon_0\varepsilon_{Si}2\phi_F}{qN_{SUB}}}.$$

(2.4)

The threshold voltage V_{TH} is given by

$$V_{TH} = V_{FB} + 2\cdot\phi_F + \gamma\sqrt{2\phi_F}$$

(2.5)

(in [7] defined as the onset of "moderate" inversion),

where ϕ_F is the bulk potential due to doping $\phi_F = \phi_t \ln\frac{N_{SUB}}{n_i}$,

(2.6)

ϕ_t the thermal potential $\phi_t = \frac{kT}{q}$,

(2.7)

γ the body effect coefficient defined as $\gamma \equiv \frac{\sqrt{2q\varepsilon_0\varepsilon_{Si}N_{SUB}}}{C'_{Ox}}$

(2.8)

and $C'_{Ox} = \left(\frac{d_{Ox}}{\varepsilon_0\varepsilon_{Ox}}\right)^{-1}$,

(2.9)

d_{Ox} and ε_{Ox} being the oxide thickness and permittivity respectively.

In order to calculate all the parameters of the MOS structure, Eq. 2.3 along with the following three equations can be solved for the four unknowns V_{Ox}, V_{Si}, Q_G, Q_I (Q_G and Q_I are the charge at the gate electrode and inverted channel respectively).

$$Q_I = \mp C'_{Ox}\cdot\gamma\cdot\sqrt{\phi_t e^{-V_{Si}/\phi_t} + V_{Si} - \phi_t + e^{-2\phi_F/\phi_t}\left(\phi_t e^{V_{Si}/\phi_t} - V_{Si} - \phi_t\right)}$$

(2.10)

$$Q_G + Q_{Ox} + Q_I = 0$$

(2.11)

$$Q_G = C_{Ox}\cdot V_{Ox}$$

(2.12)

The negative sign in Eq. 2.10 is used for $V_{Si} > 0$ (depletion or inversion) and the positive for $V_{Si} < 0$ (accumulation).

If $Q_{Ox} = 0$ *then* $Q_G = -Q_I$. Using Eqs. 2.8-2.12, Eq. 2.3 can be written as

$$V_{GB} = V_{FB} + V_{Si} \pm \gamma \cdot \sqrt{\phi_t e^{-V_{Si}/\phi_t} + V_{Si} - \phi_t + e^{-2\phi_F/\phi_t}\left(\phi_t e^{V_{Si}/\phi_t} - V_{Si} - \phi_t\right)} \qquad (2.13)$$

Equations 2.3 and 2.10-2.12 fully characterize the MOS structure. In the next section Eq. 2.3 will be rewritten for the MFIS structure. This set of equations cannot be solved analytically. Therefore the BSIM3v3 model in combination with a circuit simulator will be used for most calculations in this thesis. Additionally, in chapter 7 device simulations will be performed to verify the proposed *positive voltage erase* concept.

2.2 MFIS capacitor

The ferroelectric layer in the MFIS (Metal Ferroelectric Insulator Semiconductor) structure adds a memory function to the MOS capacitor. The regions of operation described previously are the same, since they refer to the condition of the semiconductor. Due to the hysteretic nature of the ferroelectric, the *C-V* curves are now hysteretic too and are shown in Fig. 2.3. The width of the hysteresis loop is referred to as "memory window" and can have a maximum value (for the saturated hysteresis) of $\sim 2E_C \cdot d_{Fe}$ [8], where E_C is the coercive field and d_{Fe} the thickness of the ferroelectric. Note that the direction in the *C-V* curve is clockwise, in contrast to the counterclockwise polarization hysteresis curve (e.g. Fig. 2.7). Despite the similarity with the hysteresis curve in Fig 2.7, the *C-V* curve in Fig. 2.3 (left) has a constant value in the accumulation and inversion region even when the ferroelectric is not polarized to saturation. An index to the amount of polarization is the memory window.
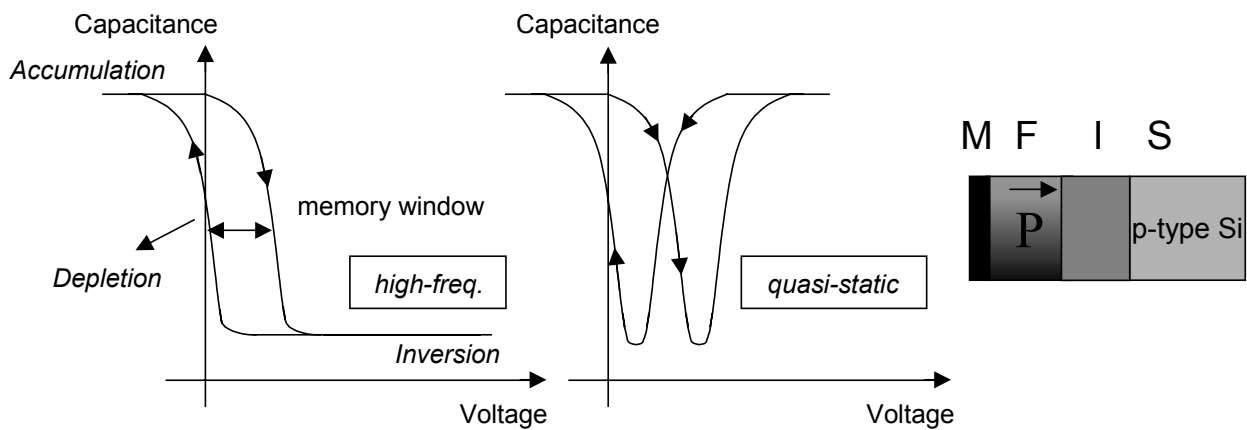

Fig. 2.3 C-V curves of an MFIS capacitor.

Equation 2.3 can be rewritten for the MFIS structure by including the voltage drop across the ferroelectric V_{Fe} (again assuming $Q_{Ox} = 0$):

$$V_{GB} = V_{FB} + V_{Si} + V_{Fe} + V_{Ox}, \text{ where} \tag{2.14}$$

$$V_{Fe} = \frac{d_{Fe}}{\varepsilon_0 \varepsilon_r} \cdot (Q'_G - P_{Fe}) \text{ and } V_{Ox} = \frac{d_{Ox}}{\varepsilon_0 \varepsilon_{Ox}} \cdot Q'_G. \tag{2.15), (2.16}$$

By introducing the effective gate voltage

$$V_{GBeff} = V_{GB} + \frac{d_{Fe}}{\varepsilon_0 \varepsilon_r} P_{Fe} \tag{2.17}$$

and the stack capacitance per area

$$C'_{Stack} = \left(\frac{d_{Fe}}{\varepsilon_0 \varepsilon_r} + \frac{d_{Ox}}{\varepsilon_0 \varepsilon_{Ox}} \right)^{-1}, \tag{2.18}$$

Eq. 2.14 can be rewritten as

$$\begin{aligned}
V_{GBeff} &= V_{GB} + \frac{d_{Fe}}{\varepsilon_0 \varepsilon_r} P_{Fe} = V_{FB} + V_{Si} + \frac{Q'_G}{C'_{Stack}} = \\
&= V_{FB} + V_{Si} \pm \gamma \sqrt{\phi_t e^{-V_{Si}/\phi_t} + V_{Si} - \phi_t + e^{-2\phi_F/\phi_t} \left(\phi_t e^{V_{Si}/\phi_t} - V_{Si} - \phi_t \right)}
\end{aligned} \tag{2.19}$$

where γ is now defined as

$$\gamma \equiv \frac{\sqrt{2q\varepsilon_0 \varepsilon_{Si} N_{SUB}}}{C'_{Stack}}. \tag{2.20}$$

The term $\dfrac{d_{Fe}}{\varepsilon_0 \varepsilon_r} P_{Fe}$ gives the voltage in the ferroelectric due to the polarization, excluding the

linear term (see Eq. 3.4). The latter is included in the term $\dfrac{Q'_G}{C'_{Stack}}$ together with V_{Ox}. ε_r is the

dielectric permittivity of the linear part of the ferroelectric layer. P_{Fe} refers to the non-linear (irreversible) part of the polarization (see also Eq. 3.3).

The flatband voltage used in Eqs. 2.14, 2.19 is not different from that in the MOS structure (Eq. 2.1) despite the ferroelectric, because of the introduction of the effective gate voltage V_{GBeff}. Without the introduction of V_{GBeff} the flatband voltage is given by [68]

$$V_{FB} = \Phi_{MS} - \frac{Q'_{Ox}}{C'_{Ox}} \pm \left(\frac{P_{Fe}}{C'_{Fe}} \right)_{(V_{GB}=0)}, \tag{2.21}$$

and is not constant but depends on the ferroelectric polarization P_{Fe} and the ferroelectric capacitance per area (including the linear and nonlinear part, see Eq. 5.14) at $V_{GB} = 0$ V. P_{Fe} in Eqs. 2.21 is equal to the ferroelectric's P_R only when the ferroelectric is polarized to saturation. The positive sign refers to P^+ and the negative to P^- (see Fig. 2.7). Unfortunately, Eq. 2.21 is of little practical use because C'_{Fe} cannot be easily determined. It is more convenient to use Eq. 2.1 for the definition of the flatband voltage, because it is easier to expand the BSIM3v3 model, where V_{FB} is defined as a constant.

2.3 The ferroelectric field effect transistor (FeFET)

Just like the MOS structure, the MFIS capacitor structure can be expanded into a 4-terminal transistor device by including two highly doped regions (source and drain) of opposite type (n-type) to that of the substrate (p-type). The result is the ferroelectric field effect transistor (FeFET) (Fig. 2.4 right) that incorporates the functionality of the MOSFET (Fig. 2.4 left) with the additional feature of non-volatility. This hybrid nature of the FeFET brings many advantages for the implementation of non-volatile memories. The FeFET will be described in more detail in chapter 3.
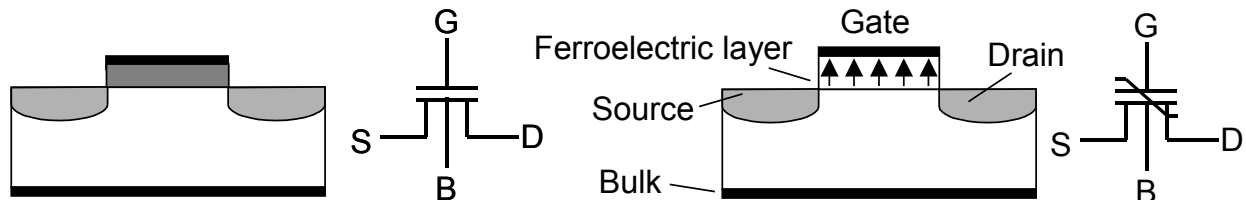

Fig. 2.4 Cross-section (left) of a MOSFET with symbol, (right) of a FeFET with symbol.

2.4 The floating gate transistor comparison

The single device memory concept is interesting, because its simple implementation requires no additional transistors as switching elements (as, for example, with FRAM). This applies to Flash memory as well and has made large memory capacities possible. An analysis of the per bit density of the various memory technologies is made in chapter 8. The FeFET has many similarities with the floating gate transistor, widely in use today in the Flash memory technology. Figure 2.5 shows the devices and their *I-V* characteristics.
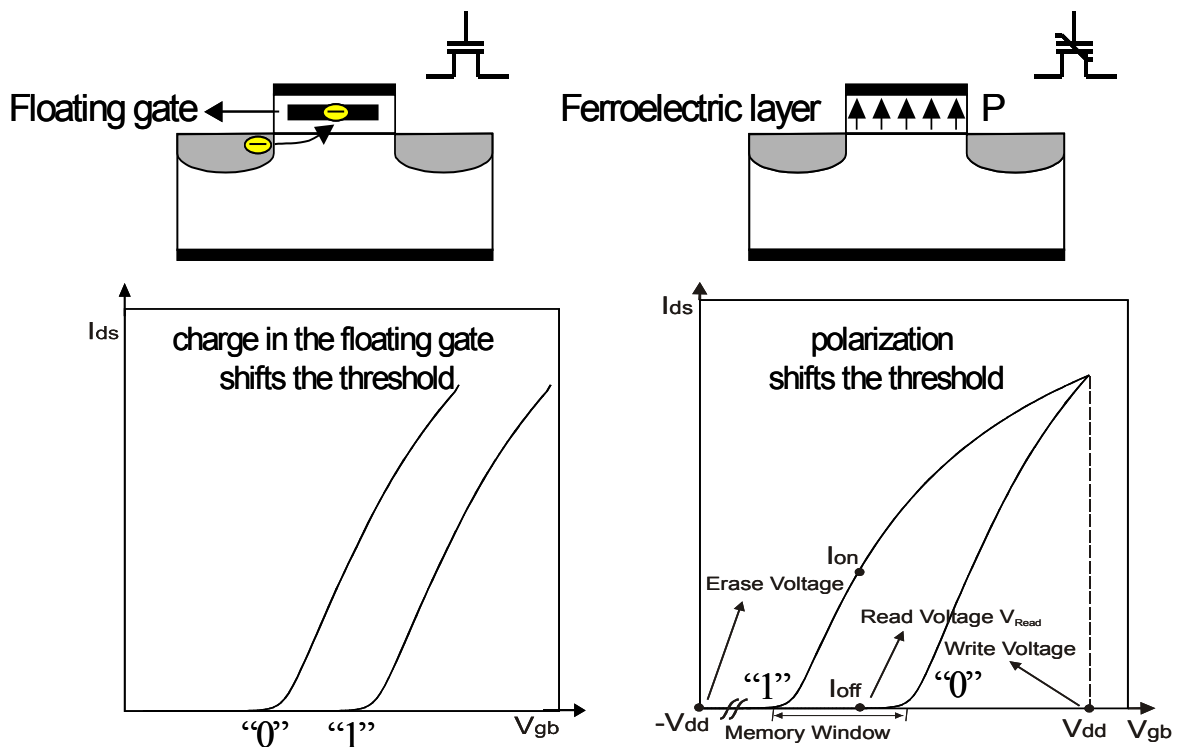

Fig. 2.5 The FeFET and floating gate transistor with their respective current characteristics.

2.4.1 Principles of the floating gate transistor

In Flash, for writing information, a high voltage is applied to the gate while grounding the drain and applying a voltage to the source. This voltage difference between source and drain causes the source electrons to accelerate (become hot, hence the term hot electrons). The high voltage applied to the gate results in an electron injection into the floating gate, thus the mechanism is termed hot electron injection or CHE[4]. For erasing, a voltage is applied to the source while grounding the gate and floating the drain (alternatively a negative voltage can be applied to the gate so that a lower voltage needs to be applied to the source). As a result, the electrons tunnel back to the source. The erase mechanism is Fowler-Nordheim tunneling and is a slower process than writing. Figure 2.6a shows the voltages applied to the floating gate transistor during write/erase. In comparison, Fig. 2.6b shows the same procedure for the FeFET.
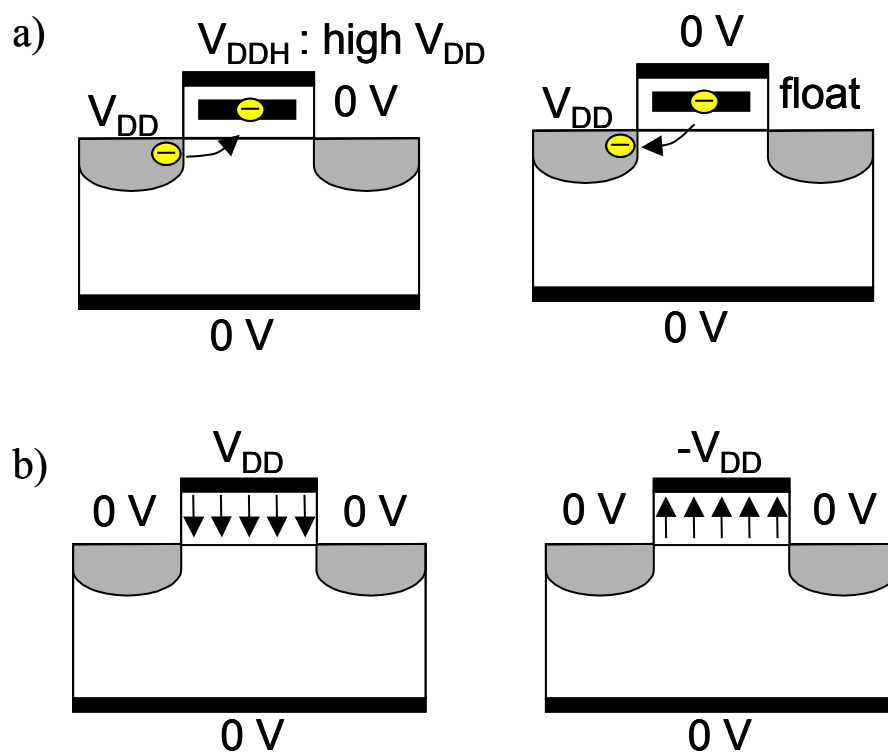
Fig. 2.6 The voltages applied to (a) the floating gate transistor and (b) the FeFET for a write (left) and an erase operation (right).

The write and erase techniques vary from memory manufacturer to manufacturer. Both Flash and EEPROM use the same device and differ only in the cell implementation, with EEPROM being more complex using 1 or 2 transistors for selection. The charge in the floating gate causes a shift in the *I-V* curve as shown in Fig. 2.5 (left). By applying a read voltage somewhere in the middle of the memory window where the I_{on}/I_{off} ratio is maximum, the two states can be distinguished (if current flows then the cell is erased "0").

---

[4] Channel Hot Electron

2.4.2 Similarities and differences between the two devices

The main differences between the two devices are the slow write/erase operation and the high voltages required for programming in Flash. Both devices have a memory window, so the read operation (current detection with a sense amplifier) is similar. In the floating gate transistor negative charge in the gate stack leads to a higher threshold voltage (Eqs. 2.1, 2.5). In the FeFET the ferroelectric polarization causes the change in the threshold voltage (Eq. 2.21). In both devices no current flows below the threshold voltage V_{TH}. Applying a gate voltage higher than V_{TH} to a FeFET that is negatively polarized can change its polarization and lead to disturbance[5], while a floating gate transistor becomes conductive without switching state. This is the reason that the NAND memory configuration cannot be applied in the case of the FeFET [8]. The writing procedure is also different, as only one voltage (V_{DD}) needs to be applied to the FeFET compared with two (V_{DD} and V_{DDH}) for the Flash transistor. Finally, both devices are of n-type, because it would not make sense to have normally-on (p-type) memory devices as this would cause higher power consumption due to leakage currents.

The FeFET is not going to be a drop-in replacement for Flash, and new programming concepts have to be applied. In chapter 7 two programming concepts will be examined and one of them will be used in the chip simulation of chapter 9.

2.5 Ferroelectric Materials

For a ferroelectric to be used in a memory device, it has to be compatible with the CMOS process that is used in the semiconductor industry. This includes compatibility with high temperatures and other materials involved in the process [76]. For a more thorough analysis on ferroelectric materials see [73], [77], [83]. In this thesis the only material properties that will be of interest are the linear dielectric permittivity ε_r, the remnant and saturated polarization (P_R and P_S respectively), and the coercive field E_C. The desired values for these parameters will now be discussed.

The two ferroelectric materials currently studied the most are $SrBi_2Ta_2O_9$ (SBT) and $PbZi_{1-x}Ti_xO_3$ (PZT). Both are used in FRAM memories and are being considered for application in the FeFET too. They are both of perovskite structure (SBT is a layered perovskite while PZT is a more classic type of ferroelectric) and thus have a high dielectric permittivity (high-k) of ~*250* in thin films [69]. A high-k material, although favored as a gate dielectric, is not a desirable property for the ferroelectric in the FeFET, because of the low-k dielectric buffer layer that is always present in the FeFET gate stack. The dielectric's lower permittivity (lower capacitance) results in a higher voltage drop across the buffer compared to that across the ferroelectric. As shown in Fig. 2.7 the two materials differ in the polarization and coercive field values. PZT has a higher remnant polarization P_R and coercive field E_C than SBT. Table 2.1 lists the parameters of the two materials for thin films.

---

[5] However, a FeFET can be designed to be less susceptible to disturbance.
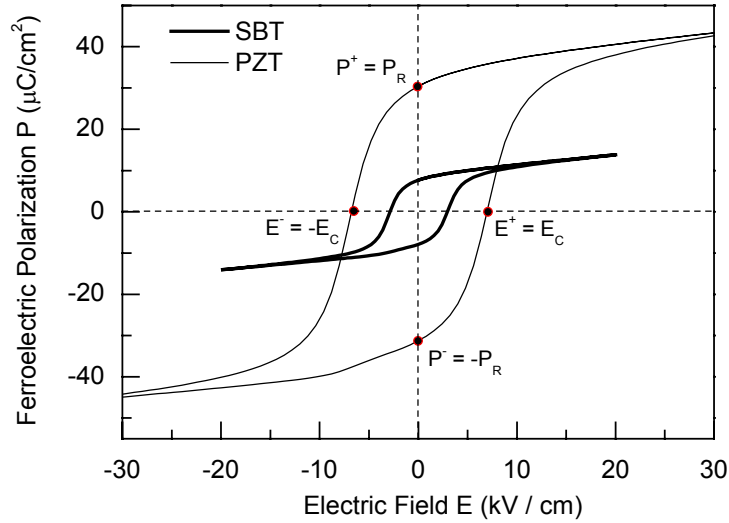
Fig. 2.7 Saturated polarization hysteresis loops for SBT and PZT (including the linear dielectric part).

Param.	Unit	PZT	SBT
P_R	$\mu C/cm^2$	32	8
P_S	$\mu C/cm^2$	40	10
E_C	kV / cm	70	30
ε_r	-	250	250

Table 2.1 Material parameters for thin film SBT [27] and [111]-oriented PZT (30:70) [69].

A ferroelectric with a high coercive field yields a higher memory window ($2E_C{\cdot}d_{Fe}$), but this is at the expense of higher voltages that are necessary for switching.

A ferroelectric with a high remnant polarization P_R leads to higher currents in the FeFET [8], but cannot be fully polarized using low voltage operation, so sub-loops are used that are less stable.

Regarding the dielectric, it should have a high permittivity ε_{Ox} and a high breakdown field. The values for *SiO$_2$* are *3.9* and *10 MV/cm* respectively.

According to Eqs. 2.22 and 2.23, the field in the oxide E_{Ox} depends mainly on the amount of polarization in the ferroelectric (since $P_{Fe} \gg \varepsilon_0 \cdot \varepsilon_r \cdot E_{Fe}$) and does not depend on the oxide thickness [9].

$$C'_{Ox} = \frac{Q'}{V_{Ox}} \ \Rightarrow\ V_{ox} = \frac{Q'}{C'_{Ox}} \ \Rightarrow\ E_{ox} = \frac{\dfrac{Q'}{\varepsilon_0 \cdot \varepsilon_{Ox}}}{d_{Ox}} = \frac{Q'}{\varepsilon_0 \cdot \varepsilon_{Ox}} = \frac{P_{Fe} + \varepsilon_0 \cdot \varepsilon_r \cdot E_{Fe}}{\varepsilon_0 \cdot \varepsilon_{Ox}}, \qquad (2.22)$$

$$\text{where}\ \ Q' = P_{Fe} + \varepsilon_0 \cdot \varepsilon_r \cdot E_{Fe}{}^6. \qquad (2.23)$$

---

[6] In this thesis Q' [C/m^2] is used in place of the electric displacement D [C/m^2].

Therefore sub-loops must be used when using a ferroelectric with a high P_R. The maximum field in the oxide when polarizing the ferroelectric is not given by Eq. 2.22. In fact, breakdown occurs while polarizing the ferroelectric and not at P^+. This is shown in Fig. 2.8. Assuming, for example, a breakdown field of *200 MV/m*, dielectric breakdown occurs for polarizations greater than *1.5 μC/cm²* and before reaching *2.5 μC/cm²* (note the clockwise orientation of the hysteresis).
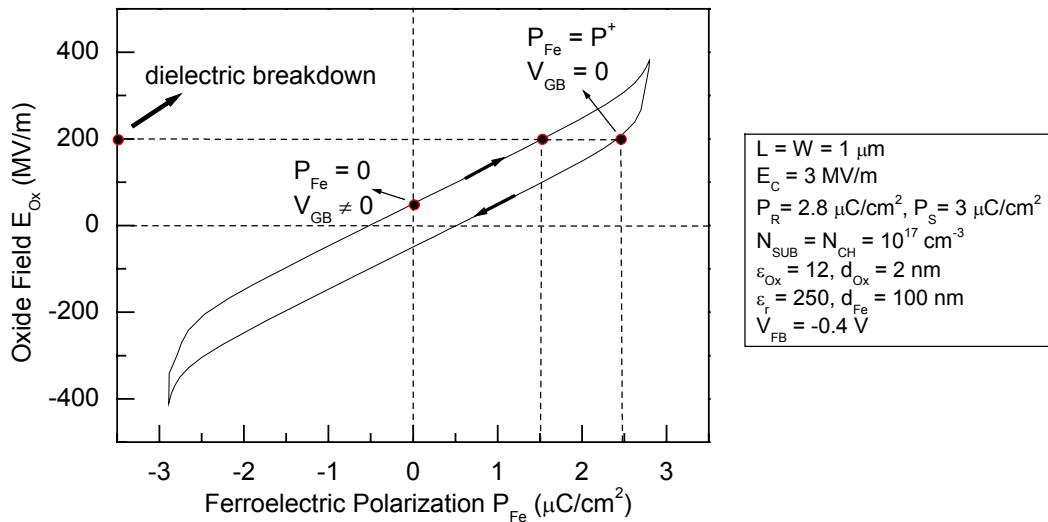

Fig. 2.8 Electric field in the gate oxide dielectric vs. polarization of the ferroelectric for an MFIS structure.

There is another boundary for P_R at the low end because at some point it affects the memory window (Fig. 2.9). This is because at some point the polarization due to the ferroelectric becomes so small that it is comparable to the polarization of the oxide dipoles (Eq. 2.23). Note that the memory window in Fig. 2.9 (right) has a maximum for a certain value of P_S. A similar dependency is that of the memory window on the thickness of the ferroelectric, as will be seen in section 5.8.
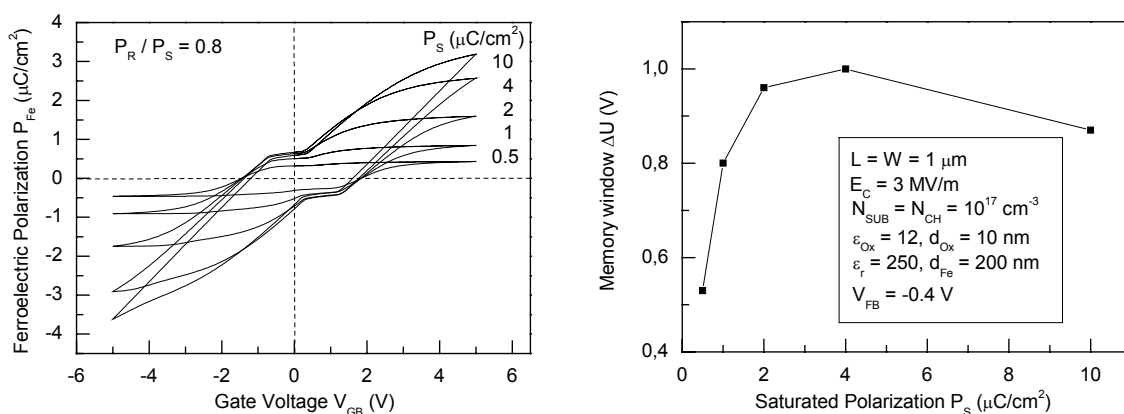

Fig. 2.9 Dependence of a FeFET's hysteresis loop (left) and memory window (right) on the ferroelectric's P_R, P_S.

2.6 Challenges for the FeFET

2.6.1 Retention loss

A challenge for the FeFET is to overcome the low values of data retention time (loss of retention). Measurements on MFIS structures and FeFETs show that the remnant polarization has a retention time of only days (at most). Figure 2.10 shows how the capacitance drops (increases) when the remnant state is the on (off) state. The following are given as reasons for the retention loss:

- The depolarization field in the ferroelectric [10],
- the injection of charge through either the ferroelectric or the oxide (leakage current) [11], [37], [84],
- the instability of hysteresis sub-loops (will not be studied further).

The first two issues will be examined in more detail in chapter 4. Sometimes the thermodynamic instability of the ferroelectric [12] is also mentioned, but according to [12] it is caused by the depolarization field.
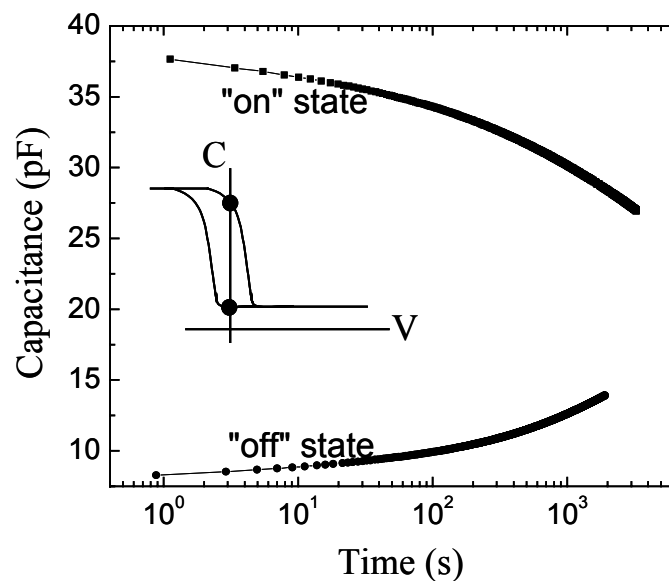
Fig. 2.10 Capacitance measurement of an MFIS structure (Pt/PZT(200nm)/STO(30nm)/Si) with time for $V_{GB} = 0$ V.

2.6.2 The growth of ferroelectrics on silicon

For all known ferroelectric oxides, the deposition directly on silicon causes inter-diffusion and chemical reactions, that result in degradation of the interface to silicon. Therefore a dielectric has to be used as an intermediate buffer layer[7]. This should ideally be amorphous, because amorphous oxides usually show better leakage current characteristics compared to crystalline ones [88], [89]. Another reason is the lattice mismatch of the silicon and the ferroelectric crystal lattice, therefore a dielectric is needed to enable a growth with the preferred orientation. Besides, because of the high temperatures usually employed during

---

[7] An oxide dielectric was considered in all simulations in this thesis.

ferroelectric deposition at oxygen atmosphere, the surface of the silicon is inevitably oxidized yielding a thin layer of SiO_2 [13]. One way to suppress the growth of SiO_2 is by using Si_3N_4 as an intermediate buffer [14].

2.6.3 High density of oxide charges

Interface states show up in a *C-V* measurement as a shifted curve. In case of mobile charges in the oxide or defects that allow charge transport from either side of the contact, a hysteresis is observed that has clock or anticlockwise orientation. In an MOS capacitor the oxide charges can be quantified (although identifying is harder), but in an MFIS structure the two hysteresis curves add up to a combined hysteresis that cannot be separated. If there are a high number of oxide charges, then the clockwise hysteresis, that is characteristic for a ferroelectric (Fig. 2.12), can be turned into an anticlockwise one. Note that the orientation of a hysteretic *C-V* curve in an MOS capacitor with oxide charges is clockwise if charge injection takes place from the silicon side and anticlockwise if charge injection takes place from the gate electrode side [15].
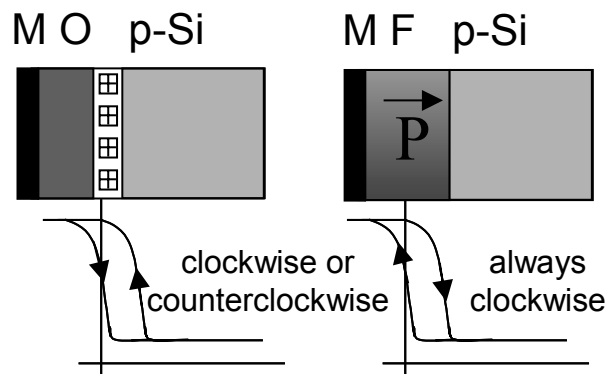

Fig. 2.12 Clockwise or anticlockwise hysteresis loops of an n-MOS capacitor (left), clockwise hysteresis of an MFS capacitor (right).

2.6.4 Fatigue – Imprint

In FRAM memories based on ferroelectric capacitors two important failure mechanisms are fatigue and imprint. Fatigue is caused after switching the polarization in the ferroelectric a large number of cycles. The result is a drop in the remnant polarization (Fig. 2.13), because of the degradation of the ferroelectric. The initial condition cannot be restored with a refresh and the damage to the ferroelectric is permanent. This is in contrast to imprint, where the ferroelectric, after being polarized at one state, tends to prefer this over the other state and shows a hysteresis that is shifted to the left (right), when polarized at the positive (negative) state (Fig. 2.14). Here a refresh can restore the initial hysteresis. In both cases the improvement of the interfaces (e.g. using oxide buffers and/or electrodes) reduces the problems to a high degree. One physical explanation of imprint is given in [21], where it is attributed to leakage current in dead layers[8] at the interfaces. Fatigue is also attributed to charge injection in [37].

---

[8] interface layers with a low permittivity that act as parasitic capacitors

Regarding the FeFET, the retention loss problem (Fig. 2.15) is the number one issue at the moment. Although both fatigue and imprint are likely to apply in the case of the FeFET too, the remedy will probably lie, as with ferroelectric capacitors, in the improvement of the interfaces.
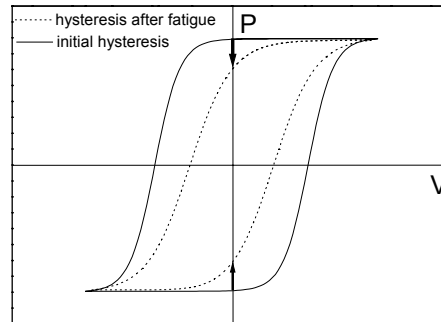

Fig. 2.13 Fatigue in ferroelectric capacitors after a high number of cycles [22].
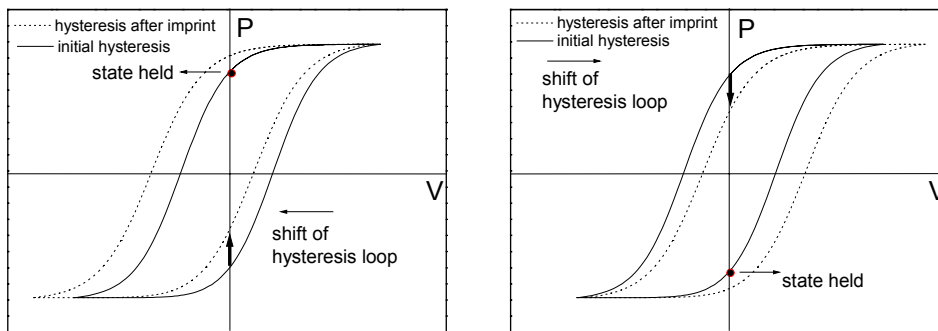

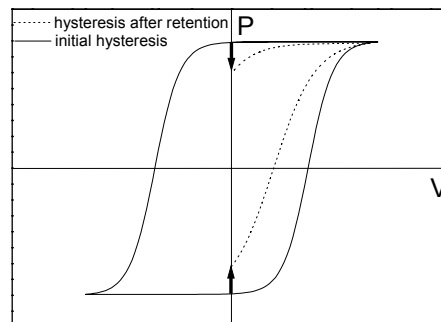*Fig. 2.14 The imprint effect in ferroelectric capacitors [22].*

Fig. 2.15 Retention loss in ferroelectric capacitors after enough time in one of the two steady states [22].

3 Modeling the FeFET

3.1 Simple FeFET model

The MOSFET in its most basic approximation has a channel resistivity that switches from zero conductivity below threshold voltage to a finite constant conductivity beyond. The drain source current I_{DS} can be described mathematically by

$$I_{DS} = \begin{cases} 0 & , V_G < V_{TH} \\[2ex] \dfrac{W}{L} \cdot \mu_n \cdot C_{ox} \cdot (V_{GB} - V_{TH}) \cdot V_{DS} & , V_{GB} > V_{TH} \end{cases} \tag{3.1}$$

where μ_n is the carrier (electron) mobility (characteristic for a semiconductor, temperature and doping dependent), L and W the channel length and width respectively, and C_{Ox} the oxide capacitance.

The ferroelectric as an ideal single-domain crystal can be approximated by a square hysteresis loop that switches state at $\pm Ec$. Figure 3.1 shows the characteristic curves of an ideal MOSFET and a ferroelectric capacitor and how they combine to give the FeFET's *I-V* characteristic. The mathematical description is then given by Eq. 3.2. Assuming a real ferroelectric yields the continuous curve of Fig. 3.1 (center).
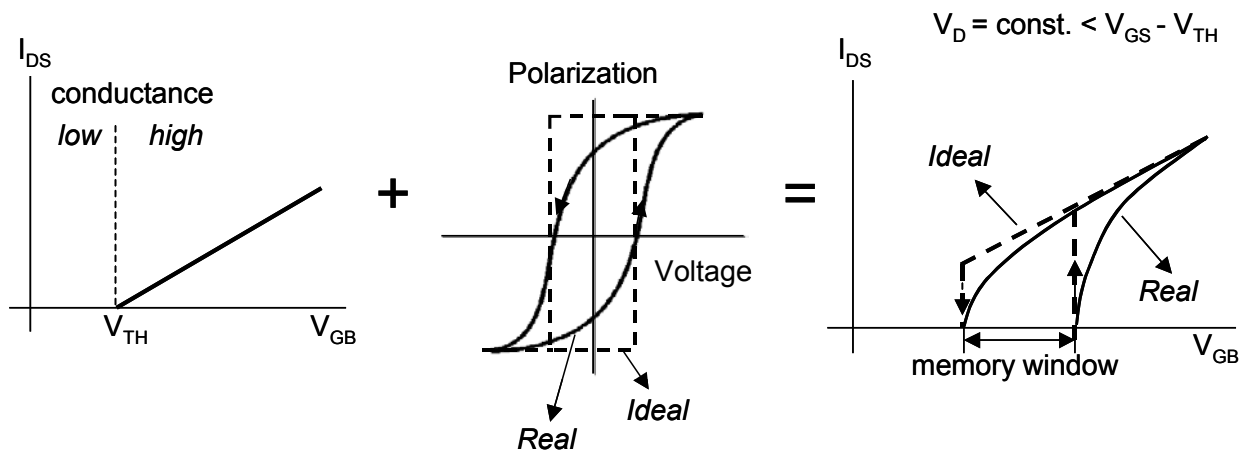

Fig. 3.1 A MOSFET and a ferroelectric capacitor characteristic combine to give the FeFET I-V. The dashed (continuous) curve represents an ideal (real) ferroelectric.

$$I_{DS} = \begin{cases} 0 & , V_{GB} < V_{TH1} \\ 0 & , V_{TH1} < V_{GB} < V_{TH2} \cap P = P^- \\ \dfrac{W}{L} \cdot \mu_n \cdot C_{Stack} \cdot (V_{GB} - V_{TH}) \cdot V_{DS} & , V_{GB} > V_{TH2} \\ \dfrac{W}{L} \cdot \mu_n \cdot C_{Stack} \cdot (V_{GB} - V_{TH}) \cdot V_{DS} & , V_{TH1} < V_{GB} < V_{TH2} \cap P = P^+ \end{cases} \tag{3.2}$$

A model of the polarization will now be described, because it is necessary for the implementation of a more accurate FeFET model.

3.2 Modeling the polarization hysteresis

Ferroelectric materials, just like ferromagnetic ones (although for different reasons), show a characteristic hysteresis curve when an electric field is applied to them. The algorithm presented here was published in [8], [16], with a more physical analysis. The description here aims only to present the mathematics behind it, which is not easily found documented. The model is known as the Preisach model [82] that uses the DFIM[9] approach [16].

The total polarization consists of a linear and a non-linear[10] part:

$$P(t) = P_{nonlin}(t) + P_{lin}(t) = -P_S + \sum_i (2 \cdot P_S \cdot A \cdot a_i) + P_{lin}(t) , \quad a_i = \{-1, +1\} \tag{3.3}$$

$$P_{lin} = \varepsilon_0 \cdot \varepsilon_r \cdot E(t) , \tag{3.4}$$

where A is given by

$$A = A(E_1, E_2) = \frac{\left(\arctan\left(\dfrac{E_1 - E_C}{\delta_1}\right) + \dfrac{1}{2} \cdot \pi\right) \cdot \left(\arctan\left(\dfrac{E_2 + E_C}{\delta_2}\right) + \dfrac{1}{2} \cdot \pi\right)}{\pi^2} \tag{3.5}$$

and δ is a constant that is material dependent and is given by

$$\delta_{1,2} = \frac{\pm E_C}{\ln\left(\dfrac{1 + \dfrac{\pm P_R}{P_S}}{1 - \dfrac{\pm P_R}{P_S}}\right)} . \tag{3.6}$$

E_C, P_S, P_R and ε_r are constant material parameters. The parameter α_i is either -1 or $+1$ (see

Eq. 3.7). Instead of δ the parameter $\sigma_{1,2} = \mp E_C \tan\left(\dfrac{\pi}{2} \cdot \dfrac{\mp P_R}{P_S}\right)$ can be used [8] (is used for the

---

[9] Distribution Function Integral Method

[10] $P_{nonlin} \equiv P_{Fe}$

simulations in this thesis). The positive sign in the expressions of δ (σ) corresponds to δ_1 (σ_2) and the negative to δ_2 (σ_1).

