Gate Carrier Injection and NC-Non-Volatile Memories

Jean-Pierre Leburton

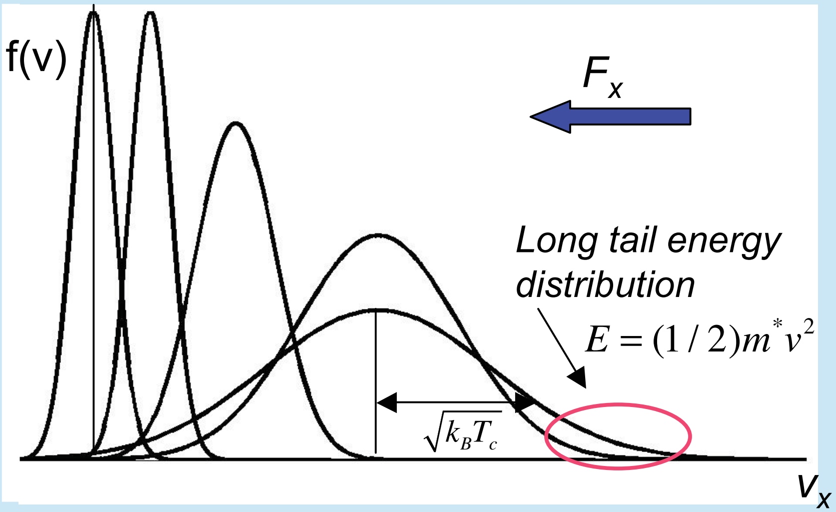
Department of Electrical and Computer Engineering and Beckman Institute

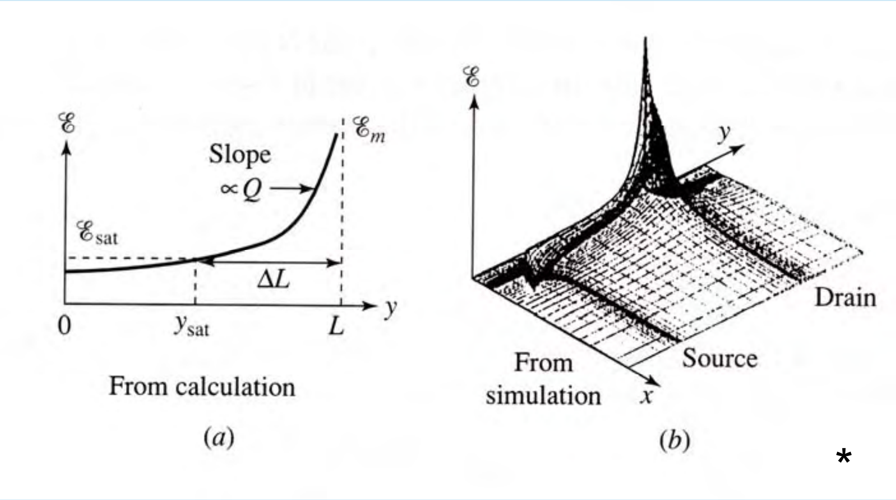
University of Illinois at Urbana-Champaign

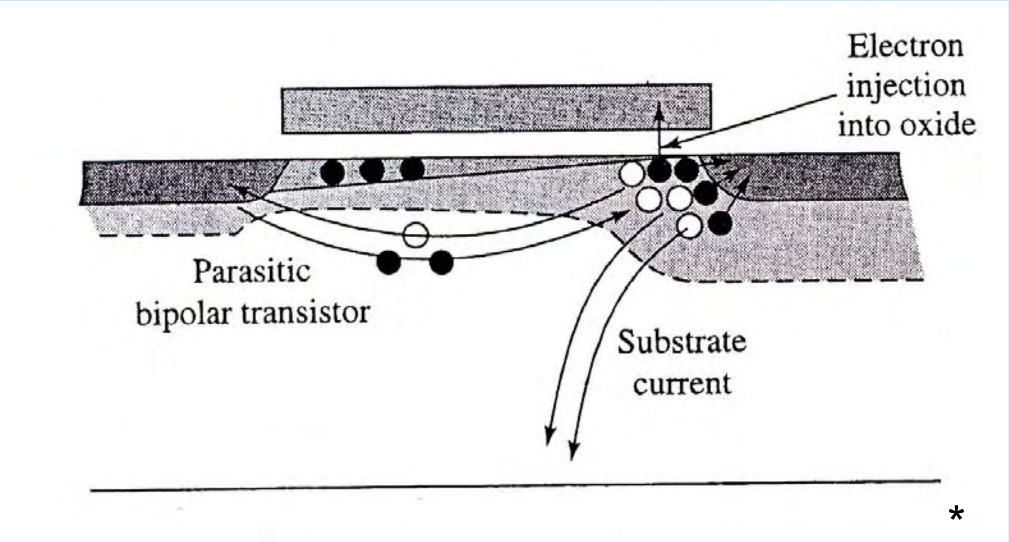
Urbana, IL 61801, USA

Hot Carrier Effects in MOSFETs

High-field/non-linear transport

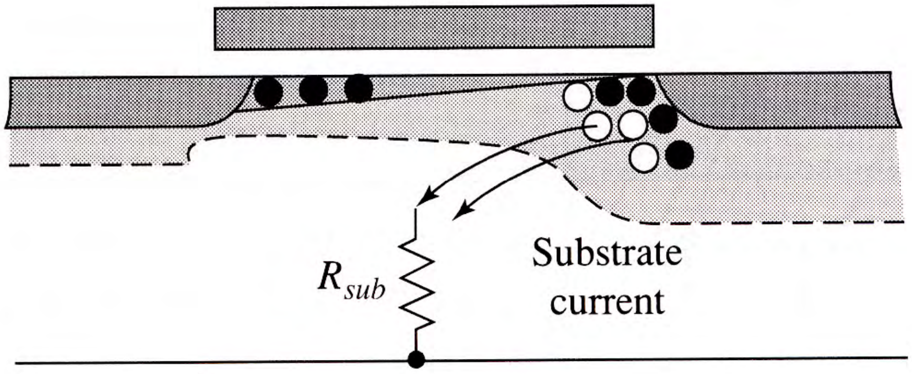




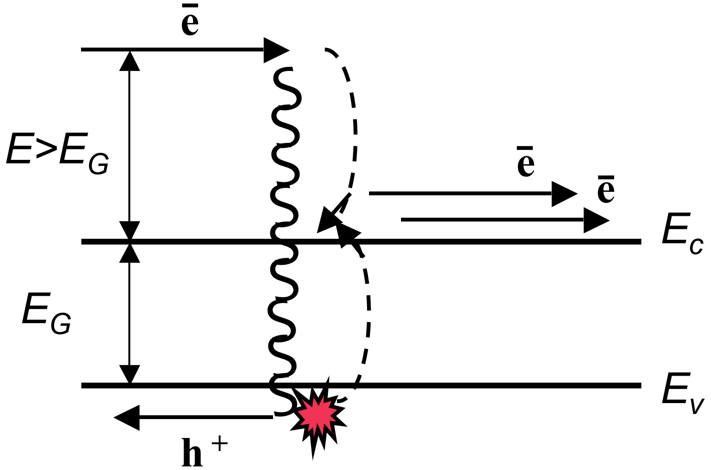


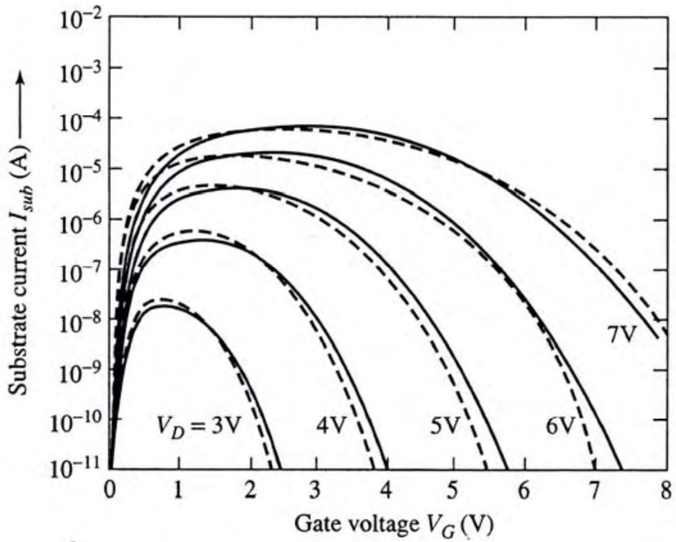
After R.S. Muller and T.I Kamins, DEIC, Wiley, 3d ed.

