Introduction

• Nano-electronic devices are based on “controlled” properties of silicon and other different materials
• To realize p-n junctions or advanced gate stack knowledge in solids physics become essential
• Electronic behavior is the results of carriers (electrons and holes) transport in silicon and in other advanced materials such as high-k dielectric, crystalline alloy etc.
• Experimental approach not only is long and expensive but also in some case not enough to physical understanding
Introduction

• A sort of “theoretical” approach must be required
• To develop a new device (MOS or Memory) different competences are needed: electronics, solid physics, chemistry, quantum physics.
• To develop a new device a sort of “ab-initio” approach is needed: the devices must work on paper before any experimental work.
• The way how to realize this paper-work is the “simulations”
Research partners in Italy

Strong collaboration:
European projects
Stages, PhD

Agrate site

- ST, Numonyx, DAI Nippon, MDM laboratory (~5000 people)
- Numonyx:
  - Technology R&D ("R2" Technology center)
    - Both R&D and initial production
    - Working 24 hours, 7 days per week
    - Advanced technology, 200 => 300 mm
  - Product R&D and business groups
Non-Volatile Memories

+ very high density
+ low cost/bit
+ high data throughput

Flash NAND
Flash NOR
PCM

Outline

- Simulation type overview
- Simulation of the process flow
- Simulation of the electronic devices
- Non volatile memories:
  - Standard floating gate approach
  - Charge trap devices
  - Advanced 3D Charge traps devices
  - Phase change devices
- Ab initio simulations
Simulations

Different kind of simulations have been employed:

• To study the “electrical” response a device simulations and analytical models are employed
• To investigate the fabrication processes a process simulations have being currently used
• To study the materials properties “ab initio” calculations based on fundamental physics have been employed

Different scale and different simulations

• DEVICES:
  • Numerical solution of PDE systems to investigate carriers transport in silicon and in different media.

• PROCESS:
  • Numerical solution of PDE systems to investigate diffusions in solid media.
  • Kinetic Monte Carlo to investigate the dynamics of multi-body physical systems.

• MATERIALS:
  • Schrödinger multi-body equations
  • Car-Parrinello molecular dynamics
Time scale

- Ab-initio ($N \sim 10^2$ $t \sim 10^{-11} s$)
- TBMD ($N \sim 10^3$ $t \sim 10^{-9} s$)
- MD ($N \sim 10^6$ $t \sim 10^{-6} s$)
- Kinetic Monte Carlo ($N \sim 10^5$-$10^7$ $t > s$)
- Continuum ($t > s$)

(courtesy of IMM Institute)

Device Simulation example
Process Simulation example

Kinetic Monte Carlo example
Ab-initio calculation example

Outline

- Simulation type overview
- Simulation of the process flow
- Simulation of the electronic devices
- Non volatile memories:
  - Standard floating gate approach
  - Charge trap devices
  - Advanced 3D Charge traps devices
  - Phase change devices
- Ab initio simulations
Fabrication: manufacturing + testing

- Wafers manufacturing
- Parametric and on-wafer (EWS) testing
- Packaging
- Final testing

- Long (~3 months) and costly process
- Thousands of input parameters/process options
- Trial and error approach inefficient

TCAD: virtual Fab + virtual bench
**Process flow**

- Sequence of a few hundreds technological steps
- Many steps are performed only on some regions:
  - Selectivity achieved through masks (lithography)
- Basic steps:
  - Introduction and activation of dopants (ion implantation and annealing)
  - Growth or deposition of insulating or metallic layers
  - Removal of insulating or metallic layers
  - Planarization
- “Front-end”:
  - P-n junctions for elementary devices (diodes, MOSFETs, Flash cells, …)
  - Lateral isolation of elementary devices
  - Active oxides and CMOS gates
- “Back-end”:
  - Connect elementary devices with metallic interconnect lines

**The implant “problems”**

- Due to the devices scaling (20nm) an accurate description of implant and post-implant approach is needed
- Dopants must be electrical activated
- Crystal damages must be recovered
Ion implantation

Crystal damage

- Defects concentrations in a crystal >1.5e22 [cm-3] -> amorphize
- Re-crystallization velocity [cm/s]:

\[ v_{\text{recris}} = 1.7 \cdot 10^8 \cdot \exp \left( \frac{2.7}{k_B T} \right) \]
If $C(x,t)$ is the concentration of a species per unit volume its flow is proportional to:

$$\vec{J} = D \cdot \nabla C + C \langle \vec{v} \rangle$$

In the case of non-stationary flow we need a balance equation:

$$\frac{\partial}{\partial t} C = \nabla \cdot \vec{J} - r + g$$

By combining balance equation and first Fick's law we obtain the second Fick's law ($D$ and $<\vec{v}>$ not dependent from $x,t$)

$$\frac{\partial}{\partial t} C = D \nabla^2 C - r + g + \langle \vec{v} \rangle \cdot \nabla C$$