The arc*tan* function (is used for all calculations in this thesis) can be replaced by the *tanh* function. The difference lies in the smoother transition of the arc*tan* function that approximates a real hysteresis more closely than *tanh*, which has a more square shape, as shown in Fig. 3.2.
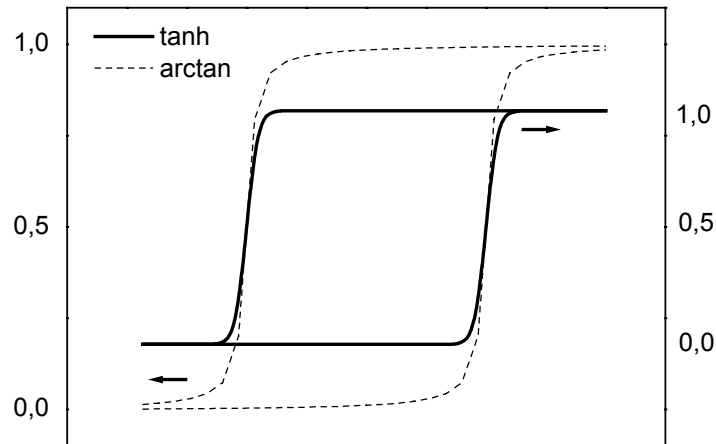

Fig. 3.2 Comparison of the arctan function with the tanh.

Equation 3.5 is a 2-variable function, whose plot is shown in Fig. 3.3 (left). E_1 is the highest and E_2 the lowest applied field. $A(E_1,E_2)$ gives the normalized polarization that has a maximum value of *1* at saturation. Moving on the diagonal path (45°) gives the values of the saturated curve (A→B) and along a path of a constant E_{min} those of a sub-loop (B→C) (Fig. 3.3 right).
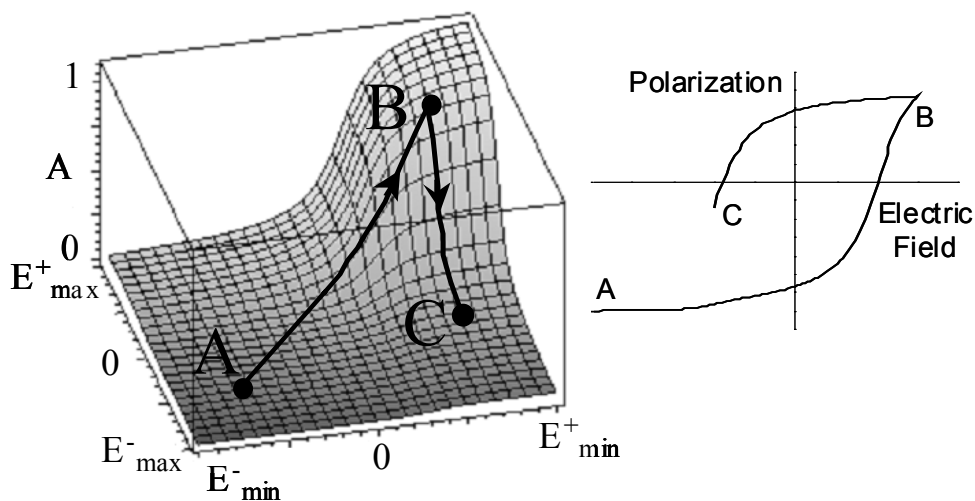

Fig. 3.3 (left) 3-D plot of the function of Eq. 3.5 that shows how the hysteresis curve is built, (right) plot of a hysteresis sub-loop after following the path drawn on the 3-D plot.

The function *A* represents the amount of dipoles that are positively polarized.

In order to account for the history of the applied voltages, two E-Field lists are built, one with the minima and one with the maxima. The block diagram in Fig. 3.4 shows the algorithm for building the two E-Field lists. Then A is calculated according to Eq. 3.7.

$$A = A(E_{max1}, E_{min1}) + A(E_{max2}, E_{min2}) - A(E_{max2}, E_{min1}) + A(E_{max3}, E_{min3}) - A(E_{max3}, E_{min2}) + ... \quad (3.7)$$
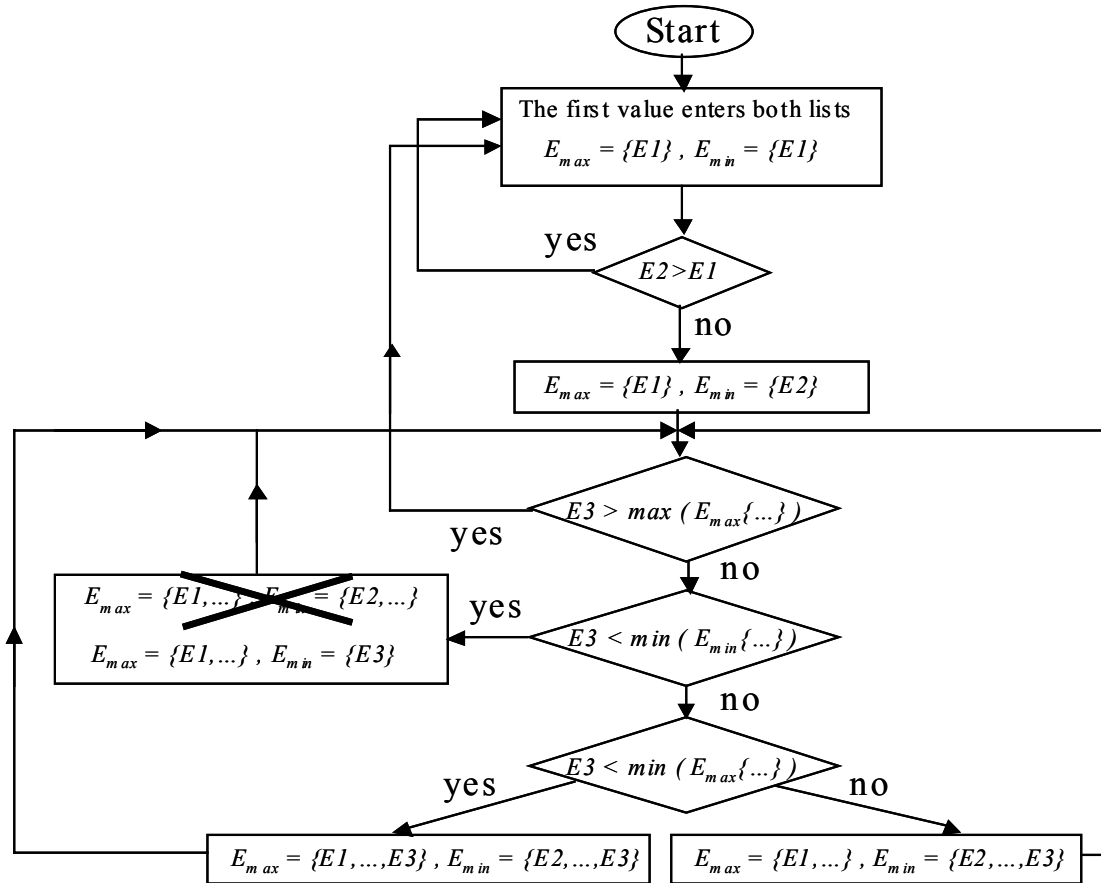
Fig. 3.4 Block diagram of the algorithm to calculate the lists of E-Field maxima and minima.

The following example should clarify how the algorithm works.

3.2.1 Example

The piece-wise-linear voltage of Fig. 3.5 is applied to a ferroelectric capacitor. At $t = t_0$ the ferroelectric is assumed unpolarized. The parameters of the ferroelectric are:

$E_C = 50\ KV/cm$, $P_R = 1\ \mu C/cm^2$, $P_S = 1.25\ \mu C/cm^2$, $d_{Fe} = 200\ nm$, $\delta = 1.36 \cdot 10^6\ V/m$.

Fig. 3.5 The voltages applied to the ferroelectric capacitor.

The non-linear part of the polarization will be calculated at all time steps. In the following calculations the voltages are displayed instead of the E-Fields for convenience (they have to be converted to E-Fields using $E_{Fe} = \dfrac{V}{d_{Fe}}$). Then Eq. 3.7 is used to calculate A and finally Eq. 3.3 to calculate the polarization (non-linear part):
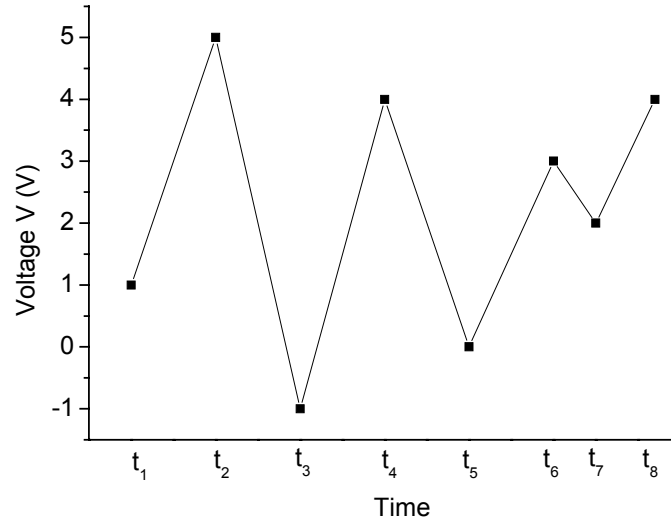
$t = t_1,\ V = 1\ V:\ V_{max} = \{1\}\ V,\ V_{min} = \{1\}$
$A = A(1,1) \Rightarrow P_{Fe} = -P_S + 2 \cdot P_S \cdot A = 0.66\ \mu C/cm^2$

$t = t_2,\ V = 5\ V:\ V_{max} = \{5\}\ V,\ V_{min} = \{5\}$
$A = A(5,5) \Rightarrow P_{Fe} = 1.16\ \mu C/cm^2$

$t = t_3,\ V = -1\ V:\ V_{max} = \{5\}\ V,\ V_{min} = \{-1\}\ V$
$A = A(5,-1) \Rightarrow P_{Fe} = -0.78\ \mu C/cm^2$

$t = t_4,\ V = 4\ V:\ V_{max} = \{5,4\}\ V,\ V_{min} = \{-1,4\}\ V$
$A = A(5,-1) + A(4,4) - A(4,-1) \Rightarrow P_{Fe} = 1.14\ \mu C/cm^2$

$t = t_5,\ V = 0\ V:\ V_{max} = \{5,4\}\ V,\ V_{min} = \{-1,0\}\ V$
$A = A(5,-1) + A(4,0) - A(4,-1) \Rightarrow P_{Fe} = 0.86\ \mu C/cm^2$

$t = t_6,\ V = 3\ V:\ V_{max} = \{5,4,3\}\ V,\ V_{min} = \{-1,0,3\}\ V$
$A = A(5,-1) + A(4,0) - A(4,-1) + A(3,3) - A(3,0) \Rightarrow P_{Fe} = 1.13\ \mu C/cm^2$

$t = t_7,\ V = 2\ V:\ V_{max} = \{5,4,3\}\ V,\ V_{min} = \{-1,0,2\}\ V$
$A = A(5,-1) + A(4,0) - A(4,-1) + A(3,2) - A(3,0) \Rightarrow P_{Fe} = 1.11\ \mu C/cm^2$

$t = t_8,\ V = 4\ V:\ V_{max} = \{5,4\}\ V,\ V_{min} = \{-1,4\}\ V$
$A = A(5,-1) + A(4,4) - A(4,-1) \Rightarrow P_{Fe} = 1.14\ \mu C/cm^2$

Figure 3.6 (left) shows the voltage maxima and minima at every calculation step and Fig. 3.6 (right) plots the results on the saturated hysteresis plot.
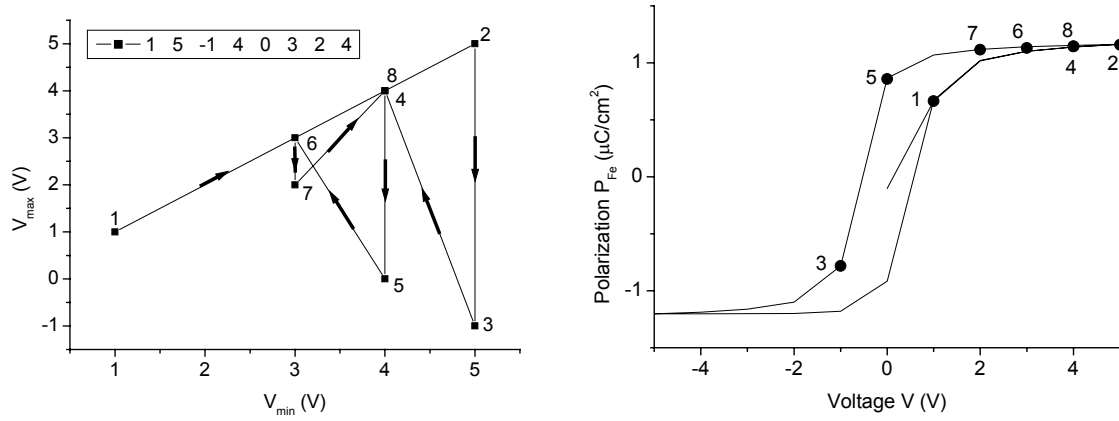

Fig. 3.6 The voltage maxima and minima at every time step (left), the polarization values calculated (right).

It should be noted that the algorithm gives better results when used with a small step-size and a large number of steps.

3.3 Parameterization of the polarization curve

The calculated hysteresis curve (Fig. 3.7 left) has a flat part in the lower branch that is due to Eq. 3.5. This can be corrected by using a set of parameters (k_1, k_2, k_3), so Eq 3.3 and 3.5 can be rewritten as follows

$$P(t) = \frac{-P_S + \sum_i (2 \cdot P_S \cdot A \cdot a_i) + P_{lin}(t)}{k_3} \ , \quad a_i = \{-1, +1\} \tag{3.8}$$

$$A = \frac{\left(\arctan\left(\frac{E_1 - E_C}{\delta_1} \right) + k_1 \right) \cdot \left(\arctan\left(\frac{E_2 + E_C}{\delta_2} \right) + k_1 \right)}{k_2} \tag{3.9}$$

Parameter k_1 adjusts the lower branch (Fig. 3.7 right). The derivative of the hysteresis curves shown in Fig. 3.8 (left), that is analogous to the capacitance, shows a hillock that should not be. Modifying k_1 can also correct the symmetry making both peaks equally high, but can shift the curve up or downwards. Then k_2 and k_3 can be used for finer adjustments and the result is shown in Fig. 3.8 (right). Figure 3.9 shows the effect of k_1 on the 3-D plot of A. The lower hysteresis branch (see also Fig. 3.3) is obtained by moving along the path of $E_1=E_2$ (45°).
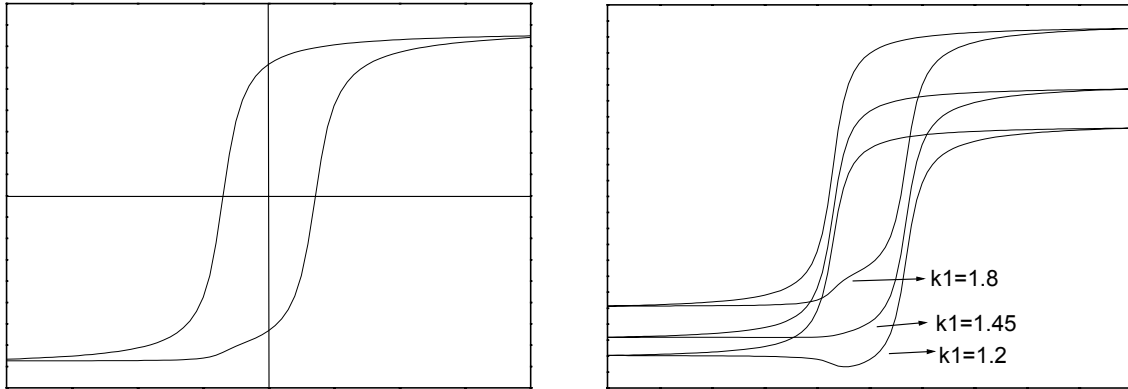
Fig. 3.7 Hysteresis plot with $k_1=\pi$, $k_2=\pi^2$, $k_3=1$ (left), hysteresis plot for different k_1 values (right).
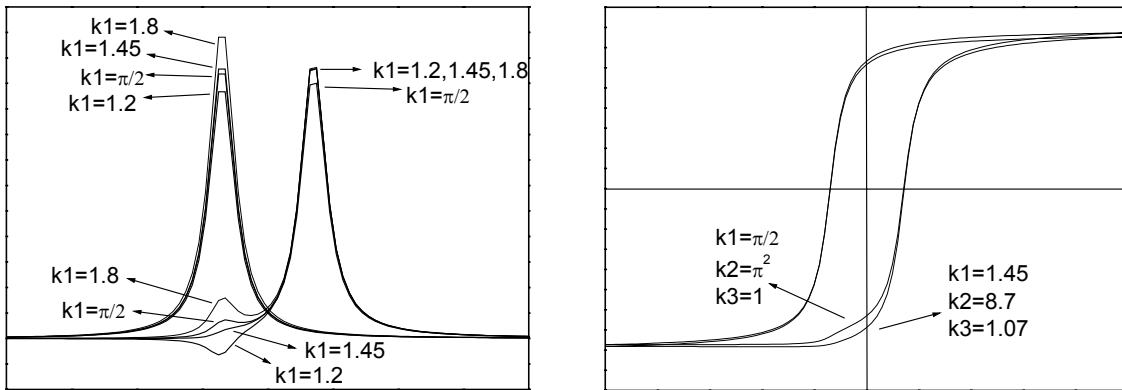

Fig. 3.8 Plot of the derivative of the hysteresis for different k_1 values (left), hysteresis plots for two sets of k parameters (right).
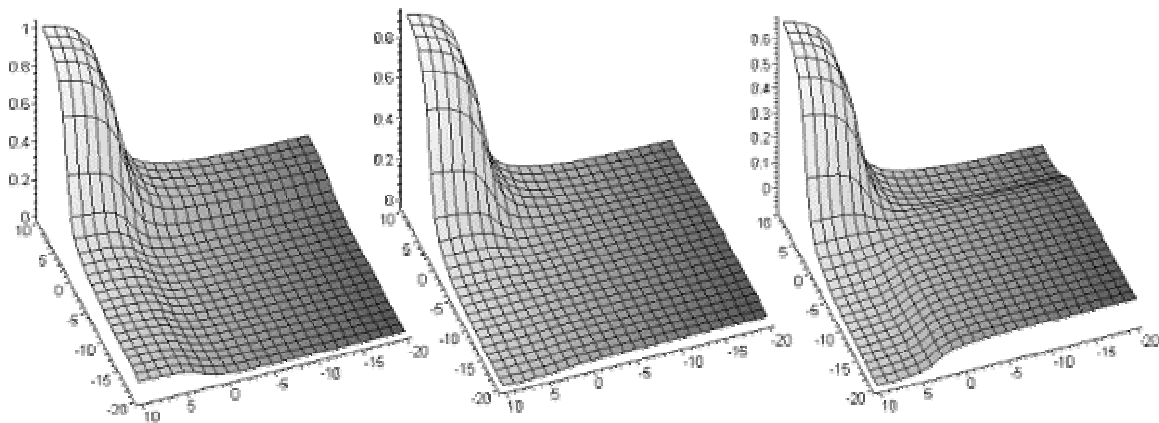

Fig. 3.9 3-D Plot of the A function for $k_1=1.2$, 1.45, 1.8 (from left to right).

3.4 The FeFET model

3.4.1 Extending the BSIM3v3

The model that is used throughout this thesis is based on the industry standard BSIM3v3 MOSFET model by the UC Berkeley [6]. It has been extended to include a ferroelectric layer by introducing the effective gate voltage V_{GBeff} and the stack capacitance C'_{Stack}

$$V_{GBeff} = V_{GB} + \frac{d_{Fe}}{\varepsilon_0 \varepsilon_r} \cdot P_{Fe} \tag{3.10}$$

$$C'_{Stack} = \left(\frac{d_{Fe}}{\varepsilon_0 \varepsilon_r} + \frac{d_{Ox}}{\varepsilon_0 \varepsilon_{Ox}} \right)^{-1} \tag{3.11}$$

and by calculating the polarization using the mathematical algorithm described in the previous section. The ferroelectric field E_{Fe} that is used to calculate the polarization is given by the recursive equation

$$E_{Fe} = \frac{-Q'_G - P_{Fe}(E_{Fe})}{\varepsilon_0 \varepsilon_r}. \tag{3.12}$$

As already mentioned, the FeFET is a MOSFET with a ferroelectric layer inserted in the gate stack, so the combination of a ferroelectric model with a MOSFET model is an effective way to model the device. This approach was described in [17] and is adopted in this thesis as well.

Although it has the same effect, it is not exactly the same as a ferroelectric capacitor connected to the gate of a MOSFET (two discrete circuit elements). An attempt to model the device this way sometimes yields no accurate results, because, depending on the MOSFET model used, the charge conservation condition is not always satisfied with certain models. It will only be used in chapter 6, where the dependence of the aspect ratio of the ferroelectric to the oxide capacitor on the depolarization field will be studied. For the rest of this thesis the approach suggested here and originally presented in [17], combining the ferroelectric capacitor model with that of the MOSFET (a single circuit element), is followed. For a high degree of accuracy BSIM3v3.1 was chosen, being an industry standard for simulations down to small device dimensions.

As mentioned in the introduction, the BSIM3 is a semi-empirical model. It does not make direct use of Eqs. 2.3, 2.10-2.12, but uses a unified equation approach for the current equation to guarantee its continuity (differentiability) in all regions of operation [6]. Besides the currents in every terminal, the voltages and charges are calculated too. The gate charge per area $Q'_G = \frac{Q_G}{L \cdot W}$ is used in Eq. 3.12 for the calculation of the ferroelectric field.

3.4.2 FeFET hysteresis curves and sub-loops

It was mentioned that the shift of the threshold voltage represents the memory function in the FeFET. Figure 3.10 (left) shows the *I-V* curves of a FeFET with additional sub-loops and intermediate thresholds. The polarization hysteresis of the ferroelectric when plotted against the gate voltage (*P-V*), gives a curve different from the well known polarization hysteresis of a ferroelectric capacitor (e.g. Fig. 3.2). This also applies to the MFIS structure (note: the *C-V* curves are still like the one in Fig. 2.3). The operation region in the center has a characteristic plateau, which is due an extended depletion region, where the polarization doesn't rise proportionally to the gate voltage. The curves in Fig. 3.10 (right) are sub-loops of the saturated hysteresis loop. As will be explained, the FeFET should not be driven to saturation unless the ferroelectric has a low saturated polarization P_S. The direction in both curves is anticlockwise. A negatively (positively) polarized FeFET must be driven to strong-inversion (accumulation) to switch state.
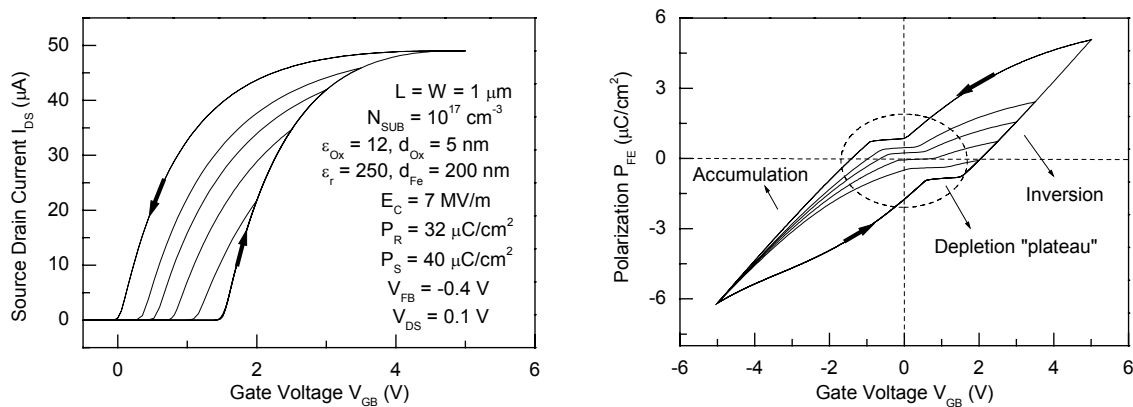

Fig. 3.10 I-V characteristic and sub-loops in a FeFET (left), P-V characteristic and sub-loops (right).

3.5 Symmetry in the P-V curve and flatband voltage effect on P-V and I-V curves

One important parameter in the FeFET is the flatband voltage as defined in Eq. 2.1. Increasing the flatband voltage causes the threshold voltage to increase. Besides this, there is an additional effect on the *P-V* and *I-V* hysteresis curves. This is studied in this section.
In the curve of Fig. 3.10 (right) the absolute values of the remnant polarization for $V_{GB} = 0\ V$ (P^+ and P^-) are about equal. This is not always so. Even a shifted hysteresis curve (*P-V*) can lead to a functional device. The *P-V* curve can be shifted by changing the flatband voltage V_{FB}. It is fixed for a given device and can change by using a different metal gate or substrate doping (see Eqs. 2.1 and 2.2). The variation of the flatband voltage has a visible effect on the *I-V* and *P-V* curves. A more negative value shifts the current curve to lower voltages and the hysteresis loop to higher polarizations. Two cases will be examined. One with a high-P_R ferroelectric that is not polarized to saturation, and one with a low-P_R ferroelectric that is driven to saturation. In both cases the ferroelectric polarization does not exceed *3 μC/cm²*.
In the first case, Fig. 3.12 (right) shows that *50 V* must be applied to the gate to fully polarize the ferroelectric, but only *7 V* will drop across the ferroelectric (Fig. 3.11 right). This puts the

sub-loops of Fig. 3.12 (left) and the "lemon-like" hysteresis of Fig. 3.11 (left) into perspective.
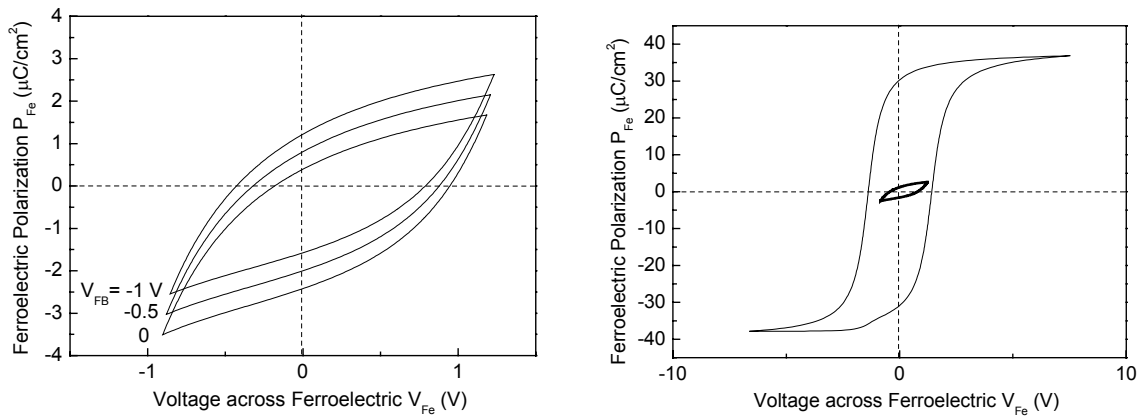

Fig. 3.11 P-V_{Fe} curves for different flatband voltages (left), the curve on the left plotted together with the saturated P-V_{Fe} curve (right).
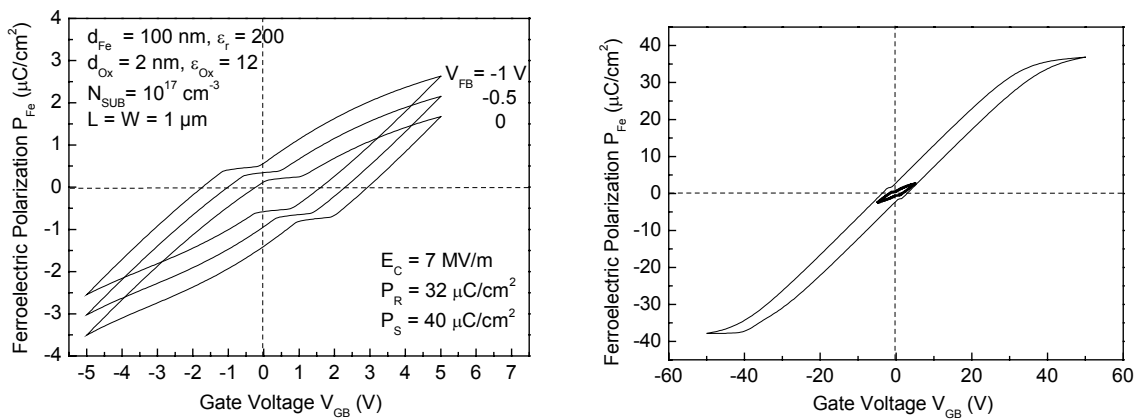
Fig. 3.12 P-V curves for different flatband voltages (left), the curve on the left plotted together with the saturated P-V curve (right).

Both polarization hysteresis curves are shifted upwards for a more negative flatband voltage, and according to Fig. 3.13, the *I-V* curve is shifted to the left. This last shift is easy to understand as, according to Eq. 2.5, the threshold increases with a higher flatband voltage. Thus, the flatband voltage is a symmetry parameter that can be tweaked for equal positive and negative remnant polarization values. However, this is not the deciding factor when designing the device. What is important is the *I-V* curve of Fig. 3.11, that gives information about the voltage thresholds and shows if the device can tolerate a disturbance voltage that is half the operating voltage, without changing its polarization. From the curves of 3.12 (left) and 3.13 it can be seen that the FeFET with V_{FB}=-1 V and equal remnant polarization values has lower voltage thresholds and is more susceptible to disturbance than the FeFET with V_{FB}=0 V. The memory window does not change much despite the shift. A more detailed explanation of disturbance in the FeFET is given in section 4.4.
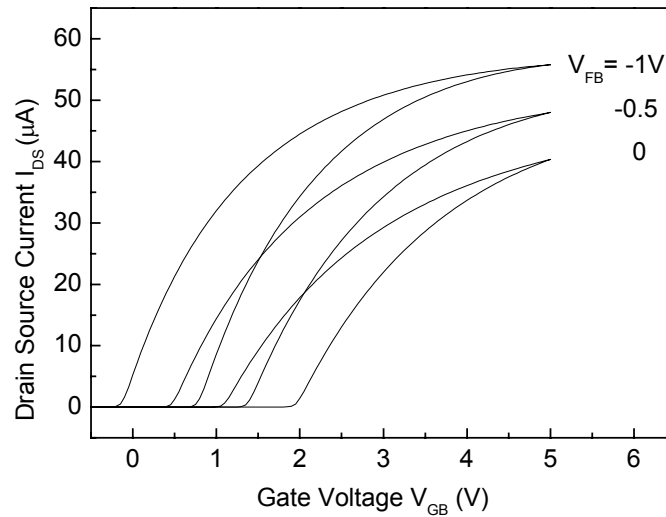
Fig. 3.13 I-V curves for different flatband voltages.

With the saturated ferroelectric the impact of the flatband voltage is a little different. Again the *I-V* curves are shifted to the right for the same reason, but the polarization curves of Fig. 3.14 (left) are shifted to the right too, instead of upwards. The reason for this is that the *P-V_{Fe}* curve is now already saturated, so there is no "room" left for a shift of the *P-V* curves to a higher polarization.
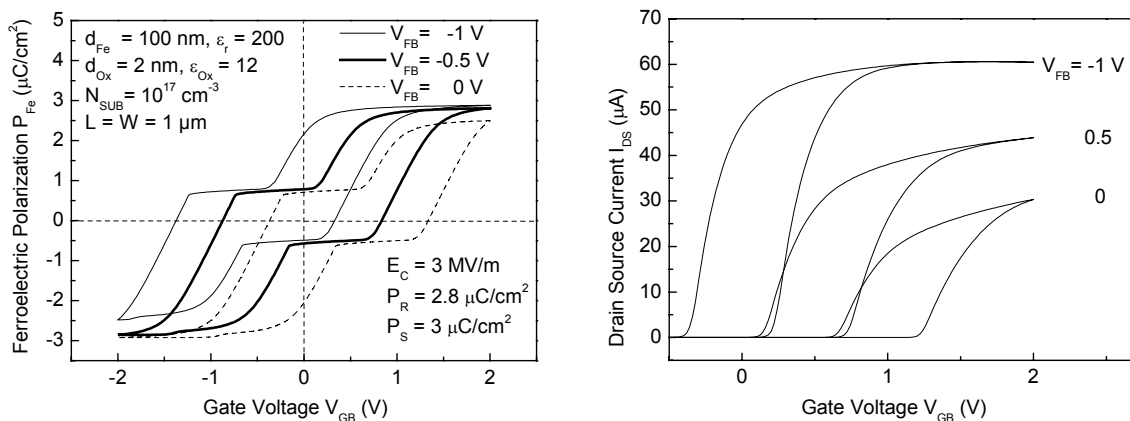
Fig. 3.14 P-V curves (left) and I-V curves (right) for different flatband voltages.

3.6 Summary

In this chapter the FeFET model was presented starting with a mathematical algorithm for the description of the polarization hysteresis. This hysteresis model was then coupled to the BSIM3v3 MOSFET model to get the FeFET model that is used in the simulations throughout this thesis. Finally, the FeFET dependence on the flatband voltage was investigated. It will be applied in section 7.5 to shift the FeFET hysteresis curve.

4 Failure mechanisms

4.1 Depolarization Field

In contrast to the ferroelectric capacitor, the remnant polarization at short circuit conditions ($V_{GB} = 0\ V$) in the FeFET is less than P_R (Fig. 4.1c). The reason for this is the depolarization field in the ferroelectric, because of the incomplete charge compensation at the interfaces (the interface to the oxide or semiconductor) [10]. This is also the reason for the plateau shown in Fig. 4.1 (right), which is characteristic of this device. As shown in Fig. 4.1b, and for the reasons presented in [8][11] the depolarization field is $0<E_{dep}<E_C$. In the ferroelectric capacitor (Fig. 4.1a) there is full charge compensation at the metallic electrodes, under ideal screening. This leads to zero electric field inside the ferroelectric[12] for $V_{GB} = 0\ V$. In the FeFET gate stack, however, the oxide or semiconductor that is in direct contact with the ferroelectric cannot provide enough charges to compensate for the surface charge of the polarized ferroelectric. The depolarization field is always opposite to the direction of the polarization and leads to its diminishing. The depolarization field is considered responsible for the low retention times in MFIS structures (Fig. 2.10). Higher retention times are achieved using an MFMIS structure, but it is not clear if this is because of better charge screening by the floating gate or improved interfaces. It has been shown in [4] that this particular structure leads to higher memory retention times. Other solutions for the reduction of the depolarization field are suggested in chapter 6.
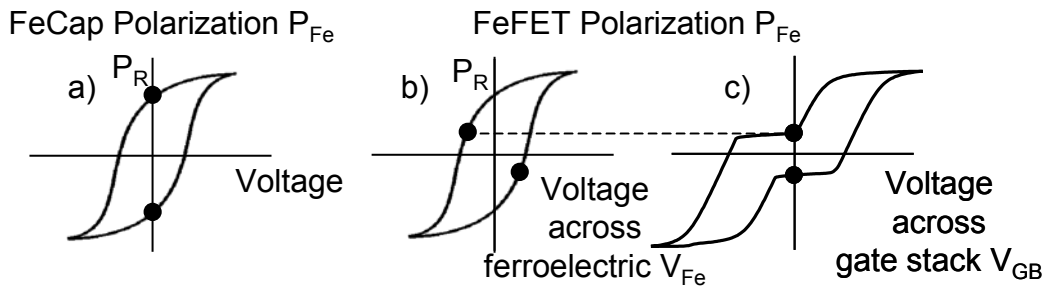
Fig. 4.1 Remnant polarization in the ferroelectric capacitor (a) and the FeFET (b and c).

The FeFET is typically described with one of its steady states shown in Fig. 4.2 (left). Here the polarization charges are shown completely compensated by dielectric dipoles. This results in a zero depolarization field and a remnant polarization of P_R. A more realistic situation is shown in the right where the polarization appears reduced (due to the depolarization field). The depolarization field is not clearly visible in this figure either.

---

[11] The depolarization field is the intersection point of the $Q_G(-V_{IS})$ and the $Q_G(V_{Fe})$ curve.
[12] Throughout this thesis the terms depolarization field and field in the ferroelectric will be synonymous ($E_{dep} \equiv E_{Fe}$).
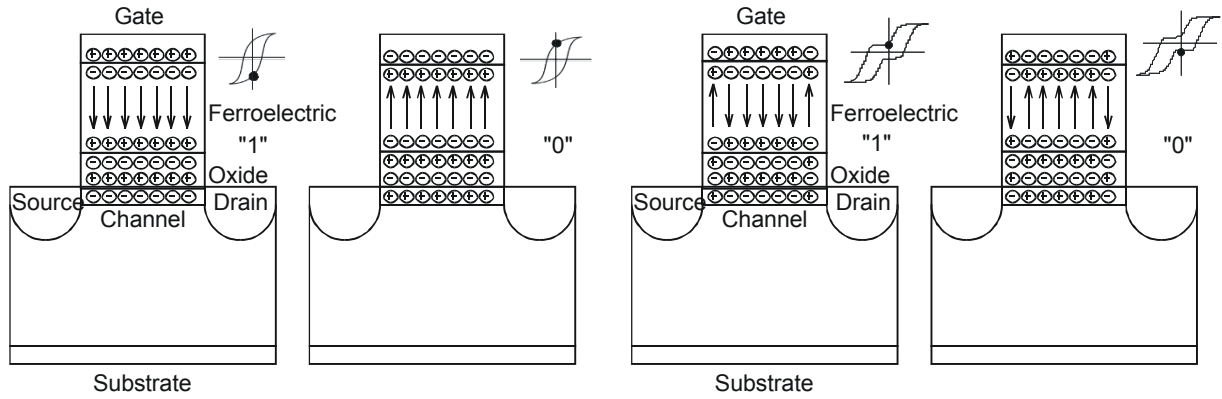
Fig. 4.2 Charge concentration in the FeFET for $V_{GB} = 0$ V, (left) $E_{dep}=0$, (right) $E_{dep} \neq 0$.

A better way to visualize the gate stack structure is to use the charge distribution and calculate from it the electric field and potential. This is shown in Fig. 4.3 for an MFM, an MFS and an MFIS structure for $V_{GB} = 0$ V. The potential difference at the terminals is due to the flatband voltage (Eq. 2.1). The charge at the interface to the ferroelectric does not fully screen the polarization charge. This gives rise to the depolarization field. In total, of course, there is charge compensation.
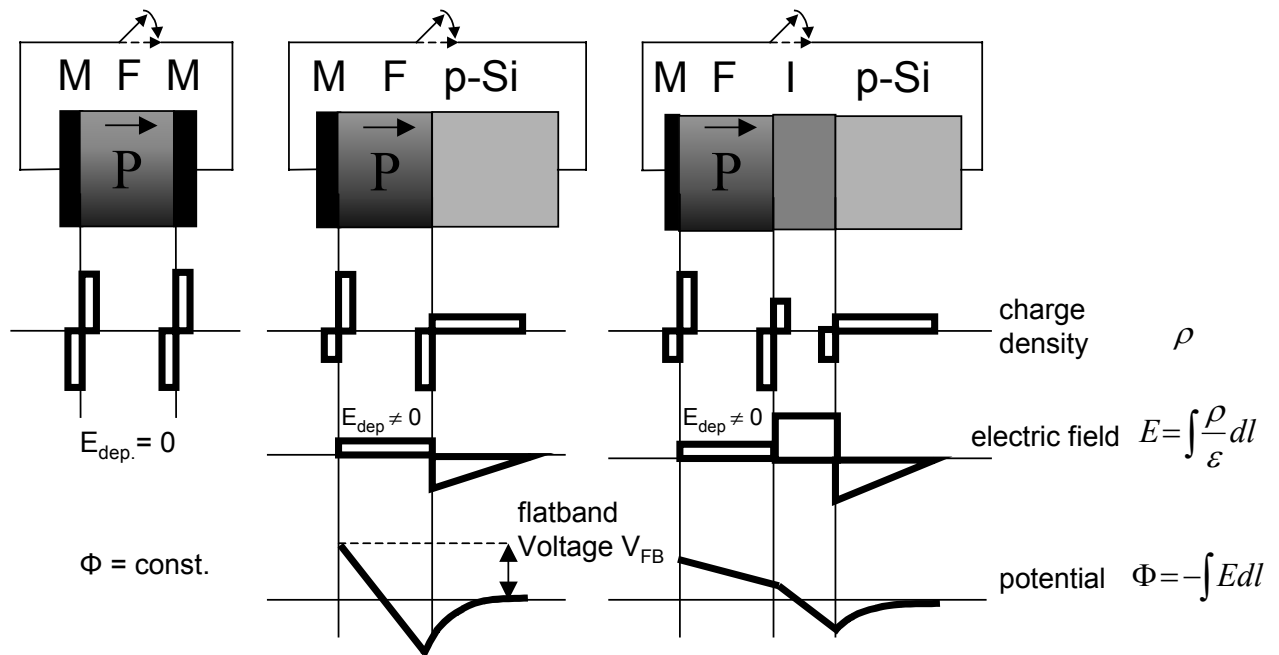
Fig. 4.3 Charge concentration, E-Field and potential distribution for $V_{GB} = 0$ V for the MFM ($E_{dep}=0$), MFS ($E_{dep}\neq0$) and MFIS ($E_{dep}\neq0$) structures.