Hot Carrier Effects: Substrate Current*

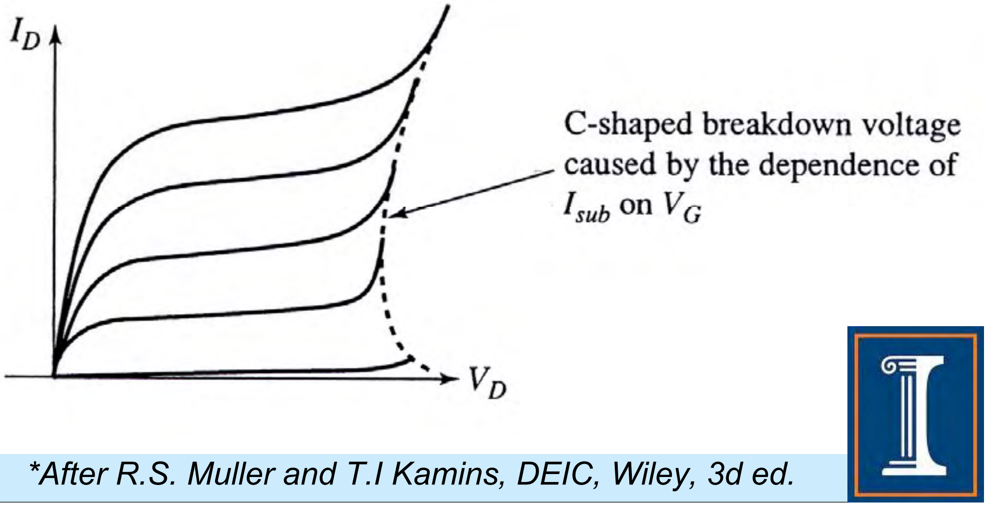


Impact Ionization





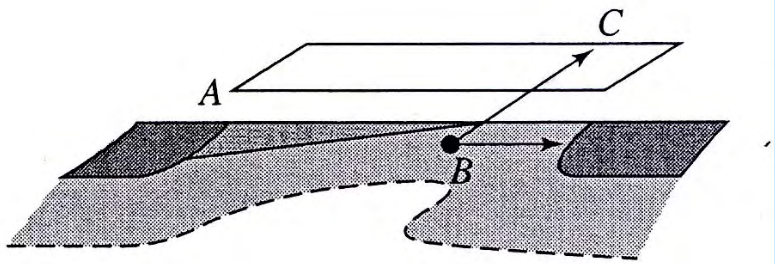
I-V characteristics



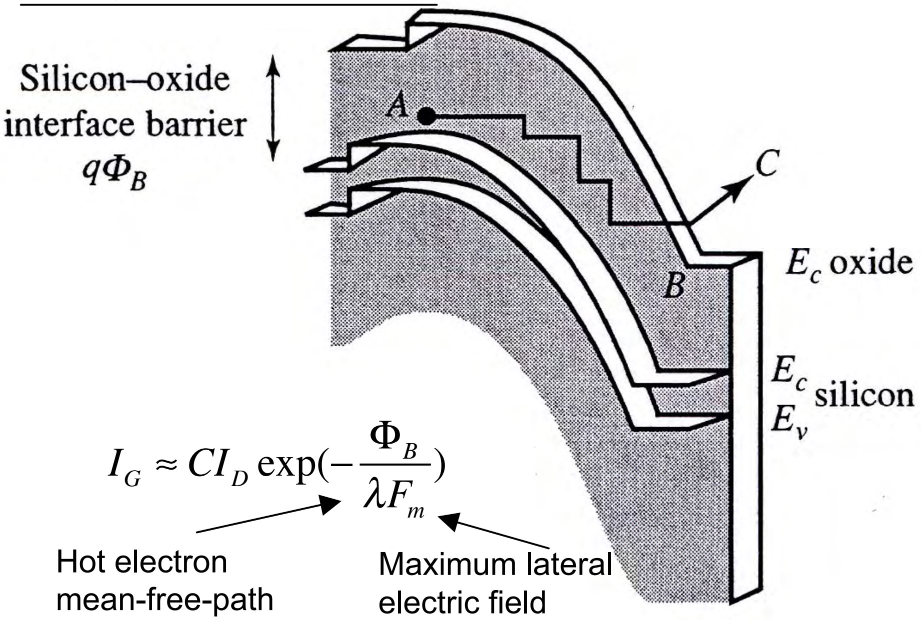
*After R.S. Muller and T.I Kamins, DEIC, Wiley, 3d ed.

Hot Carrier Injection into the Gate*

Schematic of hot carrier injection



Lucky-Electron Model

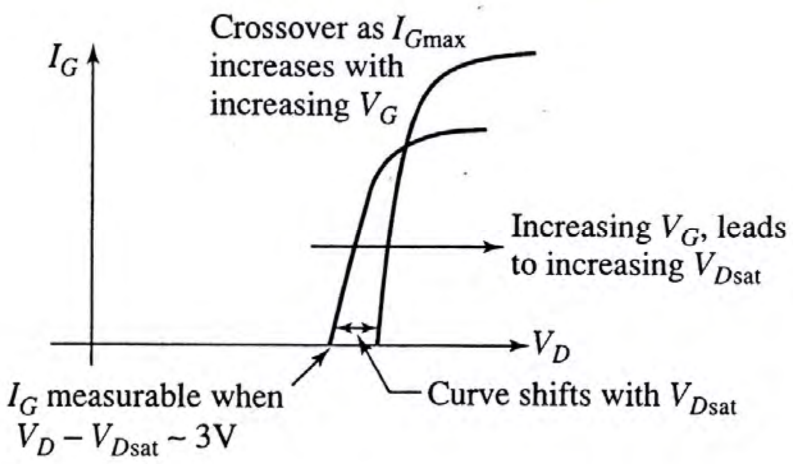


$$I_G \approx C I_D \exp\left(-\frac{\Phi_B}{\lambda F_m}\right)$$

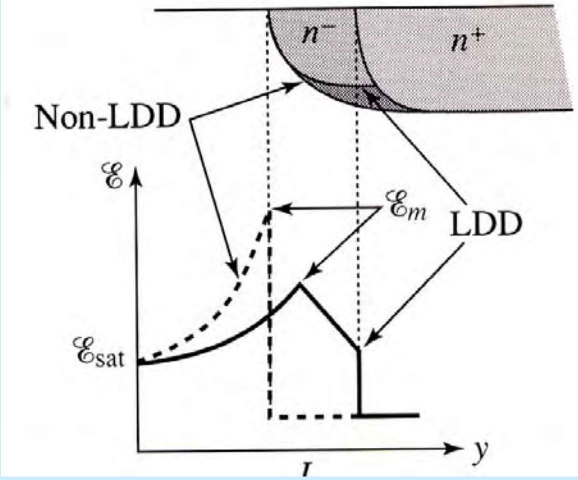
Hot electron mean-free-path

Maximum lateral electric field

Gate current vs. V_D



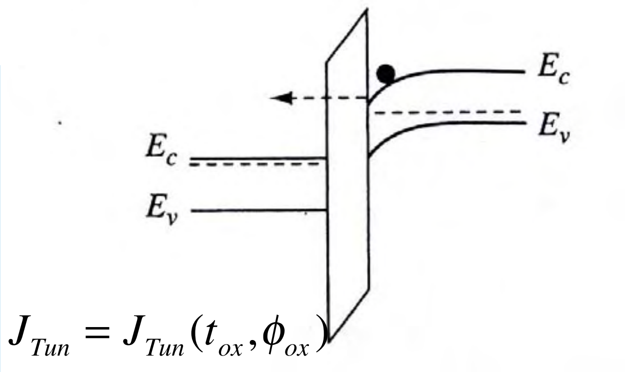
Reduction of hot carrier injection: LDD



*After R.S. Muller and T.I Kamins, DEIC, Wiley, 3d ed.

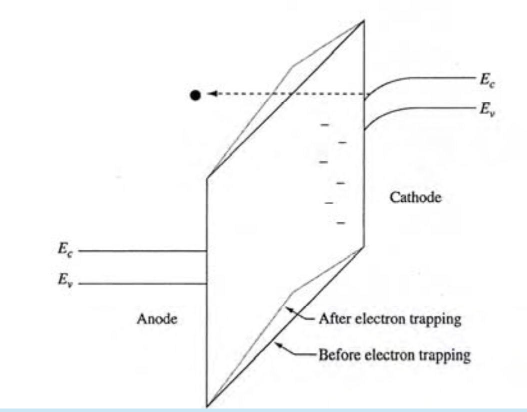
Tunneling Injection into the Gate*

Direct Tunneling

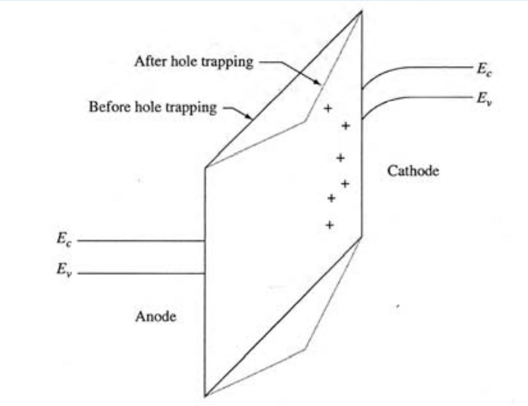


$$J_{Tun} = J_{Tun}(t_{ox}, \phi_{ox})$$

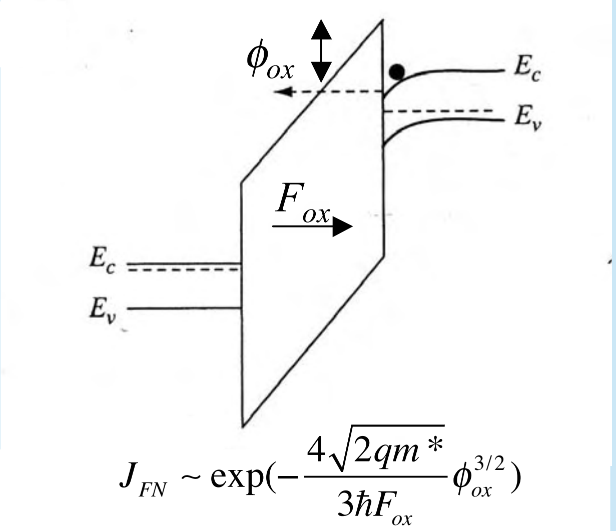
Electron trapping in SiO$_2$



Hole trapping in SiO$_2$

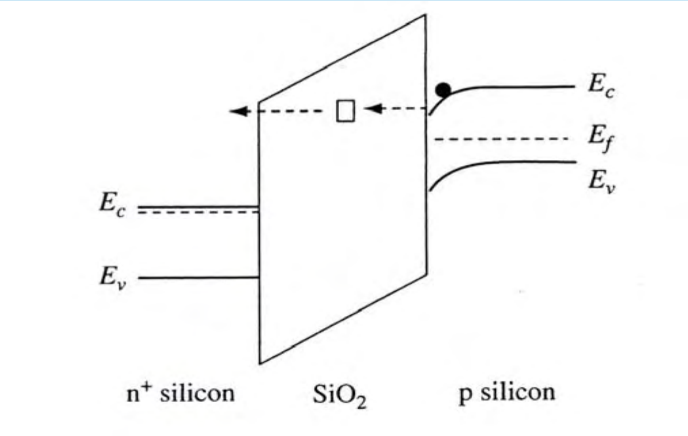


Fowler-Nordheim Tunneling



$$J_{FN} \sim \exp\left(-\frac{4\sqrt{2qm^*}}{3\hbar F_{ox}}\phi_{ox}^{3/2}\right)$$

Trap-Assisted Tunneling



→

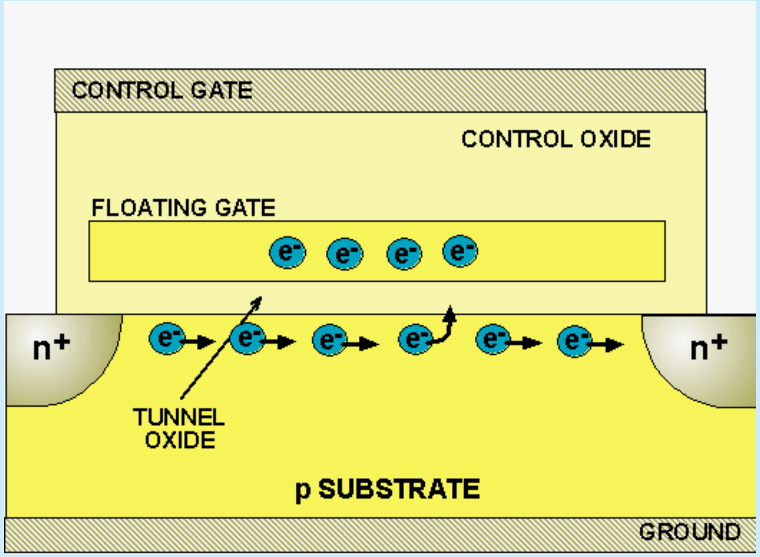
V_T- Degradation

Dissipation (gate leakage)

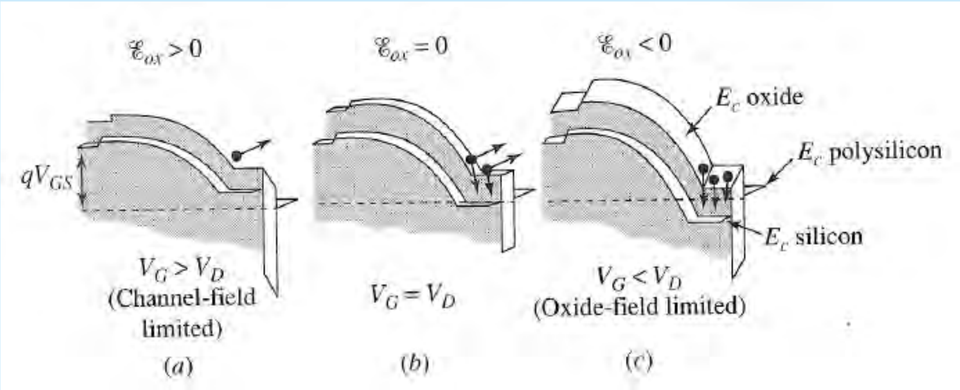
After Y.Taur and T.H. Ning, FMVD, Cambridge, 2d ed.