**Impurities diffusion mechanism in Si**

- **Direct or classical diffusion**
- **interstitial**
- **Not-direct diffusion or pair diffusion**
  - kick-out $A_s + I = A_i$
  - via vacancy $A_s + V = AV$
  - via interstitialcy $A_s + I = AI$

AI pairs
Boron Diffusion (KICK-OUT)

Substitutional Boron + Interstitial $\leftrightarrow$ Interstitial B

Annealing @ 750 C
Implant dose = 5e14 cm$^{-2}$ @ 20 keV [3]

Annealing @ 800 C
Implant dose = 5e14 cm$^{-2}$ @ 20 keV [4]

Annealing @ 800 C
Implant dose = 2e15 cm$^{-2}$ @ 20 keV [4]

Annealing @ 850 C
Implant dose = 2e15 cm$^{-2}$ @ 20 keV [4]
The “active” area

- Active area is the silicon region where the carriers transports occurs
- Different active area must be isolated

![Tilted view and cross section after trench etch](image)

Active area and stress

- Oxidation of the Si trench is the responsible of the stress
- Stress can have positive or negative effects

![Active area and stress](image)
Negative effects of the stress

Positive effects of the stress

- Stress can be employed to increase drain current

\[ I_D = \frac{\mu_n C_{ox} W}{2L} \left( V_{GS} - V_{th} \right)^2 \left( 1 + \lambda V_{DS} \right) \]

\[ C_{ox} = \frac{\varepsilon_0 \varepsilon_{ox}}{t_{ox}} \]

\[ \mu_n = \frac{q\tau}{m^*} \]

- Stress induces change in mobility
Positive effects of the stress

- Selectively grow thin high-p channel
- Low defect density layer
- Confined inversion layer
- Apply where needed
- Scalable: HOT, SiGe, Ge, III-V

"front-end" process simulation
Final cross section

Final cross section: TEM photos
Six metal layer interconnect technology

SEM cross sections (Flash array)
90nm Flash Product

64Mb 1.8V Multibank - 16.7 mm²

Outline

• Simulation type overview
• Simulation of the process flow
• **Simulation of the electronic devices**
  • Non volatile memories:
    – Standard floating gate approach
    – Charge trap devices
    – Advanced 3D Charge traps devices
    – Phase change devices
• Ab initio simulations
Device simulation

GOAL: predict electrical behavior of an elementary device

INPUT:
- geometry
doping profile

DEVICE SIMULATION
- mesh
- physical models
- parameters
- numerical solver

OUTPUT:
- I(V) curves
- Internal fields

Maxwell equation

Starting from Maxwell equation to simulate carriers evolution in Silicon

\[ \nabla \times \mathbf{E} = -\partial_t \mathbf{B}, \]
\[ \nabla \cdot \mathbf{B} = 0, \]
\[ \nabla \times \mathbf{H} = \mathbf{J} + \partial_t \mathbf{D}, \]
\[ \nabla \cdot \mathbf{D} = \rho, \]
\[ \mathbf{D} = \varepsilon \mathbf{E}, \]
\[ \mathbf{B} = \mu \mathbf{H}. \]

Considering that \( \mathbf{H}=0 \) in the most of the application the Maxwell system is overall simplified.
**Maxwell equation**

Considering electrostatic potential $\psi$ and applying divergence operator to the third equation we have:

$$\nabla \cdot \varepsilon \nabla \psi = -q \quad \text{Poisson equation}$$

$$\nabla \cdot \mathbf{J} + \partial_t \varrho = 0 \quad \text{ Continuity equation}$$

These are the starting equation for semiconductor simulation.