At the steady state, the depolarization field causes the reduction of the polarization. As Fig. 4.4 shows, the minimization of the total energy in the ferroelectric leads to the formation of domains with different orientations. In order to reduce the total energy and increase the thermodynamic stability, the ferroelectric splits into domains, thus reducing its effective polarization. Another result of the depolarization field is charge injection, that can lead to the reduction of the polarization. This will be examined in the next sections.
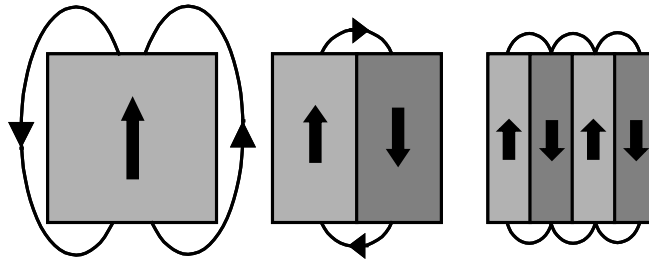
Fig. 4.4 Depolarization field causes the formation of domains to reduce the total energy [9].

4.2 Loss of Retention and Leakage current in the FeFET

As mentioned in 4.1, another reason for the loss of polarization is the leakage current in the gate stack, or rather the charge injection through the dielectric or the ferroelectric layer. This takes place when the device is at idle state ($V_{GB} = 0$ V), and is powered by the depolarization field (see Fig. 4.5 left). It is believed [18] that the leakage current leads to charge recombination at the ferroelectric-insulator interface, thus reducing the polarization of the ferroelectric (Fig. 4.5 right).
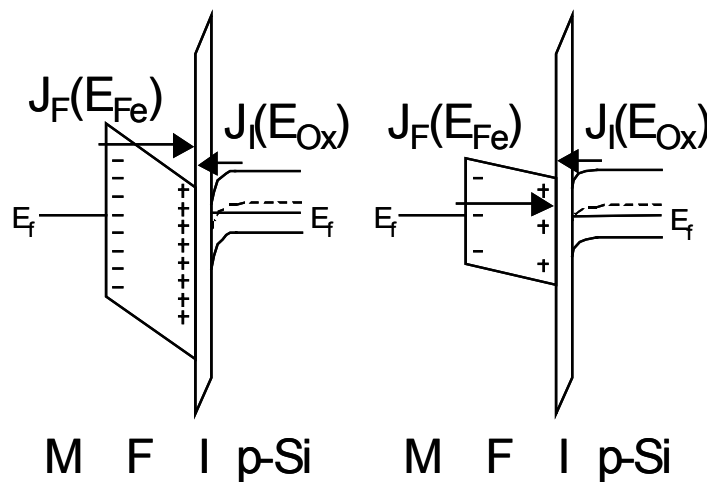
Fig. 4.5 Charge injection through the gate stack layers leads to a reduction of the polarization [18].

4.3 Charge transport in insulators

Despite the high energy bandgap in insulators and a large thickness, charge transport does occur and can lead in several cases to an unwanted leakage current. The leakage current through the gate stack in the case of the FeFET can reduce the data retention time (polarization).

The charge transport in insulators can be attributed to a number of mechanisms as can be seen in Fig. 4.6. They are divided in the injection mechanisms, such as tunnel and thermionic injection, and the transport mechanisms, such as Poole-Frenkel, hopping, drift and diffusion. The condition for drift is the existence of enough free states in the valence or conduction band, and for diffusion a carrier concentration gradient.
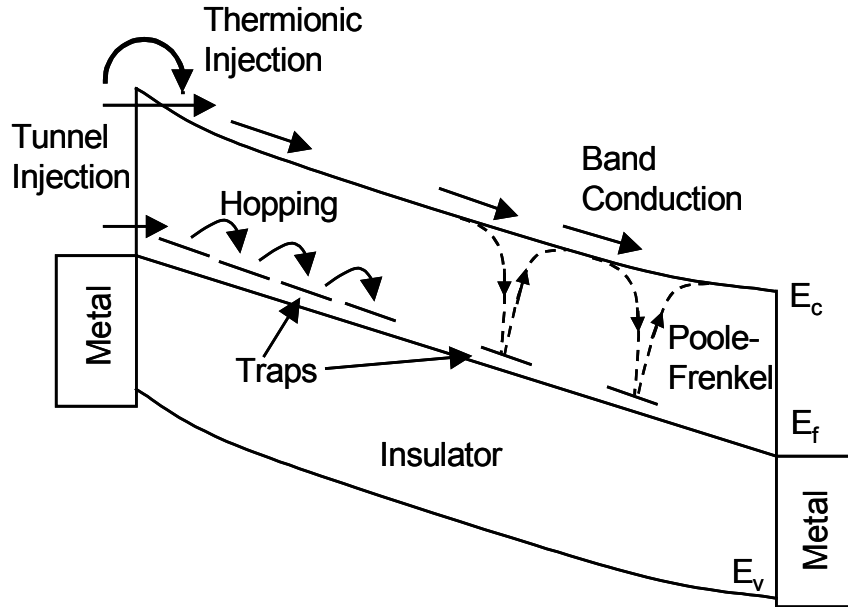
Fig. 4.6 Charge transport mechanisms in insulators [9].

Here the most important transport mechanisms in insulators are briefly described.

4.3.1 Frenkel – Poole emission

The leakage current in insulators with a high number of defects (such as ferroelectrics of perovskite type with many oxygen vacancies) can be attributed to a Frenkel-Poole emission mechanism (Fig. 4.7). The energy of the defects can be close to either the valence or the conduction band. An electron (hole), that is somehow injected into the insulator and lands on such a defect, can gain enough energy to jump to the conduction (valence) band and to the next defect. This hopping from one defect to another can result in a charge transport even in insulators with a thickness of more than *100 nm*. Equation 4.1 describes the dependence of the current density on the energy barrier Φ_B, the reduced barrier $\Delta\Phi$, the applied field E (in the calculations the field in the ferroelectric E_{Fe} was used) and the temperature T (assumed *300 K*). ε_{opt} is the optical dielectric constant (here assumed 5) at the interface where the injection takes place. It is much lower than the material's dielectric constant.
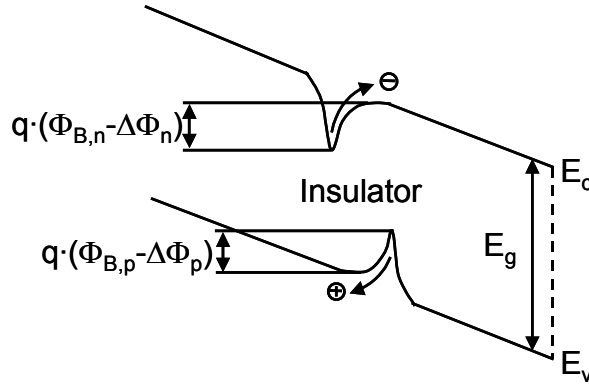
Fig. 4.7 Band diagram of an insulator under application of an electric field. This shows the energy gap reduction through the presence of defects which can enable a charge hopping transport [21].

$$J_{FP} = \sigma_{FP} \cdot E \cdot e^{\frac{-q \cdot (\Phi_B - \Delta\Phi)}{k \cdot T}}, \quad where$$

$$\Delta\Phi = \sqrt{\frac{q \cdot E}{\pi \cdot \varepsilon_{opt} \cdot \varepsilon_0}}$$

(4.1)

The hopping mechanism, shown in Fig. 4.6, where a carrier can hop from defect to defect, provided there is a high density of defects with neighboring defects being located close to each other, is not described by Eq. 4.1.

4.3.2 Schottky thermionic emission

Charge injection into the insulator can take place when a carrier gains enough energy to overcome the energy gap with the insulator. According to the Schottky emission (Fig. 4.8), the energy barrier Φ_B is reduced by $\Delta\Phi$ through the existence of mirror forces at the interface. Equation 4.2 gives the Schottky current density and barrier lowering.
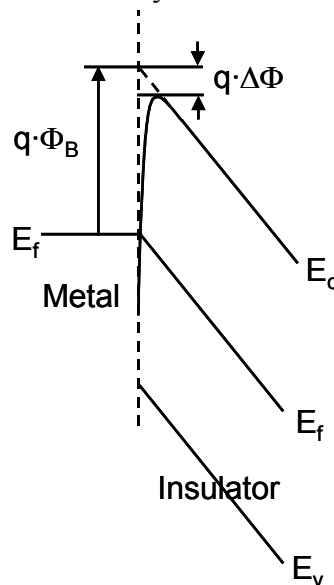
Fig. 4.8 Band diagram of an M-I (metal-insulator) contact that shows the band lowering due to the Schottky effect [21].

$$J_S = A^* \cdot T^2 \cdot e^{\frac{-q \cdot (\Phi_B - \Delta\Phi)}{k \cdot T}} \;, \; where$$

$$\Delta\Phi = \sqrt{\frac{q \cdot E}{4\pi \cdot \varepsilon_{opt} \cdot \varepsilon_0}} \;\; and \;\; A^* = \frac{4\pi \cdot q \cdot m^* \cdot k^2}{h^3} \tag{4.2}$$

A^* is the effective Richardson constant, m^* the effective electron mass (in the calculations in the next section the normal electron mass m and Richardson constant $A=120{\cdot}10^4 \frac{A}{m^2 K^2}$ was used) and h is Planck's constant. Notice the similarity in the current density expressions for J_{FP} and J_S with the former being proportional to the electric field E and the latter to the square of the temperature T. Also the $\Delta\Phi$ calculated in the Frenkel-Poole emission is double that of the Schottky emission.

4.3.3 Fowler-Nordheim Tunneling

If the thermal energy of the electrons is not enough to overcome the energy gap, they can still cross the boundary with a certain probability Θ. The resulting current density is given by Eq. 4.3.

$$J_{FN} = const \cdot E^2 \cdot \Theta = \frac{q^2}{8\pi \cdot h \cdot \Phi_B} \cdot E^2 \cdot e^{\frac{-8\pi \cdot \sqrt{2 \cdot m^*} \, (q \cdot \Phi_B)^{\frac{3}{2}}}{3 \cdot h \cdot q \cdot E}} \tag{4.3}$$

As shown in Fig. 4.9, the electrons at Fermi level E_f need to tunnel through a thickness of d_{FN} ($d_{FN} < d_{Ox}$) to enter the conduction band of the oxide.

Fig. 4.9 Band diagram of an MOS contact that is susceptible to Fowler-Nordheim tunneling.

The dielectric layer, being about an order of magnitude thinner than the ferroelectric, is more likely to give a leakage current through Nordheim-Fowler tunneling, because of the high electric field that is applied to it. Even direct tunneling is possible for thicknesses of a few *nm*.

The charge injection from either side, as in the depolarization field case, leads to incomplete charge compensation at the interfaces [18] and as a result to a reduction of the polarization (Fig. 4.5).

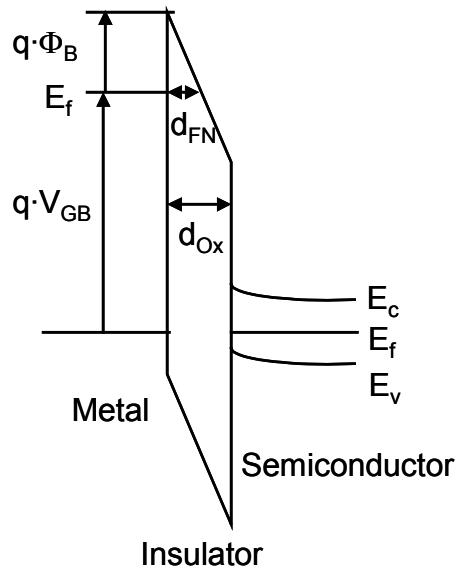The effect of the leakage current on the reduction of polarization will now be studied. It will be assumed that the ferroelectric does not present a limitation in energy states in the conduction band and does not limit the current in any way, other than by the reduction of the polarization. Also, drift and diffusion transport mechanisms will be ignored too, since the target is to calculate an upper limit of the current density (worst case calculation).

4.3.4 Loss of Retention under consideration of a Schottky emission in the ferroelectric

According to Eq. 4.2, the Schottky emission current depends on the barrier height, the electric field at the interface and the temperature. At short circuit conditions, this electric field is the depolarization field that is constant inside the ferroelectric. With a voltage applied, the electric field increases. But leakage is only an issue at off-state. Since the device spends most time being idle, only the off-state ($V_{GB} = 0\ V$) condition is of interest for retention time calculations.

The use of the FeFET model enables the calculation of a current that is self limiting and does not lead to unrealistic conditions. It will be used to calculate the retention time for different energy barriers. A similar method was used in [37].

The leakage current leads to charge recombination at the interfaces and thus to a reduced remnant polarization (Eq. 4.4). As a result the depolarizing field becomes smaller according to Eq. 3.12.

$$P_{Fe}(t_n) = P_{Fe}(t_{n-1}) - J_S(E_{Fe}, \phi_B) \cdot \Delta t , \qquad (4.4)$$

$$\text{where } E_{Fe}(t) = f(P_{Fe}) . \qquad (4.5)$$

In Eq. 4.4 the current density J_S in time step Δt ($J_S \cdot \Delta t$: charge density) causes charge recombination at the ferroelectric interface and thus the reduction of the ferroelectric polarization. Figure 4.10 shows the block diagram of the calculation. The time-step Δt can also be variable, which makes more sense for the very long simulation times applied here. Other current mechanisms can be applied too, instead of the Schottky emission. However, the Schottky limit was chosen as an upper limit to the leakage current that can take place, enabling the calculation of minimum retention times. Also, interface layers (so called "dead layers") were ignored as well as the conduction through the oxide from the other direction (see Fig. 4.5). Instead, one could use an effective barrier height Φ_{Beff} to account for the existence of interface layers or additional current transport mechanisms.

In the following table the current at $t=0$ is listed for different energy barriers. After application of the algorithm of Fig. 4.10 the retention time is calculated, defined here as the time until the polarization vanishes $(P_{Fe} \approx 0)$.

Φ_B (V)	1	1.1	1.2	1.3	1.4
J_S (A/m^2) at $t = 0$	5μ	100n	2n	50p	1p
Retention Time (s)	2500	$1 \cdot 10^5$	$1 \cdot 10^6$	$2 \cdot 10^8$	$1 \cdot 10^{10}$

Table 4.1 Schottky current for t=0 and retention times for different energy barriers (T =300 K).
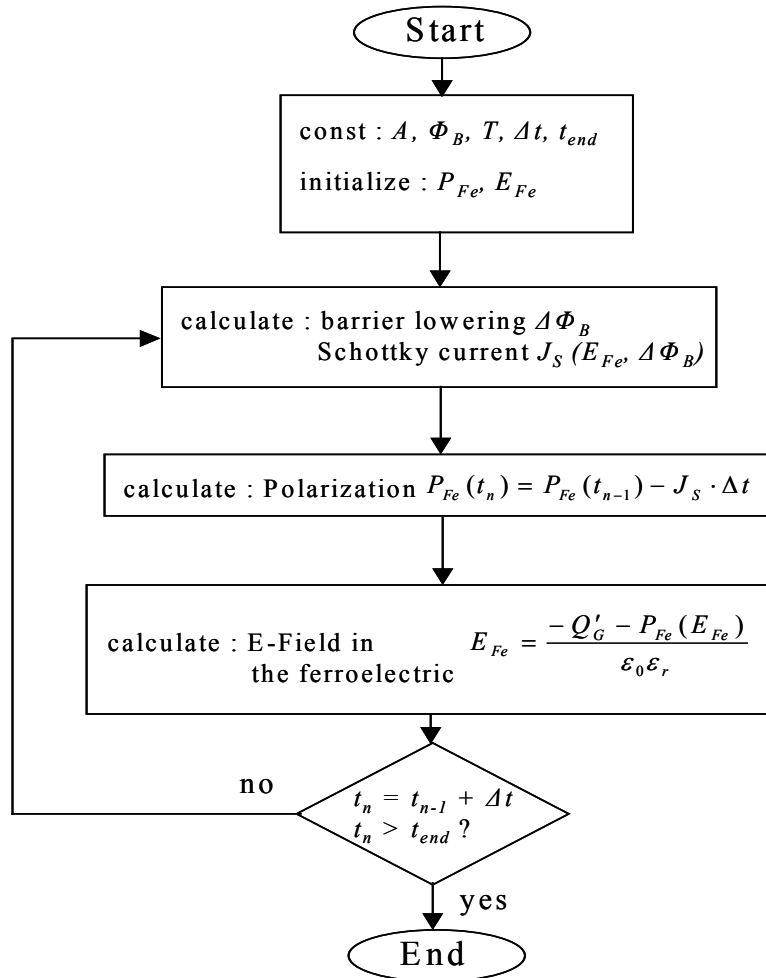
Fig. 4.10 Block diagram for the calculation of the reduction of the polarization through a recursive algorithm.

The barrier height is an important parameter in this calculation. If the metallic electrode is *Pt* (workfunction: *5.1 eV*) and the ferroelectric *PZT (3.6 eV)* then the energy barrier is $\Phi_B = \Phi_M - \Phi_{Fe} = 1.5$ *eV*. From measurements and calculations in [19], the barrier was estimated at *1.36 eV*. The highest value used in the simulation here is *1.4 eV*. The retention times calculated here, compared to reported times [4], make sense for a barrier lower than *1.3 eV*.

The FeFET used in the following simulation was that of Fig. 5.17 (low-P_R ferroelectric).

Figure 4.11 (left) shows the simulated *P-V* curve that is first initialized and left at stand-by at P^+ to slowly relax under the influence of the leakage current. Figure 4.12 shows how both the depolarization field and the current are reduced in the process, thus the polarization in Fig. 4.11 (right) does not decrease linearly and the whole process is self-limiting.
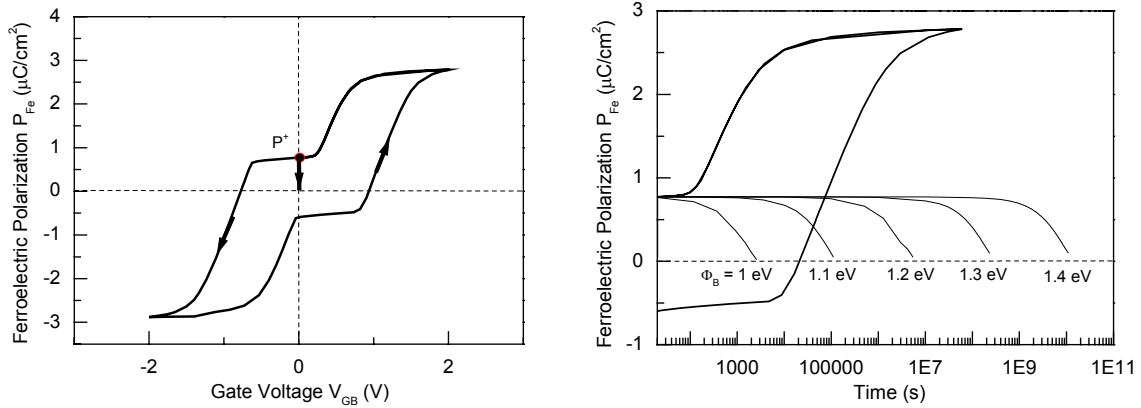
Fig. 4.11 Simulation of the reduction of the remnant polarization of a FeFET under application of the Schottky model - (left) P-V curve, (right) P-V & P-t curves.
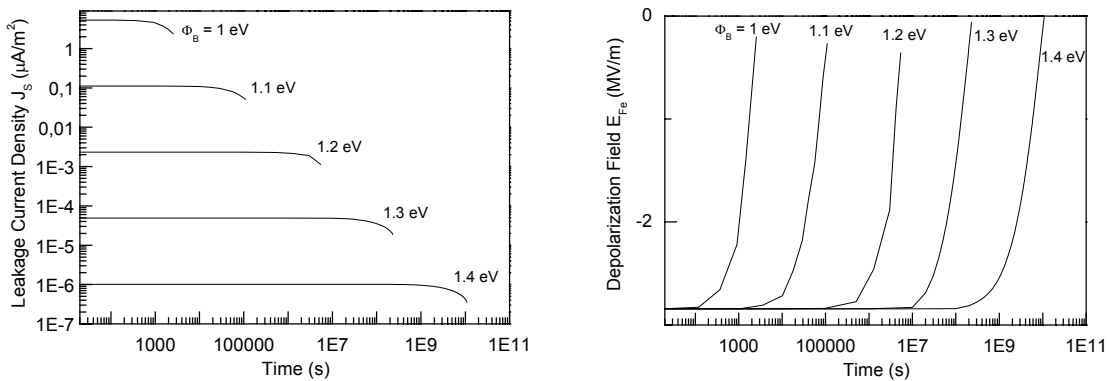
Fig. 4.12 The leakage current (left) and the depolarization field (right) are reduced with time.

Here the remnant polarization at P^+ was considered. For a FeFET with $P^+ \neq P^-$ the retention time for the state P^- should be accordingly different.

Instead of the polarization of Fig. 4.11, usually the capacitance is measured in a *C-t* measurement (Fig. 2.10). The result is that for $t = $ *Retention time* the two capacitance values at P^+ and P^- become equal. In order to directly measure the polarization, a Sawyer-Tower measurement can be performed, where the gate stack capacitor is compared to a fixed capacitor, which is much smaller than the gate stack capacitor at the measured frequency [20].

For the FeFET simulated, the electric field in the oxide (*13 MV / m*) is higher than that in the ferroelectric (*2.9 MV/m*). Despite this fact, applying the Nordheim-Fowler tunneling mechanism in the dielectric for the energy barriers in Table 4.1, does not yield a high enough current ($J_{FN} \ll J_S$) to cause any significant reduction in the remnant polarization.

Finally, applying the Frenkel-Poole emission mechanism in the ferroelectric, retention times on the order of 10^6 s (12 days) can be calculated, assuming an energy barrier of ϕ_B=0.5 eV and a conductivity σ_{FP}=200 nA/Vm, that is assumed to be field dependent [21] (in the algorithm of Fig. 4.10 the equation σ_{FP}=σ_{FP} - $J_{FP} \cdot \Delta t$ is added to the calculations).

4.4 FeFET programming and disturbance problems

The reduction (increase) of the remnant polarization below (above) a certain level can lead to false interpretation of a "1" as "0", or vice versa, during read out.

Two cases will be considered, one with a low-P_R (saturated hysteresis) and one with a high-P_R (sub-loop hysteresis) ferroelectric. First, disturbance starting from a negative, and then from a positive polarization state will be examined. The disturbance issue is important for the programming of the FeFET (chapter 7), and has to be considered when designing a FeFET.

In Fig. 4.13 the FeFET is switched from the negative to the positive state: (left) A→B→C→D→A, (right) A→B→C→D→A$^+$. When changing back to the negative state the other branch of the hysteresis is used again in counterclockwise direction and no current flows since V_G < 0 (0 < V_{TH1} < V_{TH2}).
A disturbance can occur when starting from a polarization state (at V_{GB} = 0 V) and a voltage, half the operating voltage V_{DD} is applied. It must be verified that this does not lead to a change in the polarization, either from P^+ or P^-. In Figs. 4.13-4.15 the operating voltage is V_{DD} =1.5 V (state C) and it is examined whether the application of the voltage $V_{DD}/2$ =0.75 V (state B) causes a change in the polarization or not.
Again in Fig. 4.13, starting from A we have: A→B→A (left), A→B→A$^-$ (right). The polarization at A$^-$ is slightly smaller (the absolute value) than at A and a very small current flows in the device. However, as Fig. 4.14 (right) shows, when a sub-loop hysteresis is used then the same route A→B→A$^-$ can lead to a major disturbance (polarization at A$^-$ is less than half of that at A). Thus, the hysteresis of Fig. 4.14 does not protect a device from a disturbance of $V_{DD}/2$, while that of Fig. 4.13 does.
Starting from the positive polarization state A$^+$: A$^+$→E→A$^+$. The current at state E is much larger than at state B. The ratio, however, is not high enough for the read-operation. For this reason a read voltage V_{Read} smaller than $V_{DD}/2$ must be applied. As shown in Fig. 4.15, it is at the states F and G where the maximum I_{on}/I_{off} occurs. Obviously, if the operating voltage V_{DD} is applied from state A$^+$, the polarization state will again be A$^+$ despite the sub-loop A$^+$→E→C→D→A$^+$ being driven. So, the disturbance during programming is only an issue when being at the negative state A$^-$. Unless, of course, there is a loss of retention for some reason (Fig. 4.11). In that case it is an issue for both states.

Fig. 4.13 (left) I-V curve for a FeFET during write / read / disturb, (right) P-V curve for a FeFET during write / read / disturb (ferroelectric with low-P_R).
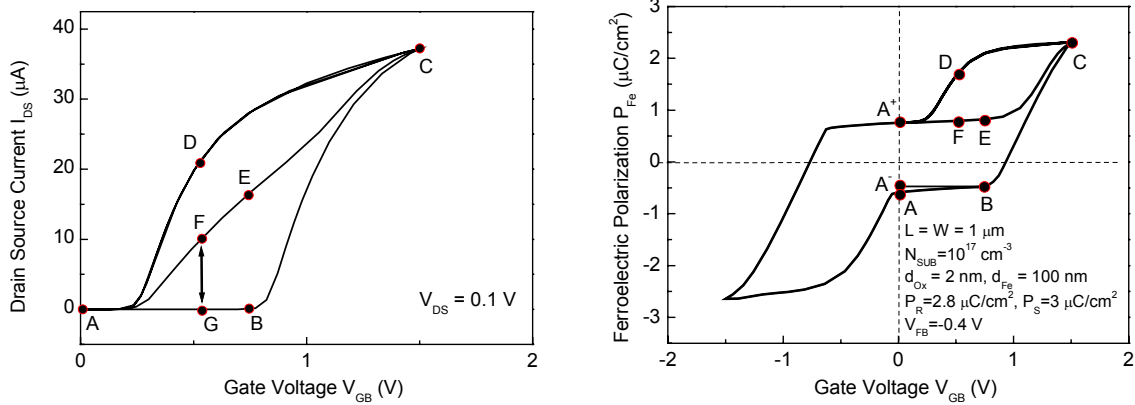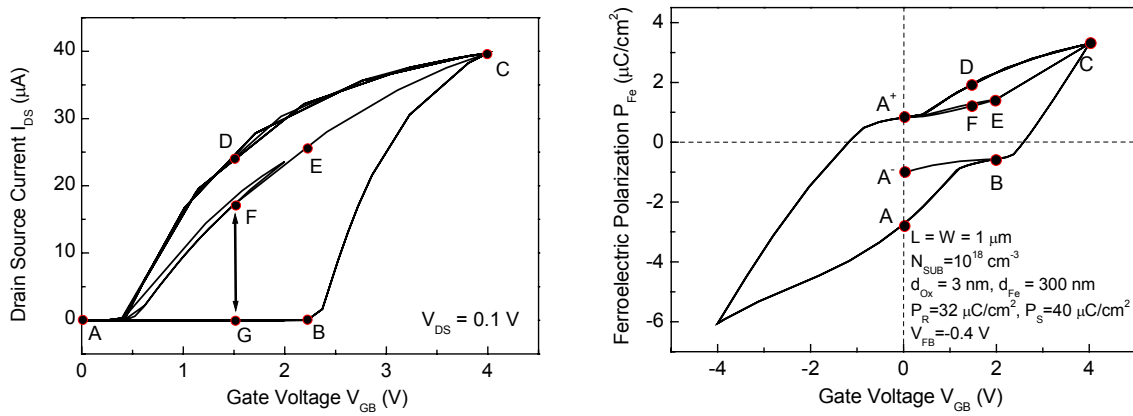

Fig. 4.14 (left) I-V curve for a FeFET during write / read / disturb, (right) P-V curve for a FeFET during write / read / disturb (ferroelectric with high-P_R).
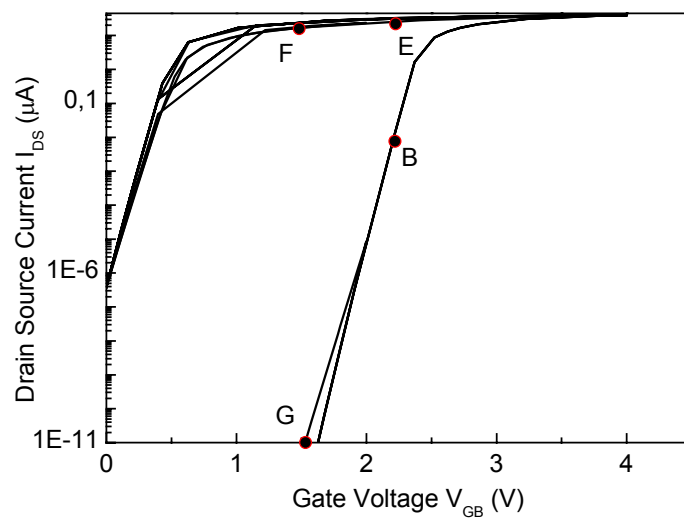
Fig. 4.15 Log-Lin I-V curve for the FeFET of Fig. 4.14 (left) (high-P_R ferroelectric).

4.5 Reduction in polarization and disturbance

In the last section it was shown how disturbance can be an issue for a badly designed FeFET. Now the effect of the loss of polarization on the read out and the reliability of the device is investigated. For this, sub-loops will be used to reduce the remnant polarization, then a read voltage will be applied and the currents will be compared. It is assumed that a FeFET that has lost an amount of its polarization, will move on a hysteresis sub-loop at read-out.

Fig. 4.16 (left) shows the polarization hysteresis of a FeFET (again the properties are those of Fig. 5.17 (low-P_R)) polarized to *2, 1.2, 1.1* and *1 V*. The read voltage is $V_{Read} = 0.7\ V$. The read currents are shown in Fig. 4.16 (right) and are also listed in Table 4.2 with the respective remnant polarization values.

Applied Read Voltage (V)	Remnant Polarization P^+ ($\mu C/cm^2$)	Percentage % of P_R (3 $\mu C/cm^2$)	Percentage % of Off – state P_R (0.77 $\mu C/cm^2$)	Read Current I_{on} (μA)
2.0	0.77	26 %	100 %	15
1.2	0.66	22 %	86 %	13
1.1	0.51	17 %	66 %	11
1.0	0.22	7%	29 %	7

Table 4.2 Reduced remnant polarization and read currents.

Fig. 4.16 Polarization reduction by driving sub-loops starting from P^+ (left) and read out currents for the reduced polarizations (right) (read voltage is 0.7 V).

It can be seen that despite a reduction in the polarization of more than 70% compared to the off-state value (more than 90% compared to P_R), the read current only drops by about half. This is as far as the positive polarization values are concerned.

To apply the same for a negative remnant polarization P^-, the ferroelectric is polarized at the negative state and a voltage smaller than V_{DD} is then applied to the gate to reduce the polarization as shown in Fig. 4.17 (left) (see also Fig. 4.14 (right) A→B→A$^-$). The voltages

applied were *0.7, 0.8, 0.85* and *0.9 V* and the polarization was reduced by up to 70 %. Then, as before, the read voltage of *0.7 V* was applied and the results are shown in Fig. 4.17 (right) and are listed in Table 4.3.

Applied Read Voltage (V)	Remnant Polarization P^+ (µC/cm²)	Percentage % of P_R (3 µC/cm²)	Percentage % of Off – state P_R (0.6 µC/cm²)	Read Current I_{off}(nA)
0.7	0.6	20 %	100 %	10
0.8	0.49	16 %	82 %	22
0.85	0.35	12 %	58 %	86
0.9	0.18	6%	30 %	730

Table 4.3 Reduced remnant polarization and read currents.

Here, the current dependence on the loss of polarization is more pronounced compared to the case with a positive P^+. The impact on the read-out is quite obvious. While the initial polarization (*0.77* and *–0.6 µC/cm²* for the positive and negative state respectively) was enough to give a current ratio of I_{on}/I_{off} of *1500,* due to the reduction of the polarization the current ratio can drop considerably. The issue is mainly with the negative P^-. While a reduction in the positive polarization P^+ by 70% still yields an I_{on}/I_{off} of *700,* the same reduction in the negative polarization P^- gives a current ratio of only *20*.
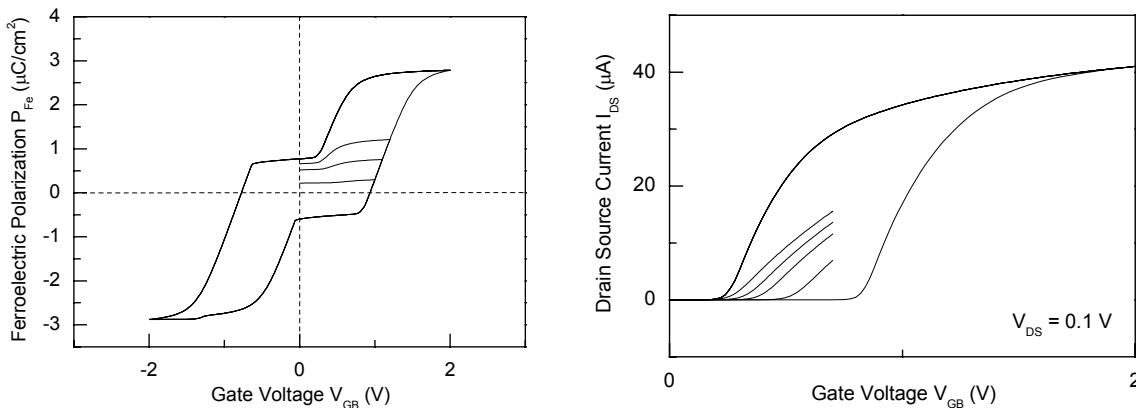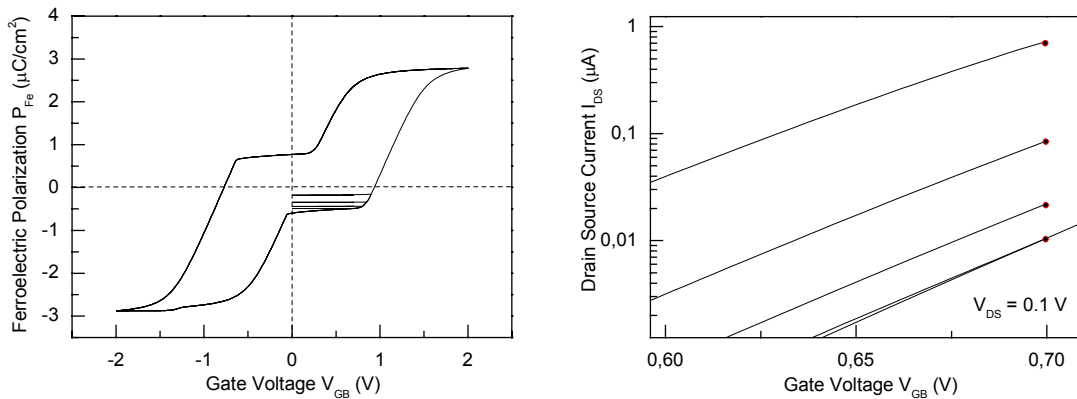

Fig. 4.17 Polarization reduction by driving sub-loops starting from P^- (left) and read out currents for the reduced polarizations (right) (read voltage is 0.7 V).

4.6 Summary

Among the failure mechanisms in the FeFET, the depolarization field is the primary reason for retention loss. At idle state, the depolarization field can cause a leakage current through the gate stack. By including leakage current models in the FeFET model, and assuming that carrier injection through the ferroelectric layer leads to charge recombination at the interface, the loss of polarization and associated retention times were calculated for different barrier heights. The loss of polarization can lead to read failure. This was examined for FeFETs polarized at "0" and "1" and it was shown that polarization loss at state "0" is more critical. Failure mechanisms will not be considered further in the rest of the thesis.

5 Scaling of the FeFET

5.1 Introduction

The trend towards always smaller and faster devices has become tradition in the semiconductor industry. It has been going on for more than 30 years since the introduction of the first integrated circuit. This has led to the doubling of the number of transistors per area every 18 to 24 months, now commonly referred to as "Moore's Law" (after Intel founder Gordon Moore), although not a real law per se but more of a trend.

By reducing the dimensions of the transistors, the building blocks of any logical circuit, it is possible to increase the processing speed and to decrease the energy dissipation per logic switch. This process is commonly known as scaling. Over the years, several approaches were proposed how to properly scale a MOSFET. There is, for example, *constant voltage scaling*, where the voltage remains unchanged, *constant field scaling,* where the field is kept constant, and *generalized scaling*, which deviates from the other two. For the scaling of the MOSFET the semiconductor industry annually publishes a set of specifications for the coming technology nodes for the forthcoming 15 years. This technology roadmap called ITRS (International Technology Roadmap for Semiconductors) is updated every year (minor update) with a major update every two years. Table 5.1 lists some transistor parameters after scaling through several process generations according to the ITRS [23].

year	Unit	2001	2004	2007	2010	2013	2016
DRAM Half Pitch	nm	**130**	**90**	**65**	**45**	**32**	**22**
Physical Gate Length	nm	65	37	25	18	13	9
Operating Voltage[**]	V	1.1	1.0	0.7	0.6	0.5	0.4
Drive Current	$\mu A/\mu m$ μA[***]	900 59	900 33	900 23	1200 22	1500 20	1500 14
Off Current	$\mu A/\mu m$ nA[***]	0.01 0.7	0.1 3.7	1 25	3 54	7 91	10 90
NOR Flash Read Current	μA	36-44	31-39	27-35	27-33	25-31	22-28
Oxide Thickness[*]	nm	1.3-1.6	0.9-1.4	0.6-1.1	0.5-0.8	0.4-0.6	0.4-0.5
Channel Doping	$10^{18}\,cm^{-3}$	4	11	23	50	130	150
Substrate Doping	$10^{18}\,cm^{-3}$	0.8-1.5	1.5-2.5	2.5-5	5-9	9-18	15-30
Channel Doping Depth	nm	21-30	15-21	10-15	7-10	5-7	3-5
Junction Doping Depth	nm	48-95	24-45	18-37	13-26	10-19	7-13

[*] is EOT (equivalent oxide thickness) to an SiO_2 dielectric [**] for high performance MOSFET
[***] adjusted for the physical gate length

Table 5.1 Parameters for current and future generation MOSFETs according to the ITRS 2002 [23].

So far it has been possible to hold on to the roadmap targets, but for the future there are doubts as dimensions reach critical values where quantum-mechanical phenomena apply. This will probably be the case sometime during the next decade. Until now, small dimensions have led to problems such as latch-up and leakage current. The first has been overcome by building trenches to isolate the devices (shallow trench isolation or STI) and as for the leakage current, one solution that is used is to deposit the silicon channel on an oxide (silicon on insulator or SOI) so that the conductivity paths towards the substrate are eliminated. The lateral dimensions are determined by the lithography. Already, some lithography methods make it possible to fabricate features with sizes in the nm region. More critical, however, seems to be the gate oxide problem that, assuming no substitute for the very convenient SiO_2 is found, will scale to sub-nm thicknesses with direct tunneling presenting a severe leakage problem. Already alternatives are being looked into with materials that have a higher dielectric constant (high-k) than that of SiO_2. This, however, is not the only criterion for the dielectric. Also of importance is a high band gap (SiO_2 has a band gap of *9.1 eV*), a high energy barrier to Si (*3.15 eV* for SiO_2), and, if possible, an amorphous structure, since it is known that amorphous oxides have better leakage current characteristics than crystalline ones. Substrate, source-drain and channel doping are also affected by scaling. Substrate and source-drain doping concentrations have to be increased with smaller device dimensions. For source-drain this means shallower p-n contacts. Channel doping is adjusted to control the threshold voltage. With decreasing sizes the gate voltage (or operating voltage) has to decrease to guarantee a constant field in the gate oxide, although sometimes the operating voltage is kept constant to stay compatible with the peripheral electronics.

This short introduction presented some of the issues with MOSFET scaling. In the following sections it will be investigated whether the scaling of the FeFET can follow the same scaling rules as MOSFET scaling or if new scaling rules are needed.

5.2 MOSFET scaling

The MOSFET is essentially a dielectric capacitor on top of a semi-conducting substrate, as shown in Fig. 5.1, which is being charged and discharged, thus modulating the field in the semi-conductor (thus the term field effect transistor or FET). That said, the field in the oxide is crucial for the performance of the device.

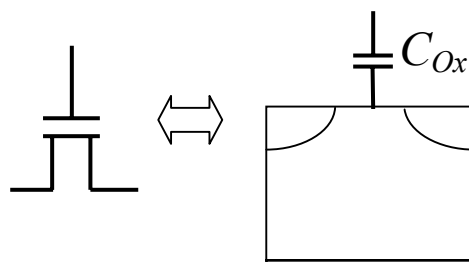*Fig. 5.1 The MOSFET as an oxide capacitor on top of a semi-conducting substrate.*

When scaling the MOSFET the first problem that is encountered by reducing the lateral dimensions (L,W) is a higher lateral field in the channel ($E_{Channel}$). To offset this field the voltage (V) is scaled, and this leads to a lower field in the oxide (E_{Ox}) that is offset by scaling the oxide thickness (d_{Ox}) as can be seen below:

$$L, W \cdot \frac{1}{k} \Rightarrow E_{Channel} = \frac{V}{L} \cdot k$$

$$\Downarrow$$

$$V \cdot \frac{1}{k} \Rightarrow E_{Channel} = const.\; ;\; E_{Ox} = \frac{V}{d_{Ox}} \cdot \frac{1}{k} \qquad (5.1\text{--}5.3)$$

$$\Downarrow$$

$$d_{Ox} \cdot \frac{1}{k} \Rightarrow E_{Ox} = const.\;, k > 1$$

This scaling approach is known as *constant field scaling*. But while in the MOSFET it is easy to keep the field in the oxide constant, by simply scaling voltage and oxide thickness by the same factor, things are not that simple in the FeFET. For this reason, the ferroelectric layer (a ferroelectric capacitor) is examined.