Injection into Floating Gates

n-channel

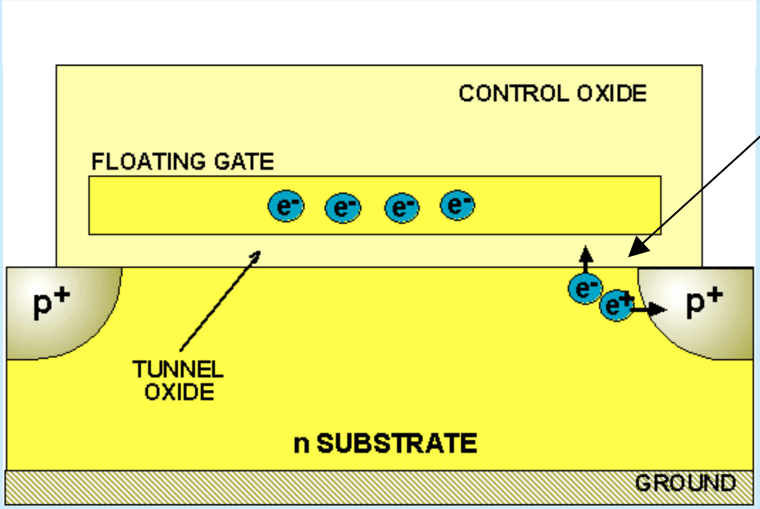


Injection by channel hot electrons (CHE)



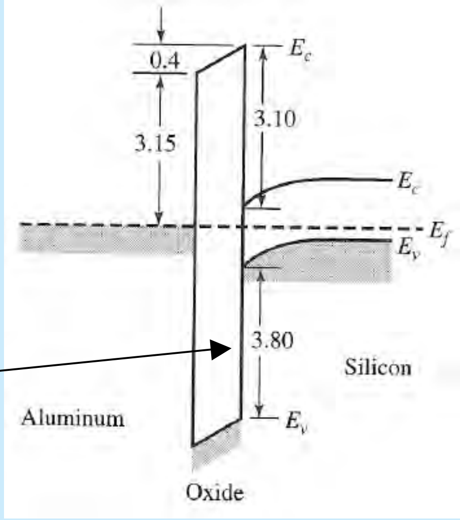
After R.S. Muller and T.I Kamins, DEIC, Wiley, 3d ed.

p-channel

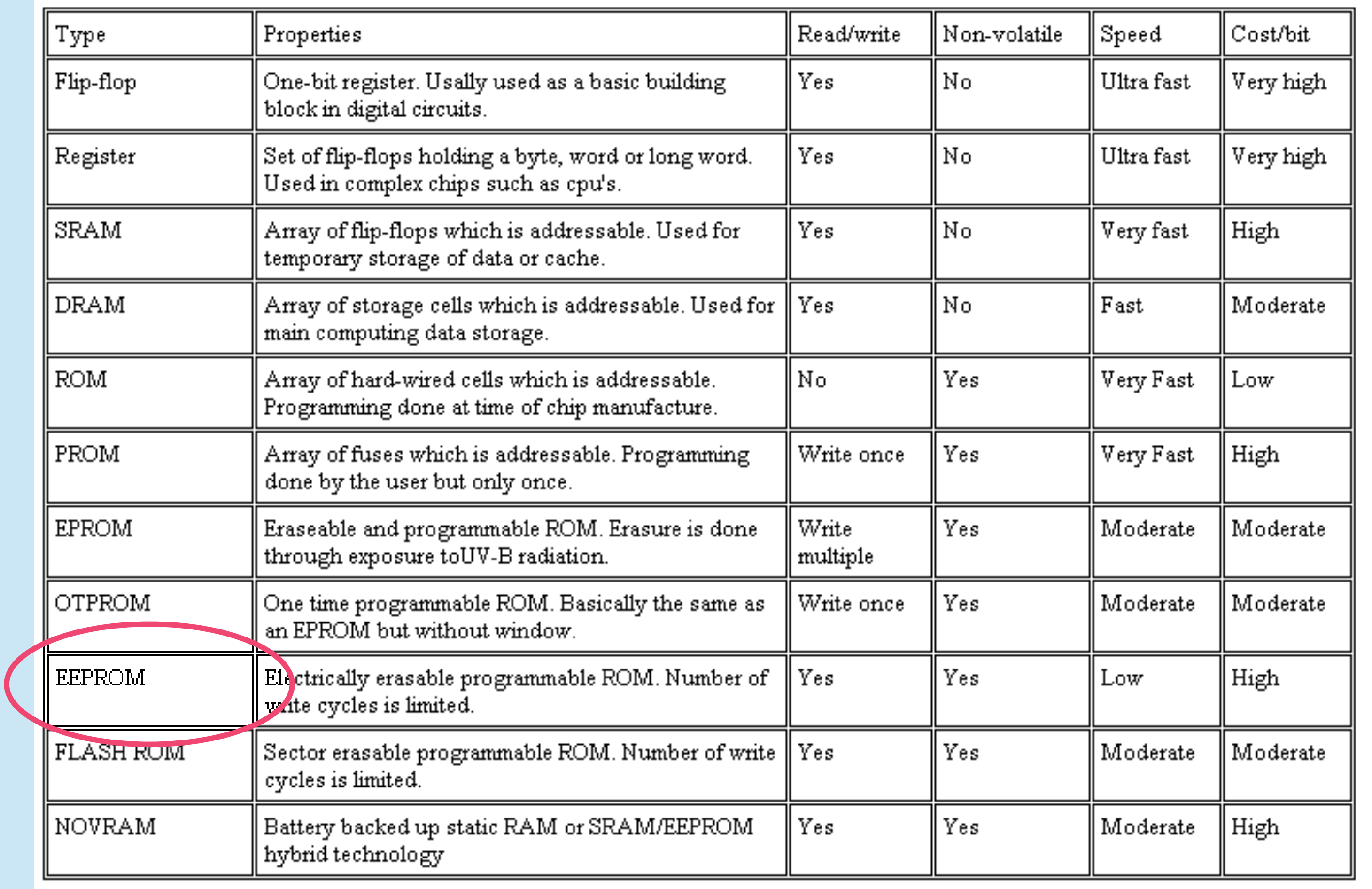


Drain-avalanche (impact ionization)

No CHE because larger oxide barrier



Solid State Memories



Type	Properties	Read/write	Non-volatile	Speed	Cost/bit
Flip-flop	One-bit register. Usally used as a basic building block in digital circuits.	Yes	No	Ultra fast	Very high
Register	Set of flip-flops holding a byte, word or long word. Used in complex chips such as cpu's.	Yes	No	Ultra fast	Very high
SRAM	Array of flip-flops which is addressable. Used for temporary storage of data or cache.	Yes	No	Very fast	High
DRAM	Array of storage cells which is addressable. Used for main computing data storage.	Yes	No	Fast	Moderate
ROM	Array of hard-wired cells which is addressable. Programming done at time of chip manufacture.	No	Yes	Very Fast	Low
PROM	Array of fuses which is addressable. Programming done by the user but only once.	Write once	Yes	Very Fast	High
EPROM	Eraseable and programmable ROM. Erasure is done through exposure toUV-B radiation.	Write multiple	Yes	Moderate	Moderate
OTPROM	One time programmable ROM. Basically the same as an EPROM but without window.	Write once	Yes	Moderate	Moderate
EEPROM	Electrically erasable programmable ROM. Number of write cycles is limited.	Yes	Yes	Low	High
FLASH ROM	Sector erasable programmable ROM. Number of write cycles is limited.	Yes	Yes	Moderate	Moderate
NOVRAM	Battery backed up static RAM or SRAM/EEPROM hybrid technology	Yes	Yes	Moderate	High

EEPROM

- Programming damages oxide
- Endurance : 10^3-10^6 cycles
- Hot-electrons or tunneling

Flash memories X HD's

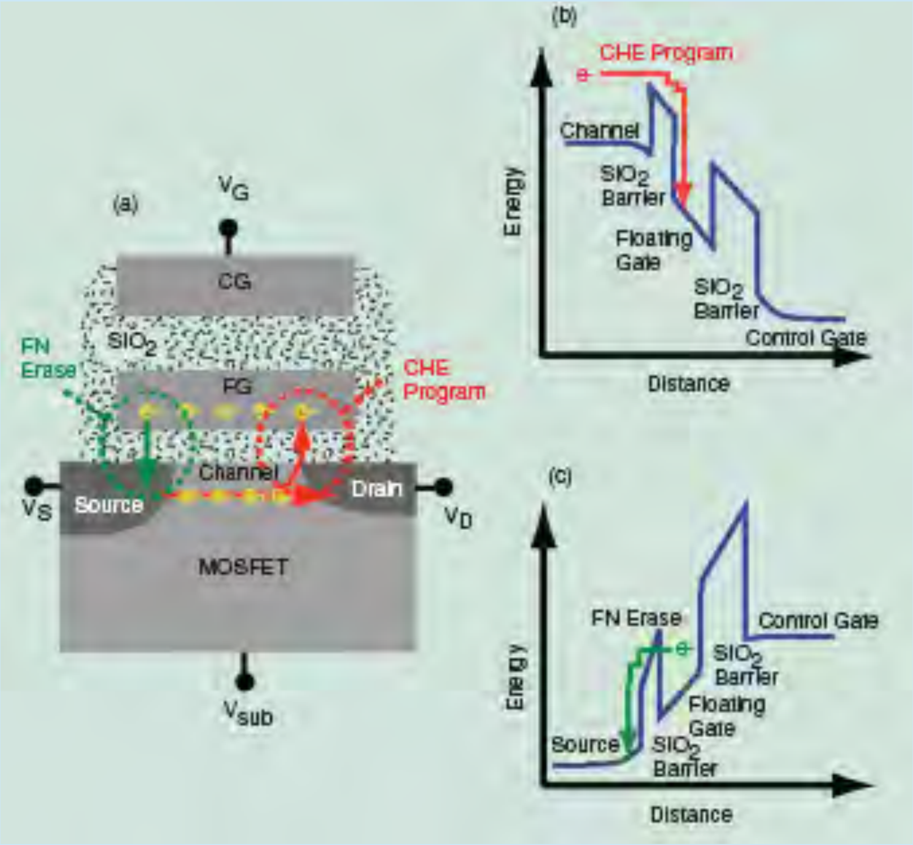
- Noiseless
- Faster access
- Smaller and lighter
- No moving parts
- Low power consumption