The charge density can be calculated as:

$$\rho = q \left( p + N_D^+ \right) - q \left( n + N_A^- \right)$$

---

**Splitting of the charge**

Because in the semiconductor positive and negative carriers are present it is convenient to consider separately $J_n$ and $J_p$ contribution to the total current flow. The current balance equation can be split

$$\nabla \cdot J_n - q \partial_t n = qR$$

$$\nabla \cdot J_p + q \partial_t p = -qR$$

$$R = R_n - G_n = R_p - G_p$$

$R$ that is the formal separation parameter can be interpreted as a net recombination term.
The drift current

Considering the drift current we have for the total J flow

\[ J_{\text{drift}} = q_p \mathbf{v}_p - q_n \mathbf{v}_n \]

If we consider the relation between velocity and Electric field:

\[ \mathbf{v}_n = -\mu_n \mathbf{E} \]
\[ \mathbf{v}_p = \mu_p \mathbf{E} \]

and combining the equation:

\[ J_{\text{drift}} = q_p \mu_p \mathbf{E} + q_n \mu_n \mathbf{E} \]

The drift-diffusion current

Considering the expression for the diffusion flow

\[ \mathbf{F}_n = -D_n \nabla n \]
\[ \mathbf{F}_p = -D_p \nabla p \]

Moreover the relation between the flux and the current and the Einstein’s relation holds

\[ J_{\text{diffusion}}^n = -q \mathbf{F}_n \]
\[ J_{\text{diffusion}}^p = q \mathbf{F}_p \]

\[ D_n = \mu_n \frac{k_B T_n}{q} \]
\[ D_p = \mu_p \frac{k_B T_p}{q} \]

Combining the equations we have the final expression of the drift-diffusion current

\[ J_n = q n \mu_n \mathbf{E} + q D_n \nabla n \]
\[ J_p = q p \mu_p \mathbf{E} - q D_p \nabla p \]
“Auxiliary” physical models

- Mobility
  - Doping
  - Temperature
  - Surface scattering
  - High-field saturation
- Generation – recombination
  - Shockley-Read-Hall, Auger
  - Tunneling
  - Impact ionization
- Band gap narrowing
- Hot carrier injection
- Stress model

Comparison of transport models

- Drift-Diffusion model: electric field distribution
- Thermodynamic model: temperature distribution
- Hydrodynamic model: electron and hole temperature distributions
MOS inversion layer

MOS Breakdown
Quantum drift-diffusion

- Self-consistent solution of Schrödinger equation with drift-diffusion model
- Local effective bandgap widening ($n_{\text{eff}}$)
- Density gradient model

Investigate breakdown mechanism (DMOS)
Optimize channel doping (Buried pMOS)

Analyze transient currents (inverter)
Outline

- Simulation type overview
- Simulation of the process flow
- Simulation of the electronic devices
- Non volatile memories:
  - Standard floating gate approach
  - Charge trap devices
  - Advanced 3D Charge traps devices
  - Phase change devices
- Ab initio simulations

NVM traditional concept

- Stacked Gate NMOS Transistor
  - Poly1 Floating Gate for charge storage
  - Poly2 Control Gate for accessing the transistor
  - Tunnel-oxide for Gate oxide
  - Oxide-Nitride-Oxide (ONO) for the inter Poly Dielectric
  - Source/Drain Junctions optimized for Program/Erase/Leakage
### NOR vs NAND

<table>
<thead>
<tr>
<th>NOR</th>
<th>NAND</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Channel Hot-Electron Programming, Tunnel Erase</td>
<td>• Tunnel Programming and Erase</td>
</tr>
<tr>
<td>• Direct access to cell</td>
<td>• Cells connected in string 32</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>• Fast Random Access</td>
<td>• Fast Program/Erase</td>
</tr>
<tr>
<td>• Byte Programming</td>
<td>• Smaller Cell</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>• Slower Program/Erase</td>
<td>• Slow Random Access</td>
</tr>
<tr>
<td>• Larger cell</td>
<td>• No Byte Programming</td>
</tr>
</tbody>
</table>

- NOR & NAND process and cell cross-sections are similar, except for the contacts.
- Cell Size: NAND ~5\(\times\)2, NOR ~10\(\times\)2
- Both capable of Multi-Level-Cell
- NOR read latency = 10's of nSec
- NAND read latency = 1-10's of nSec

---

### NOR Roadmap

- Flash Invented in mid 1980's
- NOR and NAND flash evolved from EPROM
- ~20 years & 10 Generations of ETOX® (Intel NOR) High Volume Production
- 10+ years & 6 Generations of MLC: 2bit / cell

K. Pangal, IRPS 2007

Source: Intel

Page 55 Copyright © 2008 Numonyx B.V.
NAND Roadmap

Channel hot electron
Memory main parameter

- **Writing and erasing performances:** threshold variation versus time during write or erase operation
- **Endurance:** variation of cell threshold (written and erased) after cycling operation
- **Retention:** charge loss versus time at different temperature
- **Cross talk:** influence between adjacent cell
- **Vth fluctuation:** instability of the threshold voltage
- **Others:** Program disturb, Drain disturb

Writing and erasing characteristic

- **P/E Voltages:** ±13V, ±14V… ±20V
**Endurance**

![Endurance Graph](image)