5.3 The ferroelectric layer capacitance

The ferroelectric layer can be modeled as a non-linear hysteretic capacitor in parallel with a linear dielectric capacitor as shown in Fig. 5.2. It should be mentioned that in the present analysis only the saturated loop of the hysteresis and not a complete hysteresis model is considered. Despite this, the deductions that will be made regarding the scaling of the capacitor are still valid.
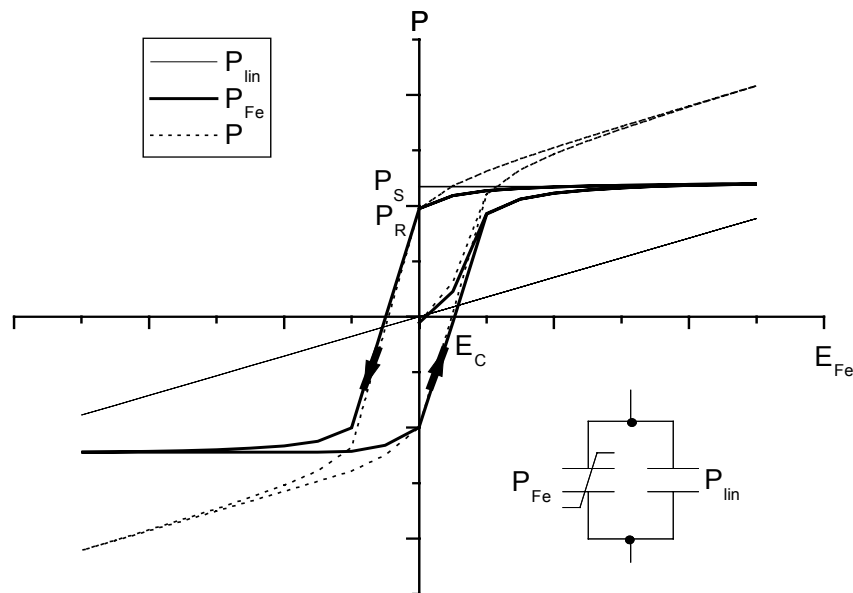

Fig. 5.2 Polarization in the ferroelectric as a sum of a linear and a nonlinear (hysteretic) capacitor.

The surface charge of the ferroelectric is given by

$$Q = (P_{Fe} + P_{lin}) \cdot A , \qquad (5.4)$$

where
$$P_{lin} = \varepsilon_0 \cdot \varepsilon_r \cdot E_{Fe} \tag{5.5}$$

is the linear part of the polarization, ε_r the dielectric constant of the linear part of the ferroelectric capacitor and E_{Fe} the field in the ferroelectric defined as

$$E_{Fe} \equiv \frac{V_{Fe}}{d_{Fe}}. \tag{5.6}$$

The nonlinear part P_{Fe}, shown in Fig. 5.2, can be described with [24]

$$P_{Fe} = \pm P_S \cdot \tanh\left(\frac{\pm E_{Fe} - E_C}{2 \cdot \delta}\right), \tag{5.7}$$

where
$$\delta = \frac{E_C}{\ln\left(\dfrac{1 + \dfrac{P_R}{P_S}}{1 - \dfrac{P_R}{P_S}}\right)} \tag{5.8}$$

is a constant, E_C the coercive field strength, P_R the remnant polarization and P_S the saturated polarization.

The positive sign in the expression for P_{Fe} is for the ascending and the negative for the descending hysteresis branch. In order to keep the surface charge Q (note: in the MOSFET it is the gate charge Q_G (Eq. 3.12) that defines the state of the channel) constant the field E_{Fe} has to remain constant as can be seen in Eqs. 5.4-5.8. Another combination that leads to a constant Q is to scale d_{Fe} and ε_r by factor $1/k$ and E_C by k ($k>1$). Then from Eqs. 5.4-5.8 it follows that:

$$
\begin{aligned}
E'_{Fe} &\to E_{Fe} \cdot k \\
P'_{lin} &\to P_{lin} \\
\delta' &\to \delta \cdot k \\
P'_{Fe} &\to P_{Fe} \\
Q' &\to Q
\end{aligned}
\tag{5.9-5.13}
$$

But modifying the dielectric permittivity (linear part) of the ferroelectric requires a new ferroelectric to be used. This is not a parameter that can be tweaked at will, since the choice of materials is limited.

The total capacitance of the ferroelectric C_{Fe} has a voltage dependency. It can be calculated by the following equation

$$C_{Fe} = C_{lin} + C_{nonlin} = C_{lin} + \frac{Q_{Fe}}{V_{Fe}} = \varepsilon_0 \cdot \varepsilon_r \cdot \frac{A}{d_{Fe}} + \frac{\pm P_S \cdot \tanh\left(\frac{\pm \frac{V_{Fe}}{d_{Fe}} - E_C}{2 \cdot \delta}\right) \cdot A}{V_{Fe}}, \quad (5.14)$$

where $A = L \cdot W$ (length · width) is the area of the gate capacitor. The C_{Fe} vs. V_{Fe} curve is shown in Fig. 5.3 and is a characteristic "butterfly" curve with the minimum value being C_{lin} and the maximum occurring at $E_{Fe} = E_C$. Equation 5.14 has a singularity at zero voltage. This is because Eq. 5.14 is only a mathematical approximation.

A constant field E_{Fe} in the ferroelectric can be achieved if both d_{Fe} and V_{Fe} are scaled by the same factor (see Eq. 5.6) and this leads, according to Eq. 5.14, to the scaling of C_{Fe} by the inverse factor:

$$d_{Fe}, V_{Fe} \cdot \frac{1}{k} \Rightarrow C_{Fe} \cdot k, \quad k > 1 \qquad (5.15)$$
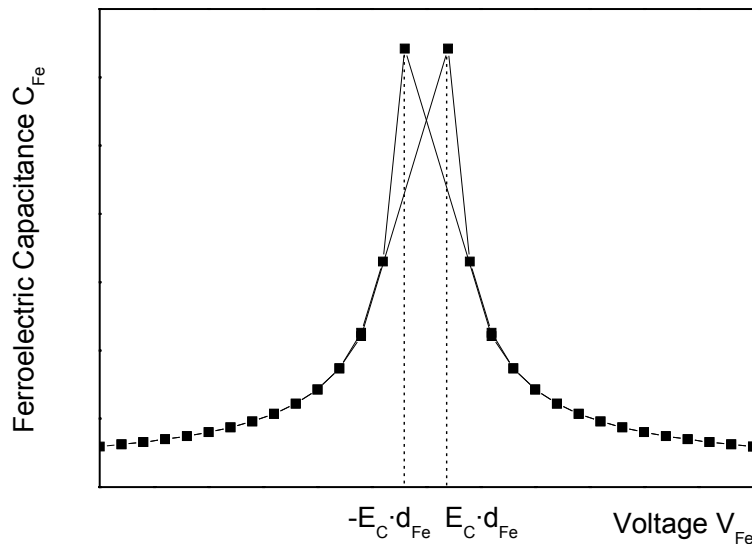
Fig. 5.3 Voltage dependency of the ferroelectric capacitance.

5.4 The FeFET gate stack

Scaling V_{Fe} in a ferroelectric capacitor is not a problem since this is the voltage applied. In the FeFET, however, V_{Fe} cannot be scaled linearly with the gate voltage, because the gate stack is a voltage divider that does not consist of only linear capacitors (not V_{Fe} is applied but V_{GB}). Figure 5.4 shows the different capacitances in the gate stack. The ferroelectric in series with the oxide layer capacitance C_{Ox} makes the total capacitance of the gate stack:

$$C_{Stack} = (C_{lin} + C_{nonlin}) \| C_{Ox} = C_{Fe} \| C_{Ox} \qquad (5.16)$$

Both capacitors are assumed to have the same area.
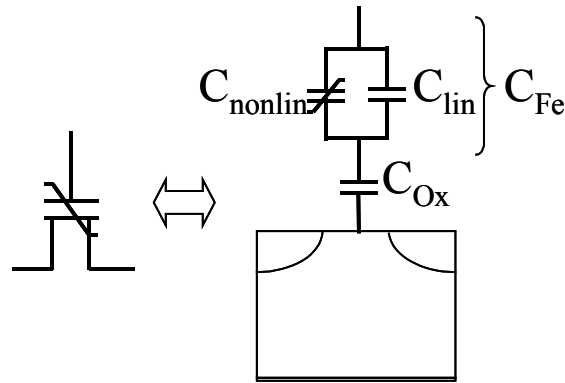
Fig. 5.4 Capacitances in the FeFET gate stack.

When adjusting the thickness of the layers, attention must be paid not to exceed critical field values in the oxide and the ferroelectric. The fields in the oxide and the ferroelectric are as follows:

$$E_{Ox} = \frac{\dfrac{V}{\dfrac{\varepsilon_{Ox} \cdot d_{Fe}}{\varepsilon_{Fe} \cdot d_{Ox}} + 1}}{d_{Ox}} < \frac{V}{t_{Ox}} \quad , \quad E_{Fe} = \frac{\dfrac{V}{\dfrac{\varepsilon_{Fe} \cdot d_{Ox}}{\varepsilon_{Ox} \cdot d_{Fe}} + 1}}{d_{Fe}} \qquad (5.17), (5.18)$$

where V is the gate voltage and ε_{Fe}[13] the effective ferroelectric permittivity. The voltage drop in the silicon depletion region ($0 \leq V_{Si} \leq V_{Si,max} = 2 \cdot \dfrac{k \cdot T}{q} \cdot \ln \dfrac{N_{SUB}}{n_i}$, *for V>0*) was ignored and the flatband voltage assumed 0 for simplicity. Because the applied voltage is across the entire gate stack there is less threat of dielectric breakdown than in the MOSFET. The maximum field in the ferroelectric should be high enough to switch enough dipoles in order to yield a large "memory window", defined as the difference between the two voltage thresholds (Fig. 5.5).

---

[13] ε_{Fe} represents both the linear and non-linear ferroelectric part and is voltage dependent, in contrast to ε_r which represents only the linear part and is constant.

Fig. 5.5 I_{DS} vs. V_{GB} (I-V) curves of FeFETs with different memory windows.

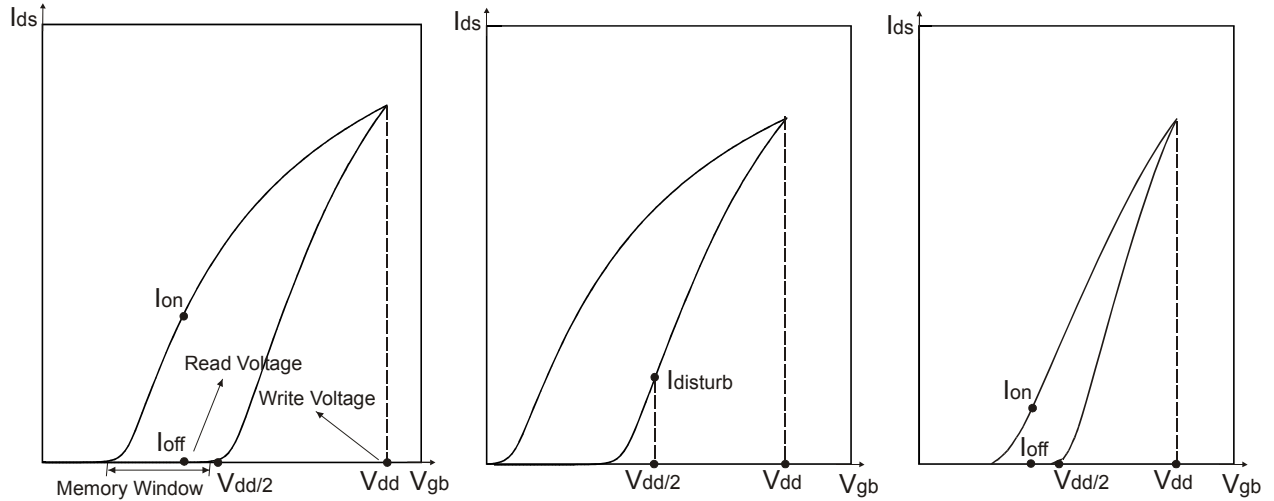From Eqs. 5.17 and 5.18 it is apparent that if both thicknesses are scaled with one factor the fields are scaled by the inverse factor. Since it is important that the higher voltage drop be across the ferroelectric and because of the fact that most ferroelectrics are of perovkite structure with a high dielectric permittivity, the ferroelectric must have a higher thickness than the dielectric. Ferroelectrics with a high remnant polarization P_R cannot be driven to saturation using low voltage operation and must be polarized along a hysteresis sub-loop. This is the case e.g. with PZT or SBT (see section 5.10 for the determination of an ideal low-P_R ferroelectric).

Besides the scaling of the write voltage, which is equal to the operating voltage V_{DD}, the read and erase voltages also need to be scaled. The read voltage is chosen such that the ratio I_{on} / I_{off} is maximum (see Fig. 5.5). The erase-voltage that can be used to reverse the polarization is $V_{Erase} = -V_{DD}$.

5.5 The FeFET memory window

The maximum value for the memory window, for a saturated hysteresis, is $\Delta U_{max} \approx 2 \cdot E_C \cdot d_{Fe}$ [8]. According to this equation the memory window is proportional to the ferroelectric thickness d_{Fe} and the coercive field E_C. Thus, a thicker ferroelectric or a larger sub-loop will yield a larger memory window. A larger E_C[14] does only lead to a larger memory window if the ferroelectric is saturated. Otherwise, with the same applied voltage, a smaller sub-loop is driven and the memory window is smaller. Figure 5.5 shows some typical *I-V* curves. To obtain these plots the drain voltage is kept constant and the gate voltage is swept between V_{DD} (the write-voltage) and -V_{DD} (the erase-voltage). The current at V_{Read} is equal to either I_{on} or I_{off} depending on the state of the device (actually, according to Fig. 4.13, the read current is at state F and not D as shown in Fig. 5.5, but the I_{on}/I_{off} ratio is not much different, so it is adequate to study only the saturated loop). The drain voltage can be set to V_{DD} or lower.

Besides the memory window size that is important for a high I_{on}/I_{off} ratio, the position is also important. As shown in Fig. 5.5, an *I-V* with an equally large memory window positioned too far to the left will give a high disturbance and change the polarization of the FeFET after

---

[14] In this thesis E_C refers to the coercive field of the saturated hysteresis, which is a constant material parameter. The symbols E^+ or E^- can be used for the unsaturated hysteresis.

applying the disturbance voltage $V_{DD}/2$. In the right curve of Fig. 5.5, an *I-V* curve with a small memory window gives a small I_{on}/I_{off} ratio, although not susceptible to disturbance. An exception to this last case (small memory window) is when the ferroelectric is polarized to saturation (remember it must have a low P_R), because then the current curves have a higher transconductance (see Fig. 5.17).

The FeFET model will be used to simulate the effect of scaling, after modification of various parameters, on the memory window and the device operation in general. The curves that will be presented are the devices' *I-V* curves (source-drain current vs. gate voltage) that are characteristic for a transistor device (and in our case the FeFET) and reveal several properties such as threshold voltage, transconductance, operating current, leakage current and the I_{on}/I_{off} ratio.

5.6 Specifications and requirements for the FeFET

For the scaling simulations the material properties were assumed fixed (for the FeFET PZT was chosen as the ferroelectric and CeO_2 as the gate oxide dielectric – see Table 5.2 for the material parameters). Although, mathematically, both the dielectric constant and the oxide thickness are freedom factors in the scaling of the dielectric, it is easier to make a thinner oxide than to find a new material with a higher permittivity and otherwise similar properties. After all, SiO_2 is scheduled to be replaced as a gate oxide by a high-k dielectric for the first time around the *45 nm* generation.

Table 5.1 listed some of the parameters that are modified during the scaling of the MOSFET. All of the transistor parameters specified in that roadmap exist in the FeFET too. For simplicity only the oxide and ferroelectric thickness, the operating voltage and the substrate doping will be considered here. They will be adjusted to achieve the desired operation of the FeFET. In the case of the substrate doping a uniform concentration will be assumed. For the rest of the BSIM parameters the default values (level 49) are used.

Two cases will be considered based on the voltages used and this will lead to two scaling approaches. In the first, both the write and the read voltage will be kept constant during scaling. In the second, the write voltage will be the operating voltage specified in the current ITRS (2002 update) and the read voltage can be freely chosen. Starting with the *130 nm* technology node (referred to in the roadmap as half pitch length but here as channel length), the voltages used will be ≤ *1.1 V* (operating voltage at *130 nm* for a high performance MOSFET). Using such low voltages makes polarizing the ferroelectric hard, because of the high coercive field of PZT and the voltage drop across the dielectric. According to the ITRS, the specification for the read current I_{Read} for NOR Flash memory from the *130 nm* to the 22 *nm* technology node is between *22* and *44 μA*. The requirements set here for the FeFET will be more moderate: $I_{on} > 10\mu A, I_{on}/I_{off} > 1000$.

5.7 Constant gate stack scaling of the FeFET

When the write voltage is kept constant and the gate stack is not modified (see Fig. 5.6), the ferroelectric is polarized to the same degree, because the voltages (electric fields) applied to the capacitors in the gate stack also remain constant. The only parameter that is left for adjustment is the substrate doping, that is increased to account for small size effects.
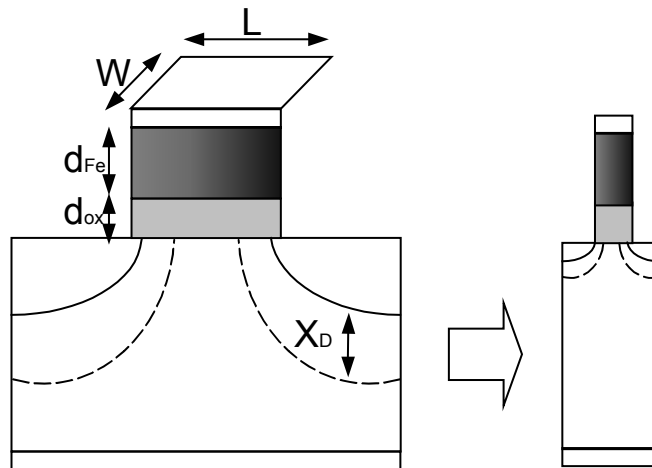
Fig. 5.6 Scaling of the ferroelectric field effect transistor using constant gate stack (X_D is the maximum depletion depth).

This is done in Fig. 5.7, that shows the simulated *I-V* curves of the scaled FeFETs. It can be seen that the memory window remains almost constant and that is because the polarization hysteresis loop does not change much. For constant field scaling, the drain voltage should also be scaled to keep the lateral field in the channel constant.
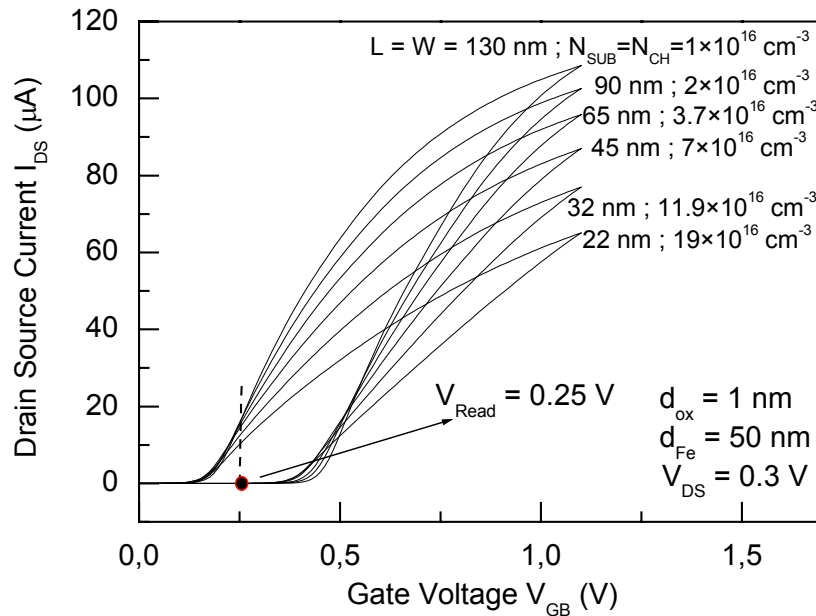
Fig. 5.7 I-V curves for the FeFET devices scaled using constant gate stack scaling.

Table 5.2 shows the parameters of the scaled FeFETs. The I_{on}/I_{off} ratio drops considerably, so that in the last technology node the target of having a read-current $I_{on} > 10 \, \mu A$ and a current ratio of more than 1000 could not be met. One solution is to make the specifications more flexible and allow for a higher operating voltage, so that a higher degree of polarization can

be reached, which translates to a larger memory window and finally a higher I_{on}/I_{off} ratio. Starting with a *130 nm* FeFET, the layers' thicknesses with $d_{Fe} = 50\ nm$ and $d_{Ox} = 1\ nm$ were chosen after parameter sweeping, for the ferroelectric between *20* and *200 nm* and the dielectric between *1* and *5 nm*, to satisfy the conditions of a maximum memory window, a maximum "on/off " current ratio and a minimum leakage current[15] (at the positive polarization state P^+). One downside to this approach is that scaling makes the device incompatible to core voltage logic of the same technology node. Thus, if the *130 nm* FeFET can be driven from the output of a logic device of the same technology, the *90 nm* FeFET, which requires *1.1 V,* is incompatible to *90 nm* CMOS *(1 V)*. Overall, provided that the voltages used are not a problem for memory design, this scaling approach is the easiest, because it only requires modifying one parameter (besides the lateral dimensions) and most importantly, no scaling in the vertical direction.

Parameter	Description	Values					
$L = W$ *(nm)*	Channel Length, Width	130	90	65	45	32	22
$N_{SUB} = N_{CH}$ *(cm⁻³)*	Doping Concentration	$1 \cdot 10^{16}$	$2 \cdot 10^{16}$	$3.7 \cdot 10^{16}$	$7 \cdot 10^{16}$	$1.2 \cdot 10^{17}$	$1.9 \cdot 10^{17}$
d_{ox} *(nm)*	Oxide Thickness	1					
d_{Fe} *(nm)*	Ferroelectric Thickness	50					
V_{DD} *(V)*	Operating Voltage	1.1					
V_{Read} *(V)*	Read Voltage	0.25					
V_{DS} *(V)*	Drain Source Voltage	0.3					
ε_{Ox}	Oxide Permittivity	12					
ε_r	Linear Ferroelectric Permittivity	250					
E_C *(kV/cm)*	Coercive Field	70					
P_S *(µC/cm²)*	Saturated Polarization	40					
V_{FB} *(V)*	Flatband Voltage	-0.4					
P_R *(µC/cm²)*	Remnant Polarization	32					
I_{on} *(µA)*	Read Current	16	16	16	15	13	10
I_{on} / I_{off}	Current Ratio	16000	8000	4000	1900	1100	700
I_{leak} *(nA)*	Leakage Current at "on-state"	1.3	1.9	2.5	3.6	3.9	3.8

(Input Parameters / Output Parameters — bracketed at right)

Table 5.2 Parameters for the FeFETs scaled with constant gate stack scaling.

5.8 Variable gate stack scaling of the FeFET

In this approach the specifications of the ITRS for the operating voltage will be used, so that the devices are compatible with the respective MOSFETs in each technology node. The use of

---

[15] Leakage current in this chapter refers to the off-state source drain current, unlike in chapter 4 where it refers to the current through the gate stack.

lower voltages in each scaling step means the amount of polarization becomes less with each step, so the gate stack must be modified to allow for a larger voltage drop across the ferroelectric. The non-linearity of the gate stack was examined in section 5.3. In order to choose the right parameters, the dependence of the FeFET functionality on each of them is examined by simulating the *I-V* curves of a FeFET. Without loss of generality, we will use the *130 nm* FeFET as a reference.

Figure 5.8 shows the dependence on the substrate doping concentration ($N_{SUB} = N_{CH}$). A higher doping shifts the *I-V* curve to the right and yields a smaller memory window. A lower doping yields a larger memory window at the expense of lower threshold voltages and thus a higher leakage current.
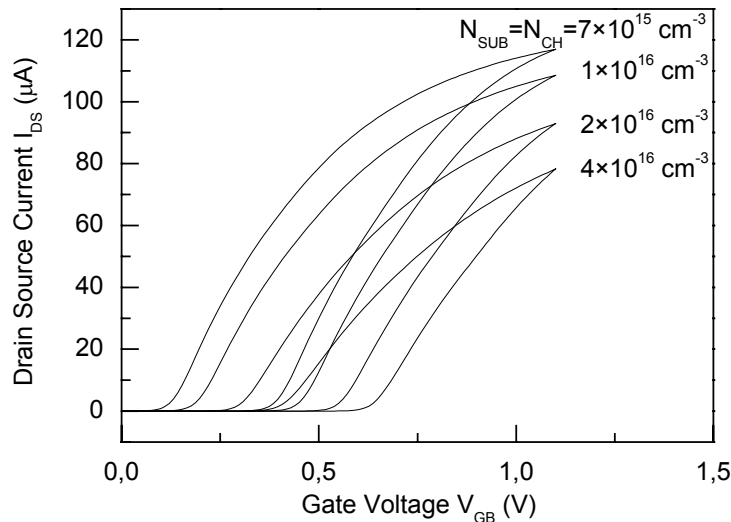
Fig. 5.8 I_{DS} vs. V_{GB} curves for the 130 nm FeFET of Table 5.2 for different doping concentrations.

Figure 5.9 shows that the oxide thickness has a similar effect as the doping concentration. A higher thickness reduces the oxide capacitance and causes a higher voltage drop across the oxide and a smaller across the ferroelectric, hence the smaller hysteresis.

Fig. 5.9 I_{DS} vs. V_{GB} curves for the 130 nm FeFET of Table 5.2 for different oxide thicknesses.

The ferroelectric thickness has a different effect. For a certain value a maximum memory window is reached. Beyond that value the memory window shrinks again, as shown in Fig. 5.10 (left). This was also described in [8]. Also, scaling the layer thicknesses by keeping the ratio constant will not yield acceptable results (Fig. 5.10 right) since the voltage divider is non-linear.

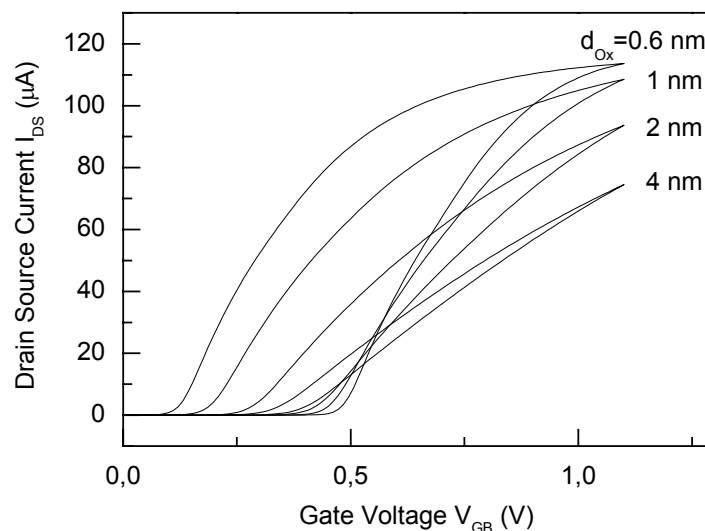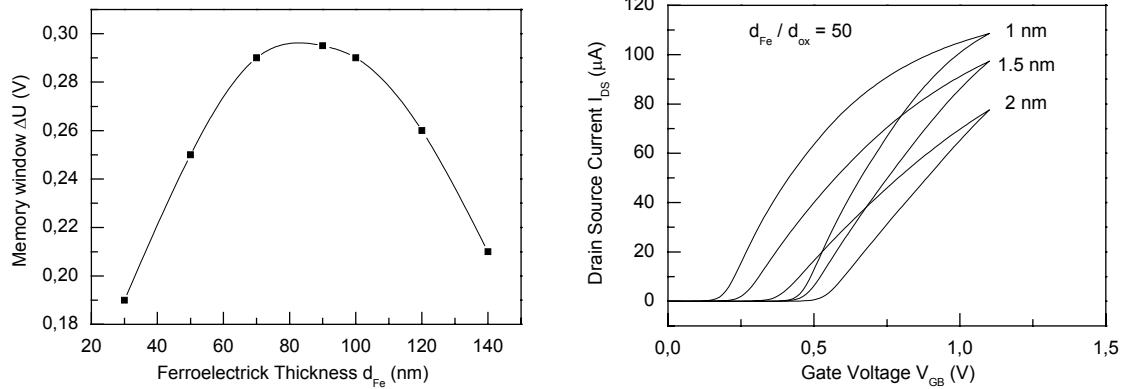*Fig. 5.10 Dependency of the memory window on the ferroelectric layer thickness (left), I_{DS} vs. V_{GB} curves for the 130 nm FeFET of Table 5.2 for different oxide and ferroelectric thicknesses (the thickness ratio is kept constant) (right).*

From the studied parameter dependence and after the necessary parameter sweeping, a set of parameters was chosen to get the *I-V* curves of Fig. 5.11. The FeFET parameters are listed in Table 5.3.

Fig. 5.11 I_{DS} vs. V_{GB} curves for the FeFET devices scaled using variable gate stack scaling.

The *130 nm* FeFET is the same as in last section. The different I_{on} listed in Table 5.3 for the *130 nm* FeFET was calculated at a higher read voltage and shows that a higher read current is

at the expense of a lower I_{on}/I_{off} ratio and vice versa. In the next scaling step the operating voltage was reduced from *1.1 V* to *1 V*. To achieve a similar polarization curve the ferroelectric thickne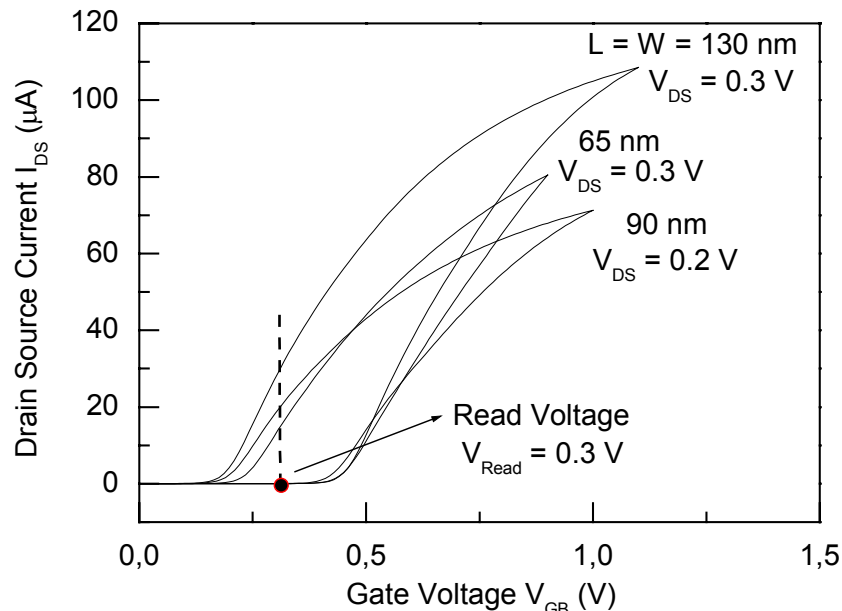ss was increased to *60 nm* and the substrate doping was doubled. Also, the drain voltage, applied during reading, was reduced to *0.2 V*. The read-voltage was chosen such that the current ratio condition was satisfied. In the *65 nm* technology node the operating voltage of *0.7 V* was not big enough to achieve a high enough polarization. Figure 5.12 shows the polarization hysteresis loops (actually sub-loops) for the three calculated scaling steps. The constant polarization plateau in the middle extends to about *1 V* (from about *–0.5 V* to *0.5 V*). This makes the device inoperable for very low voltages.

Parameter	Description	Values		
$L = W$ *(nm)*	Channel Length, Width	130	90	65
d_{ox} *(nm)*	Oxide Thickness	1	1	0.6
d_{Fe} *(nm)*	Ferroelectric Thickness	50	60	60
V_{DD} *(V)*	Operating Voltage	1.1	1.0	0.9
V_{Read} *(V)*	Read Voltage	0.32	0.3	0.31
V_{DS} *(V)*	Drain Source Voltage	0.3	0.2	0.3
N_{SUB}, N_{CH} *(cm^{-3})*	Doping Concentration	$1 \cdot 10^{16}$	$2 \cdot 10^{16}$	$4 \cdot 10^{16}$
I_{on} *(µA)*	Read Current	32	18	16
I_{on} / I_{off}	Current Ratio	2900	1000	1200
I_{leak} *(nA)*	Leakage Current at "on-state"	1.3	1	0.2

(Input Parameters / Output Parameters)

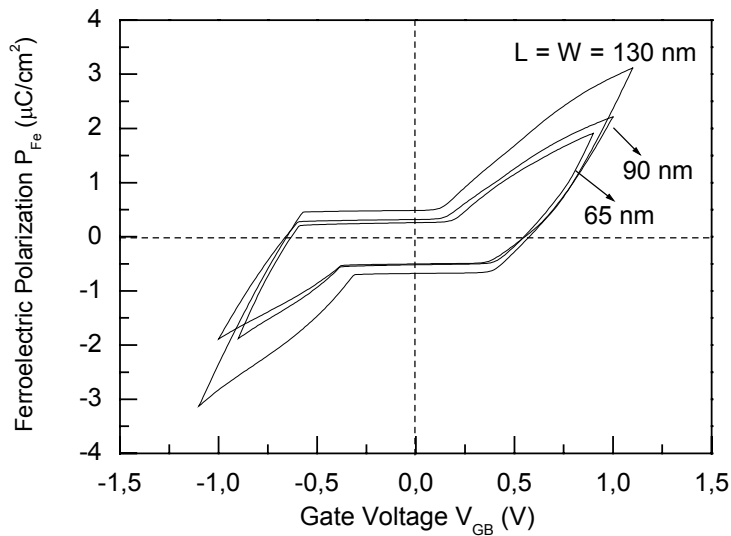*Table 5.3 Parameters for the FeFETs scaled with variable gate stack scaling.*

Fig. 5.12 Polarization vs. Gate Voltage (P-V) curves for the FeFETs scaled using variable gate stack scaling.

The lowest voltage for which the FeFET can achieve acceptable operation, under the requirements set here, is *0.9 V*. This still demands a very thin dielectric layer of *0.6 nm*. This should probably be an EOT (equivalent oxide thickness) of a high-k dielectric (compared to CeO_2). In this last calculation the operating voltage differs from the one specified in the roadmap for this technology node. This particular requirement in this approach means modifying the gate stack, which makes this a multi-parameter problem compared to the easier approach of last section. For all the complexity the gains are not that big since only one scaling step could be calculated. Figure 5.13 compares the two approaches based on the I_{on}/I_{off} ratio and clearly the first one can potentially lead to very small functional devices.
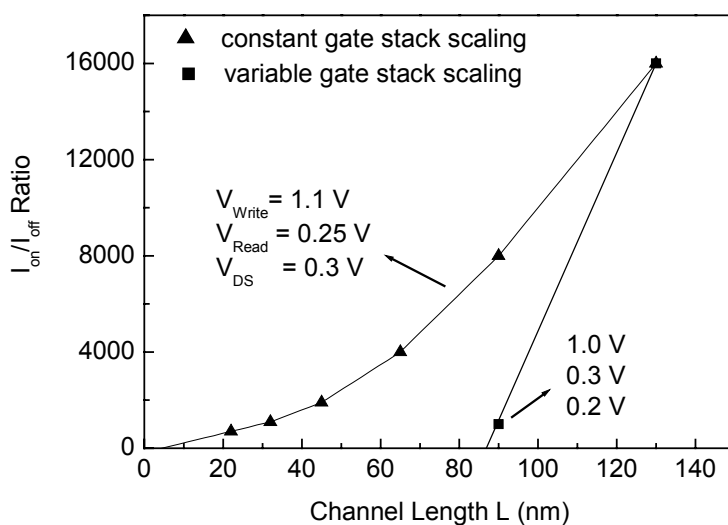

Fig. 5.13 I_{on}/I_{off} ratio of the scaled FeFETs for the two scaling approaches.

In Fig. 5.14 the currents calculated with the *constant gate stack scaling* approach are displayed.
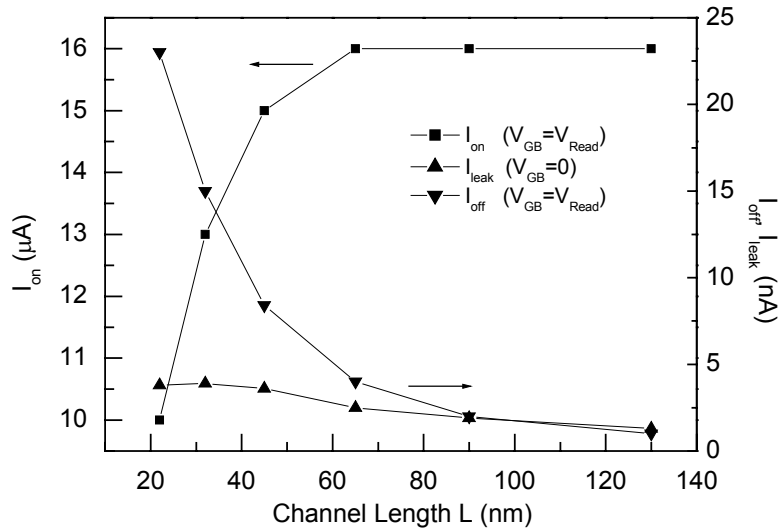

Fig. 5.14 I_{on}, I_{off} and I_{leak} (I_{DS} at P^+, $V_{GB} = 0$ V) *currents for the FeFET scaled with constant gate stack scaling.*

5.9 Finite size effects in ferroelectrics

It is known that ferroelectrics exhibit finite size effects, when the thickness is comparable to the correlation length [25]. They also have generally a lower permittivity than in bulk form, and the coercive field is thickness-dependent too [81]. Further, below a certain film thickness it is believed that the ferroelectric state cannot be sustained and vanishes [12]. It is thus important to know if this is also an issue in FeFETs. In section 5.8 a variable gate stack method was used to scale the FeFET for use with a smaller operating voltage. A thinner ferroelectric enables lower voltage operation because the of the memory window dependence on the coercive field strength and the ferroelectric thickness:

$$\Delta U_{max} \approx 2 \cdot E_{Fe} \cdot d_{Fe} \tag{5.19}$$

Figure 5.10 (left) showed the dependence of the memory window on the ferroelectric layer thickness. Figure 5.15 shows how the *I-V* curves of a FeFET scale from *5 V* operation down to *4 V* and *3 V* by thinning only the thickness of the ferroelectric.
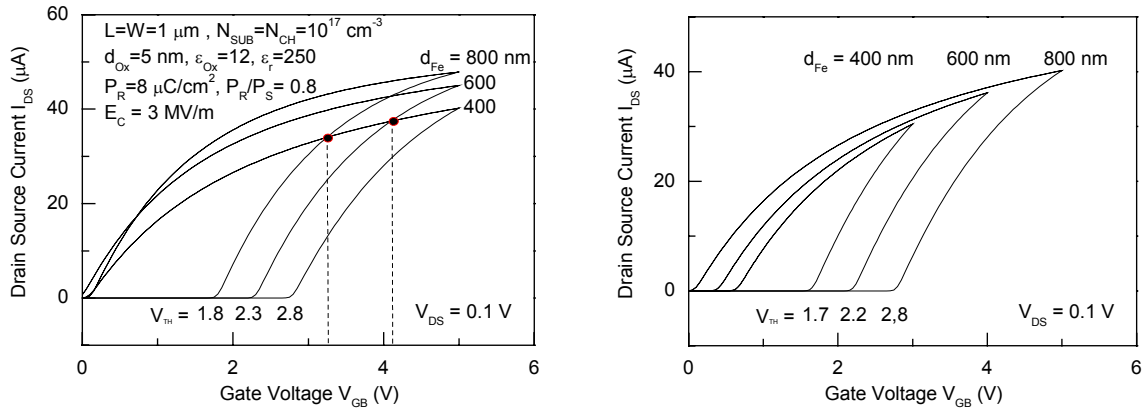
Fig. 5.15 Shifting the threshold voltage of the FeFET by only scaling the thickness of the ferroelectric (the ferroelectric properties are those of SBT).

First, a FeFET with a wide memory window is designed for *5 V* operation ($V_{TH} > V_{DD}/2$). Then the ferroelectric is made thinner and the thickness, for which the two *I-V* curves intersect at ~ *4 V,* is chosen for the FeFET operating with *4 V*. The same procedure is applied for the FeFET operating with *3 V*. This can continue until the memory window becomes too small for a functional FeFET (see Fig. 5.15 right).

Scaling only the ferroelectric is one way (albeit, not a very effective one, as the specifications of section 5.6 cannot be met) to scale the FeFET and the question of size effects comes to mind. Moreover, as seen in section 5.3, this is how a ferroelectric capacitor is scaled. However, the film thickness determined in section 5.8 is at a minimum *50 nm* (see Table 5.3) and nowhere near *10 nm* or below, where a super-paraelectric limit is believed to exist. The lateral size, however, does get to smaller sizes (< *50 nm*). According to [26] BaTiO$_3$ exhibits ferroelectric properties down to a layer thickness of *2.4 nm* (~6 unit cells). For lower thicknesses, ferroelectricity disappears, because of the depolarization field. In [70] PbTiO$_3$ grains of *20 nm* lateral size were measured with three-dimensional piezoresponse force microscopy (PFM) and found to be ferroelectric.
Should the ferroelectric material parameters change below a certain size, they should be considered in the scaling approach.