Applications

- Digital cameras
- Portable devices
- Removable data storage

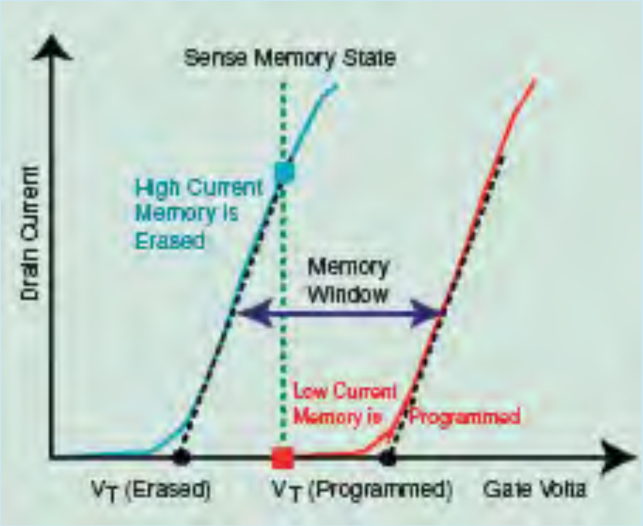
Flash Memory Device: Basic Operation*

ETOX: Hot electron-tunneling combined



But leakage through defects!!!

V_T-shift



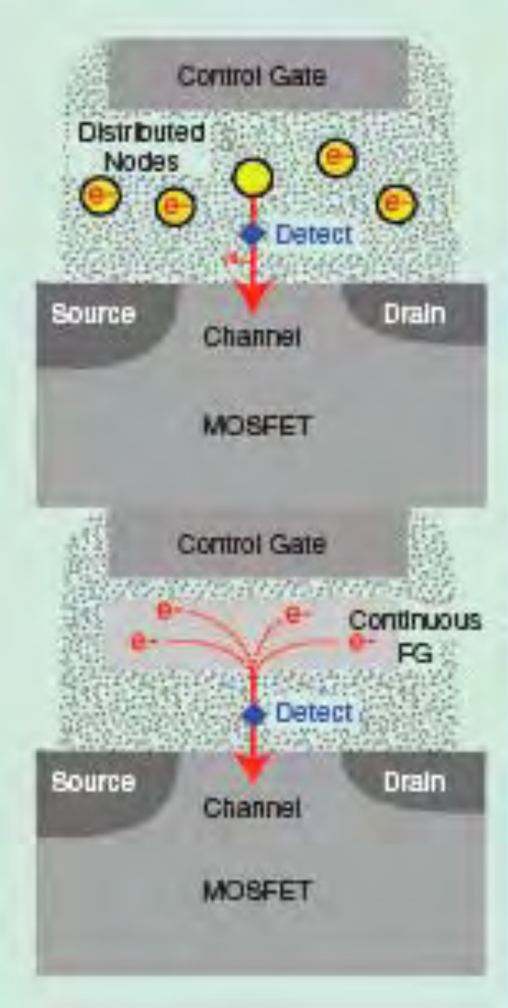
- FG electrically disconnected
- Data stored in form of charge packages
- Transport mechanisms (CHE)
- FN tunnelling (oxide damage)
- Memory cells altered individually
- Data storage sensed by conductance
- Non-volatile storage
- Down scaling X retention time

---

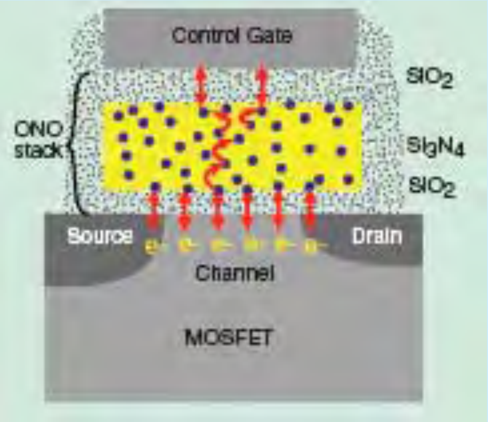
* A. Thean and J.P.Leburton, IEEE Potentials, 21(4) 35, (2002)

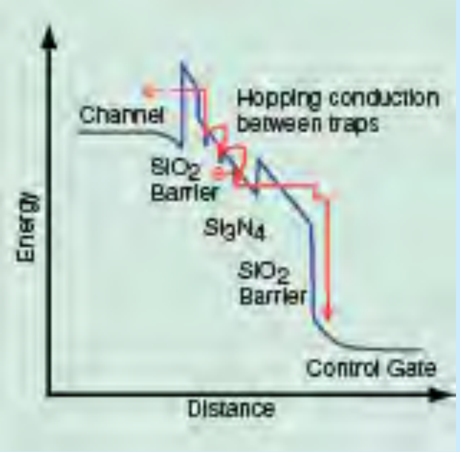
Novel Memory Cells (Leakage Reduction)*

Individual nodes in dielectrics



SONOS



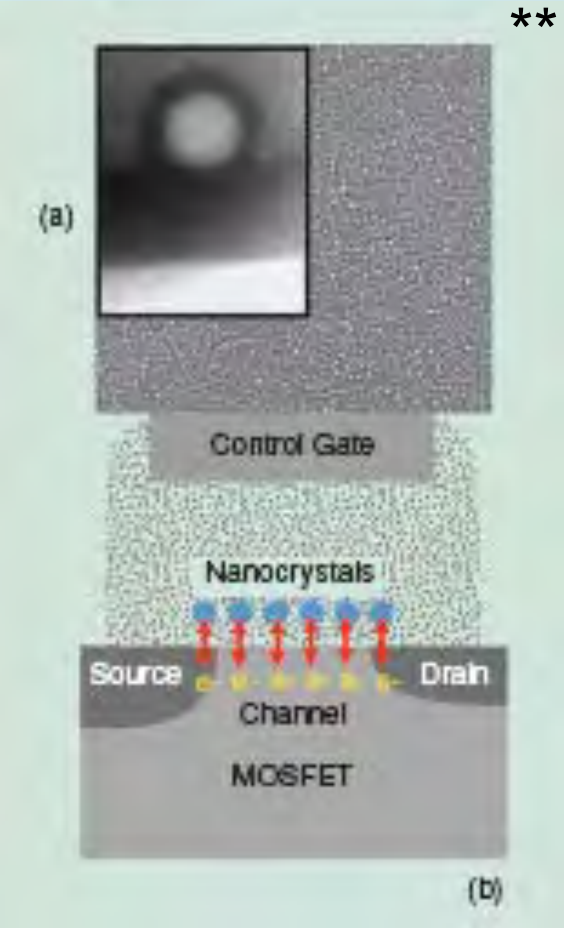


J.P. Leburton, IWSG-2009, IITB, India

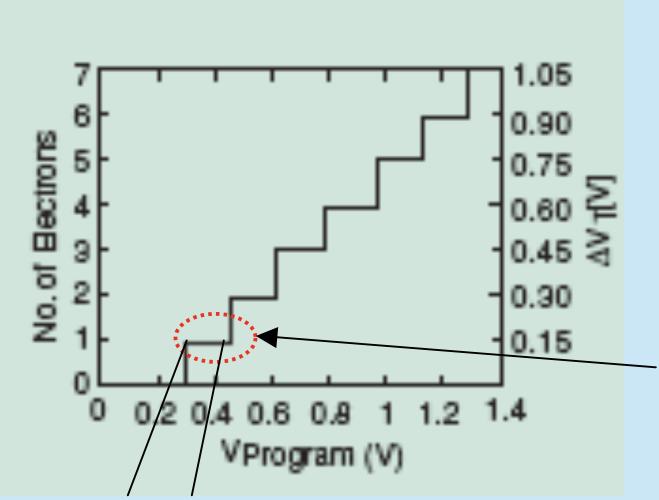
* A. Thean and J.P.Leburton, IEEE Potentials, 21(4) 35, (2002)

Nanocrystal Memories*

NC memory device structure



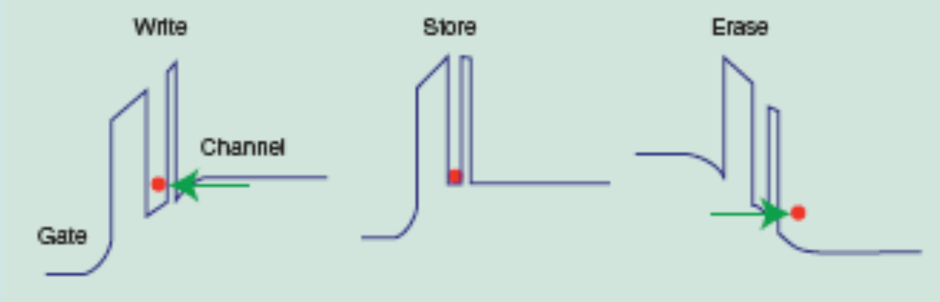
Single electron charging***



$\Delta E<e^2/2C$: Coulomb blockade

$\Delta V_G=e/C$; C:NC capacitance

NC memory operation principle*



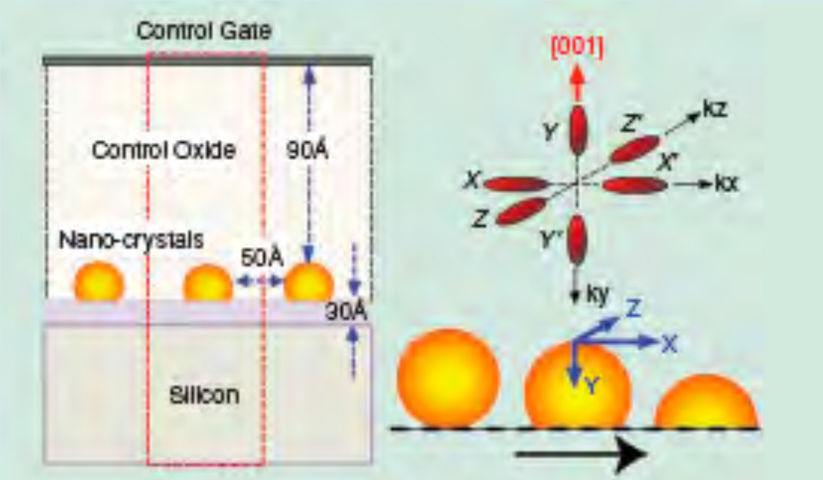
** Courtesy Motorola Inc.

* S. Tiwari et al. IEDM Tech Dig., 521, Dec. 1995.