**Retention**

![Retention Graph](image)
Cross talk

\[ \alpha_n = 2\alpha_x + 2\alpha_y + 4\alpha_{xy} \]

Total noise

\[ \alpha_n = 2\alpha_x + \alpha_y + 2\alpha_{xy} \]

Cross talk requires realistic 3D structure

- Gate & Source cross sections imported from TEM
**V_{th} distribution for Multi-Level Cell**

![Diagram of V_{th} distribution for Multi-Level Cell](image)

**V_{th} fluctuation due to discrete dopants**

![Diagram of V_{th} fluctuation due to discrete dopants](image)

- **Programming operation:** discrete traps statistically distributed in number and position into the nitride.
- **Reading operation:** discrete dopants statistically distributed in substrate gives non-uniform substrate inversion, enhancing percolative S-D conduction.

---

**Interference problem**

- V_{th} erase 00
- V_{th} pgm 01
- V_{th} pgm 10
- V_{th} pgm 11

---

**Threshold voltage shift statistical dispersion**

**Virgin threshold distribution**

**Threshold shift distribution due to 1 electron**
Current crowding

Vth fluctuation due to the random Telegraph Noise (1/2)

- Threshold voltage variations induced by trapping/detrappping events have been inserted through a local threshold shift:

\[ \Delta V_{TH} \approx \frac{q}{C_{OX} S} \]

\[
V_{CC(@I_{READ})} \propto V_{TH} + \sqrt{\frac{2I_D T_{ox}}{\mu_{eff} \epsilon_{ox}} \left( \frac{L}{W} \right)_{eff}}
\]
Random Telegraph Noise (2/2)

Empty trap  Full trap

• Depending on trap position along the channel and underlining NUC, local $\Delta V_{TH}$ can have different impact on total current

Outline

• Simulation type overview
• Simulation of the process flow
• Simulation of the electronic devices
• Non volatile memories:
  – Standard floating gate approach
  – Charge trap devices
  – Advanced 3D Charge traps devices
  – Phase change devices
• Ab initio simulations
Flash scaling issues

Main issues:
- Device electrostatic integrity → poor coupling factor (poor W/E), low Ion (enhanced access time), high Ioff (shrunk operating window), electron number fluctuation (increased dispersion)
- Enhanced interference between nearby cells → read fail, program disturbance
- Process complexity → strong investments to assure sufficient fault tolerance

S.J. Baik et al.

Nitride based memory

High-k as top dielectrics improve P/E efficiency but introduce reliability constrains
Charge trap stack

• Metal gate
  – High workfunction to reduce electron injection in nitride during erase (erase saturation)
  – Good thermal stability and adhesion properties

• Blocking dielectric
  – High-k to reduce field
  – Low leakage (retention) and trapping (cycling)

• Trapping layer
  – Many deep traps

• Bottom dielectric
  – Low leakage (retention) and trapping (cycling)

The difference between the traditional approach

Trapping layer is not a constant-potential region
Outline

- Simulation type overview
- Simulation of the process flow
- Simulation of the electronic devices
- Non volatile memories:
  - Standard floating gate approach
  - Charge trap devices
  - Advanced 3D Charge traps devices
  - Phase change devices
- Ab initio simulations
3D architecture

- Planar Cell programming optimization
- Advanced Cell structures investigation: Multi-Gate and 3D-stacking

Multi-gate solution

3D-stacking solution

J. Jeong, SAMSUNG, VLSI 2009
B. Katsumoto, Toshiba, VLSI 2009

Fin-Fet architecture
Fin-Fet confinement effect

- FIN-FET needs to take into account quantization due to the narrow active area (typically in the range of 10 nm).

MACARONI

Geometrical parameters:
- $R_{\text{ox}} = 8$ nm
- $R_{\text{si}} = 16$ nm
- $R_{\text{tun}} = 20$ nm
- $R_{\text{n}} = 24.5$ nm
- $R_{\text{top}} = 29.5$ nm
- $L_g = 40$ nm
- $L_{\text{ox}} = 40$ nm
- $L_b = 160$ nm
- $W_b = 80$ nm
- Oxide is not shown
Macaroni (matrix view)

- PitchX = 160 nm
- PitchY = 240 nm
- matrix is composed by 18 cell and two level
- Different Z planes

Macaroni reading current

- Virgin cells
- Programmed cell
- Programmed region
Outline

- Simulation type overview
- Simulation of the process flow
- Simulation of the electronic devices
- Non volatile memories:
  - Standard floating gate approach
  - Charge trap devices
  - Advanced 3D Charge traps devices
  - Phase change devices
- Ab initio simulations