5.10 The ideal ferroelectric

As already mentioned, high-P_R ferroelectrics should be polarized along a hysteresis sub-loop. The reason is that breakdown in the underlying oxide will occur, if it is attempted to polarize them to saturation. Additionally, sub-loops are believed to be generally unstable, and the polarization diminishes, if it is not previously saturated.
In this section, the parameters of an ideal (fictitious) ferroelectric are sought, that can be fully polarized (saturation) and yield an acceptable memory window comparable with that of a high P_R ferroelectric. The maximum value for P_R, is examined which, in combination with a coercive field E_C, makes it possible to reach polarization saturation without causing breakdown in the dielectric. The linear dielectric constant will be assumed *250*.

Regarding the ferroelectric parameters the following should be noted:

The P_R/P_S ratio is important for a high I_{on}/I_{off} ratio. A more tetragonal hysteresis gives a higher I_{on}/I_{off}. P_R should also not be too small, as in that case its effect on the charge in the channel would be minimal and it would yield a small memory window as a result. From the simulation of Fig. 2.9, a value of $\sim 2 \ \mu C/cm^2$ should be considered a minimum for P_S.

The maximum value for the memory window is $\Delta U_{max} = 2 \ E_C \cdot d_{Fe}$. Although a high memory window is required, E_C should not be too high as that would require higher voltages to switch polarization. It shouldn't also be too low, because then the memory window would be too small. As an example, to obtain a *1 V* memory window for a ferroelectric with thickness d_{Fe} =*200 nm* driven to saturation, E_C should be *2.5 MV/m*.

The voltage drop across the capacitors in the gate stack is:

$$V_{Ox} = V_{GB} \cdot \frac{C_{Fe}}{C_{Fe} + C_{Ox}} \quad and \quad V_{Fe} = V_{GB} \cdot \frac{C_{Ox}}{C_{Fe} + C_{Ox}} \qquad (5.20), (5.21)$$

As an oxide CeO_2 (ε_r =*12* [27]) is assumed. The target is to have at short circuit conditions (V_{GB} = *0 V*) the same remnant polarization as with the high-P_R ferroelectric. From Fig. 5.16 (left) P_R and E_C are chosen. Due to the asymmetry in the sub-loop hysteresis P^- and E^- can be chosen. Last, P_S is chosen such that the P_R/P_S ratio is high, so saturation can be reached with low voltages.

Figure 5.16 shows the results of the simulation embedded with the high-P_R simulation. The hysteresis is more quadratic with a higher P_R/P_S ratio of *0.93*. The coercive field (where the hysteresis intersects with the x-axis is lower than with a high P_R but as can be seen in Fig. 5.17 the memory window is comparable. This is because the high-P_R material is not fully polarized. Had this been the case, the memory window would have been

$$\Delta U = 2E_C \cdot d_{Fe} = 2 \cdot 100 \ nm \cdot 7 \ MV/m = 1.4 \ V, \qquad (5.22)$$

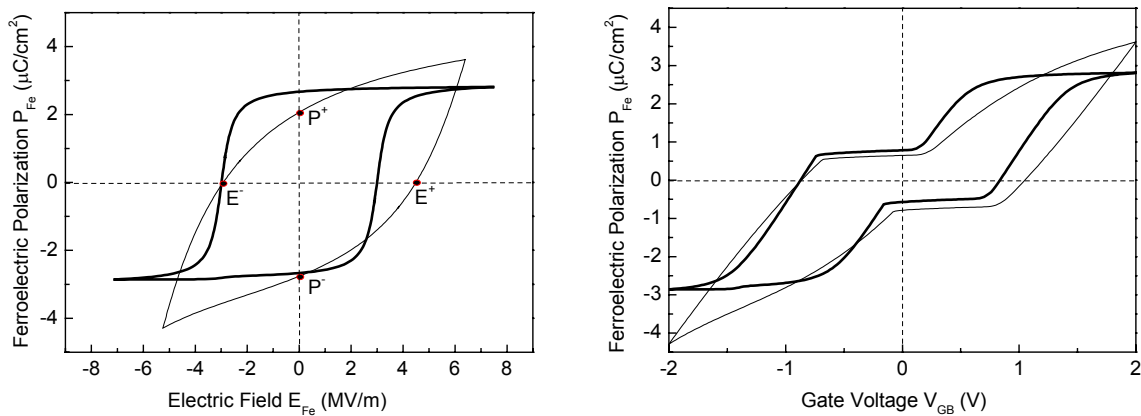which is too high for *2 V* operation and would require a higher voltage to operate.

Fig. 5.16 P-E_{Fe} (left) and P-V (right) curves for a ferroelectric with a low P_R (thick line) compared with those for a high-P_R ferroelectric (thin line).

The material parameters for the simulations are shown in Fig. 5.17. The two materials are used for most simulations in this thesis. One ferroelectric with such low polarization values as these calculated in this section is $YMnO_3$ [45].
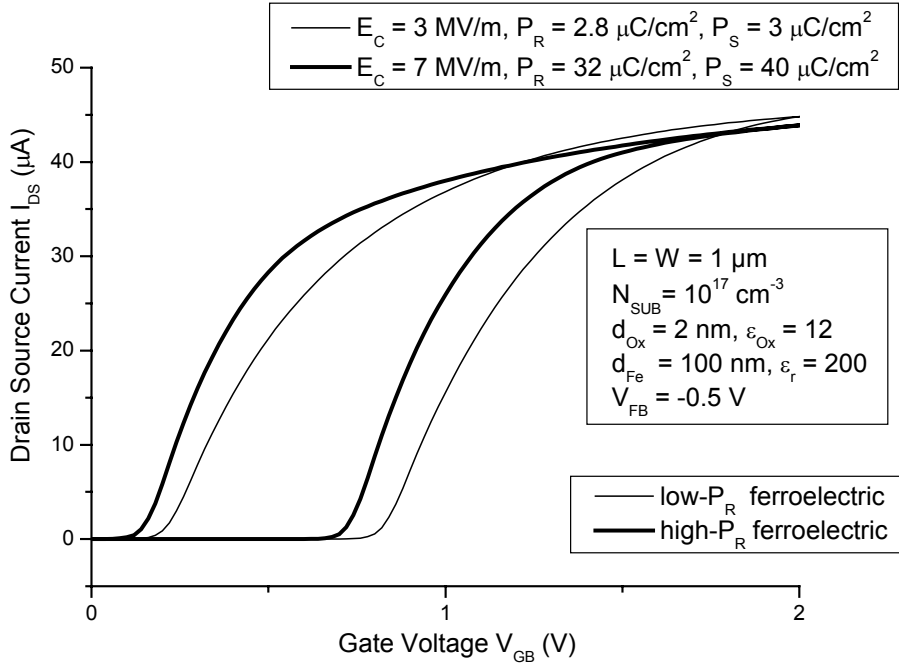

Fig. 5.17 I-V curves for a ferroelectric with a low P_R compared with those for PZT, both having an equally large memory window.

5.11 Multilevel Cells

So far in this chapter the scaling of the FeFET was studied with the aim of a higher integration that would lead to a higher bit density per area. Another way to accomplish this, is by using multilevel cells to store more than 1 bit per FeFET. The same principle has been in use in Flash for some time [28], and it is conceivable that it could be applied to the FeFET as well. Storing *2 bits/cell* effectively doubles the storage density, with minimum technology overhead.

The change of state in the FeFET can be seen as a shift of the threshold voltage. To define four different states *(2^2 = 4)*, four threshold voltages are needed. This is easily achieved using sub-loops as shown in Fig. 5.18. In order to write the information, one of three voltages would be applied (the fourth state is the erased cell). Four different states are defined by applying voltages *V_1, V_2, V_3* and *$-V_1$* (to reverse the polarization) setting the threshold voltage to *V_{T1}, V_{T2}, V_{T3}* and *V_{T4}* respectively.
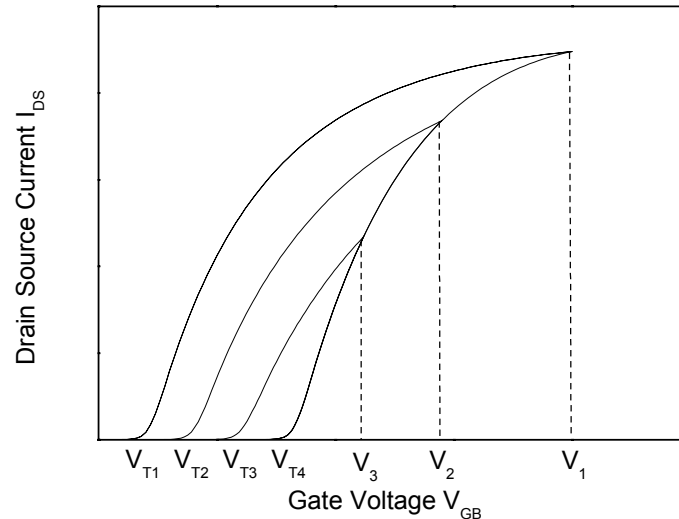
Fig. 5.18 I-V curves with sub-loops that define the different states in a multi level FeFET cell.

The reading can be done in two ways:

The first is more time consuming as it requires testing (three comparisons) at each read voltage (between two thresholds) if the FeFET is conducting, starting from the lowest. When the condition is satisfied and the device is "On" the test is finished. Otherwise, the device is "Off".

The second is to apply the highest of the read voltages and infer from the current the state. The read voltages should be chosen such, that the current is proportional to the amount of the polarization. This would make the detection easier. The current would then be compared with three references (if not equal the fourth state is the one) using three sense amplifiers and the necessary logic circuitry.

Ideally a FeFET with a high transconductance ($g_m \equiv \dfrac{\partial I_{DS}}{\partial V_{GB}}$) is best suited for a multi level cell as it yields higher I_{on}/I_{off} ratios at the different states.

Of course multilevel cells work against the scaling trend, because a smaller device that operates at a lower voltage has a smaller memory window, and the storage of multiple bits is not without a smaller noise margin. Storing more than *2 bits/cell* requires even more thresholds (2^n for *n bits/cell*) and this means more sub-loops. This can go on as long as the memory window and the I_{on}/I_{off} ratio allow. Because of the instability of sub-loops and as long as a way around this is not found, the implementation of multilevel cells in the FeFET will remain only of theoretical value.

5.12 Summary

In this chapter, two different scaling approaches were suggested for the FeFET, based on the operating voltage used. The first approach keeps the vertical field constant by leaving the gate stack unchanged, uses a constant operating voltage and requires only to modify the substrate doping concentration. In the second approach, the operating voltage is reduced to that specified in the ITRS and thus the gate stack has to be modified. The second approach is more complex, because more parameters must be modified than in the first. More importantly,

because of the reduction of the operating voltage, this approach cannot be extended beyond a few technology nodes. It should be noted that the simulated devices were scaled to meet the restrictive specifications set, and the parameters calculated should be regarded in that context. Next, the desired parameters of a hypothetical ferroelectric were discussed and it was shown that it should ideally have a low remnant polarization of about 3 $\mu C/cm^2$, a quadratic hysteresis (high P_R/P_S ratio) and a low coercive field of about *3 MV/m* for low voltage operation. Finally, FeFETs that can store more than one bit were discussed, as an alternative to physical scaling. Their feasibility is dependent on the stability of hysteresis sub-loops.

6 Reducing the depolarization field

The reason for the depolarization field, as already mentioned, lies in the incomplete charge screening of the ferroelectric surface charge. This is believed to lead to a reduction of the remnant polarization and to be responsible for the low retention times. In this chapter, two possibilities will be suggested to reduce the depolarization field. The first is through the increase of the substrate doping concentration, and the second through the increase of the oxide area compared to the area of the ferroelectric.

6.1 High substrate doping

Figure 5.8 in the last chapter showed the dependence of the *I-V* curve on the substrate doping concentration. A higher concentration shifted the curve to the right and led to a smaller memory window, because of the smaller voltage drop across the ferroelectric. However, only values smaller than 10^{18} cm^{-3} were investigated. Figure 6.1 (left) shows the effect of higher doping concentrations on the polarization curve, and Fig. 6.1 (right) on the FeFET *I-V* curves (see also Fig. 5.8).
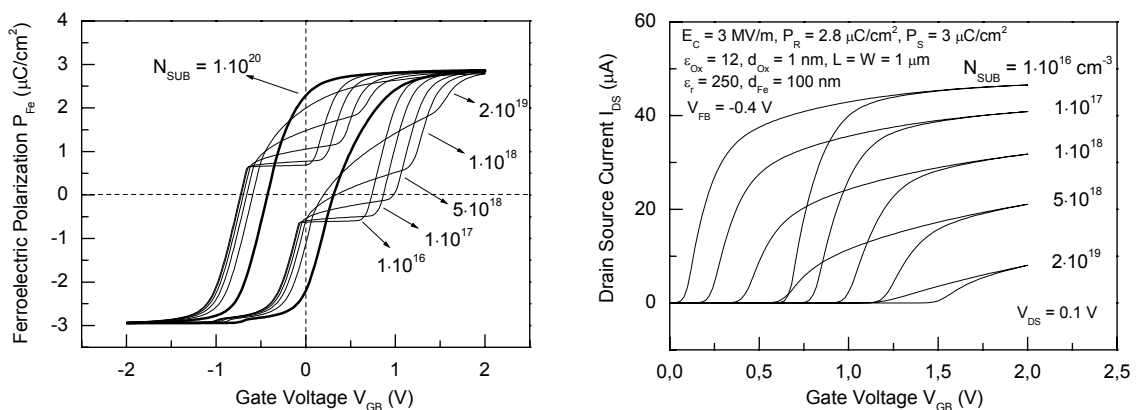

Fig. 6.1 The effect of the doping concentration on the (left) polarization hysteresis, (right) the FeFET I-V curves.

While the low doped FeFETs show a *P-V* curve with a plateau, with increased doping concentration it turns into a slope, increasing at the same time the P^+ value until, for a concentration of 10^{20} cm^{-3}, the *P-V* curve is identical to the P-V_{Fe} curve. The depolarization field at this point is close to zero. This means that the voltage drop across the dielectric is almost zero and the gate voltage is applied across the ferroelectric layer alone. This is not a desired condition as the MOSFET underneath the ferroelectric is not controlled. Smaller doping concentrations appear more reasonable, as the resulting FeFET functionality is not suppressed. Figure 6.2 shows how the depolarization field and the remnant polarization (they are interdependent) depend on the substrate doping.
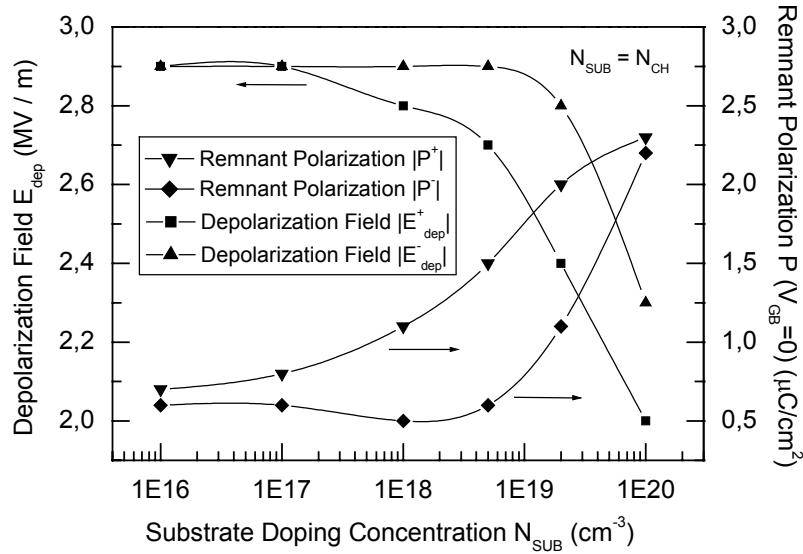
Fig. 6.2 Depolarization field and remnant polarization vs. the substrate doping concentration.

It is possible that a higher doping concentration could result in a better charge screening of the polarization surface charge and subsequently to a reduction in the depolarization field. However, what the simulation here shows is that the voltage drop across the oxide is reduced to very low levels (less than *0.1 V*) for a depolarization field that is *30%* of $E_{dep,max}$. At this point the *I-V* curve of the FeFET is shifted far to the right along with the threshold voltages. A reasonable expectation is a reduction of the depolarization field in the order of *10%* without hampering the device functionality.

Usually, the substrate doping is increased for scaling to smaller dimensions to reduce the maximum depletion region and counter small channel effects. Unfortunately here, the increase in substrate doping also brings a shift in the threshold voltage (Eq. 2.5) so that a higher operating voltage is needed, although the device dimensions are not changed. This has to be taken into consideration in view of FeFET scaling.

6.2 Investigating the case of $A_{Fe} \neq A_{Ox}$

In the analysis until now the area of the gate stack was always the same, both for the ferroelectric and dielectric layer, and equal to the channel dimensions $A_{Fe} = A_{Ox} = A = W \cdot L$. This is probably the easiest approach, since the same lithography mask can be used to create the structures. It is, however, interesting to study the case where the areas of the two layers are not the same (Fig. 6.4). It was not mentioned in chapter 5, but increasing the surface of the ferroelectric layer and thinning the oxide layer would enable a FeFET operation with smaller voltages, as Eqs. 6.1-6.4 show, but would unfortunately be impractical and more difficult to fabricate.

$$V_{Ox} = V \cdot \frac{C_{Fe}}{C_{Fe} + C_{Ox}} \quad and \quad V_{Fe} = V \cdot \frac{C_{Ox}}{C_{Fe} + C_{Ox}}$$

$$A_{Fe} \cdot k \Rightarrow C_{Fe} \cdot k$$

$$d_{Ox} \cdot \frac{1}{k} \Rightarrow C_{Ox} \cdot k \qquad (6.1\text{-}6.4)$$

$$V, d_{Ox} \cdot \frac{1}{k}, A_{Fe} \cdot k \Rightarrow V_{Ox}, V_{Fe} = const., k > 1$$

Instead, the case where $A_{Fe} < A_{Ox}$, will be examined. In order to simulate this structure, the model of a ferroelectric capacitor connected in series to the transistor's gate electrode (Fig. 6.3) is used, because the FeFET model of section 3.4.1 assumes equal surfaces. The transistor model HSPICE54 (equivalent to BSIM4, that allows the modification of the oxide permittivity in contrast to BSIM3) is used, although other transistor models can be used too. Unfortunately, not every transistor model can be combined with a ferroelectric capacitor to build a FeFET model because there are issues with the convergence of the simulation (charge conservation).
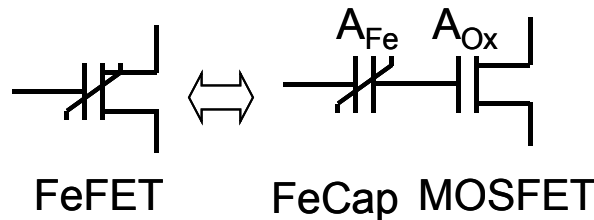
Fig. 6.3 A ferroelectric capacitor connected to the gate of a MOSFET used as a FeFET model for the case where $A_{Fe} \neq A_{Ox}$.

The parameters of the following FeFETs simulated are listed in Table 6.1 (channel dimensions are specified inside the graphs).

Parameter	unit	Value
d_{Ox}	nm	1
d_{Fe}	nm	100
ε_{Ox}	-	12
ε_r	-	250
P_S	$\mu C/cm^2$	3
P_R	$\mu C/cm^2$	2.8
E_C	MV/m	3
N_{CH}	cm^{-3}	10^{17}
N_{SUB}	cm^{-3}	10^{17}
V_{FB}	V	-0.4

Table 6.1 Parameters for the FeFETs with $A_{Fe} < A_{Ox}$.

Fig. 6.4 A FeFET with $A_{Fe} > A_{Ox}$ (left) and $A_{Fe} < A_{Ox}$ (right).

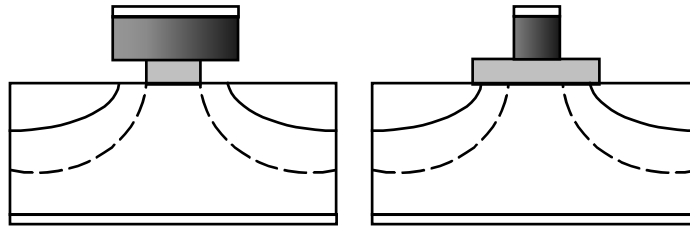Figure 6.5 (left) shows the simulation of the FeFET for four different area ratios. It can be seen that the remnant polarization increases with a higher area ratio $A_{Ratio}=A_{Ox} / A_{Fe}$. Figure 6.5 (right) shows this dependence. The remnant polarization reaches a maximum, which is the material's P_R for the saturated hysteresis. Also, the plateau around $0 V$ disappears completely. This simulation could perhaps also explain the polarization hysteresis curves in [10] (*P-V* of an MFIS structure with TGS on Si) that also do not show any plateau around $0 V$.
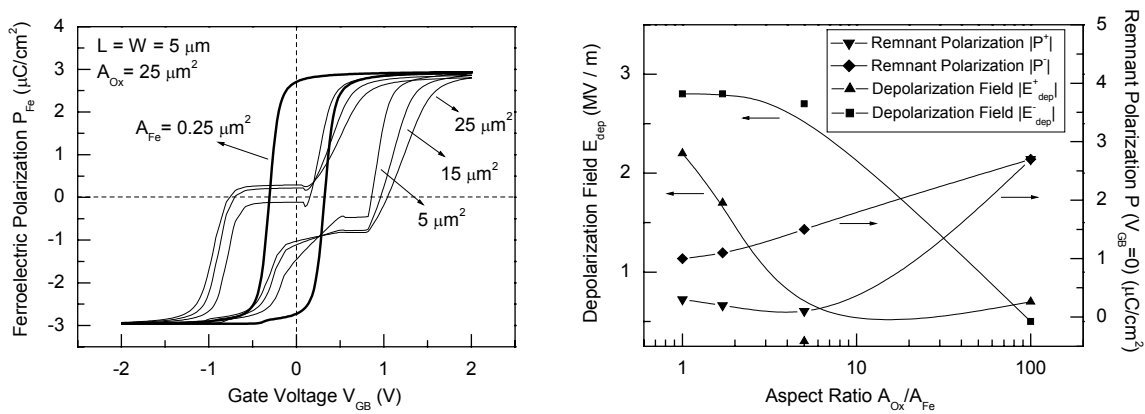
Fig. 6.5 The effect of different area ratios A_{Ratio} on the hysteresis curve (left), depolarization field and remnant polarization vs. the area ratio (right).

In Fig. 6.6 (left) only the width of the channel was modified, and the hysteresis curves are again distorted. The dependence of the depolarization and the remnant polarization on the area ratio is plotted in Fig. 6.6 (right). Thus, it is not the area ratio alone that affects the depolarization field.
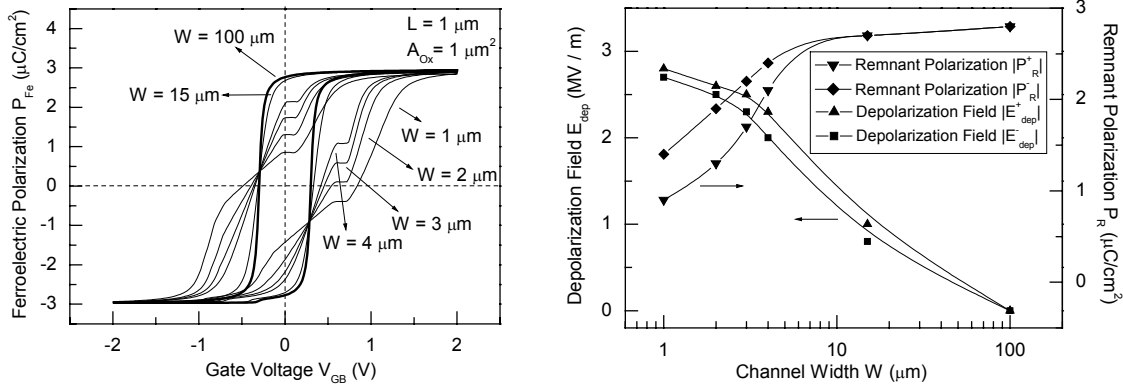
Fig. 6.6 (left) The effect of different area ratios A_{Ratio} (only the width was modified this time) on the hysteresis curve, (right) depolarization field and remnant polarization vs. the area ratio.

6.3 Explanation of the polarization curves

In section 4.1 the depolarization field was attributed to incomplete charge screening of the surface charge of the ferroelectric at the interface to the oxide/semiconductor. In the case where $A_{Fe} < A_{Ox}$, there is sufficient charge in the oxide surface to achieve screening of the ferroelectric surface charge. This, however, is not the reason for the simulation results. Just as with increasing the doping concentration, the voltage drop across the ferroelectric increases at the expense of that across the oxide, as the oxide capacitance becomes much larger than the ferroelectric one ($A_{Fe} << A_{Ox} \Rightarrow C_{Fe} << C_{Ox}$). Again, a functional device with zero depolarization field and a reasonable hysteresis curve is unfortunately not possible. The fact that the depolarization field in the MFIS structure is reduced has an impact on retention measurements as reported in [4], [29] and [30].

One other way in which using different areas for the gate stack layers can prove useful, is that the oxide can be made thicker. Equations 6.1-6.4 show that the voltage divider ratio can be kept constant by decreasing the oxide thickness and increasing the surface of the ferroelectric. Conversely, a thicker oxide and a ferroelectric with a smaller surface (both scaled with the same factor) lead to equally scaled capacitances as shown by the following equations.

$$A_{Fe} \cdot \frac{1}{k} \Rightarrow C_{Fe} \cdot \frac{1}{k}$$

$$d_{Ox} \cdot k \Rightarrow C_{Ox} \cdot \frac{1}{k} \tag{6.5), (6.6}$$

$$\Rightarrow const.\ voltage\ divider\ ratio\,, k > 1$$

This can be seen in Fig. 6.7 where a FeFET with *W = 15 μm, L = 1 μm* and $A_{Ox} = 1\ μm^2$ (see also Fig. 5.9) is simulated for different oxide thicknesses. The device with the *10 nm* thick oxide gives a very acceptable *I-V* characteristic and has high remnant polarization values (low depolarization field) at the same time. Note also that the memory window does not change much (Fig. 6.7 right).
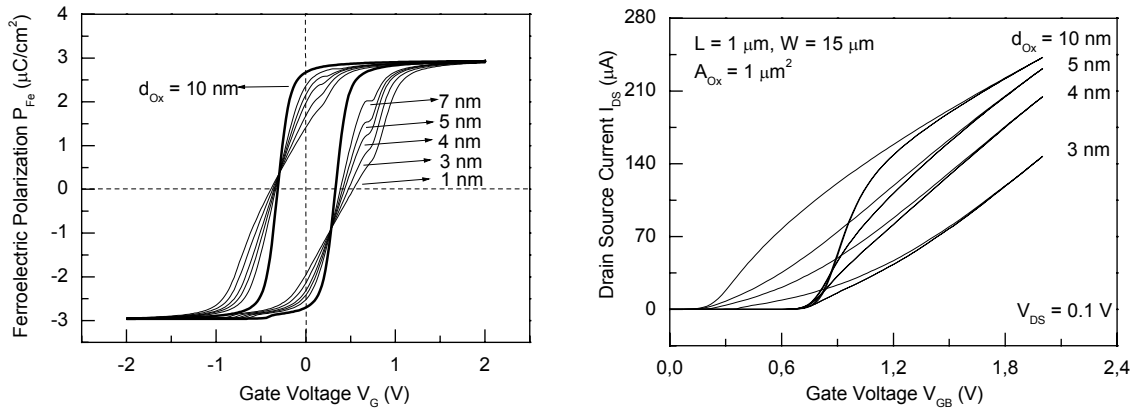
Fig. 6.7 (left) Hysteresis curves for a FeFET with A_{Ratio} = 15 and different oxide thicknesses, (right) I-V characteristics of the FeFET.

The case where the area of the ferroelectric is larger than that of the oxide (Fig. 6.4 left), leads to a smaller value of P_R because the ferroelectric capacitance becomes larger and the voltage drop across the ferroelectric smaller. The fabrication of such a structure is more difficult.

Another case was shown in Fig. 2.9, where the memory window dependency on the saturated polarization of a ferroelectric (P_R / P_S = 0.8) was examined. Although, the depolarization field drops considerably (Fig. 6.8) and the polarization plateau is less distinct, the memory window is reduced too. Besides, the ferroelectric's saturated polarization is not a parameter that can be modified without having to use a different ferroelectric.
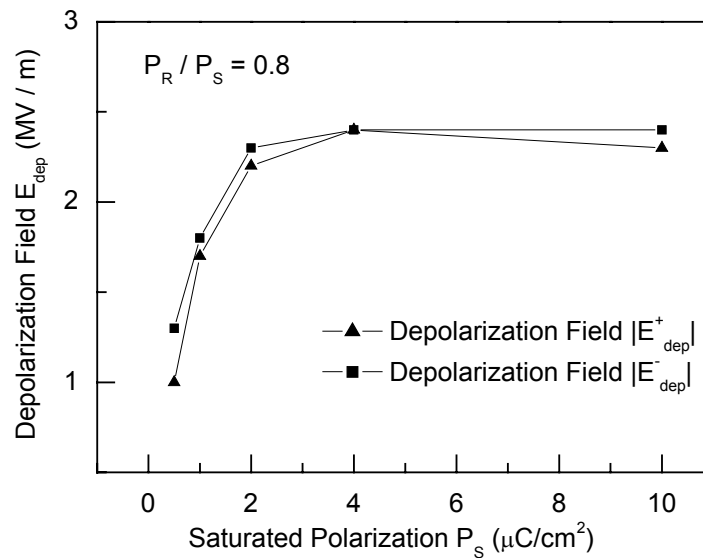
Fig. 6.8 Dependence of the depolarization field on the saturated polarization for the FeFET of Fig. 2.9.

6.4 Summary

The case $A_{Fe} \neq A_{Ox}$ although leading to interesting results, is of less practical importance in view of the scaling prospects of the FeFET. Having a higher value of remnant polarization

would increase data retention and it seems the bigger the A_{Ratio} the higher the resulting retention times. In order to reduce the depolarization field, a high area ratio of about $A_{Ratio} \sim 100$ is needed. Unfortunately, a FeFET with channel dimensions $L \times W \, \mu m^2$ and a $L/10 \times W/10 \, \mu m^2$ ferroelectric is not easy to fabricate in small sizes.

7 FeFET programming concepts

Two programming techniques for using FeFETs in memory applications will now be discussed. The first was also suggested in [8] and is based on Flash programming for mass memory application. The other alternative, that was presented in [31], makes the FeFET based memory a viable DRAM alternative.

In Flash memory technology before writing information into a cell it is first erased. This methodology is passed over to the FeFET because of the similarity of the devices. Because an erase operation is needed prior to every write, the write operation lasts double the time of a single write operation (the erase operation can be seen as a write of the opposite polarization).

There are two ways of erasing information in the FeFET (Fig. 7.1). The first applies a negative voltage to the gate electrode to switch the ferroelectric polarization (*negative gate erase*). The other applies a positive voltage to source and drain while grounding the gate (*positive voltage erase*).
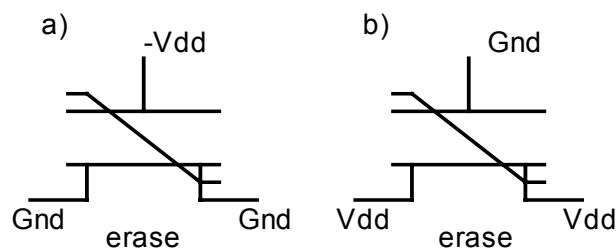

Fig. 7.1 Reversing the polarization in a FeFET using (a) negative gate erase and (b) positive voltage erase.

Based on these two erase mechanisms, two different programming techniques are derived. One difference between the two is that with the *negative gate erase*, the substrate doping does not have to be considered at all. In case of the second erase method the substrate doping concentration is the key for enabling *positive voltage erase*.

Another way to differentiate between programming approaches in general is through a different array structure. Basically, three such configurations are the most popular in Flash memories, these being NAND, NOR and AND in that particular order. In [8] it was shown that both NAND and NOR cannot be applied to the FeFET for disturbance reasons, leaving AND as the only choice. For the rest of this thesis the AND configuration will be used, except in the next chapter where the different architectures are compared in terms of packing density and efficiency.

In the AND architecture the FeFETs are arranged as shown in Fig. 7.2. The MOSFETs in the first row are "pass" transistors that short BL1 with BL1b during a "write operation", so that $V_{Source} = V_{Drain}$. This is one benefit of the AND architecture.
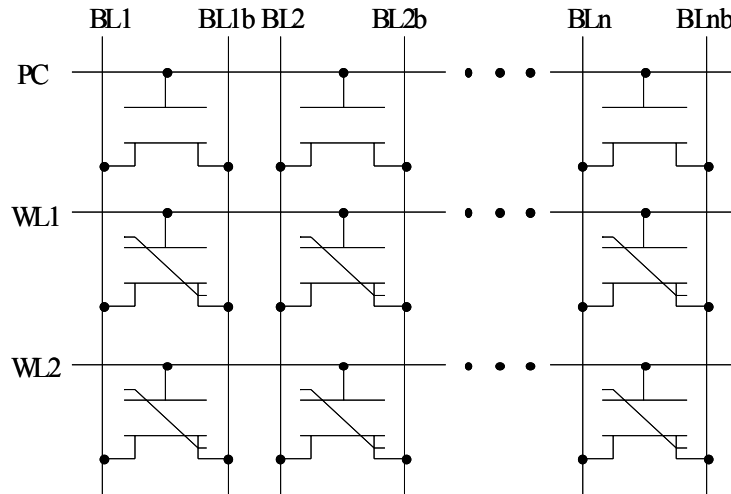
Fig. 7.2 The AND memory architecture (BL: bitline, WL: wordline, PC: parallel connect).

7.1 Negative gate erase

Two programming schemes that have been suggested using *negative gate erase* [32] are shown in Fig. 7.3.
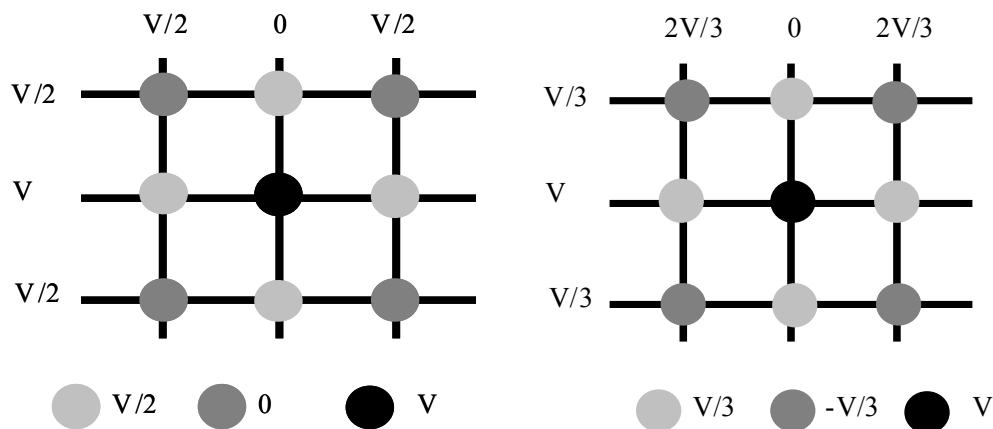
Fig. 7.3 *V/2 (left) and V/3 (right) programming schemes using negative gate erase [32].*

The horizontal voltage is applied to the wordlines and the vertical to both bitlines (source and drain are at the same potential through the pass transistor). Figure 7.3 only shows the write operation. For the erase operation the wordline is charged to a negative potential with all other terminals grounded. Actually, the source and drain potential does not matter ("X") during erase, as the gate potential will always be more negative than the silicon surface potential (even if $V_S = V_D = -V_{DD}$, this case would bias the p-n junctions in the forward direction) leading to a polarization reversal as shown in Fig. 7.4.

Fig. 7.4 When a negative voltage is applied to the wordline, all cells in that wordline (sometimes referred to as page) are erased (substrate is at 0 V).

The only exception is when the substrate potential is reduced to $-V_{DD}$ as proposed in [33] by the use of p-wells. This results in a higher technological complexity, lower packing density and also speed deterioration (charging a p-well to 0 and $-V_{DD}$ is slow due to its large diffusion capacitance).

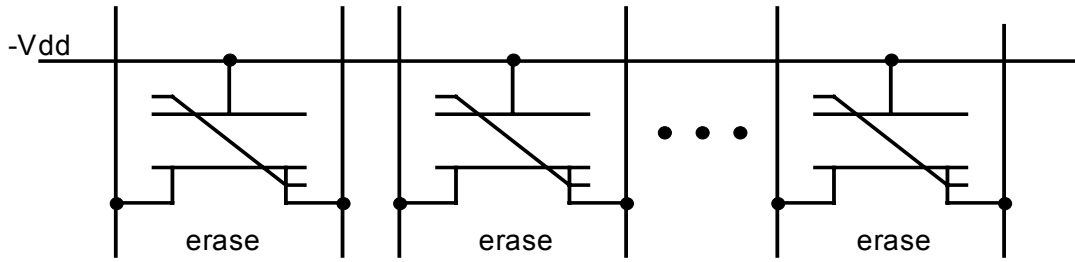During read and erase, the pass transistor is disabled. The potential to which the bitline is charged, is chosen in such a way that the current through the conducting FeFET will be large enough to discharge the bitline as shown in Fig. 7.5 (the speed of the discharge mainly depends on the bitline capacitance and of course whether the FeFET is set to "1" or "0").
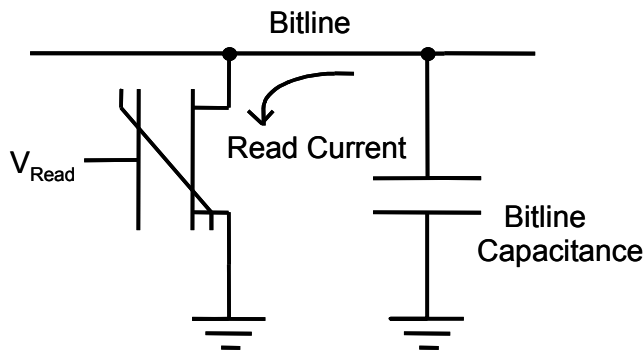
Fig. 7.5 Bitline discharge through the conducting FeFET.

The voltages applied for the programming scheme in Fig. 7.3 (left), known as V/2 scheme because of the $V_{DD}/2$ voltage used to protect the neighboring cells from disturbance, are listed in Table 7.1. Because of this, the FeFET must be designed so that a $V_{DD}/2$ pulse at the gate does not cause a change in its polarization (see section 4.4). An alternative was proposed with the V/3 programming scheme of Fig. 7.3 (right), where the disturbance is only $V_{DD}/3$, but affecting all cells besides the wordline and bitline being accessed.

	PC	WL1	WL2	BL1	BL1b	BL2	BL2b
Write	V_{DD}	V_{DD}	$V_{DD}/2$	0	0	$V_{DD}/2$	$V_{DD}/2$
Erase	X	$-V_{DD}$	0	X	X	X	X
Read	0	V_{read}	0	V_{DS}	0	V_{DS}	0

Table 7.1 The voltages applied in the V/2 programming scheme for the three operations.

A third possibility is shown in Fig. 7.6. Here only a single wordline is accessed. The cells that have $0\ V$ applied to the gate and V_{DD} to source and drain are not disturbed if an appropriate

substrate doping is applied (10^{16} cm^{-3}), as suggested in [8]. The advantage is a lower power dissipation since only one wordline is charged. Table 7.2 lists the voltages applied in this scheme. Only the write operation differs from that of the *V/2* and the *V/3* programming rules.
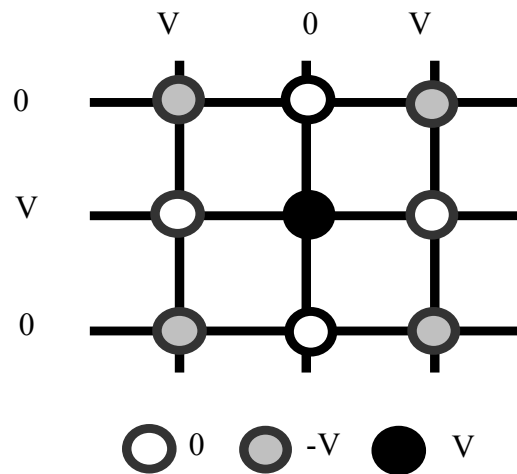
Fig. 7.6 Programming rule using a –V disturbance based on an appropriate doping concentration [8].

	PC	WL1	WL2	BL1	BL1b	BL2	BL2b
Write	V_{DD}	V_{DD}	0	0	0	V_{DD}	V_{DD}
Erase	X	$-V_{DD}$	0	X	X	X	X
Read	0	V_{read}	0	V_{DS}	0	V_{DS}	0

Table 7.2 The voltages applied in the V/2 programming scheme for the three operations.

7.2 Positive voltage erase

Thus far, there was a need for a separate erase operation and it was realized by the application of a negative voltage to the FeFET gate. By using the *positive voltage erase* in Fig. 7.1b it is possible to do away with a separate erase operation. This programming scheme is shown in Fig. 7.7 and the voltages are listed in Table 7.3.

Fig. 7.7 Programming rule requiring no separate erase operation.