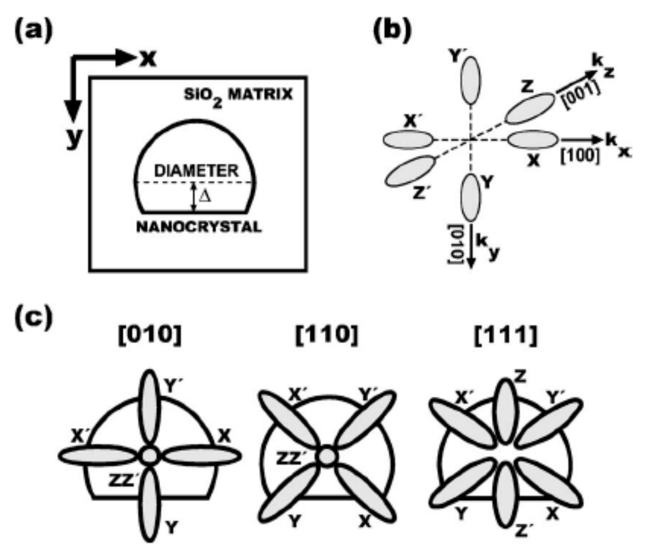
*** A. Thean and J.P. Leburton, IEEE EDL 20, 286, 1999.

NC Memory Device: QM Modeling*

Simulated structure



Crystallographic orientations



Schroedinger Equation (effective mass approx.)

$$\left[-\frac{\hbar^2}{2}(\partial_x,\partial_y,\partial_z)\underbrace{\begin{pmatrix} m_{xx}^{-1}(\vec{r}) & m_{xy}^{-1}(\vec{r}) & m_{xz}^{-1}(\vec{r}) \\ m_{yx}^{-1}(\vec{r}) & m_{yy}^{-1}(\vec{r}) & m_{yz}^{-1}(\vec{r}) \\ m_{zx}^{-1}(\vec{r}) & m_{zy}^{-1}(\vec{r}) & m_{zz}^{-1}(\vec{r}) \end{pmatrix}}_{\hat{M}_{v,T}^{-1} = \hat{\Theta}_T^{-1}\hat{M}_v^{-1}\hat{\Theta}_T}\begin{pmatrix}\partial_x \\ \partial_y \\ \partial_z\end{pmatrix} + V(\vec{r})\right]\Psi_{v,n}(\vec{r}) = E_{v,n}\Psi_{v,n}(\vec{r})$$

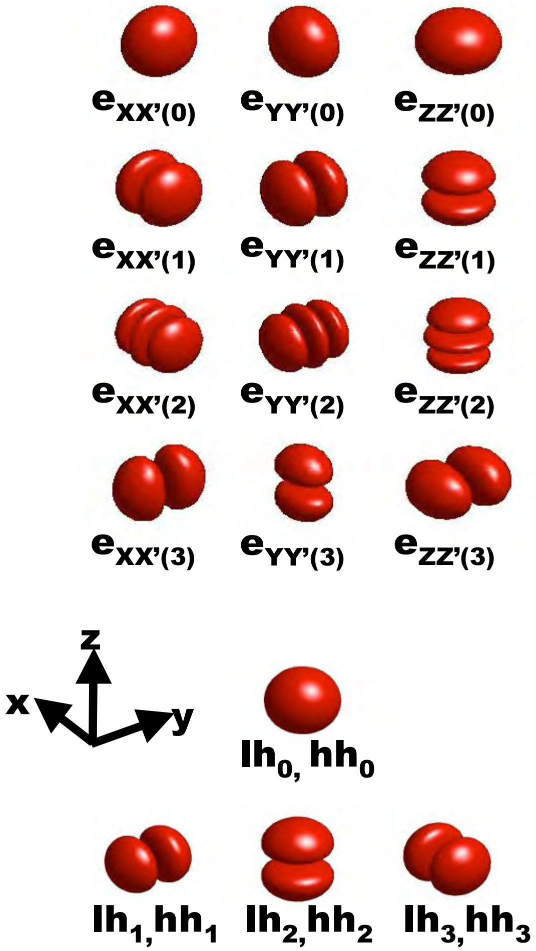
$$with \quad \hat{M}_v^{-1} = \begin{pmatrix} m_1^{-1}(\vec{r}) & 0 & 0 \\ 0 & m_2^{-1}(\vec{r}) & 0 \\ 0 & 0 & m_3^{-1}(\vec{r}) \end{pmatrix}$$

Rotation matrix

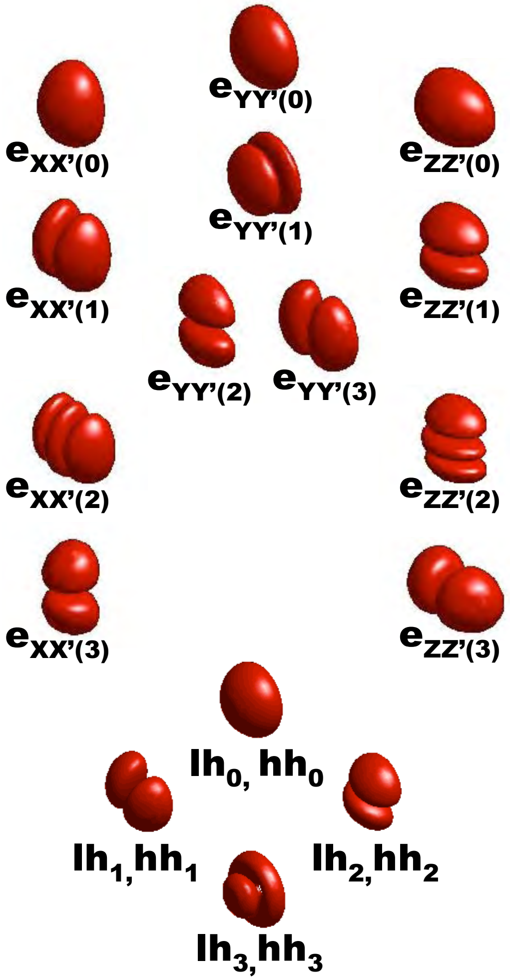
J.S. de Sousa et al., APL 82, 2685 (2003)

Electronic Orbitals*

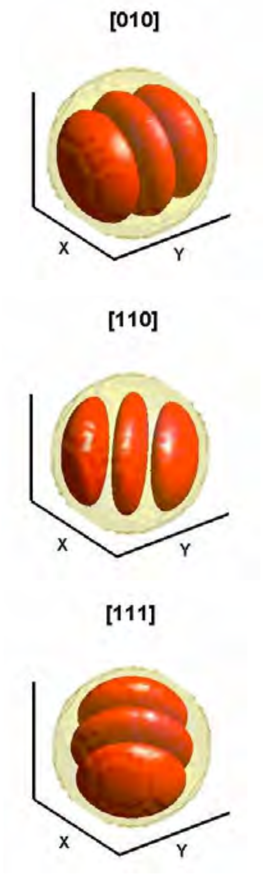
SPHERICAL NC



HEMISPHERICAL NC



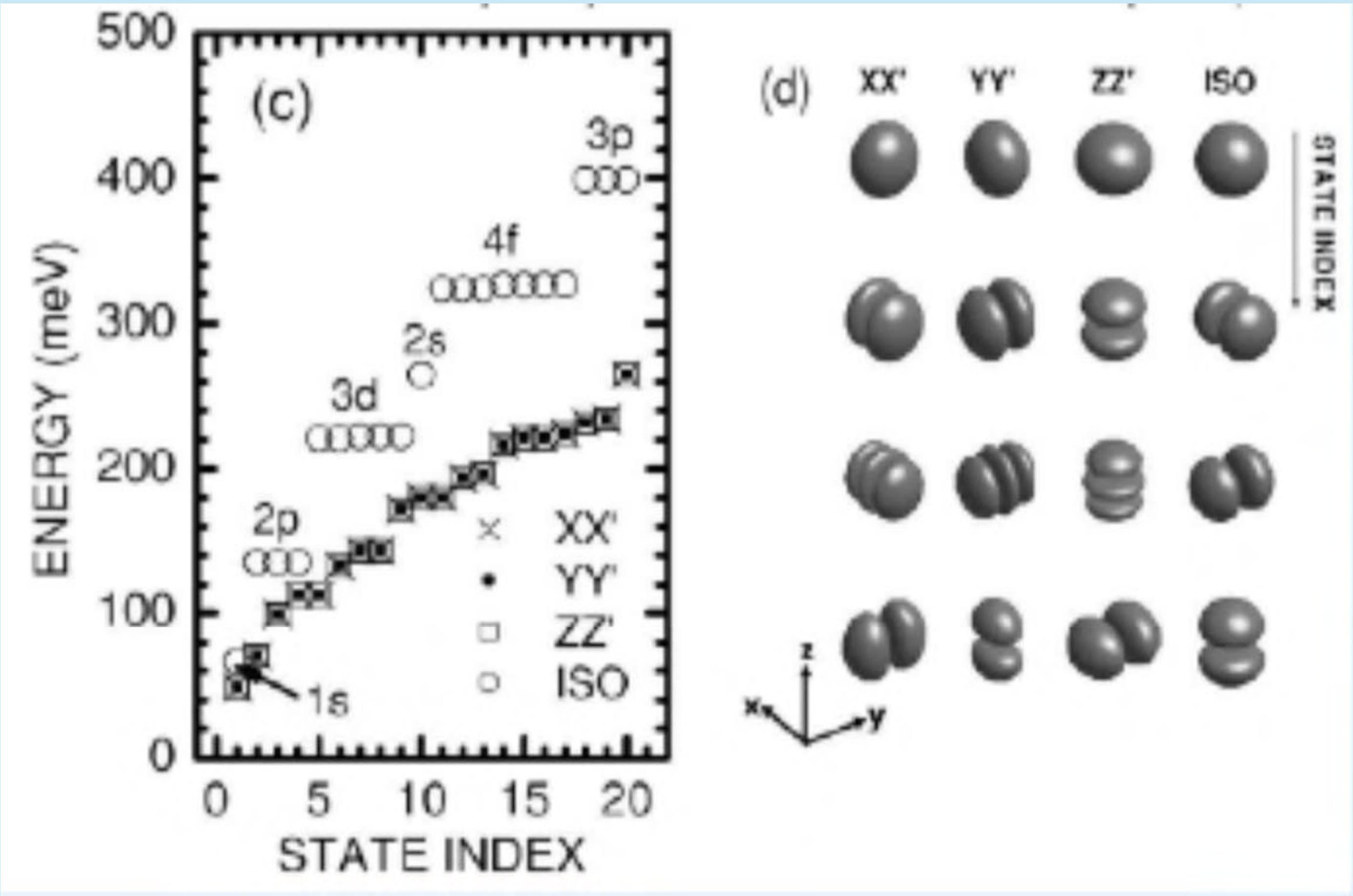
CRYSTALLOGRAPHIC ROTATION EFFECT



---

*J.S. de Sousa et al., APL 82, 2685 (2003)