Phase change material (Ge$_2$Sb$_2$Te$_5$)

- Polycrystal (f.c.c. or hexagonal)
- Amorphous
- Different structure
- Refractive index
- Electric conductivity

Solid-state memories
PCM Technology Concept

- **Storage mechanism**
  amorphous / poly-crystal phases of a chalcogenide alloy, usually Ge$_2$Sb$_2$Te$_5$ (GST)

- **Reading mechanism**
  resistance change of the GST
  amorphous → high resistance (~ $10 \, \Omega \, cm$) → reset state
  crystalline → low resistance (~ $10 \, m\Omega \, cm$) → set state

- **Writing mechanism**
  self-heating due to current flow (Joule effect)
  melting temperature ($T_m \approx 630^\circ C$, for $t_{\text{RESET}} \approx 10$-$100\,\text{ns}$)
  crystallization temperature ($T_x \approx 400^\circ C$, for $t_{\text{SET}} \approx 100$-$1000\,\text{ns}$)

---

PCM cell working principle
Experimental characteristics

Current-voltage (I-V) curve

Programming curve

Band structure model

Crystal

Amorphous

Donor-like Traps (C$_3^+$)

Acceptor-like Traps (C$_1^-$)

Localized states ("Lone Pairs", C$_2^0$)
Electro-thermal simulation

I=500 μA

Programming curve

Top contact
GST (red=amorphous)
Bottom contact
Outline

• Simulation type overview
• Simulation of the process flow
• Simulation of the electronic devices
• Non volatile memories:
  – Standard floating gate approach
  – Charge trap devices
  – Advanced 3D Charge traps devices
  – Phase change devices
• Ab initio simulations

Different scaling
**Born-Oppenheimer approximation**

\[
H_{tot}(x_1, \ldots, x_N; R_1, \ldots, R_M) = E_{tot}(x_1, \ldots, x_N; R_1, \ldots, R_M)
\]

\[
H_{tot} = T_e + T_n + V_{en} + V_{ee} + V_{nn}
\]

**Separazione dei moti elettronici e nucleari**

\[
H_{elec}(x, R) \Psi_{elec}(x, R) = E_{elec}(R) \Psi_{elec}(x, R)
\]

\[
(T_a + E_{elec}(R)) \Psi_{nucl}(x, R) = H_{nucl}(x, R) \Psi_{nucl}(x, R) = E_{tot} \Psi_{nucl}(x, R)
\]

- Permette di definire un'equazione di Schrödinger elettronica e una nucleare.
- La parte elettronica dipende in modo parametrico dalle coordinate nucleari.
- La soluzione dell'eq. di S. elettronica definisce un potenziale entro cui si muovono in nuclei (PES).
- L’approssimazione di BO è solitamente molto buona (es. $H_2$ ⇒ errore $\approx 10^{-4}$).

**Electronic solution**

I metodi QM risolvono l’equazione di Schrödinger elettronica (per una data geometria nucleare)

\[
H_{elec}(x, R) \Psi_{elec}(x, R) = E_{elec}(R) \Psi_{elec}(x, R)
\]

\[
H_{elec} = T_e + V_{ee} + V_{en} + V_{nn}
\]

**Energia cinetica elettronica**

\[
T_e = -\sum_{i}^{N} \frac{1}{2} \nabla_i^2
\]

**Repulsione elettrone-elettrone**

\[
V_{ee} = \sum_{i}^{N} \sum_{j \neq i}^{N} \frac{1}{|x_i - x_j|}
\]

**Interazione elettrone-nucleo**

\[
V_{en} = -\sum_{i}^{N} \sum_{a}^{M} \frac{Z_a}{|R_a - x_i|}
\]

**Repulsione nucleo-nucleo**

\[
V_{nn} = \sum_{a}^{M} \sum_{a'}^{M} \frac{Z_a Z_{a'}}{|R_a - R_{a'}|}
\]

B. Civalleri 2007
AB-initio DFT dimer simulation

- Calculation of the vibrational energy of different dimer:
  - Si-Si
  - Si-N
  - N-N

Beta Silicon Nitride crystal traps

\[ \text{Si-N} \leftrightarrow \text{e}^- + \text{h}^+ \]
Available stage topics @ Numonyx Agrate

- Charge Trap memories
  - 3D modeling for advanced devices
  - Statistical approach on programming mechanisms
  - Material modelling with ab-initio simulations
  - Numerical implementation of advanced physical models

- Phase Change Memories
  - Conduction mechanisms for amorphous phase
  - Extension of thermo-electrical and phase change modelling to next technology nodes