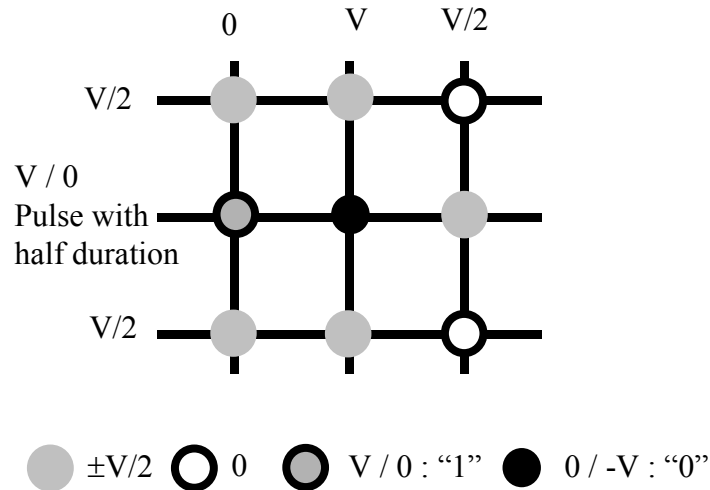|	PC	WL1	WL2	BL1	BL1b	BL2	BL2b	BLn
Program	V_{DD}	V_{DD} *(half pulse)*	$V_{DD}/2$	0	0	V_{DD}	V_{DD}	$V_{DD}/2$
Read	0	V_{Read}	0	V_{DS}	0	V_{DS}	0	0

Table 7.3 The voltages applied in the positive erase programming scheme.

The difference is that the erase is now incorporated in the write operation (in Table 7.3 termed "Program"). The read operation remains the same. Here, in contrast to Fig. 7.6, the polarization in the FeFET is reversed via the voltage difference "$-V$". This will be explained more precisely in section 7.4.

In Fig. 7.8 the arrangement of Fig. 7.7 is shown in device schematic level. During the write operation one word (in one wordline) can be written at a time (random access). The bitline is charged with *GND* for the cells to be written (BL1) and V_{DD} for those to be erased (BL2). During the first half of the programming cycle, the FeFET at WL1,BL1 is set at "1", but that at WL1,BL2 is not modified. During the second half, the FeFET at WL1,BL2 is erased ("0") because the WL1 pulse drops to zero.

For the unselected cells the bitlines (BLn) are charged with $V_{DD}/2$. All wordlines apart from the one accessed (WL1) are charged with $V_{DD}/2$ as well. This is at the expense of higher power dissipation, but on the other hand it is necessary to guarantee disturbance free programming. Figure 7.9 shows the simulated bit and wordline voltages, where the capacitance and resistance of the wordlines and bitlines have been ignored for simplicity. The cells whose state is not changed are marked in Fig. 7.8 with "X".
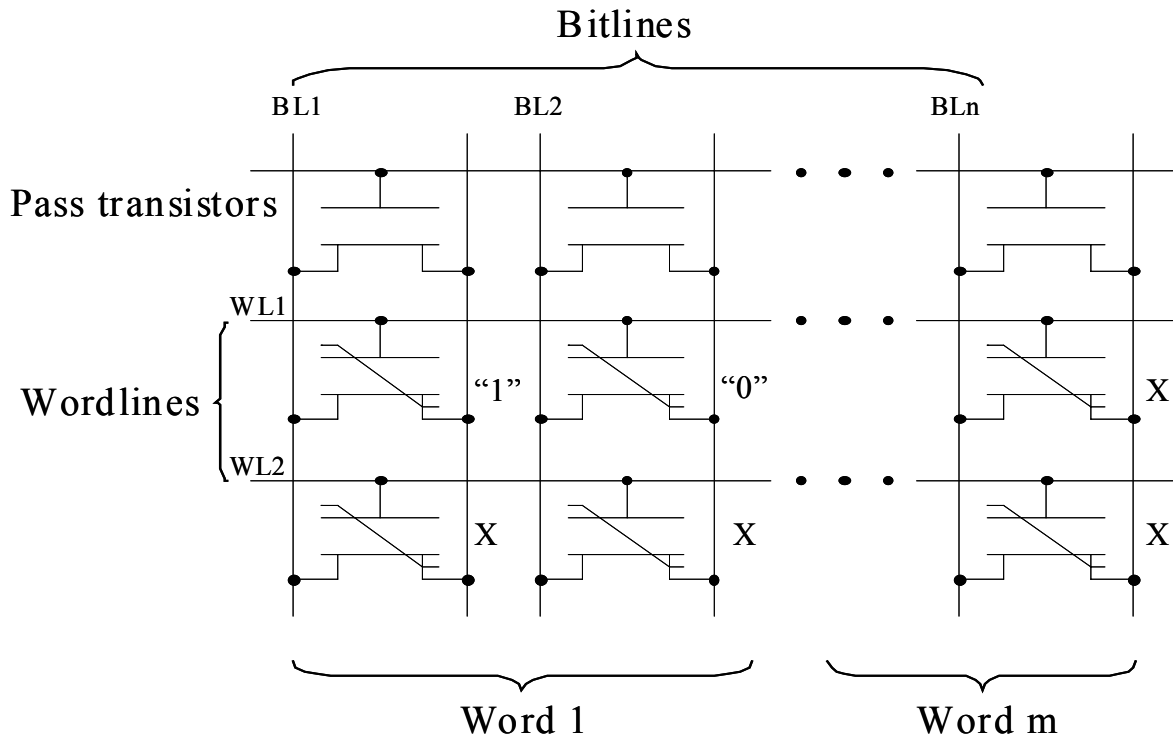
Fig. 7.8 FeFET memory matrix arranged in an AND structure.
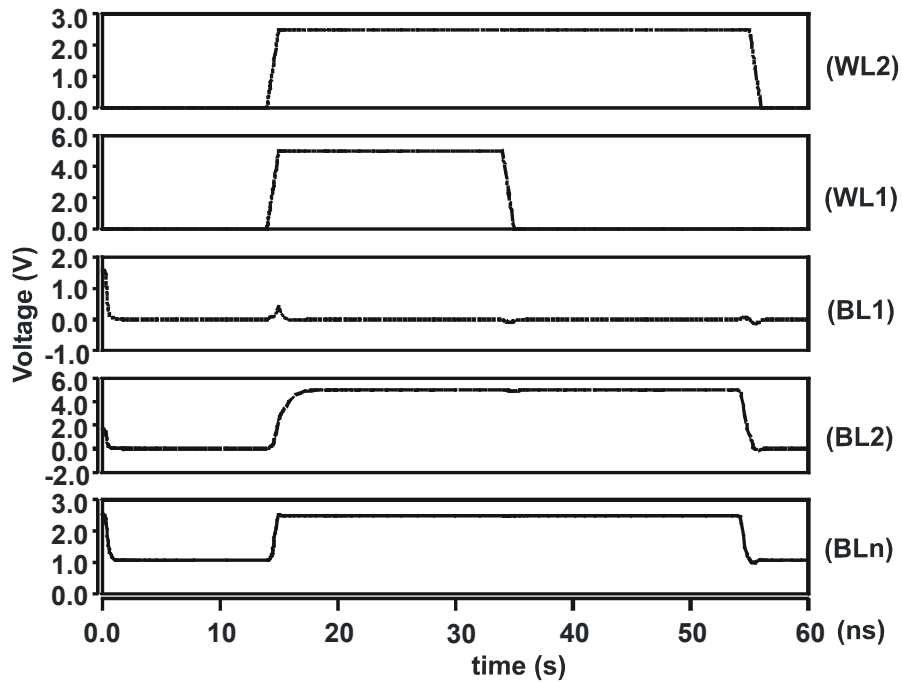
Fig. 7.9 Voltages applied during programming.

The advantages of this programming scheme are higher speed operation through elimination of a separate erase operation and the possibility to access a single word without erasing the entire row, thus enabling true random access operation. Another advantage lies in the fact that a single voltage polarity is necessary for this memory to work. This will be discussed in the next section.

7.3 Positive vs. Negative Voltages

In order to generate a negative potential on-chip, in case it is not externally supplied, a circuit known as charge pump must be used. In this circuit a capacitor is charged by a high frequency signal. This is a common practice in many memory chips that require voltages higher than the operating voltage $V_{DD}$. It can be easily integrated into any digital circuit design based on standard CMOS technology. One issue with charge pumps is the voltage rippling at the output, which can be reduced using a larger output capacitor, but then increases the time until steady state is reached. Including one into a chip increases the complexity. Additionally, the voltage cannot be generated instantly [38]. These are some of the issues that can be avoided when using the *positive voltage erase* programming. Another issue is that applying a negative voltage pulse to a wordline inevitably leads to the erasing of all FeFETs in the wordline (Fig. 7.4), regardless whether or not they are addressed by a bitline. The reason is that the voltage in the ferroelectric layer is negative and there is no way to protect a cell from being erased unless the substrate potential is made negative too. Doing so, however, would disable the functionality of the logic electronics in the memory circuit since the substrate is shared with the rest of the circuitry.

Positive voltage complying with standard CMOS voltage levels (*0 … V$_{DD}$*) can drive the FeFET. A drawback of this approach is the amount of polarization reversal (Fig. 7.17), which can lead to a smaller memory window.

7.3.1 Switching a negative voltage

The MOSFET transistor can be used as a switch to pass any voltage between *0 V* and *V$_{DD}$*. Actually the n-type MOSFET is more suitable to pass *0 V* and the p-type *V$_{DD}$*. But, in order to switch a potential lower than the ground (substrate) potential, a single MOSFET is not sufficient. One possibility is to use a voltage divider consisting of a p-type MOSFET and an ohmic resistor (Fig. 7.10). To precisely set the desired resistance of the p-MOSFET an accurate voltage must be generated and applied to the gate, which requires further overhead. For the resistor a diffusion region can be used which takes up a lot of space.
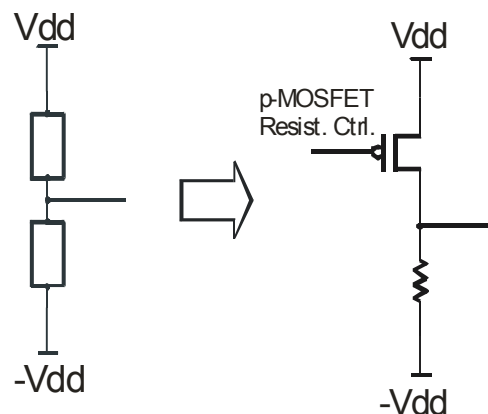

Fig. 7.10 Voltage divider using a p-MOSFET and a resistor to switch a negative voltage.

7.3.2 Generating a voltage between *0 V* and *V$_{DD}$*

To generate a voltage between *0 V* and *V$_{DD}$* a simple voltage divider consisting of two n-MOSFETS can be used (Fig. 7.11). It is important to carefully adjust the sizes of the transistors to achieve the desired ratio. If the voltage is to be used to drive a large load, the devices' width must be high enough to guarantee a low channel resistance.
Another alternative is a so called "bandgap generator". This enables voltages to be generated with high precision since the generated voltage is a multiple of the semiconductor bandgap, which is constant for a given semiconductor at constant operation conditions. However, the fact that this concept requires the use of a BiCMOS (Bipolar-CMOS) process makes it an attractive option only for initially complex designs.
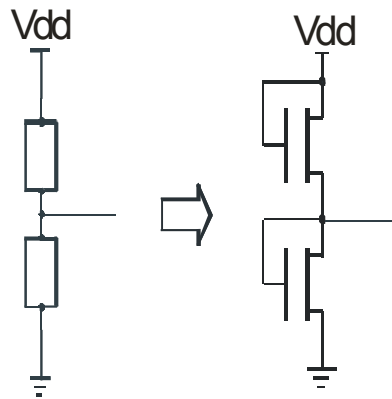

Fig. 7.11 Voltage divider utilizing MOSFETs connected as resistors [35].

7.4 Setting the FeFET to accumulation

In order to reverse the polarization, the FeFET must be driven to accumulation. In this section the possibility of causing accumulation using only positive voltages will be examined. First, using device simulation the conditions that cause accumulation in the MOSFET will be investigated. The results will then be expanded to the FeFET. One way to reach accumulation in a MOSFET is to set $V_{GB} < 0$ (or more precisely $V_{GB} < V_{FB}$, see Fig. 2.1). Another possibility is shown in Fig. 7.12a where $V_{GB} = 0$. This case is now explained in more detail.

7.4.1 Positive voltage erase using a low doped substrate

Here a low uniformly doped substrate *(10^{14} cm^{-3})* is used, and a voltage is applied to source and drain *(5 V)* while the gate and bulk are grounded. The voltage distribution shows that the potential at the oxide interface is approximately *3.5 V*, thus the field in the oxide is reversed, that is $V_{Ox} = V_G - V_{Si} = 0 - 3.5 = -3.5 V$ (in the FeFET this would result in the polarization reversal). Figure 7.12b shows the same simulation with a voltage only applied to the source.
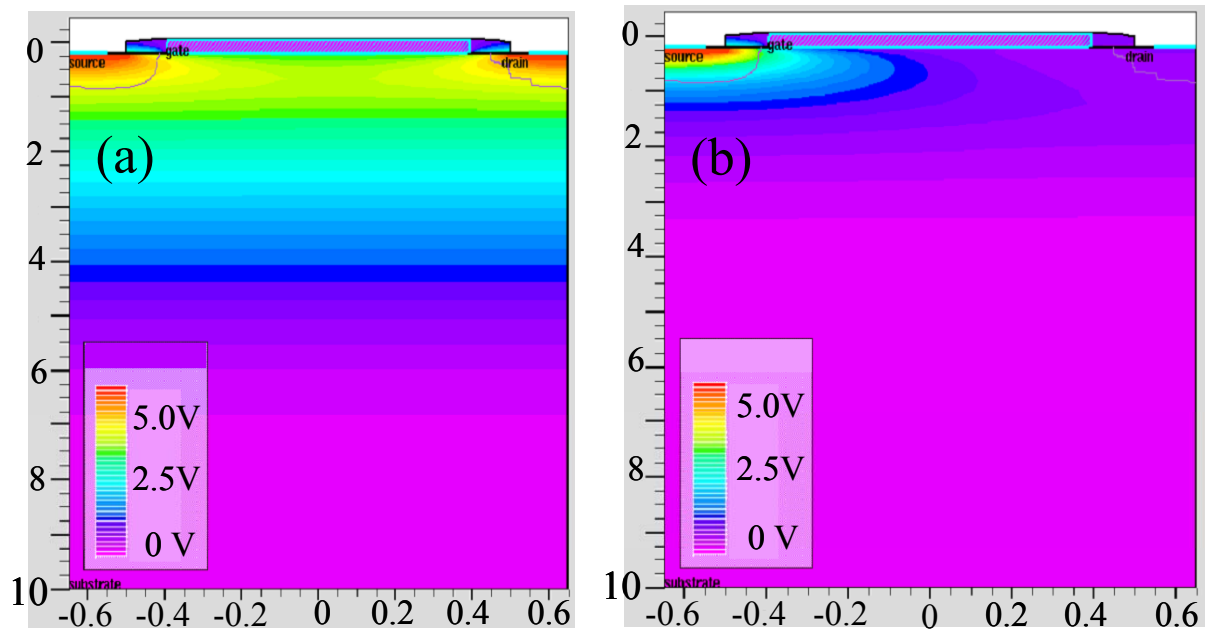
Fig. 7.12 Device simulation of the potential distribution in the MOSFET after applying (a) V_G= V_B=0 V and V_S= V_D=5 V, (b) V_G= V_B= V_D =0 V and V_S =5 V (dimensions are in μm)[16].

What happens during the erase operation is shown in Fig. 7.13. It shows the accumulation of holes in the channel as well as in the bottom of the substrate. Note that the substrate doping is only 10^{14} cm^{-3}, whereas MOSFETs have typically a substrate concentration of about 10^{17} cm^{-3}. Because of this low doping concentration it is possible to set the MOSFET to accumulation with the voltages specified. The same applies for the FeFET. This is now further studied using the model of section 3.4.1.

Fig. 7.13 Device simulation of the erase operation showing the hole concentration in the semi-conductor (dimensions are in μm).

---

[16] The device simulations in Figs. 7.12, 7.13 were performed with ATLAS [38].

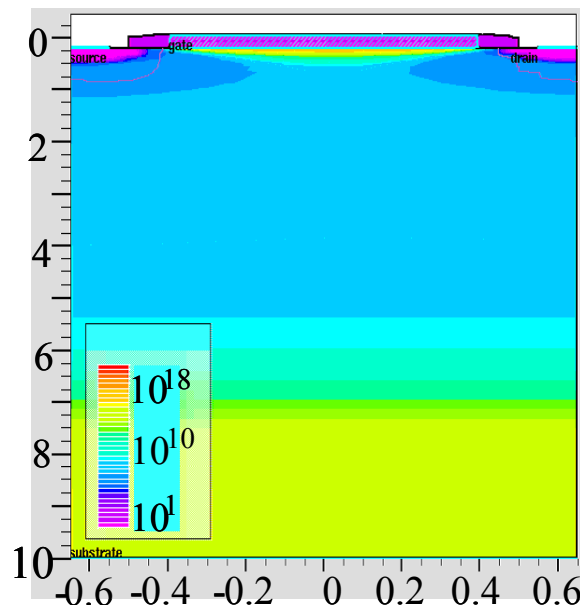In Fig. 7.14 a simulation of the gate charge shows that the accumulation region can be shifted by varying the doping concentration. The simulation was performed using the BSIM3v3 model with three device parameters controlling the accumulation region shift (channel doping N_{CH}, substrate doping N_{SUB} and channel doping depth X_T).
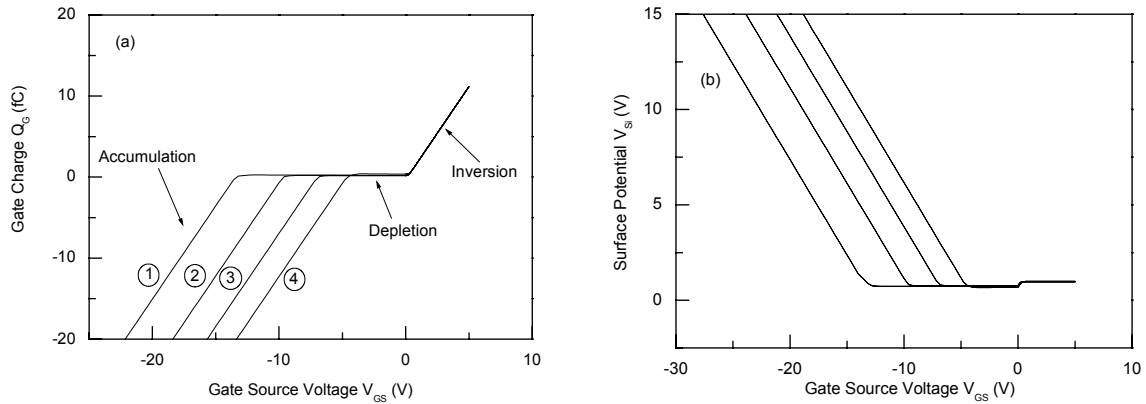

Fig. 7.14 Dependence of the (a) gate charge, (b) surface potential on the applied gate-source voltage V_{GS} (V_S=0 for V_{GS}>0, V_G=0 for V_{GS}<0) by varying the substrate doping concentration.

The left figure shows the gate charge vs. the gate source voltage. Note that for V_{GS}>0, V_S=0 and for V_{GS}<0, V_G=0. In the right figure the surface potential is plotted, showing basically the same effect. The parameters varied in the simulation of Fig. 7.14 are listed in Table 7.4.

Curve	$N_{CH} (cm^{-3})$	$N_{SUB} (cm^{-3})$	$X_T (nm)$
1	10^{17}	$5.0 \cdot 10^{14}$	100
2	10^{17}	$2.5 \cdot 10^{14}$	100
3	10^{17}	$1.0 \cdot 10^{14}$	100
4	10^{17}	$1.0 \cdot 10^{14}$	110

Table 7.4 Doping parameters varied to achieve a shift of the accumulation region to the right.

The BSIM3v3 simulation is not totally in agreement with the device simulation performed with ATLAS [38]. According to the device simulation, the doping concentration of the channel N_{CH} and the channel doping depth X_T have no impact, in contrast to the BSIM3v3 simulation. In any case, the common denominator is the low doping concentration of the bulk ($10^{14} cm^{-3}$).

7.4.2 Positive voltage erase using a high doped substrate

Figure 7.14 shows that the onset of accumulation is about *10 V* higher. Further lowering the substrate doping does not shift the charge curve any more to the right. Sweeping the doping parameters independently over a wide range from 10^{14} to 10^{19} reveals for $N_{SUB} = 5 \cdot 10^{18}\ cm^{-3}$ and $N_{CH} = 1 \cdot 10^{16}\ cm^{-3}$ a further possibility to shift the onset of the accumulation region to the right. This is an unusual doping profile. Figure 7.5 (left) shows the usual doping profile in a MOSFET (gaussian doping profile) and (right) the doping profile suggested here.
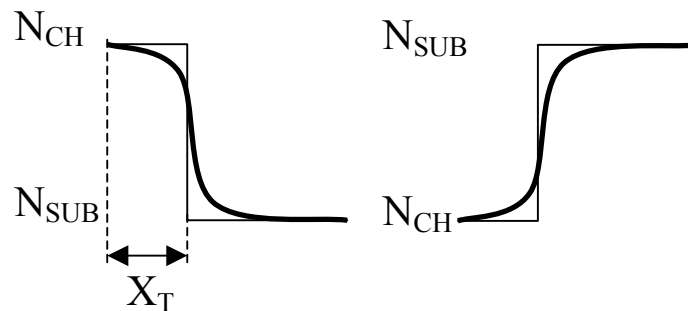
Fig. 7.15 The usual doping profile applied in a MOSFET (left), an unusual doping profile that enables positive voltage erase (right). According to BSIM3v3 both can be used for positive voltage erase.

The result of this doping profile is seen in the charge characteristic of Fig. 7.16. It can be seen that the gate charge varies almost linearly with the gate source voltage, leaving only a small depletion region. As a result, the surface potential for $V_{GS} = -5\ V$ ($V_{GB} = 0\ V$) rises from *3.5 V* in the last case to almost *5 V* here. The effect this has on the programming of the FeFET is shown in the next section.

Fig. 7.16 Shift of the accumulation region by using a doping concentration, where $N_{SUB} > N_{CH}$.

7.4.3 Positive voltage erase and the memory window

In Fig. 7.17a the current characteristic, and in Fig. 7.17b the polarization hysteresis (*P-V*) curves are plotted, after applying a *positive voltage erase* with a substrate doping of $N_{SUB} = 10^{14}\ cm^{-3}$, $N_{CH} = 10^{17}$.

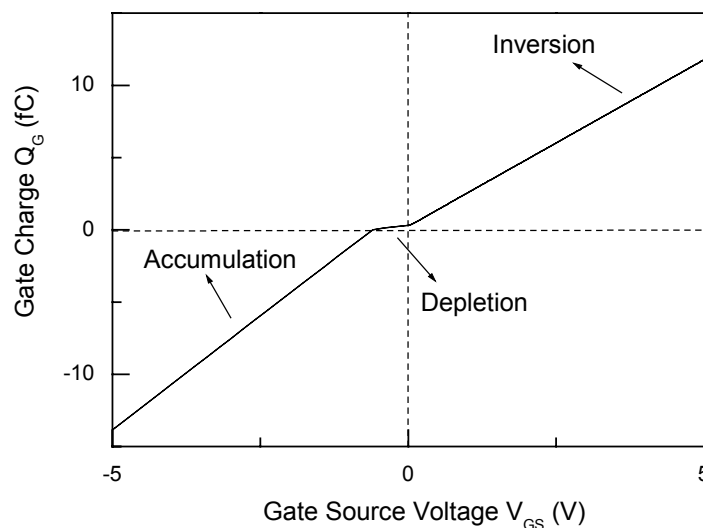The thin solid line shows the *I-V* curve in case of the *negative gate erase (V_{GB}=-5V)*. The thick solid line shows the extent to which the polarization is reversed with the *positive voltage erase*. It is far from being a complete reversal. The amount of disturbance after application of a $V_{DD}/2$ (from the negative polarization state P^-) is shown with the dashed curve. The polarization and the memory window are further reduced at the same time. Despite all that, the remaining memory window gives an I_{on}/I_{off} current that is sufficient for a memory application.

Fig. 7.17 (a) Current and (b) polarization curves when using the positive voltage erase scheme with the low substrate doping concentration. the solid line in (a) shows the reduced memory window after the positive erase, and the dashed after applying a gate voltage of 2.5 V.

The same case with the other doping profile of Fig. 7.15 is shown in Fig. 7.18a and b. The graph shows that the polarization is reversed to about $-2~\mu C/cm^2$ compared to about $-3~\mu C/cm^2$ for the case of $V_{GB} = -V_{DD}$ (thin solid line), enabling a large memory window and minimizing disturbance problems. The disturbance still leads to a big reduction in the remnant polarization, but the memory window is minimally affected. The difference between Figs. 7.17 and 7.18 is traced to the surface potential being *3.5 V* and *5 V* in the first and second case, respectively, for V_{GS}=-5 V (V_{GB}=0 V).
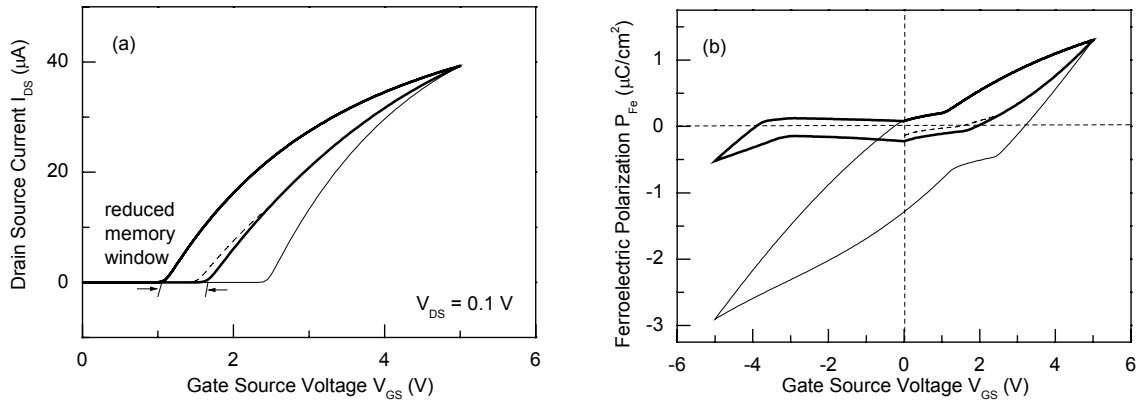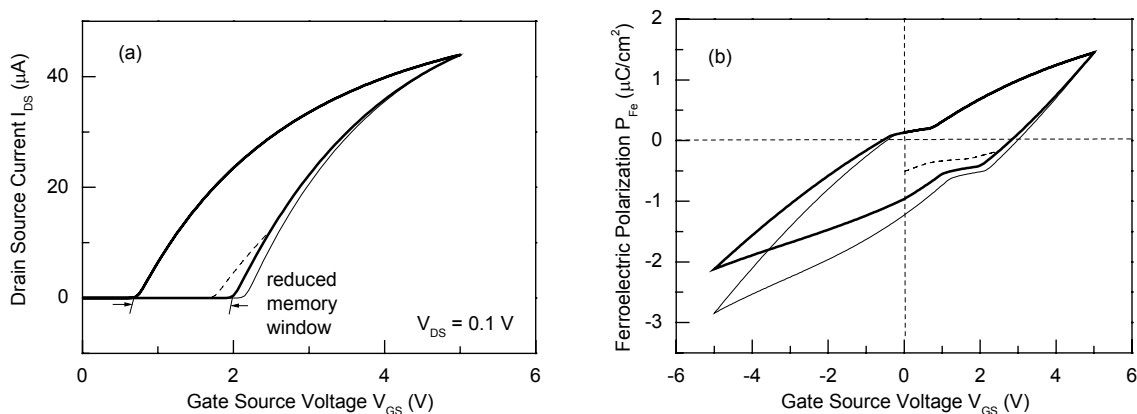

Fig. 7.18 (a) Current and (b) polarization curves when using the positive voltage erase scheme with the high substrate doping concentration (the solid line in (a) shows the reduced memory window after the positive erase, and the dashed after applying a disturb voltage of 2.5 V).

The properties of the FeFET used in the simulation are summarized in Table 7.5. The material parameters for the ferroelectric were based on PZT [69]. The default values were used for the rest of the BSIM3v3.1 (Level 49) parameters. For the FeFET in Fig. 7.18 the doping parameters were $N_{SUB}=5{\cdot}10^{18}$ cm^{-3} and $N_{CH}=10^{16}$ cm^{-3}, and for the FeFET in Fig. 7.17 the substrate doping concentration was uniform $N_{SUB} = N_{CH} = 10^{14}$ cm^{-3}.

Parameter	unit	Value
W	μm	1
L	μm	1
d_{Ox}	nm	10
d_{Fe}	nm	360
ε_{Ox}	-	12
ε_r	-	200
P_S	$\mu C/cm^2$	40
P_R	$\mu C/cm^2$	32
E_C	MV/m	7
N_{CH}	cm^{-3}	10^{17}
N_{SUB}	cm^{-3}	10^{14}
V_{FB}	V	-0.4

Table 7.5 Parameters for the FeFET of Fig. 7.17.

At this point it should be mentioned that the *positive voltage erase* in combination with the second doping profile could not be verified with device simulations.

7.5 FeFET Scaling and positive voltage erase

In chapter 5 FeFET scaling for devices with sub-micron channel dimensions was examined. The operating voltages used were smaller than *1.5 V* and *negative gate erase* was suggested for polarization reversal. To apply *positive voltage erase* one has to look at the *P-V_{GS}* (as opposed to *P-V_{GB}* for *negative gate erase*) hysteresis loop. Figure 7.17b shows that *V_{GS}< -4 V* (*V_{GB} = 0 V*) has to be applied to reverse the polarization. Thus, this doping profile makes it impossible for a FeFET to scale to smaller dimensions (because the voltage cannot be scaled any more). Another option is to use the doping profile shown in Fig. 7.15 (right). Because of the almost linear dependence of the gate charge on *V_{GS}* (Fig. 7.16), it is possible to scale the erase voltage.

Of the suggested scaling methods in sections 5.7 and 5.8 *constant gate stack scaling* requires *negative gate erase*, because FeFET programming ("write" and "erase") is done by accessing only the gate electrode, while *positive voltage erase* works by applying a voltage to source and drain. This means that *variable gate stack scaling* has to be used in combination with *positive voltage erase*, and thus the gate stack has to be adjusted for lower voltage operation and a similar approach to that of section 5.8 has to be followed to determine the necessary

FeFET parameters. The next simulation shows that the gate stack layer thicknesses can be adjusted for a *0.18 μm* FeFET that operates with *2 V*. All other FeFET parameters are listed in Table 7.5 (doping is that of Fig. 7.18). The remnant polarization values are negative at both states (P^+ and P^-) (Fig. 7.19b). The hysteresis can be shifted with a more negative flatband voltage (as shown in section 3.5) which also shifts the voltage thresholds (Fig. 7.19a).
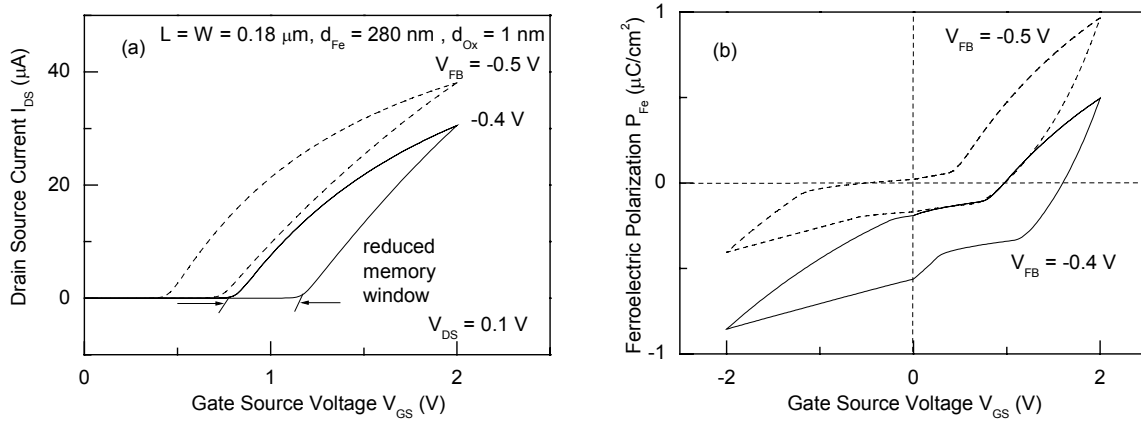

Fig. 7.19 (a) Current and (b) polarization curves using the positive voltage erase scheme for a 0.18 μm FeFET.

7.6 Examining the case $V_{Read} = 0$ V

In all cases discussed so far, the read voltage was always $V_{Read} > 0$ *V*, that is, the FeFET was not conducting at $V_{GB} = 0$ *V*, and $0 < V_{TH1} < V_{TH2}$. In section 2.4.2 it was suggested that a normally-on FeFET would not make sense, due to possible power consumption at idle state. Besides, floating gate transistors are not normally-on either. In some publications [90], however, this case is proposed. Figure 7.20 shows a FeFET *I-V* characteristic that enables read out without applying a read voltage to the gate. Note that $V_{TH1} < 0 < V_{TH2}$. For the write and erase operations, either of the programming schemes presented so far can be applied. In Fig. 7.20 *positive voltage erase* is used. The advantage with $V_{Read} = 0$ *V* is less power dissipation during read out, since no wordline must be charged.
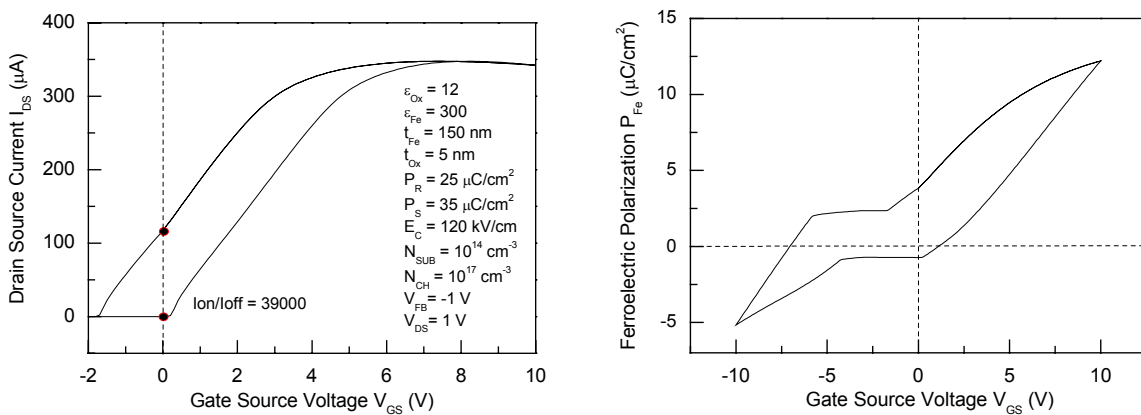

Fig. 7.20 Current characteristic (left) and polarization hysteresis (right) of a FeFET that uses the positive voltage erase method. No voltage needs to be applied to the gate for read out ($V_{Read} = 0$ V).

However, there is an issue at read out that makes it difficult to determine the state of a FeFET. Figure 7.21 shows what happens during read out. The charging of the bitlines to the voltage V_{DS} is applied to all FeFETs in one column. In case more than one devices of that column are conducting, the bitline can discharge in more than one ways and there is no way to tell which device is conducting. In order to make read out possible a negative voltage must be applied to the gate of the non-addressed devices to disable them, but this would increase power consumption. Also, the voltage should not be too high, to avoid driving the FeFETs to accumulation and causing an erase. For a simpler read out, however, it is best if the FeFET does not conduct at $V_{GB} = 0\ V$.
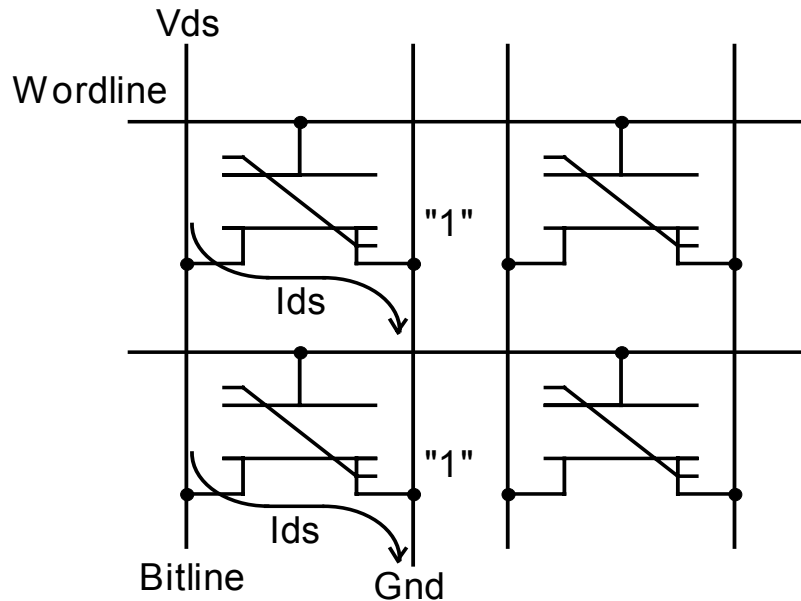
Fig. 7.21 Read out operation without applying a gate voltage ($V_{Read} = 0\ V$) leads to the discharging of the bitline through all conducting FeFETs in one column.

7.7 Summary

Two different programming concepts have been presented : the "*negative gate erase*" and the newly introduced "*positive voltage erase*". The latter leads to an easier implementation of a memory design through the elimination of a separate erase operation. In this new programming concept, the doping concentration of the substrate plays a major role. Further, it was shown that *positive voltage erase* can be combined with the *variable gate stack scaling* method. Finally, it was shown that the read operation in a memory application is simpler if the FeFET is not conducting without a voltage applied to its gate ($V_{GB} = 0\ V$).

8 Memory configurations and integration density

In the last chapter the AND matrix configuration was chosen instead of NAND or NOR because the presence of the pass transistor makes it possible to apply the same voltage to source and drain during a programming operation. In Flash technology, among the three configurations, it is NAND that dominates the market, because it enables a higher integration density. Table 8.1 gives a comparison of the different technologies. According to the table, no technology is superior in all areas. While NAND Flash has the highest packing density of any memory technology, it is not suitable for code execution because of its slow serial read access operation. It is, however, very well suited for data storage applications due to its ability to store simultaneously large blocks of memory.

Technology	NAND Flash	NOR Flash	AND Flash	DRAM
Cell size	$4F^2$	$10F^2$	$8F^2$	$8F^2$
Read access	Medium	Fast	Fast	Fast
Write/erase access	Fast	Slow	Fast	Fast / no erase
Purpose	Data storage	Code execution	Data Storage	Code Execution

Table 8.1 A comparison of the matrix configurations used in Flash technology and DRAM (F is the minimum feature size) [34],[42],[43].

The three configurations will now be examined individually in detail. Of the different circuit views that will be shown, the planar view is the one that is actually drawn by the layout designer. Figure 8.1 shows the planar layout of a MOSFET and the section view.
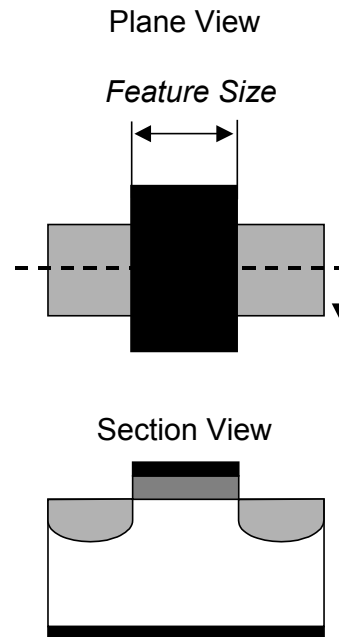
Plane View

Feature Size

Section View

Fig. 8.1 Plane and section view of a MOSFET device – The MOSFET is defined at the cross point of the diffusion box (horizontal) and the polysilicon box (vertical).

8.1 The AND configuration

The AND configuration uses additional MOSFETs as pass transistors, one in every bitline. For the FeFET memory application this is very convenient since it enables to set source and drain in a FeFET to the same potential. It is named after the logic gate AND, whose output goes to ground if one of the inputs is *0 V*. In this context the bitline (output) goes to ground (discharges) if one of the FeFETs in one column (input) is conducting (a conducting Flash transistor is actually at state "0"). Figure 8.2 shows the implementation in schematic level. Using two metallizations it can be realized as shown in Fig. 8.3 (left). This is the most straightforward implementation. The cross-section of the layout in Fig. 8.3 (right) shows the different layers that are used in a *0.12 μm* CMOS process technology as specified in [35]. The 3-D view of the structure in Fig. 8.4 shows the complexity of the process.