Energy Spectra: Effective Mass Anisotropy



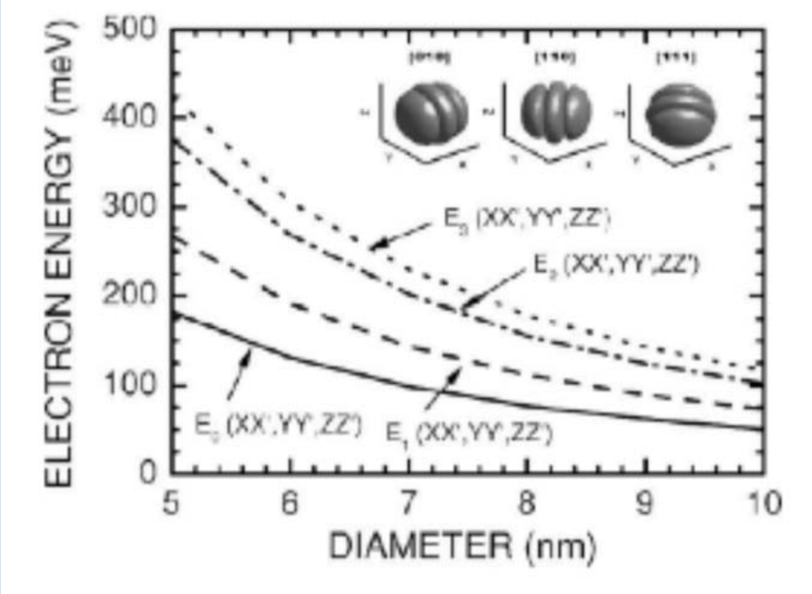
Spherical nanocrystal

D = 10 nm

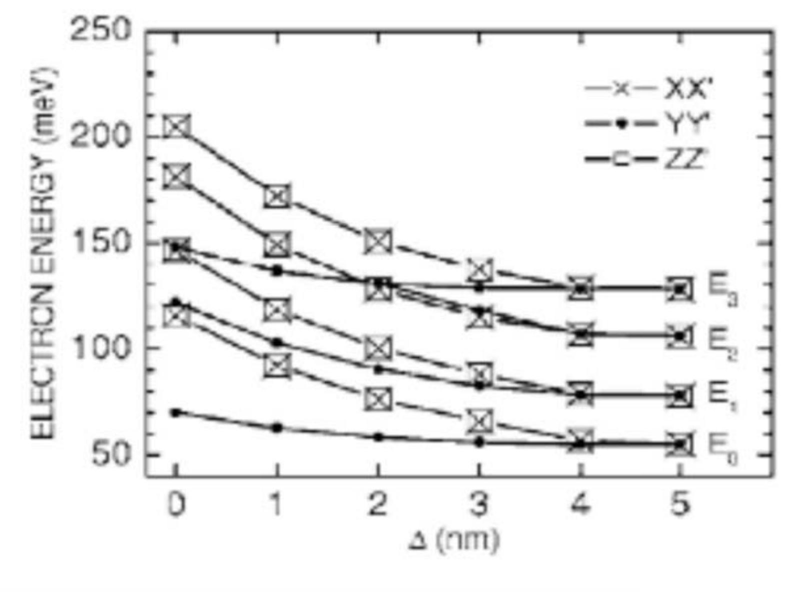
$$\boxed{1/m_{iso} = 2/(3m_t) + 1/(3m_l)}$$

Energy Spectra: Size and Shape Effects *

Spherical Quantum Dots



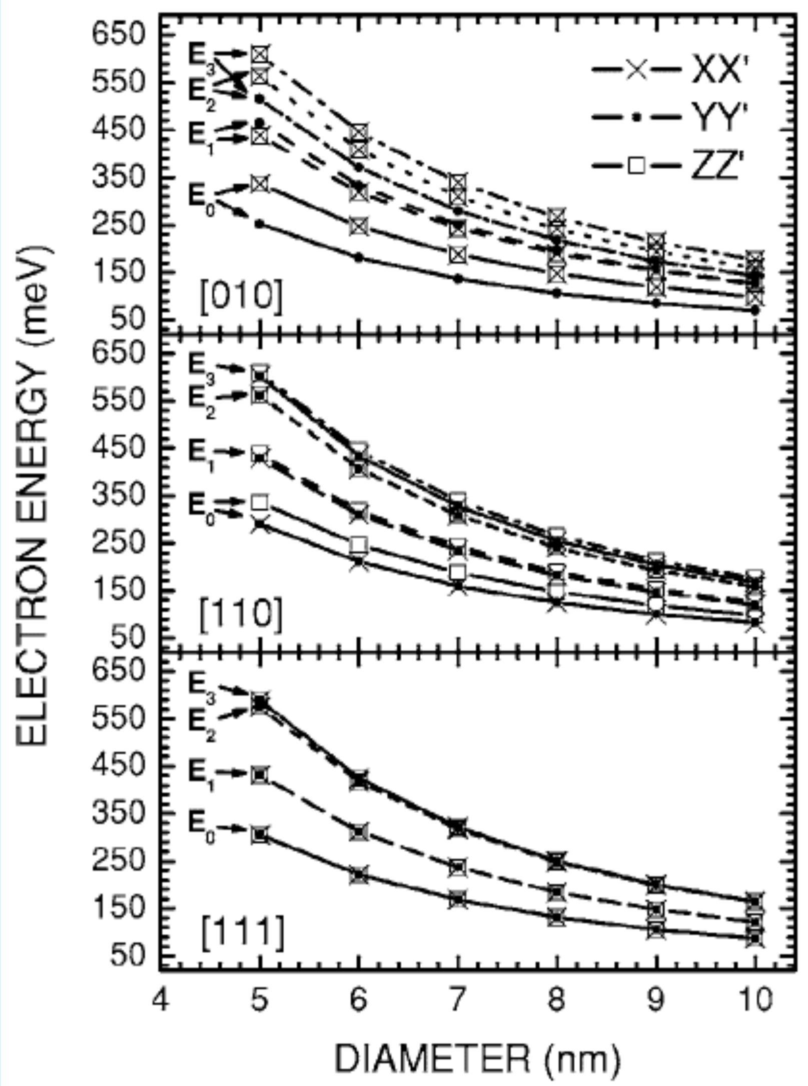
Truncated Nanocrystals

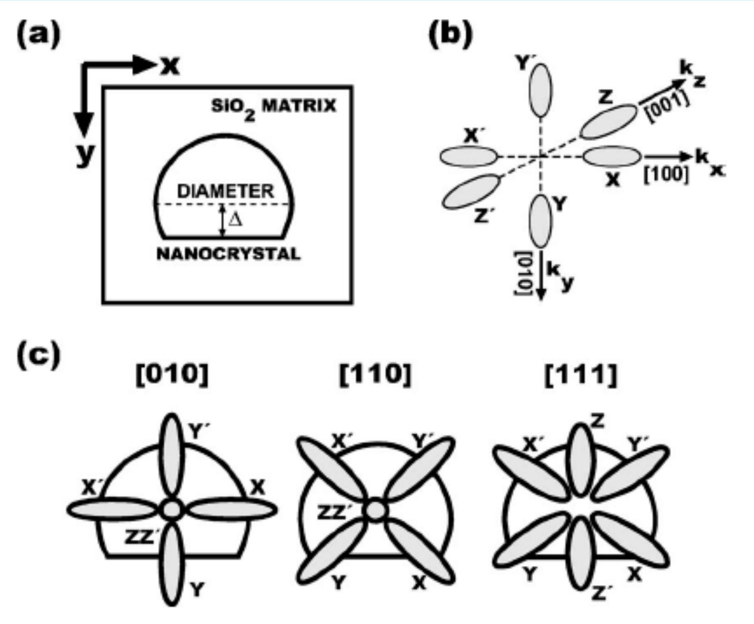


- Degeneracy among energy valleys
- Orbitals orientation follow the rotation of the effective mass tensor

- Lifting of energy valleys degeneracy
- Accidental degeneracies

Crystallographic Orientation Effects *

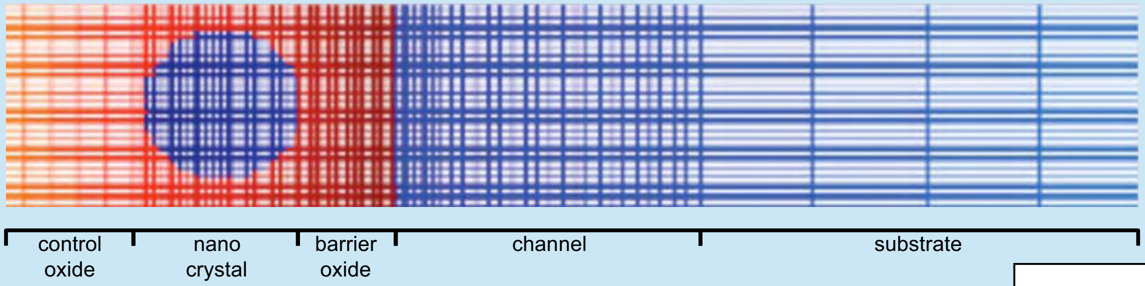




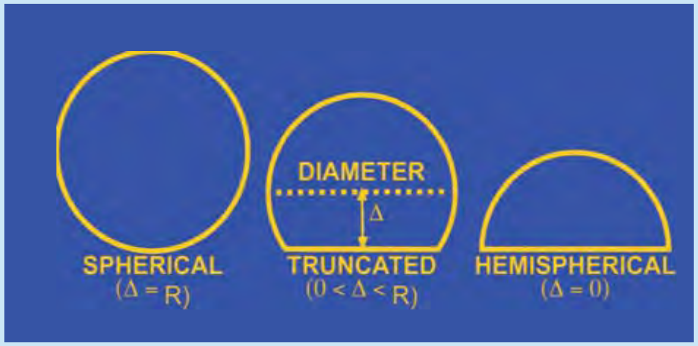
- Different crystalline orientations are responsible for accidental degeneracy
- $\Delta E_n < k_B T$ (room temperature) for the [010] orientation
- Minibands appear for the [110] orientation
- Despite of the non-symetrical shape, energy valleys degeneracy is recovered for the [111] orientation

Self-Consistent Device Modeling*

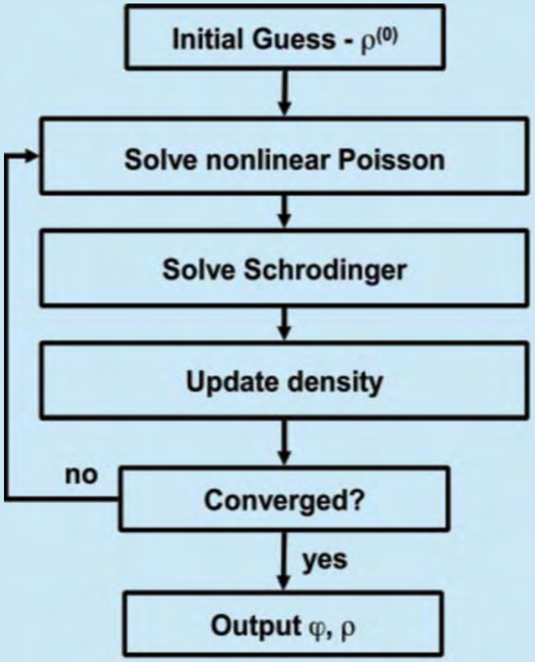
non-uniform grid (n_x=33, n_y=133, n_z=33)



- QD embedded in a MOS device
- Metallic gate
- Substrate thickness ~ 2µm
- Si band structure: effective mass anisotropy, energy valleys degeneracy and crystallographic orientation



Fully 3D Iterative Scheme



Initial Guess - $\rho^{(0)}$

Solve nonlinear Poisson

Solve Schrodinger

Update density

Converged?

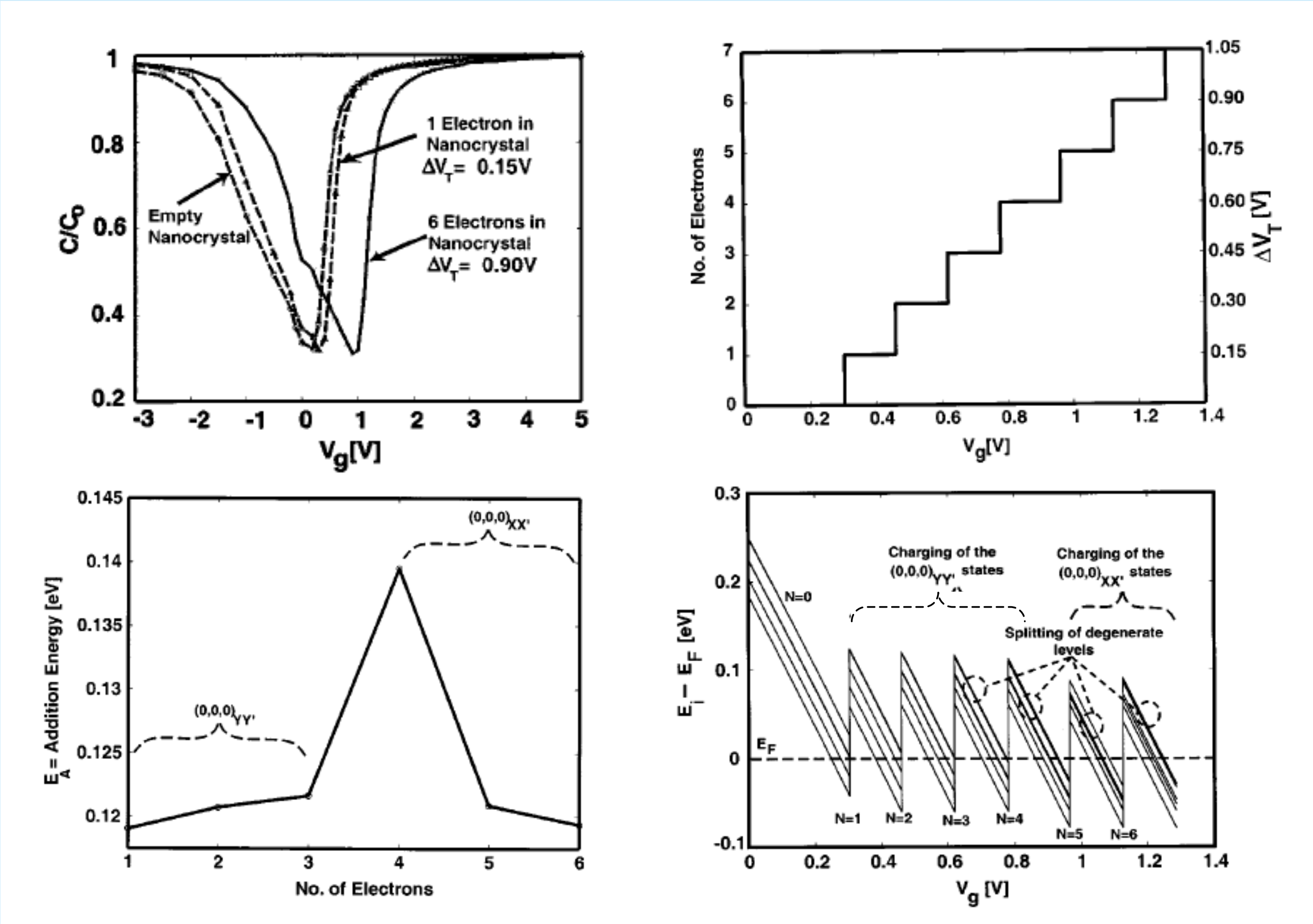
no

yes

Output φ, ρ

*A. Thean and J.-P. Leburton, J. Appl. Phys. 89, 2808 (2001)

Single Electron Charging: Statics*



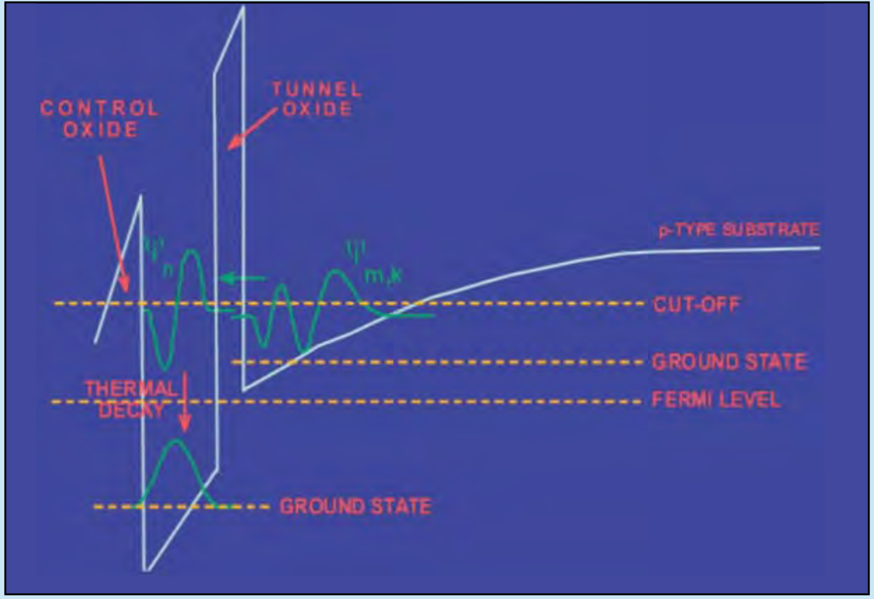
Spherical nanocrystal

D = 12.5 nm

*A. Thean and J.-P. Leburton, J. Appl. Phys. 89, 2808 (2001)

Data Operation Modeling: Dynamics*

Data programming



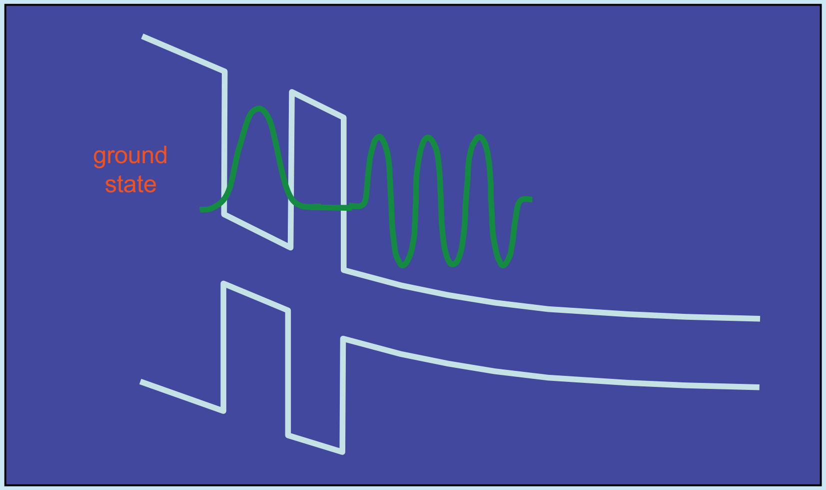
Bardeen Hamiltonian approach

$$\tau_{mn}^{-1} = \frac{2\pi}{h}\sum_{k_{//}} |M_{mn}|^2 f_{2D}(E_m)(1-f_{2D}(E_n))\delta(E_m - E_n)$$

$$M_{mn} \approx -\frac{h^2}{2m^*}\int_\Omega dr[\varphi_m^*\nabla\varphi_n - \varphi_n\nabla\varphi_m^*]$$

$$\tau^{-1} = \sum_{m,n}\tau_{mn}^{-1}$$

Data erase and retention



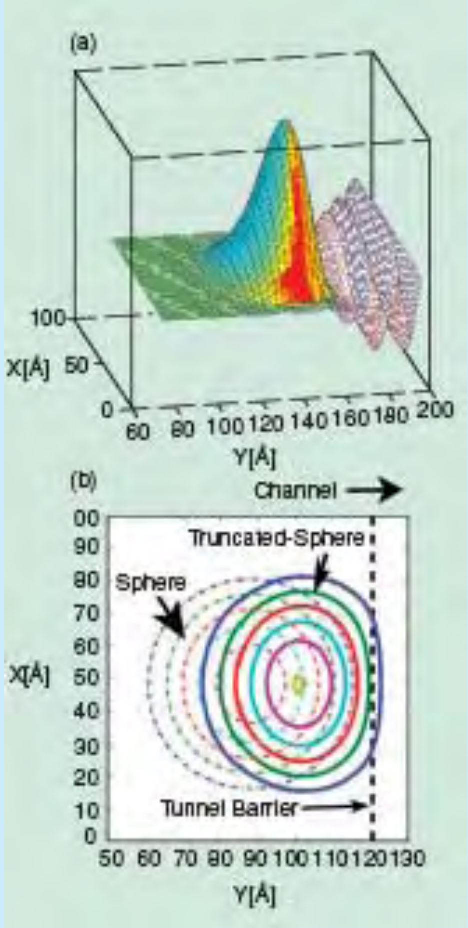
$$\varphi_n(x,y-y_0,z) = \varphi_n(x,y_0,z)e^{ik_y(y-y_0)}$$