One way to make the layout more space efficient is to use diffusion regions for the bitlines as suggested in [34]. The reason this is more space (area) efficient is because the metal-diffusion contact in Fig. 8.3 takes up more space compared to the layout of Fig. 8.5. At the same time the process complexity is reduced as shown in Fig. 8.6. The disadvantage is the higher capacitance and the much higher resistivity of the diffusion bitlines compared to the metallic ones. This leads to higher RC times and thus to slower access times that have to be taken into account in the memory design process.
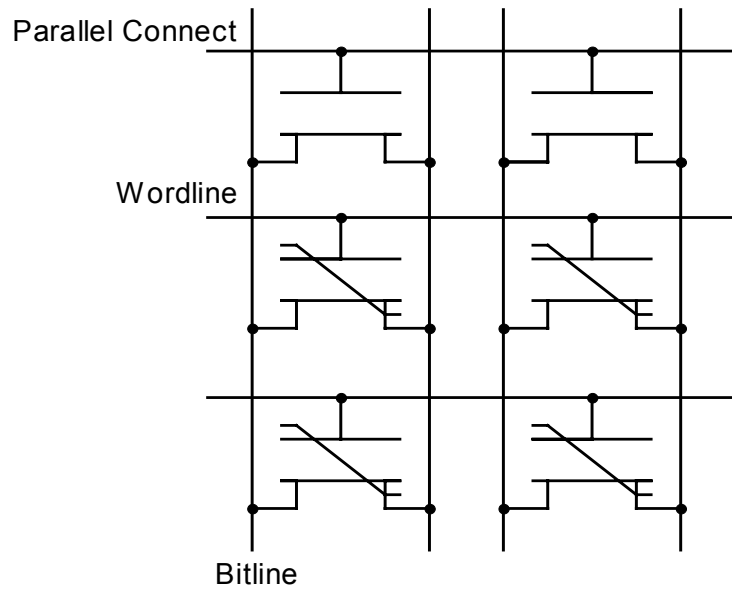
Fig. 8.2 Schematic of the AND array configuration.
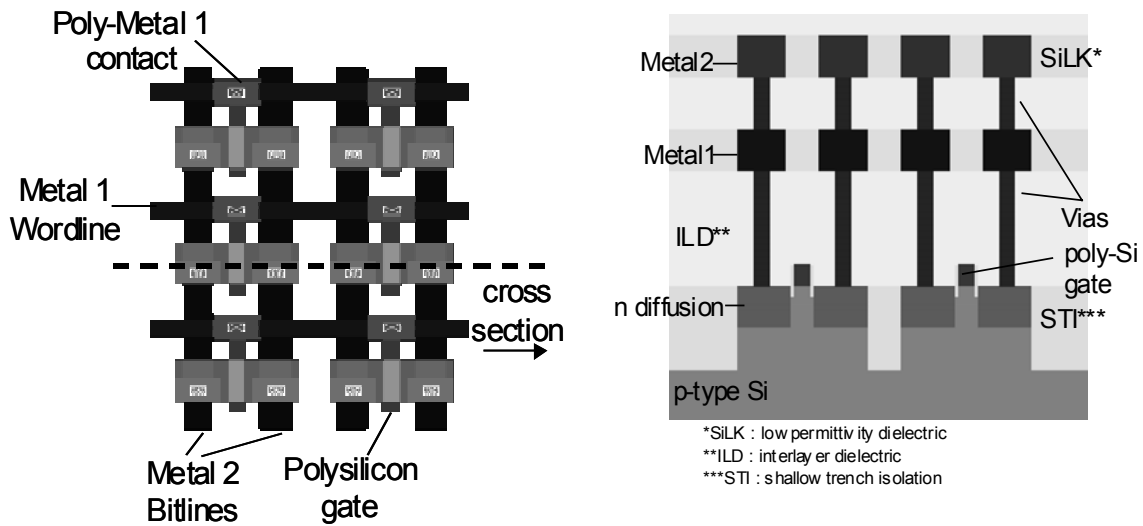
Fig. 8.3 Layout and cross-section of the AND configuration using metallic. For the layouts a 0.12μm process technology provided by [35] was used.
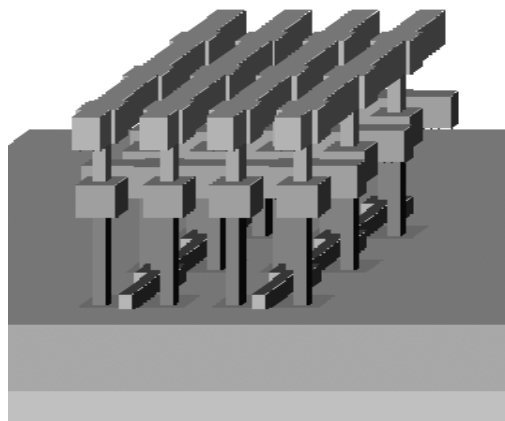
Fig. 8.4 3-D view of the AND structure of Fig. 8.3 using 2 metallizations. The wordlines are implemented with the 1st metallization and the bitlines with the 2nd.
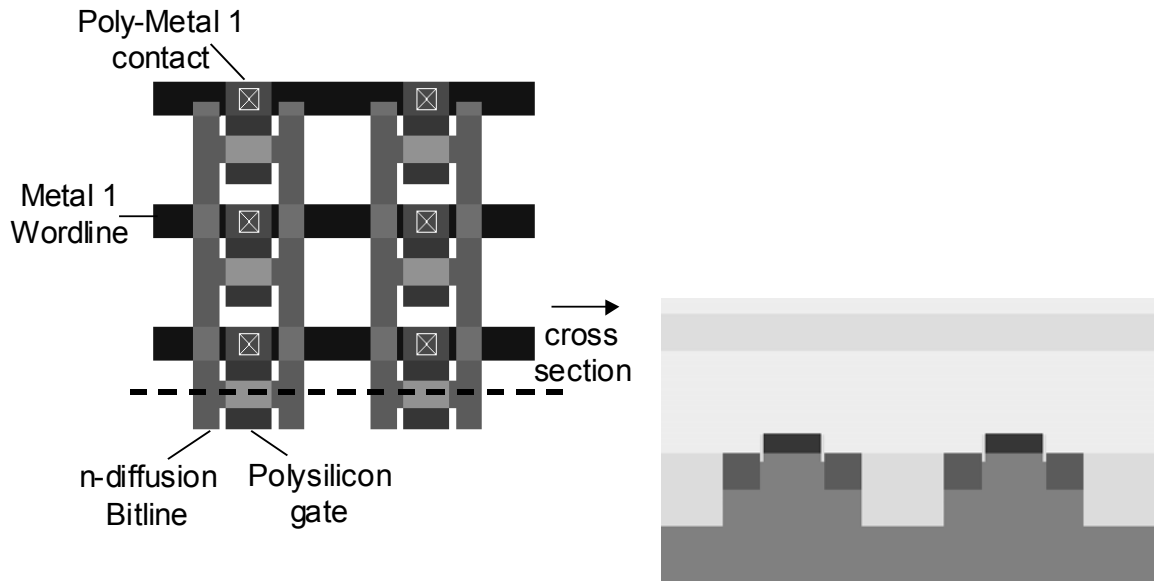
Fig. 8.5 Layout and cross-section of the AND configuration using diffusion bitlines.
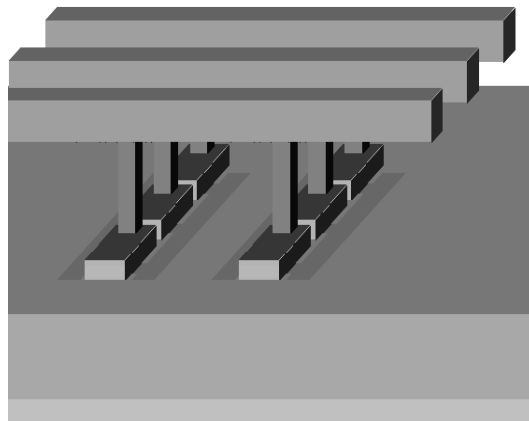
Fig. 8.6 3-D view of the AND structure of Fig. 8.5 using metallic wordlines and diffusion bitlines.

The diffusion lines can be used as local bitlines within blocks that connect to a global bitline as shown in Fig. 8.7 and this way the RC delay times can be kept low [8]. The same principle can be applied to wordlines too.
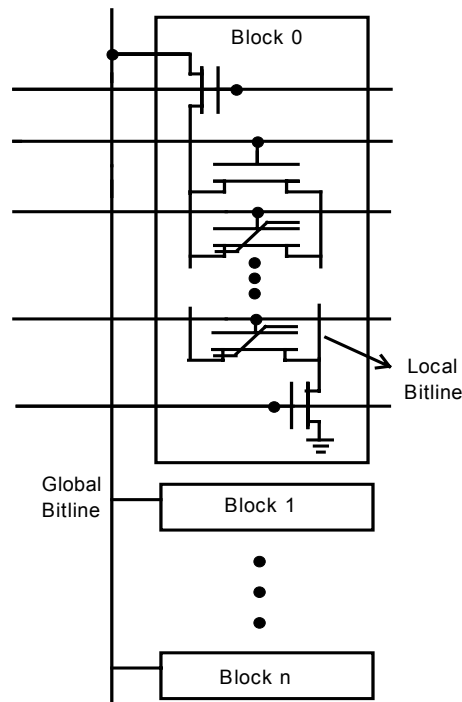
Fig. 8.7 Block diagram showing the bitline divided in local and global bitlines with a transistor connecting each block to the global bitline [8].

It should be mentioned that in such memory designs the design rules of a specific technology are not strictly followed [35], in order to produce a design that is as tightly packed as possible but still functional. Having that in mind and ignoring any possible design rules violations, the layout of Fig. 8.8 can be designed with a unit cell size of $4F \times 2F = 8F^2$, using diffusion bitlines. It is hard to see how this could be further reduced, unless using multi level cell (MLC) technology to reduce the effective cell size. Despite this, in [36] Hitachi (the only adopter of the AND Flash technology) claims for AG-AND[17] (assist gate-AND) a cell size of only $4F^2$ that can be further reduced to remarkable $2F^2$ using MLC.
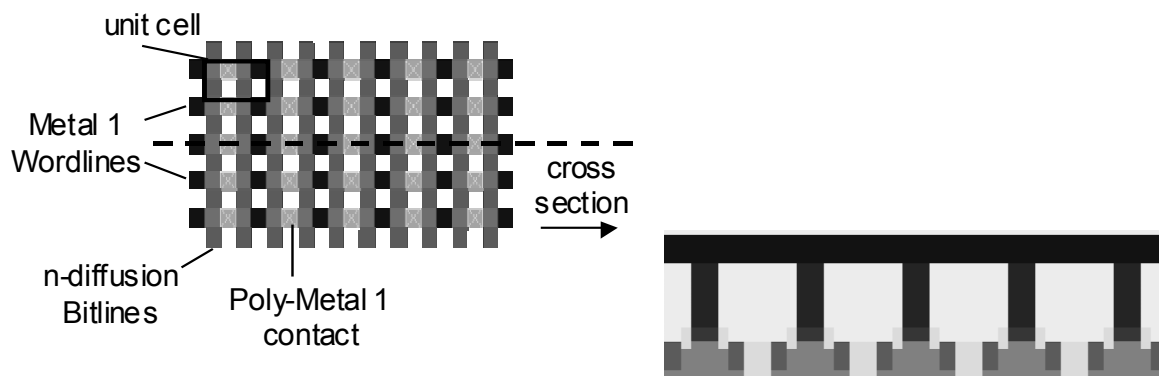
Fig. 8.8 Layout and cross-section of the AND configuration using diffusion bitlines.

---

[17] This is achieved by using assist gates in place of doped source and drain regions that, at the same time, prevent interference between neighboring cells. Normally, a trench is used for isolation (STI).

According to [34], the cell size and the read access times for AND are between those of NAND and NOR, so it is a good compromise. For FeFET application, on the other hand, AND seems to be the only way.

8.2 The NOR configuration

The NOR Flash architecture was the first of the three to be applied. It is easily derived from the 1T-1C DRAM cell structure as shown in Fig. 8.9 (see also Fig. 1.2 for the FRAM and FeFET parallels), and thus the addressing is similar. It is very well suited for code execution due to the random access mode and fast read times.
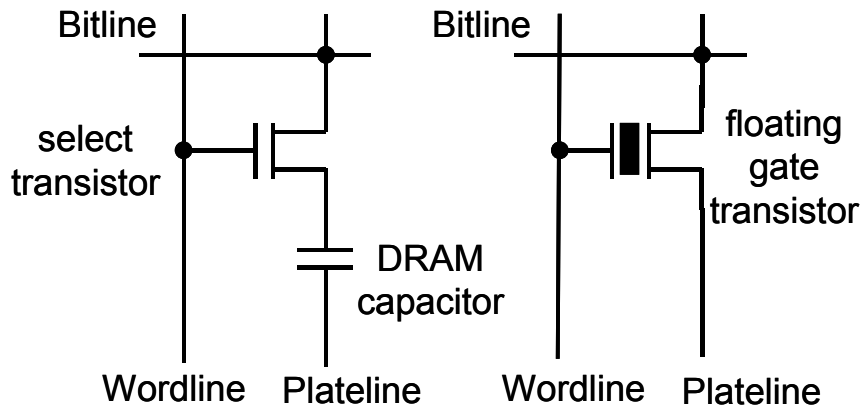
Fig. 8.9 Schematic of a DRAM (left) and a NOR Flash (right) basic cell.

The array structure is shown in Fig. 8.10. The difference between AND and NOR is that in the latter the transistors' source connect to the plate line potential.
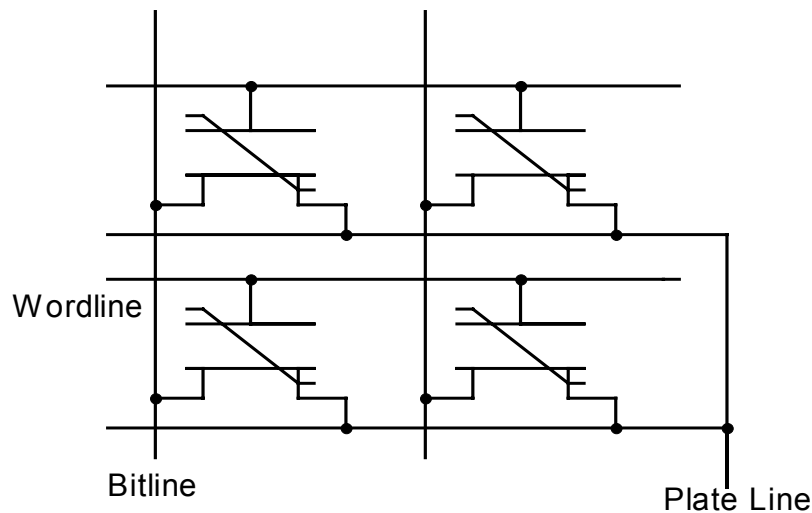
Fig. 8.10 Schematic of the NOR array configuration.

A single cell has a "±" shape. Two "±" shaped cells connected is the basic structure as shown in Fig. 8.1. Repeating this structure in the X and Y direction gives the complete memory array

layout. The unit cell size is about $5F \times 2F = 10F^2$. Being more than twice as large as NAND, has limited its application only for code execution, that requires fast read access.
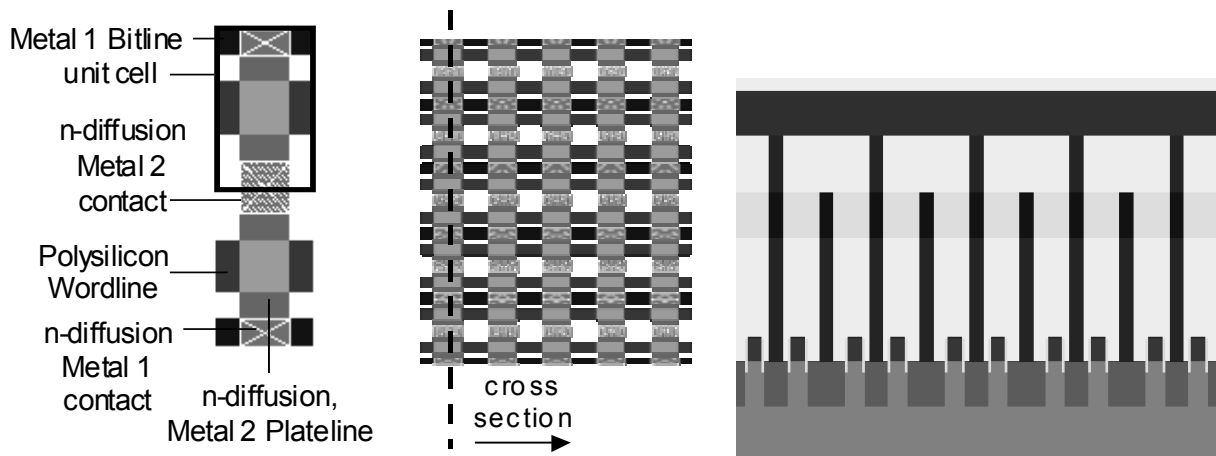
Fig. 8.11 "±" shaped cell, layout and cross-section of the NOR configuration.

8.3 The NAND configuration

Finally, NAND is the architecture with the highest packing density with a unit cell area of $2F \times 2F = 4F^2$. The schematic is shown in Fig. 8.12. Figure 8.13 shows one possibility to realize such a dense transistor array. Diffusion lines are used for the bitlines, and polysilicon for the wordlines. Metallic bitlines can be used to connect groups of bitlines, as in Fig. 8.7 for the AND configuration.
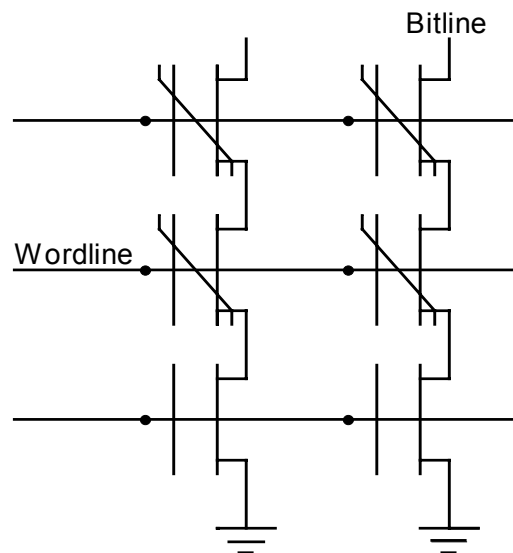
Fig. 8.12 Schematic of the NAND array configuration.
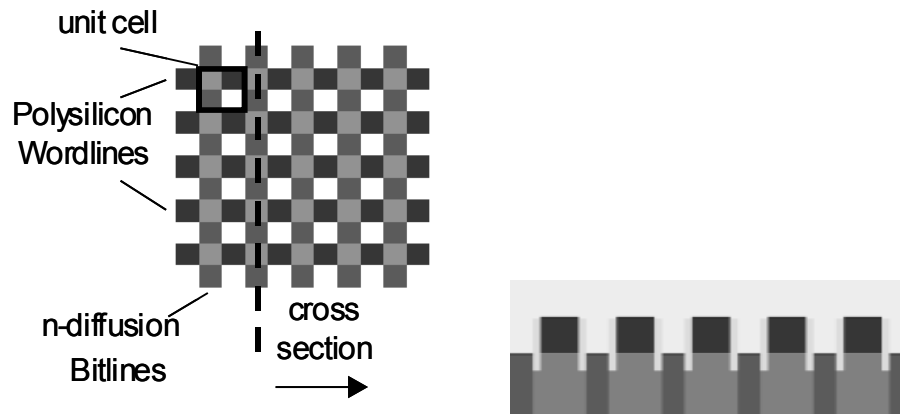
Fig. 8.13 Layout and cross-section of the NAND configuration.

The high density is achieved through the transistors sharing the source and drain area. The drawback of NAND is its slow serial read access. The reason for this is the small current that passes through all devices in the bitline during the read operation (after the transistors at the bottom of Fig. 8.12 are enabled).

8.4 Summary

None of the array configurations is clearly superior in all areas. The AND configuration, although less established in the Flash technology, has big potential in targeting the FeFET memory application and can enable both a high integration density ($8F^2$) and fast memory access (random access mode).

9 FeFET memory chip simulation

In this last chapter the design of a FeFET memory chip in schematic level will be presented, based on conclusions from previous chapters. The *positive voltage erase* method will be the programming method used and the FeFET array will have the AND configuration.

The 1-Kbit (Memory array: 32×32 bit) FeFET Chip has a word length of 16 bit. A total of 64 words can be stored / read out (1 word per cycle). About *50 ns* are needed for a write operation and another *50 ns* for a read operation, although both operations can be further optimized using a timing scheme that is adjusted according to the RC times (matrix size). A block diagram of the chip is shown in Fig. 9.1. The individual blocks of the memory chip are the memory devices (FeFETs), wordline and bitline drivers, the sense amplifiers, the row and column decoder. They will now be described separately.
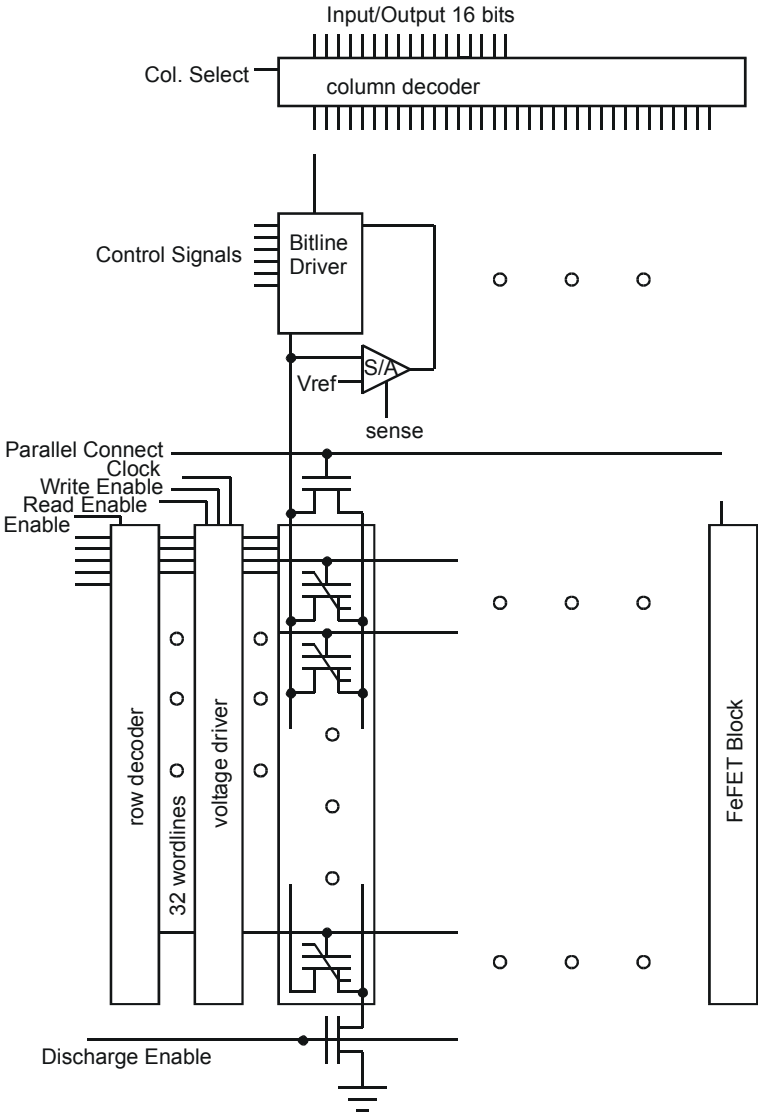
Fig. 9.1 Block diagram of a 1-Kbit memory circuit using FeFETs.

9.1 The FeFET chip building blocks

9.1.1 The row decoder

The 32 bitlines are controlled with the five inputs of the Row-Decoder and an "Enable" signal. An implementation that uses 6-input AND logic gates, is shown in Fig. 9.2 (left). However, it is more common to break the row-decoder down to smaller ones (Fig. 9.2 lower right hand), using more logic gates with fewer inputs, because of the high resistance of the series connected n-MOSFETs in the AND logic gate (Fig. 9.2 upper right hand).
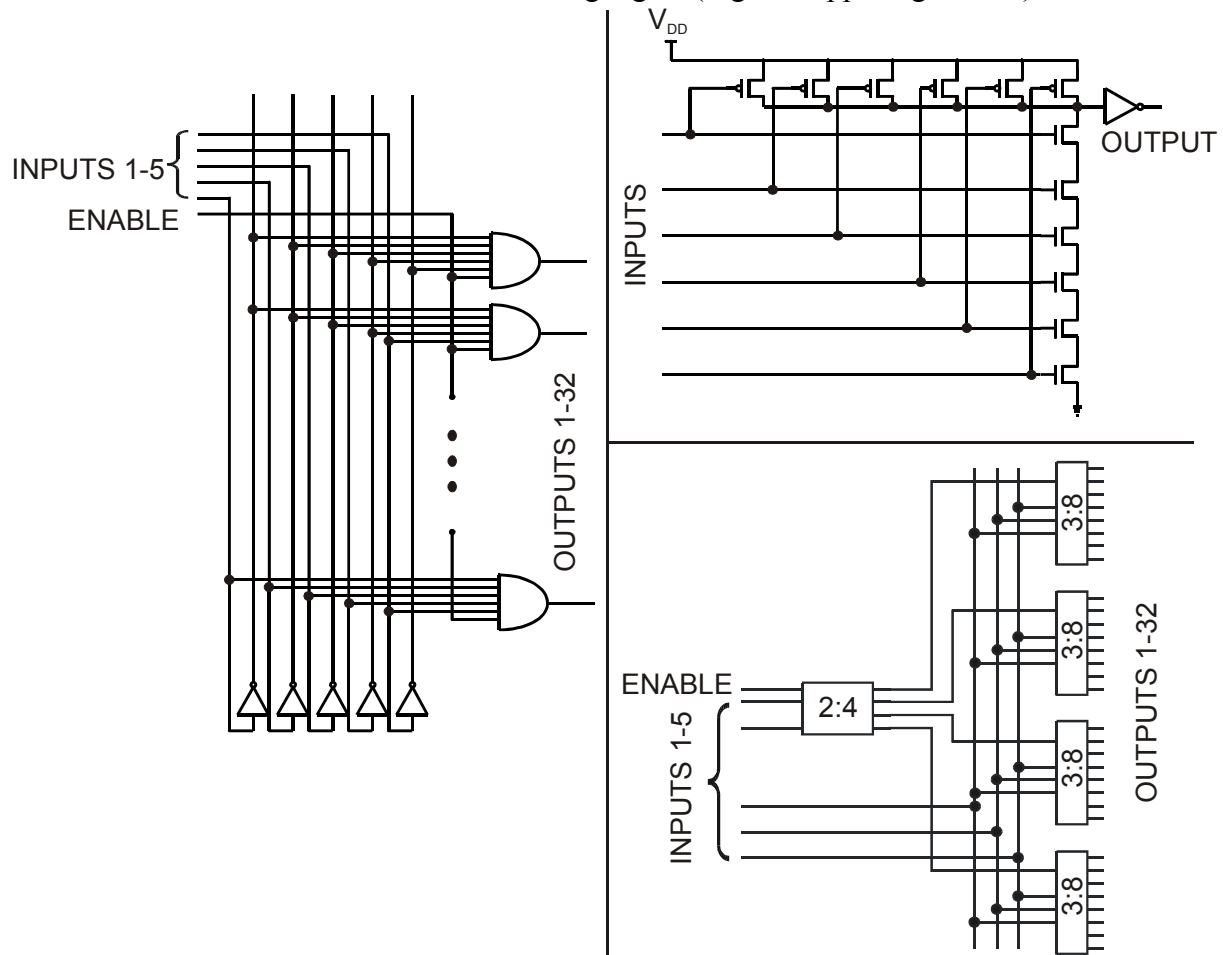
Fig. 9.2 Schematic of the row-decoder (left) consisting of 6-input AND gates (upper right hand), 5:32 row decoder consisting of one 2:4 decoder and four 3:8 decoder (lower right hand).

9.1.2 The wordline driver

After the row decoder follows a wordline voltage driver that generates the necessary voltages to drive the wordline. It has two additional inputs that set the output voltage according to the type of operation selected (read, program) and a clock input that is needed to enable half duration pulses for programming. Figure 9.4 shows the wordline driver schematic. The clock has a period of $T = 40\ ns$ (frequency is $f = 1/T = 25\ MHz$). The operating voltage is $V_{DD} = 5\ V$ and the read voltage $V_{Read} = 2\ V$.

The output of the wordline driver is one of the three signals in Fig. 9.3. The third signal is a $V_{DD}/2$ full duration pulse for protecting a cell that is not being addressed, from being erased by a charged bitline. Figure 9.5 shows a simulation of the input and output signals. During the write (read) operation the output is V_{DD} (V_{Read}) in WL1 (the addressed wordline) for $t=T$ and $V_{DD}/2$ ($0\ V$) for $t=2T$ in all other wordlines. During the read operation the output is V_{Read} in WL2 (the addressed wordline) for $t=2T$ and $0\ V$ in all other wordlines.

Fig. 9.3 Wordline driver conditions for a read/write/protect.
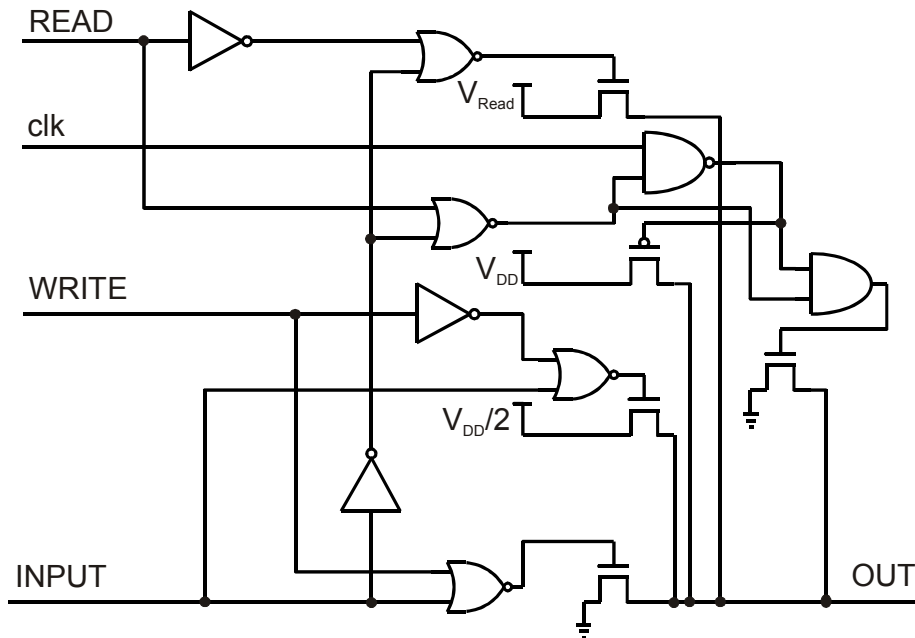
Fig. 9.4 Schematic of the wordline driver.
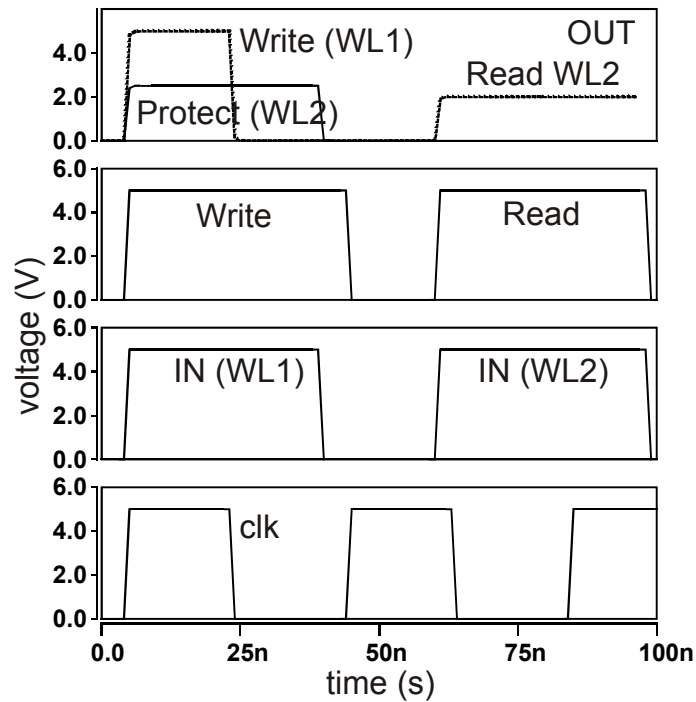
Fig. 9.5 Simulation of the wordline driver input and output signals. First, the first wordline is written and then the second read out.

9.1.3 The column decoder

The Column-Decoder is bi-directional (Fig. 9.6). It consists of transmission gates, that conduct both ways. An inverter selects which of the 2 words (16 bits) is to be written / read out.
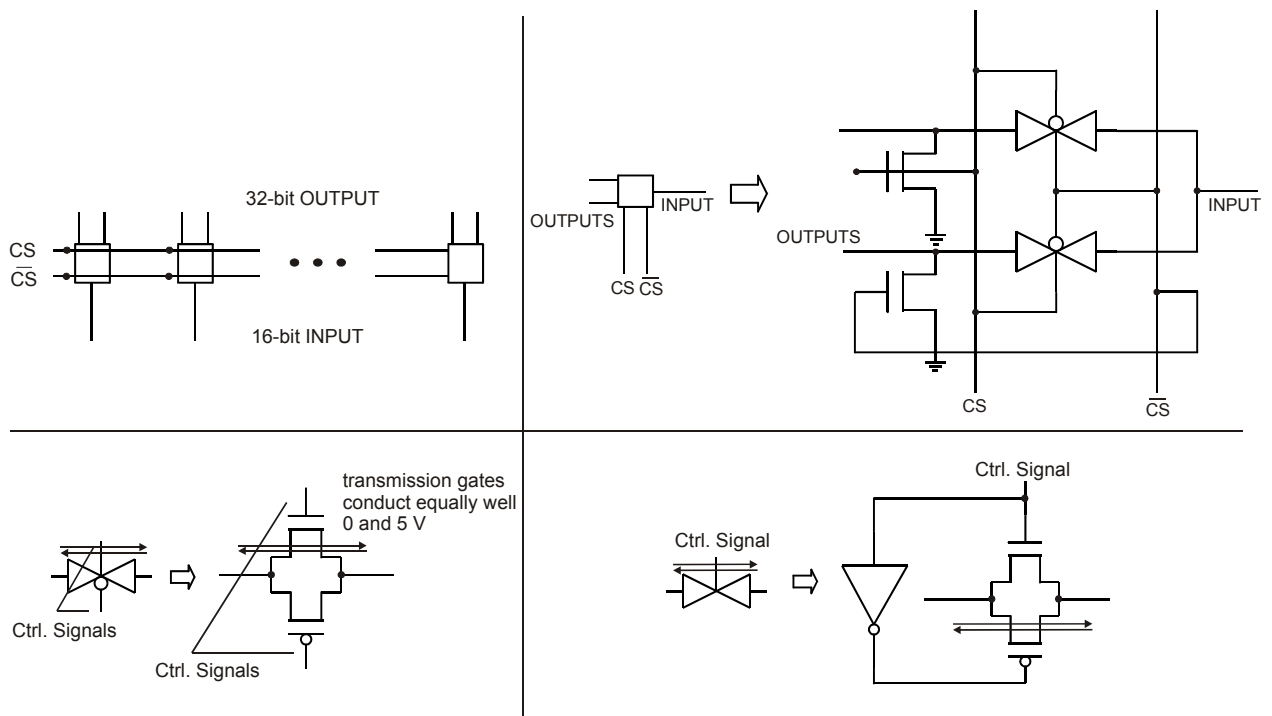
Fig. 9.6 Schematics of the column decoder (upper left hand), a 2-1 column decoder using two transmission gates and two pull-down MOSFETs (upper right hand), a transmission gate (lower left hand) and a transmission gate with inverter (lower right hand).

9.1.4 The bitline driver

The bitline driver has the purpose of pre-charging the bitline to the drain voltage V_{DS} during the read out operation, or during the write operation to either the operating voltage (for the cells to be erased), *0 V* (for the cells to be written) or $V_{DD}/2$ (in order to protect from disturbance the unselected cells). It also interfaces with the read amplifier as can be seen in Fig. 9.7.
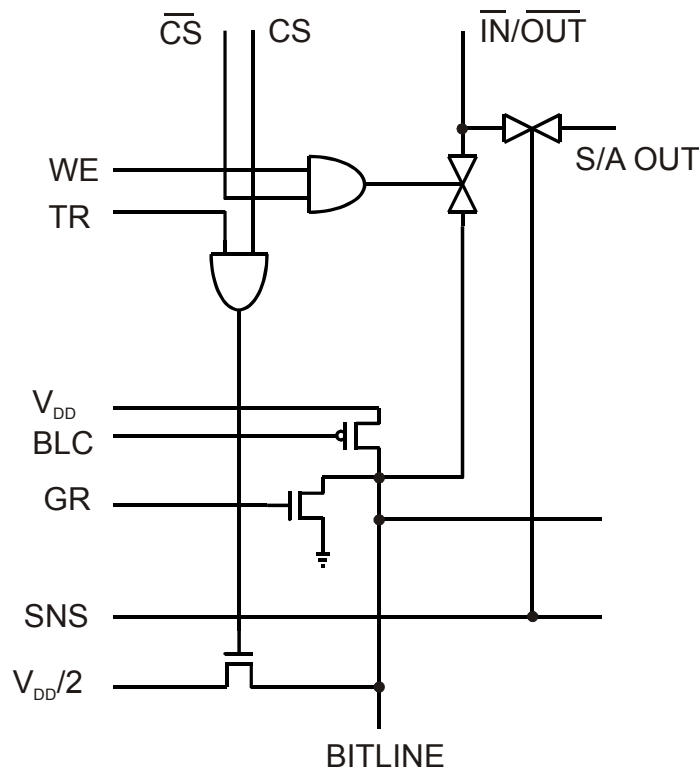
Fig. 9.7 The bitline driver that connects the memory matrix to the column decoder and interfaces with the sense amplifier.

9.1.5 The sense amplifier

For read out the sense amplifier of Fig. 9.8 is used. Voltage sensing is applied, that is the input voltage is compared with a constant reference voltage. The sense amplifier consists of two cross-coupled inverters that have two stable states and two outputs. One of them is used as the amplifier output. The transmission gates are enabled first (EN) and then the SNS (sense) input activates the cross-coupled inverters before the amplifier reaches a steady state. A simulation of the amplifier operation is shown in Fig. 9.9. The reference voltage is *3 V* and is compared first with *2 V (Vout="0")* and then with *5 V (Vout="1")*. The output is active as long as the SNS input is enabled.
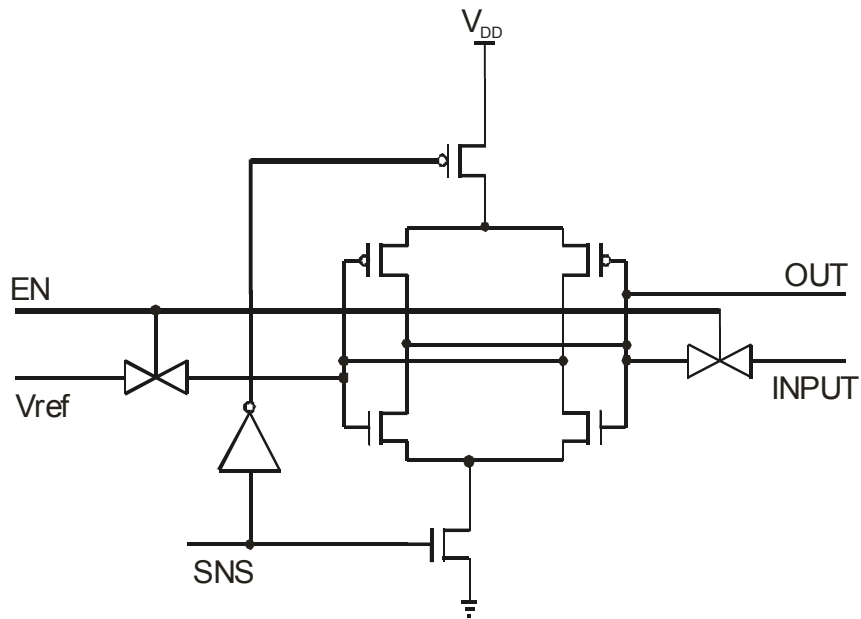
Fig. 9.8 Sense amplifier using two cross coupled inverters.
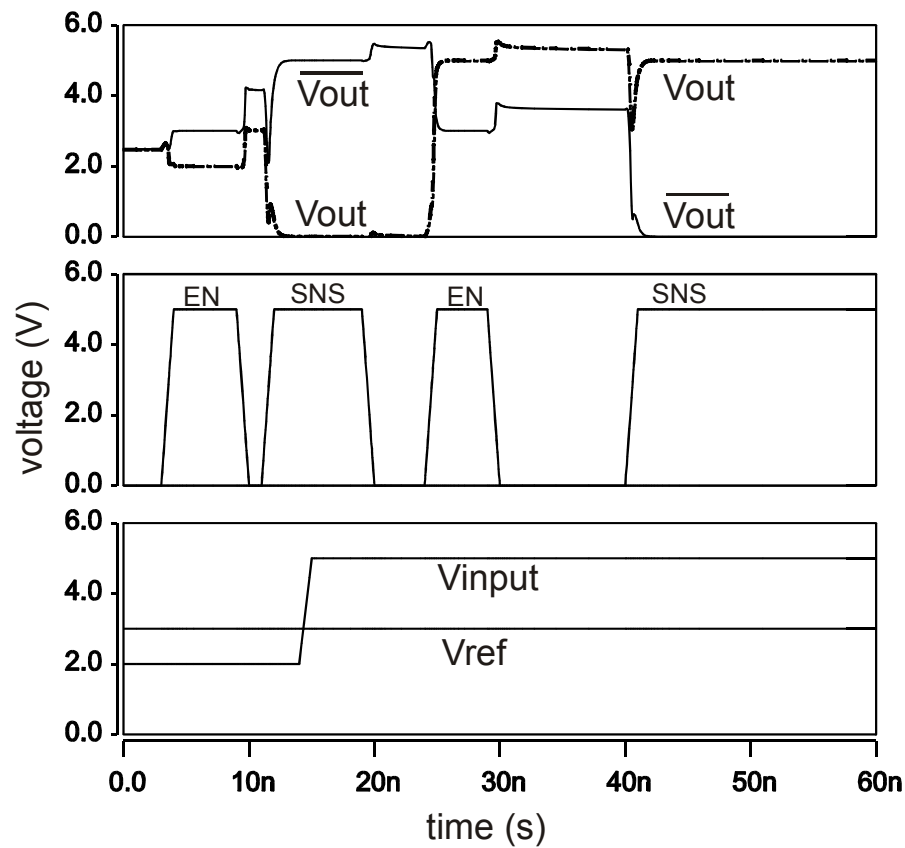
Fig. 9.9 Simulation of the sense amplifier of Fig. 9.8, comparing the signals Vinput and Vref.

The various blocks of the design are listed in Table 9.1 along with the number of transistors they consist of.

Function Block	Number of MOSFET, FeFETs
FeFET Array	$32^2 + 32 = 1056$
Row Decoder	458
Wordline Drivers	$32 \cdot 37 = 1184$
Bitline Drivers	$32 \cdot 23 + 2 = 738$
Sense Amplifiers	$32 \cdot 16 = 512$
Column Decoder	$16 \cdot 6 + 2 = 98$
Total	4046

Table 9.1 The building blocks of the designed 1-Kbit memory chip (the two additional transistors in the bitline driver and the column decoder are for the inverters, and in the FeFET array the pass transistor are included).

Having described the chip sub-systems, the read and write operations will now be explained, in more detail including the control signals and the bitline driver as opposed to the FeFET-centric approach in chapter 7. Only two wordlines and bitlines (there are a total of 32×32 in a Kbit chip) will be examined. In Fig. 9.10 (left) one bitline is shown. Instead of a pass transistor, a transmission gate (4 transistors) is used, because it passes equally well V_{DD} and *GND*. A single pass transistor can be used if the FeFETs are less susceptible to disturbance like for example the FeFET in Fig. 7.18. The FeFET used in the following simulations is that in Fig. 7.17. The row-decoder, column-decoder and wordline driver will not be included in the description as they have been described separately in previous sections. The input and output are inverted, that is a *0 V* is interpreted as a "write" (or a FeFET at state "1" during read out) and a 5 *V* is interpreted as an "erase" (or a FeFET at state "0" during read out). Table 9.2 lists the various control, input/output signals.

| Signal	Full name	Description
CS	Column select	Selects one of two words to access
WE	Word enable	Is input to gate AND1 that controls the transmission gate TG2
TR	Trigger	Is input to gate AND2 that controls T3 to apply $V_{DD}/2$ to bitline
BLC	Bitline charge	Controls T1 that applies V_{DD} to bitline
GR	Ground	Controls T2 that discharges the bitline to ground
ENABLE	Enable	Enables the transmission gates in sense amplifier
SNS	Sense	Applies the sensing pulse in the sense amplifier
PC	Parallel connect	Connects FeFET source and drain to bitline potential
DS	Discharge select	Controls T4 that discharges the bitline to ground

Table 9.2 Control signals with short description.

In the following simulation the write operation is from *0-50 ns* and the read operation from *50-100ns*. As in Fig. 7.8, it will be attempted to write a "1" and a "0" to the FeFETs of the first wordline and read it afterwards. The complete diagram of the input/output signals is shown in Fig. 9.10 (right). The Bitline signals are shown in Fig. 9.12 for different bitline capacitances.

9.2 Write Operation

First the bitlines are charged, before the wordlines can be addressed. The reverse order would cause disturbance problems.

The input pulse $\overline{IN}$ is applied at the input. TR is "1"[18] during writing. CS is "0" (first word is selected) and so T3 will remain "0" (AND2 output is "0") for the entire first word (first 16 bitlines). PC is "1", so TG3 is enabled, and any voltage applied to any wordline will be applied to the gate of all FeFETs in that wordline. $\overline{CS}$ is "1" and after WE becomes "1", transmission gate TG2 (AND2 output is "1") passes the input pulse to the bitline. Because $\overline{IN1}$ is "0" ($\overline{IN2}$ is "1") the bitline is not charged (charged to "1").

WL1 is charged to V_{DD} ("write" or "erase") and WL2 to $V_{DD}/2$ ("protect"). This way the first FeFET of WL1 is "written", the second (BL2) "erased", and those of WL2 "protected". At the end of the write operation GR becomes "1" to discharge the bitline to ground.

9.3 Read-out Operation

The read operation starts with BLC charging the bitline to a potential (here V_{DD}). Only one inverter is used to enable the p-type transistors T1 for all bitlines. After the bitline is charged, the WL to be read out is charged to V_{Read} (here *2 V*) and at the same time DS becomes "1" to enable T4 to discharge the bitline through a conducting FeFET. The reading pulse could be synchronous with DS but, because of the way the wordline driver works, its duration is the same as the input pulse to the wordline driver (Fig. 9.5). Next, EN goes to "1" and the transmission gates in the S/A are enabled. Finally, the sense pulse SNS makes the amplifier lock to either "1" or "0" by comparing the reference voltage (here $V_{DD}/2$) to the bitline potential. The bitline potential is "0" ("1") for the FeFET of BL1 (BL2) whose state was "1" ("0"). So the S/A of BL1 outputs "0" ($\overline{\mathrm{Vref} > "0"} \to "0"$) and that of BL2 "1" ($\overline{\mathrm{Vref} > "1"} \to "1"$). SNS also enables TG1 so that the output of the S/A is passed to the output.

---

[18] For the signal levels "1" refers to *5 V* (V_{DD}) and "0" to *0 V*.
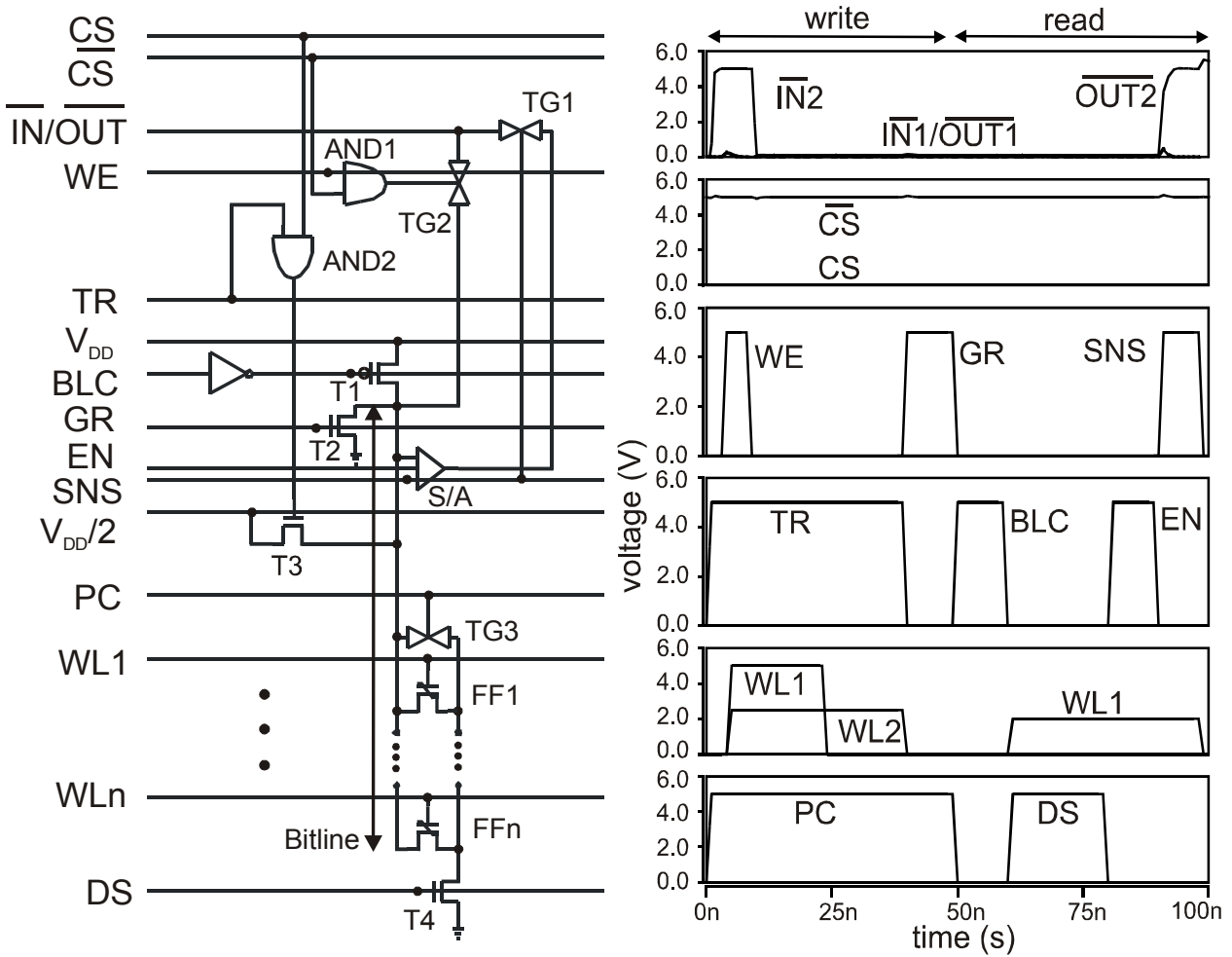
Fig. 9.10 Schematic of (left) a complete bitline, (right) the input and output signals.

9.4 Effect of the bitline capacitance on the read-out operation

The resistance and capacitance values for the bitlines and wordlines were extracted using a layout program [44] for the CMOS *0.12 μm* technology specified in the program. Here are the values used:

	Capacitance	Resistance
n-diffusion	350 aF/μm^2	25 Ohm/square
1st metallization	200 aF/μm^2	0.05 Ohm/square
2nd metallization	180 aF/μm^2	0.05 Ohm/square

Table 9.3 Surface capacitances and resistances for a 0.12 μm CMOS technology [44].

Now, assuming chip dimensions (bitline length × wordline length) of *100 μm × 250 μm* gives the following cases for the resistances and capacitances of Fig. 9.11.

	2 Metal.	1 Metal., n-dif. BL
C_{BL}	18 fF	35 fF
C_{BLC}	3.5 fF	3.5 fF
R_{BL}	5 Ohm	2.5 kOhm
C_{WL}	50 fF	50 fF
C_{WLC}	8.5 fF	8.5 fF
R_{WL}	12.5 Ohm	12.5 Ohm

Table 9.4 Bitline and Wordline resistances, capacitances and coupled capacitances.

The simulation for both cases of Table 9.4 doesn't make much difference to the results as shown in Fig. 9.12. This means that the high bitline resistance doesn't affect the discharging of the bitline as much as the bitline capacitance. The simulation was repeated varying only the bitline capacitance C_{BL} and the results are shown in Fig. 9.12. The bitline discharges in very short time (a few *ns*), if the bitline capacitance is small. After increasing the bitline capacitance from *18 fF* to *200 fF* (for a bitline 11 times longer) the bitline discharges only to about *2 V*. Setting the reference voltage in the S/A to *3.5 V* would be enough to ensure a correct read-out, albeit with a higher noise margin. Note again that the FeFET at BL1,WL1 is set to "1" (thus, BL1 discharges at read out) and the one at BL2,WL1 to "0" (thus, BL2 does not discharge at read out).
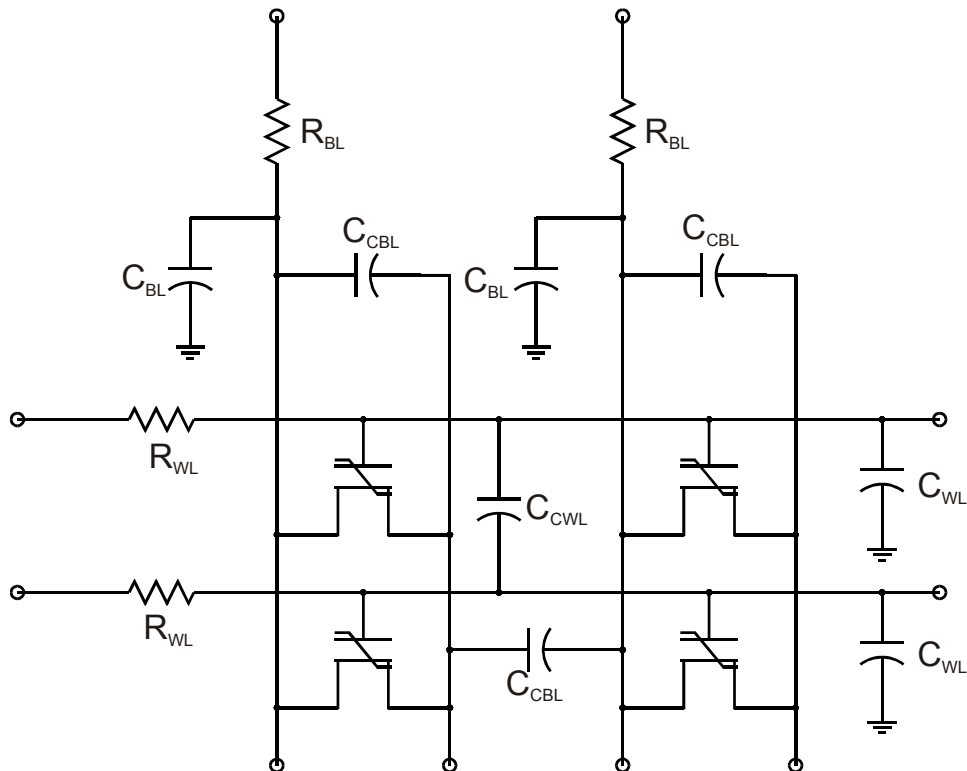
Fig. 9.11 Bitline and Wordline resistances and capacitances and coupled capacitances.
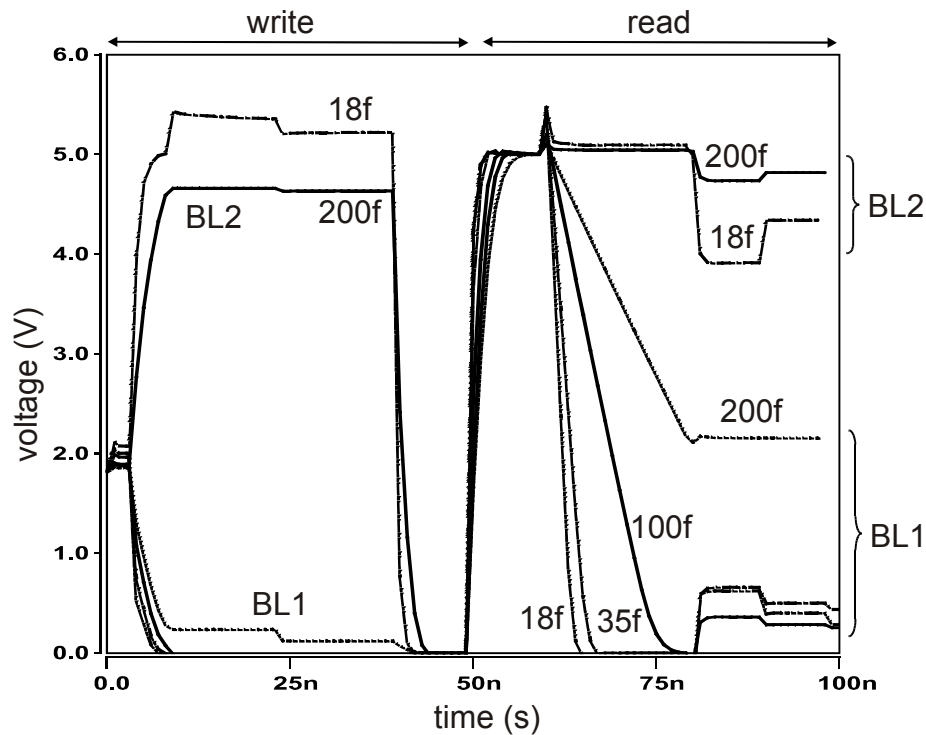
Fig. 9.12 Simulation of the bitline potential for different bitline capacitances leads to a slower read out (capacitances are in F).

During the write operation one word (16 bits in one row) can be written at a time. A total of 64 words are stored and can be read out in timeframes of *50 ns*. This timeframe can be further optimized for a particular chip size and is limited by RC constants due to the parasitic bitline and wordline capacitors. After all, RC times are the limiting factor to the speed of operation of any memory chip. Still, with *50 ns/word*, data rates of *40 MB/s* are possible. Another way to increase the data rate is by using a bigger word length. The simulation time (1 write and 1 read cycle) for the 1-Kbit chip is in the order of a few hours on a SUN BLADE 100 workstation.

9.5 Summary

The implementation of a FeFET based 1-Kbit memory chip in schematic level was detailed. The new programming concept contributed to the simplicity of the design and to high data rates, as a result of the elimination of a separate erase operation. For bigger memories, increasing the FeFET array leads to higher RC times and thus to slower read/write access times that must be taken into account in the design process.

Summary and outlook

The ferroelectric field effect transistor as a non-volatile memory device is a promising candidate for a future generation memory device, offering several advantages compared to other alternatives, as shown in this thesis.

The low data retention times of the FeFET were discussed and attributed to the reduction of the polarization due to the depolarization field in the ferroelectric and leakage current through the gate stack. The leakage current leads to local charge compensation and eventually to the reduction of the remnant polarization. This current was modeled assuming a worst case Schottky limit, and minimum retention times were calculated.

The aspect ratio of the ferroelectric area to the dielectric was also studied. Simulations showed that the remnant polarization can be increased, and at the same time the depolarization field reduced, by choosing a ferroelectric layer with a much smaller area than the oxide layer. Although not the ideal solution to the FeFET, it has been reported that it leads to higher retention times.

The prospects of device miniaturization were then investigated and two approaches were suggested for scaling the FeFET to smaller dimensions. The first is *constant gate stack scaling*, which keeps the gate stack unchanged and thus the gate voltage does not have to scale either. The other is *variable gate stack scaling* and requires the gate stack to be modified by scaling the layers with different factors. With *constant gate stack scaling* it was possible to simulate a scaled device down to *22nm* (end of the ITRS roadmap 2002). The difference from *constant field scaling* (used in MOSFET scaling), is that the drain voltage is not scaled (lateral field increases). The memory window does not change much, even after scaling through several technology nodes. In *variable gate stack scaling*, besides the gate and drain voltages, the gate stack layers' thicknesses are modified too. To determine the optimal thicknesses the parameter dependence was simulated. With this scaling approach, the on/off current ratio drops considerably with every scaling step, due to the decreasing operating voltage and the resulting smaller hysteresis sub-loop. As a result, *variable gate stack scaling* can only be applied for a few scaling nodes compared to the more promising *constant gate stack scaling*.

A new programming concept was suggested by introducing a *positive voltage erase* method as opposed to *negative gate erase*. For this new concept a different substrate doping is necessary. The main advantage is a faster write access due to the elimination of a separate erase operation. The downside is that the power consumption is higher. The programming concept was used in a memory design with FeFETs arranged in the AND matrix structure. This particular configuration is convenient, because it enables the FeFET source and drain to be set to the same potential.

The memory integration density was then examined and it was shown that the AND architecture can be made to have a very dense structure, similar to the DRAM memory density.

Finally, the new programming concept suggested was realized in a 1-Kbit memory chip at schematic level, including the FeFET matrix and the peripheral electronics. Each building block was described separately, and the chip functionality was verified by simulation of one write and one read operation.

The concept of the FeFET as a memory device holds a lot of promise for non-volatile memories with faster access times. Despite the known problem of short retention times, some companies keep it in their roadmaps for the time when FRAM scaling runs out. Until then, it is hoped that improved processing and new materials will solve the existing problems. The

scaling of the FeFET should then only be limited by the limits of lithography (with *constant gate stack scaling*), provided short channel leakage can be dealt with, perhaps with the introduction of a back gate. As for switching speeds, the inherent fast switching of the ferroelectric is in line with the fast charging and discharging of dielectric capacitors in DRAM memories. The combination of the fast memory access with the high integration density and the scaling prospects of the FeFET make it a promising candidate for next generation non-volatile memories.

Still, research is needed at this point to increase retention times. For that, the origins of retention loss must be further examined, for example, as suggested here by limiting the leakage current through the use of better buffers and less leaky ferroelectrics, and by reducing the depolarization field by, for example, using different aspect ratios for the oxide and the ferroelectric layers in the gate stack.

References

[1]   I. M. Ross, US Patent 2,791,760, (1957).

[2]   J. L. Moll, Y. Tarui, Trans. on Electron Devices, **ED-10**, 338, (1963).

[3]   S.-Y. Wu, A New Ferroelectric Memory Device, Metal-Ferroelectric-Semiconductor Transistor, IEEE Trans. on Electr. Dev., **21**, 8, (1974).

[4]   T. Y. Tseng, S. Y. Lee, Improvement in retention time of metal-ferroelectric-metal-insulator-semiconductor structures using MgO doped $Ba_{0.7}St_{0.3}TiO_3$ insulator layer, Applied Physics Letters, **83**, 5, (2003).

[5]   H. Ishiwara, Current Status and Prospects of FET-Type Ferroelectric Memories, FED Journal, 11, (2000).

[6]   BSIM3v3.2.2 Manual, UC Berkeley, (1999).

[7]   Y. Tsividis, Operation and Modeling of the MOS Transistor, McGraw-Hill, (1999).

[8]   M. Ullmann, Ferroelektrische Feldeffekttransistoren: Modellierung und Anwendung, doctoral thesis, Universität der Bundeswehr Hamburg, (2002).

[9]   Hermann Kohlstedt, Hiroshi Ishiwara, in: Nanoelectronics and Information Technology, R. Waser (Ed.), Wiley-Vch, (2003).

[10]   P. Würfel, I. P. Batra, Depolarization Effects in Thin Ferroelectric Films, Ferroelectrics, **12**, pp. 55, (1976).

[11]   M. Takahashi et. al., Analysis and Improvement of Retention Time of Memorized State of Metal-Ferroelectric-Insulator-Semiconductor Structure for Ferroelectric Gate FET Memory, Jpn. J. Appl. Phys., **40**, pp. 2923, (2001).

[12]   I. P. Batra, B. D. Silverman, Thermodynamic Stability of Thin Ferroelectric Films, Solid State Communications, **11**, pp. 291, (1972).

[13]   T. Yamaguchi et. al., Improvement of Memory Characteristics of Metal-Ferroelectrics/Insulating Buffer Layer/Semiconductor Structures by Combination of Pulsed Laser Deposited $SrBi_2Ta_2O_9$ Films and Ultra-Thin SiN Buffer Layers, Jpn. J. Appl. Phys., **39**, pp. 2058, (2000).

[14]   J.-P. Han et. al., Buffer layer dependence of memory effects for $SrBi_2Ta_2O_9$ on Si, Integrated Ferroelectrics, **34**, (2001).

[15]   M. Takata et. Al, Fundamental Characteristics of New Non-Volatile Memory with Extremely High Density Metal Quantum Dots, ECS International Semiconductor Technology Conference, (2002).

[16]   H. Goebel et. Al, Distribution Function Integral Method for Modeling Ferroelectric Devices, ISSDM, (1999).

[17]   M. Ullmann et. al., A BSIM3v3 and DFIM Based Ferroelectric Field Effect Transistor Model, IEICE Trans. Electron., **E83-C**, pp. 1324, (2000).

[18]   T. P. Ma, J.-P. Han, Why is Nonvolatile Ferroelectric Memory Field-Effect Transistor Still Elusive?, IEEE Electron Devices Leters, **23**, 7, (2002).

[19]   J.C. Shin et Al., Characterization of Pb(Zr,Ti)O3 Thin Films Prepared by Metal-Organic Chemical-Vapor Deposition Using a Solid Delivery System, Journal of the Korean Physical Society, (1999).

[20]   Ali Sheikholeslami, Transient Modeling of Ferroelectric Capacitors for Semiconductor Memories, MSC thesis, University of Toronto, (1994).

[21]   M. Grossman, Imprint: An Important Failure Mechanism of Ferroelectric Thin Films in View of Memory Applications, Dissertation RWTH Aachen, (2001).

[22]   http://www.fujitsu.com/services/microelectronics/technical/fram/

[23]   International Technology Roadmap for Semiconductors, http://public.itrs.net/, (2002).

[24]   S.L. Miller and P. J. McWhorter, Physics of the ferroelectric non-volatile memory field effect transistor, Journal of Applied Physics, **72**, pp. 5999, (1992).

[25]   R. Kretschmer and K. Binder, Phys. Rev., **B20**, pp. 1065, (1979).

[26]   J. Junquera, P. Ghosez, Critical thickness for ferroelectricity in perovkite ultrathin films, Letters to Nature, **422**, pp. 506, (2003).

[27]   T. Haneder et. al., Gate Stacks for Low Voltage Ferroelectric Field Effect Devices Based on $Pt/SBT/CeO_2/Si(100)$. MRS Proceedings Ferroelectric Thin Films IIX, (1999).

[28]   Intel Strataflash, http://www.intel.com/design/flash/

[29]   S.-M. Yoon, E. Tokumitsu, H. Ishiwara, Improvements of Memory Retention Characteristics in Ferroelectric Neuron Circuits Using a $Pt/SrBi_2Ta_2O_9/Pt/Ti/SiO_2/Si$ Structure-Field Effect Transistor as Synapse Device, Jpn. J. Appl. Phys., **39**, pp. 2119, (2000).

[30]   S.-K. Kang, H. Ishiwara, Data Retention Characteristics of Metal-Ferroelectric-Metal-Insulator-Semiconductor Diodes with $SrBi_2Ta_2O_9$ Ferroelectrics and Al_2O_3 Buffer Layers, Jpn. J. Appl. Phys., **41**, pp. 2094, (2002).

[31]   M. Fitsilis et. Al, New Concept for using Ferroelectric Transistors in Nonvolatile Memories, Integrated Ferroelectrics, **60**, pp. 45-58, (2004).

[32] H. Ishiwara, T. Shimamura, E. Tokumitsu, Proposal of a Single-Transistor-Cell-Type Ferroelectric Memory Using an SOI Structure and Experimental Study on the Interference Problem in theWrite Operation, Jpn. J. Appl. Phys., **36**, pp. 1655, (1997).

[33] Y.S.Yang et. al., Design of a single-transistor-type ferroelectric field effect transistor memory, Journal of the Korean Physical Society, **40**, pp. 701 (2002).

[34] Memory Report, ICE, (1997).

[35] E. Sicard, S. D. Bendhia, Deep sub-micron CMOS circuit design simulator in hands, http://intrage.insa-tlse.fr/~etienne/Microwind/book.html, to be published.

[36] Hitachi, http://www.hitachi.com/New/cnews/031211_031211.pdf

[37] A. K. Tagantsev et. al., Polarization fatigue in ferroelectric films : Basic experimental findings, phenomenological scenarios, and microscopic features, Journal of Applied Physics, **90**, 3, (2001).

[38] ATLAS User's Manual, Silvaco International, (2000).

[39] M.-D. Ker, C.-Y. Chang, H.- C. Jiang, Design of Negative Charge Pump Circuit with Polysilicon Diodes in a 0.25-µm CMOS Process, IEEE Asia-Pacific Conference on ASIC. Proceedings, 145, (2002).

[40] Kijima, Jpn. Journal of Applied Phys., **40**, pp. 2977, (2000). http://www.chipidea.com/esscirc2003/papers/essderc/oral_presentations/032_183.pdf

[41] T. Li, S. T. Hu, B. Ulrich and D. Evans, Mat. Res. Soc. Symp. Proc., **688**, C4.3, (2002).

[42] Toshiba,http://www.semicon.toshiba.co.jp/eng/prd/memory/pdf/nandflash_what_e.pdf

[43] S. H. Lee et. al., Improvement of Data Retention Time using DRAM cell with Metallic Shield Embedded (MSE)-STI for 90nm Technology Node and beyond, ISSDERC– ISSCIRC Proc. Paper, (2003).

[44] Microwind, http:/www.microwind.org

[45] S. Parashar et. al., Electrical properties of ferroelectric YMnO3 films deposited on n-type Si(111) substrates, Journal of Phys. D: Appl. Phys., **36**, (2003).

[46] Sinharoy S. et. al., Growth and the microstructural and ferroelectric characterization of oriented $BaMgF_4$ thin films, IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, **38**, Issue: 6, pp. 663, (1991).

[47] Fujisaki, Y.; Iseki, K.; Ishiwara, H., Long retention performance of a MFIS device achieved by introducing high-k $Al_2O_3/Si_3N_4/Si$ buffer layer, Mater. Res. Soc. Symposium Proceedings, **786**, pp. 297, (2004).

[48] T. A. Rost, H. Lin, and T.A. Rabson, Appl. Phys. Lett., **59**, pp. 3654 (1991).

[49]   D. R. Lampe et al., Ferroelectrics, **133**, pp. 61, (1992).

[50]   Nakamura Y. Nakao, A. Kamisawa, and H. Takasu, Integrated Ferroelctrics, **9**, pp. 179 (1995).

[51]   Y. Nakao, T. Nakamura, A. Kamisawa, and H. Takasu, Integrated Ferroelectrics, **6**, pp. 23, (1995).

[52]   B.-E. Park, S. Shioriki, E. Tokumitsu, and H. Ishiwara, Jpn. J. Appl. Phys., **37**, pp. 5145, (1998).

[53]   T. Hirai, K. Teramoto, K. Hagashima, H. Koike, and Y. Tarui, Jpn. J. Appl. Phys., **34**, pp. 4163, (1995).

[54]   E. Tokumitsu et al., IEEE Electron. Dev., **18**, pp. 160, (1997).

[55]   R.A. McKee, F.J. Walker, and M.F. Chisholm, Phys. Rev. Lett., **81**, pp. 3014, (1998).

[56]   E. Tokumitsu, G. Fuji, and H.Ishiwara, Appl. Phys. Lett., **75**, pp. 575, (1999).

[57]   S.-M. Yoon, E. Tokumitsu, and H. Hiroshi, IEEE Electr. Devices, **20**, pp. 526, (1999).

[58]   S.-M. Yoon, E. Tokumitsu, and H. Hiroshi, IEEE Electr. Devices, **47**, pp. 1630, (2000).

[59]   E. Tokumitsu, G. Fuji, and H. Ishiwara, Jpn. J. Appl. Phys., **39**, pp. 2125, (2000).

[60]   H. Ishiwara, Mat. Res. Soc. Symp. Proc., **596**, pp. 427, (2000).

[61]   H. Ishiwara, in: Handbook of Thin Film Devices, Maurice H. Francombe (Ed.), Academic Press, 2000.

[62]   Z. Yu et Al., J. Vac. Sci. Technol., **B 18**, pp. 2139, (2000).

[63]   B H. Park, B. S. Kang, S. D. Bu, T. W. Noh, J. Leeand, W. Jo, Nature, **401**, pp. 682, (1999).

[64]   T. Kijima et. al., Jpn. Journal of Applied Phys., **40**, pp. 2977, (2001).

[65]   T. Li, S. T. Hu, B. Ulrich, and D. Evans, Mat. Res. Soc. Symp. Proc., C4.3, pp. 688, (2002).

[66]   Ramtron International, http://www.ramtron.com.

[67]   Cypress Semiconductor, http://www.cypress.com.

[68]   T. Mihara, US Patent 5,515,311, (1996).

[69]   B. Tuttle et. al., Highly Oriented, Chemically Prepared Pb(Zr,Ti)O$_3$ Thin Films, J. Am. Ceram. Soc., **76**, pp. 1537, (1993).

[70]   A. Roelofs et. al., Piezoresponse force microscopy of lead titanate nanograins possibly reaching the limit of ferroelectricity, Applied Physics Letters, **81**, 27, pp. 5231, (2002).

[71]   Saber mixed mode simulator, Synopsys Inc., http://www.synopsys.com/products/mixedsignal/saber/saber_ds.html.

[72]   Bart Van Zeghbroeck, Principles of Semiconductor Devices, http://ece-www.colorado.edu/~bart/book/, to be published, (2004).

[73]   N. Braithwaite and G. Weaver, Electronic Materials, 2nd Ed., Butterworth-Heinemann, (1998).

[74]   32. IFF-Ferienkurs, Neue Materialien für die Informationstechnik, Forschungszentrum Jülich, (2001).

[75]   W. Liu, MOSFET Models for SPICE Simulation, John Wiley & Sons, (2001).

[76]   U. Hilleringmann, Silizium-Halbleitertechnologie, Teubner Studienbücher, (1999).

[77]   F. Jona, G. Shirane, Ferroelectric Crystals, Dover Publications, (1993).

[78]   T. Nakamura, Y. Nakao, A. Kamisawa, H. Takasu, A Single-Transistor Ferroelectric Memory Cell, ISSCC, (1995).

[79]   S.-M. Yoon and H. Ishiwara, IEEE Trans. on Electr. Dev., **48**, 9, pp. 2002, (2001).

[80]   S. Sinharoy et. al., Growth and characterization of ferroelectric BaMgF$_4$ films, J. Vac. Sci. Technol., A 9 (3), pp. 409, (1991).

[81]   Chandra P. et. al., Thickness-dependence of the coercive field in ferroelectrics, preprint at http://xxx.lanl.gov/abs/cond-mat/0206014l, (2002).

[82]   I. D. Mayergoyz, Mathematical models of hysteresis, Springer, (1991).

[83]   D. Damjanovic, Ferroelectric, dielectric and piezoelectric properties of ferroelectric thin films and ceramics, Rep. Prog. Phys., **61**, pp. 1267, (1998).

[84]   A. K. Tagantsev and I. A. Stolichnov, Injection-controlled size effect on switching of ferroelectric thin films, Applied Phys. Letters, **74**, 9, pp. 1326, (1999).

[85]   K. Eisenbeiser et. al., Field effect transistors with SrTiO$_3$ gate dielectric on Si, Applied Phys. Letters, **76**, 10, pp. 1324, (2000).

[86]   H. K. Kim, N. A. Basit, Inter. J. of High Speed Electr. and Sys., **10**, 1, pp. 39, (2000).

[87]   H. Sugiyama et. al., Jpn. J. Appl. Phys., **39**, pp. 2131, (2000).

[88]   I. Kashiwagi, A study on dependence of amorphous or crystalline rare earth gate oxides on Si surface orientation, Master thesis, Tokyo Institute of Technology, (2003).

[89]   G. D. Wilk et. al., High-k gate dielectrics: Current status and materials properties considerations, J. of Appl. Phys., **89**, 10, pp. 5423, (2001).

[90]   R. Schroeder et. al., All-Organic Permanent Memory Transistor using an Amorphous, Spin-Cast Ferroelectric-like Gate Insulator, Adv. Mater., **16**, 7, pp. 663, (2004).

Symbol List

Symbol	Description	Unit
A	Area of capacitor, transistor channel	m^2
A	Amount of positive Polarization	-
A	Richardson constant	A/m^2K^2
A^*	Effective Richardson constant	A/m^2K^2
A_{Fe}	Area of ferroelectric	m^2
A_{Ox}	Area of oxide	m^2
A_{Ratio}	Area ratio A_{Ox} / A_{Fe}	-
C_{Fe}	Ferroelectric capacitance	F
C_{lin}	Linear part of ferroelectric capacitance	F
C_{nonlin}	Nonlinear part of ferroelectric capacitance	F
C_{Ox}	Oxide capacitance	F
C_{Stack}	Stack capacitance	F
d_{Fe}	Ferroelectric thickness	m
d_{Ox}	Gate Oxide thickness	m
E	Electric field	V/m
E^+	Positive coercive field	V/m
E^-	Negative coercive field	V/m
E_C	Coercive field	V/m
E_c	Conduction band energy	V/m
$E_{Channel}$	Lateral E-Field in the channel	V/m
E_{dep}	Depolarization field	V/m
E_{Fe}	Field in the ferroelectric	V/m
E_f	Fermi energy of the semiconductor	J
E_g	Energy gap of the semiconductor	J
E_i	Intrinsic energy	J
E_{Ox}	Field in the gate oxide	V/m
E_{max}	Maximum electric field	V/m
E_{min}	Minimum electric field	V/m
E_v	Valence band energy	J
E_{vacuum}	Vacuum energy	J
g_m	Transconductance	S
I_{DS}	Drain source current	A
I_{on}	Current for FeFET at "1" for $V_{GB}= V_{Read}$	A
I_{off}	Current for FeFET at "0" for $V_{GB}= V_{Read}$	A
I_{Read}	Read current	A
I_{leak}	Leakage current	A
J_{FP}	Frenkel-Poole current density	A/m^2
J_S	Schottky leakage current density	A/m^2
J_F	Leakage current density in the ferroelectric	A/m^2

Symbol	Description	Units
J_I	Leakage current density in the insulator	A/m^2
J_{FN}	Fowler-Nordheim current density	A/m^2
L	Channel length	m
L_{eff}	Effective channel length	m
m^*	Effective electron mass	m
N_{CH}	Channel doping	cm^{-3}
N_{SUB}	Substrate doping	cm^{-3}
n_i	Intrinsic doping	cm^{-3}
P	Total Polarization ($P_{Fe} + P_{lin}$)	C/m^2
P^+	Positive remnant polarization	C/m^2
P^-	Negative remnant polarization	C/m^2
P_{Fe}	Non-linear part of ferroelectric Polarization	C/m^2
P_{lin}	Linear part of ferroelectric polarization	C/m^2
P_{nonlin}	Nonlinear part of ferroelectric polarization	C/m^2
P_R	Remnant polarization	C/m^2
P_S	Saturated polarization	C/m^2
Q	Charge	C
Q_G	Gate charge	C
Q_I	Inversion charge (channel charge)	C
Q_{Ox}	Oxide charge	C
T	Temperature	K
T	Period	s
t	Time	s
V	Voltage	V
V_B	Bulk voltage	V
V_{DD}	Operating voltage	V
V_D	Drain voltage	V
V_{DS}	Drain source voltage	V
V_{Erase}	Erase voltage	V
V_{FB}	Flatband voltage	V
V_G	Gate voltage	V
V_{GB}	Gate bulk voltage	V
V_{GBeff}	Effective gate voltage	V
V_{GB}	Gate source voltage	V
V_{max}	Maximum applied voltage	V
V_{min}	Minimum applied voltage	V
V_{Ox}	Voltage across oxide	V
V_{Fe}	Voltage across ferroelectric	V
V_{Read}	Read voltage	V
V_S	Source voltage	V
V_{Si}	Surface potential	V
V_{TH}	Threshold voltage	V
W	Channel width	m
X_D	Depletion depth	m
X_T	Channel doping depth	m
γ	Body effect coefficient	$V^{0.5}$
ΔU	Memory window	V
ρ	Charge density	C/m

δ	*Ferroelectric material constant*	*V/m*
σ	*Ferroelectric material constant*	*V/m*
σ_{FP}	*Conductivity*	*A/Vm*
ε_{Fe}	*Ferroelectric permittivity (lin+nonlin)*	-
ε_{Ox}	*Oxide dielectric permittivity*	-
ε_{opt}	*Optical dielectric permittivity*	-
ε_r	*Linear part of ferroelectric permittivity*	-
ε_{Si}	*Permittivity of Silicon*	-
μ_n	*Electron mobility*	m^2/Vs
ϕ_F	*Bulk potential*	*V*
ϕ_t	*Thermal potential*	*V*
$\Delta\Phi$	*Barrier lowering*	*V*
Φ_B	*Energy barrier*	*V*
Φ_{Fe}	*Ferroelectric workfunction*	*V*
Φ_{MS}	*Metal-semiconductor contact potential*	*V*
Φ_M	*Metal workfunction*	*V*
Φ_S	*Semiconductor workfunction*	*V*
χ	*Electron affinity*	*V*

Constants

Symbol	*Name*	*Value*
h	*Planck's constant*	*6.626 m^2kg/s*
k	*Boltzmann constant*	*1.38 10^{-23} J/K*
q	*Electron charge*	*1.6 10^{-19} C*
m	*Electron mass*	*9.1 10^{-31} kg*
ε_0	*Permittivity of free space*	*8.854 10^{-12} F/m*

Acknowledgements

Lebenslauf

Persönliche Daten

Name :   Fitsilis
Vorname :   Michael
Geburtsdatum und –ort :   27.12.1976, Karditsa, Griechenland
Staatsangehörigkeit :   Griechisch

Schulbildung

1982 - 1990	Karditsa
1990 - 1993	Lüdenscheid
1993 - 1994	Düsseldorf Schulabschluss (griechisches Lyzeum)
09/1994	Aufnahmeprüfung

Studium

10/1994 - 11/1999 Studium der Elektrotechnik und Informationstechnik, Fachrichtung Nachrichtentechnik, Aristoteles Universität von Thessaloniki
Diplomarbeit : „Entwicklung von Software zur Untersuchung der Ausbreitung von Radio Signalen"

Promotion

09/2001 - 08/2004 Arbeiten zur Dissertation am Institut für Festkörperforschung, Elektrokeramische Materialien (EKM), Forschungszentrum Jülich

Praktikum

07/1998 - 09/1998 Deutsche Telekom, Nürnberg, „ISDN-Capi Programmierung"
07/1999 - 02/2000 Hewlett-Packard, Böblingen, Abteilung Network Support Lab

Berufliche Tätigkeit

03/2000 - 05/2001 Wissenschaftlicher Mitarbeiter am Institut für elektrotechnische Grundlagen der Informatik (IEGI), Universität Karlsruhe : „Untersuchung von Nichtlinearitäten in Supraleitenden Resonatoren"