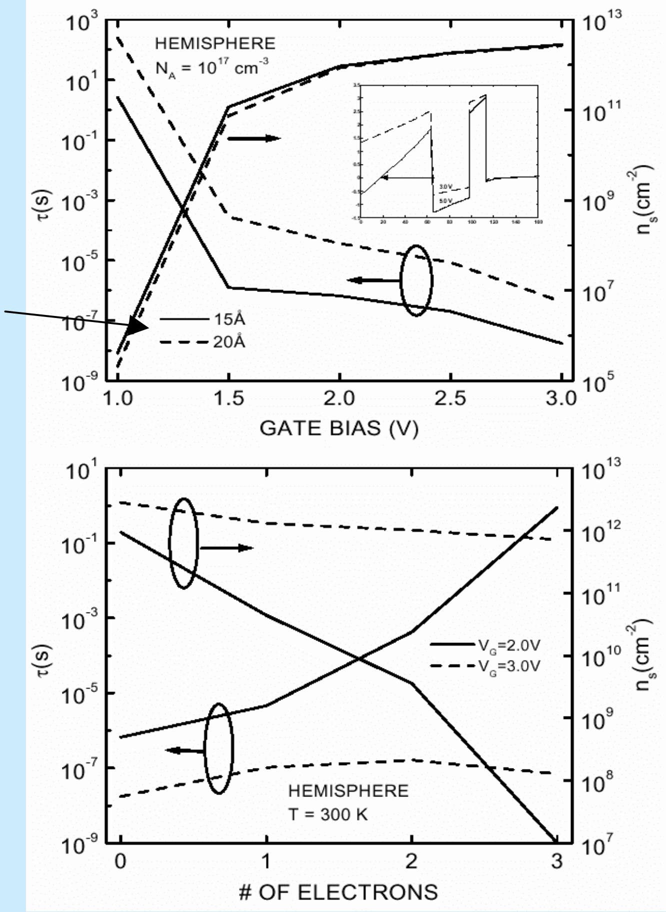
$$k_y = \frac{\sqrt{2m_{Si}(E_n-E_c)}}{h} \qquad \frac{1}{m_{ox}}\frac{\partial\varphi}{\partial y}\bigg|_{y_0^-} = \frac{1}{m_{Si}}\frac{\partial\varphi}{\partial y}\bigg|_{y_0^+}$$

$$\tau_R = \frac{h}{2\,\mathrm{Im}(E_n)}$$

*J. S. de Sousa et al, J. Appl. Phys. 92, 6182 (2002)
*J. S. de Sousa et al, Appl. Phys. Lett. 82, 2685 (2003)

Charging Time Dynamics*





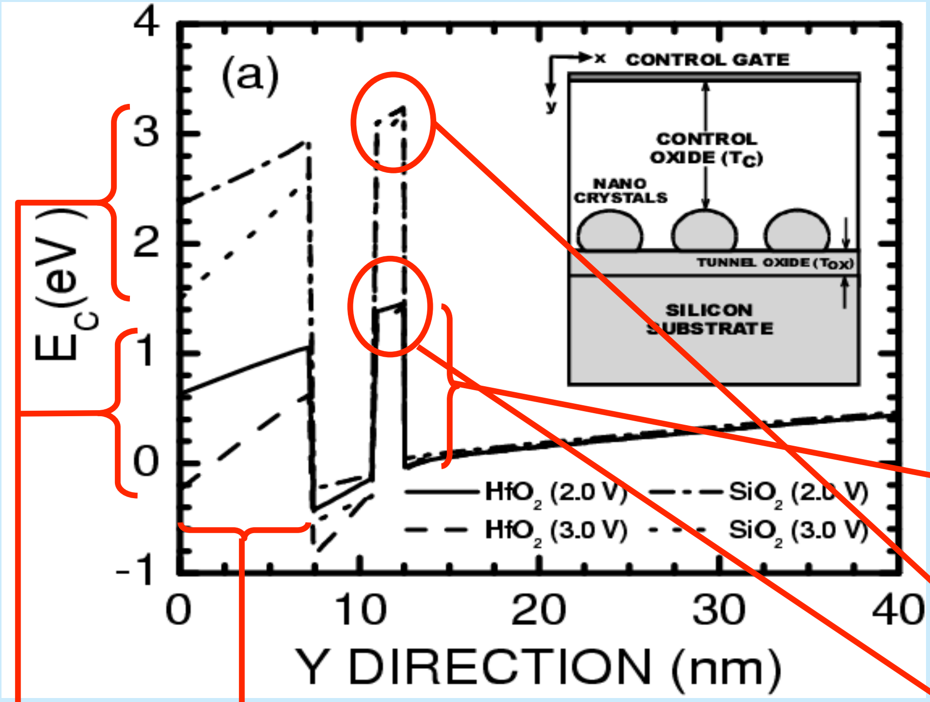
Tunneling barrier thickness

D=7nm

- Practical programming times (≤100 ns) are only achieved by combining very thin oxide barriers (≤ 20Å) and V_G >2.0V (consistent with experiment)
- Correlation between the average charging time and the number of electrons in the channel

*J. S. de Sousa et al, J. Appl. Phys. 92, 6182 (2002)

High-K Oxides:* Electrostatics



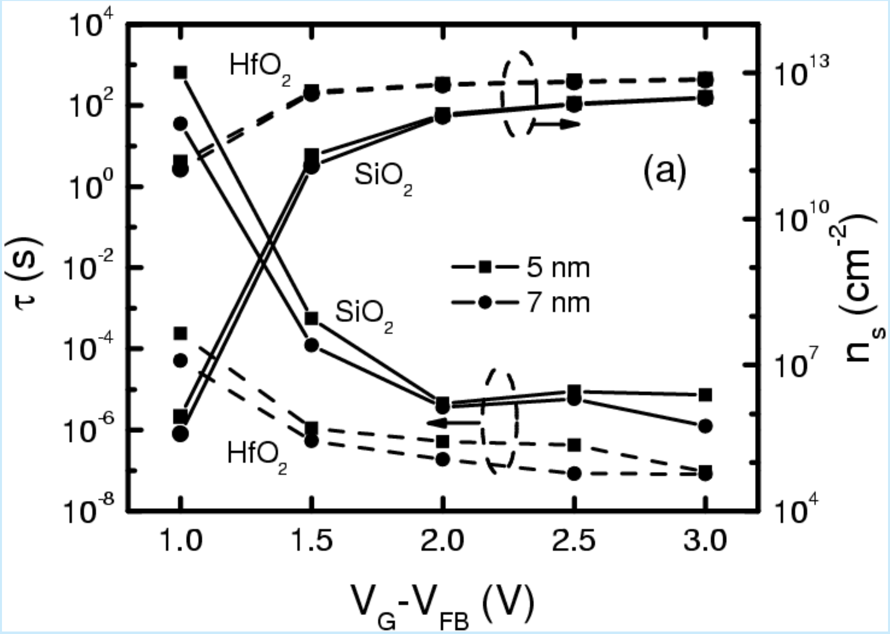
1st consequence: redistribution of the electrostatic potential (EP) across the device

Smaller HfO_2 ΔE_C may favor FN tunneling through the gate compromising data write and retention for V_G>2.5V. Thus, T_C must be increased (>20nm)

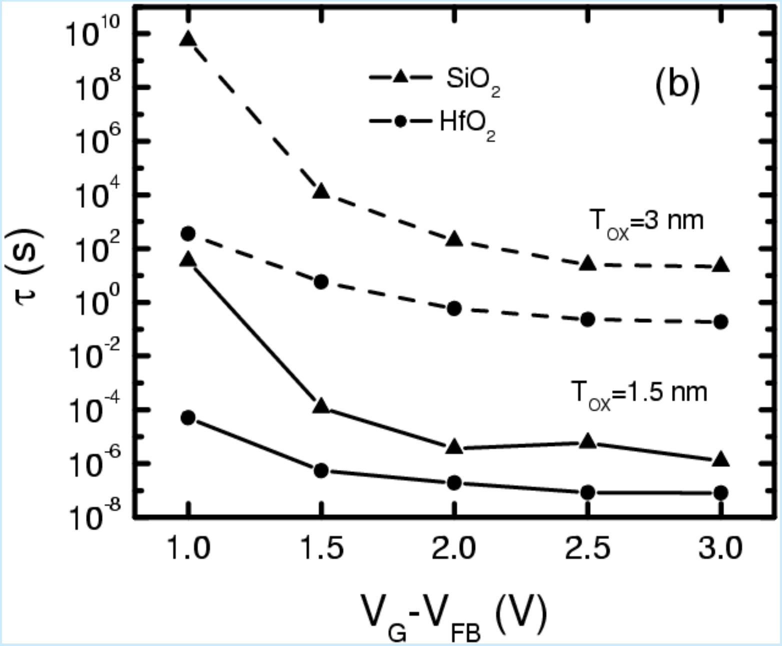
Concerns on the dielectric breakdown: F(HfO_2)=10MV/cm and F(SiO_2)=20 MV/cm. Quality of the oxide becomes crucial !!

EP drop in the oxide layer is larger forSiO_2 than for HfO_2

High-K Oxides: Programming



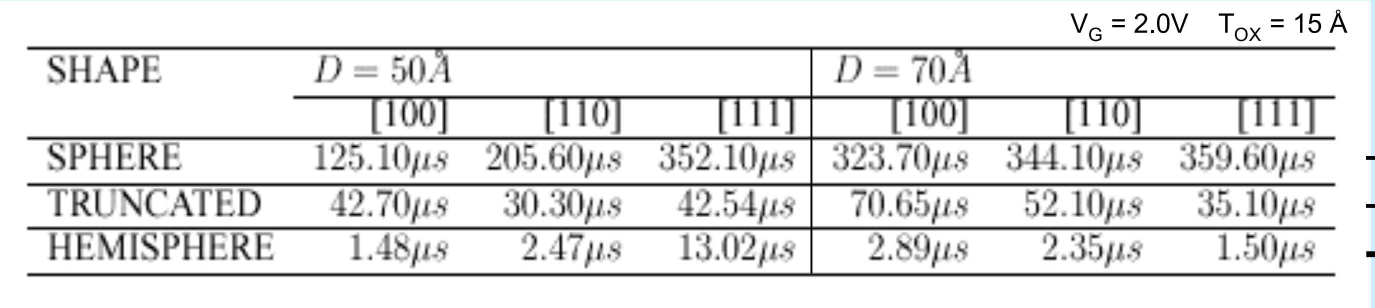
High-k materials increases write performance, but also decrease retention time (device reliability). Strategy: increase tunneling oxide thickness !



Main advantage: we can increase the tunneling oxide and still obtain good performances because of the smaller ΔE_C !

A tough problem: Data retention

The faster data are written ...



V_G = 2.0V T_{OX} = 15 Å

SHAPE	$D = 50\text{Å}$			$D = 70\text{Å}$		
	[100]	[110]	[111]	[100]	[110]	[111]
SPHERE	$125.10\mu s$	$205.60\mu s$	$352.10\mu s$	$323.70\mu s$	$344.10\mu s$	$359.60\mu s$
TRUNCATED	$42.70\mu s$	$30.30\mu s$	$42.54\mu s$	$70.65\mu s$	$52.10\mu s$	$35.10\mu s$
HEMISPHERE	$1.48\mu s$	$2.47\mu s$	$13.02\mu s$	$2.89\mu s$	$2.35\mu s$	$1.50\mu s$

J. S. de Sousa et al, J. Appl. Phys. 92, 6182 (2002)
J. S. de Sousa et al, Appl. Phys. Lett. 82, 2685 (2003)

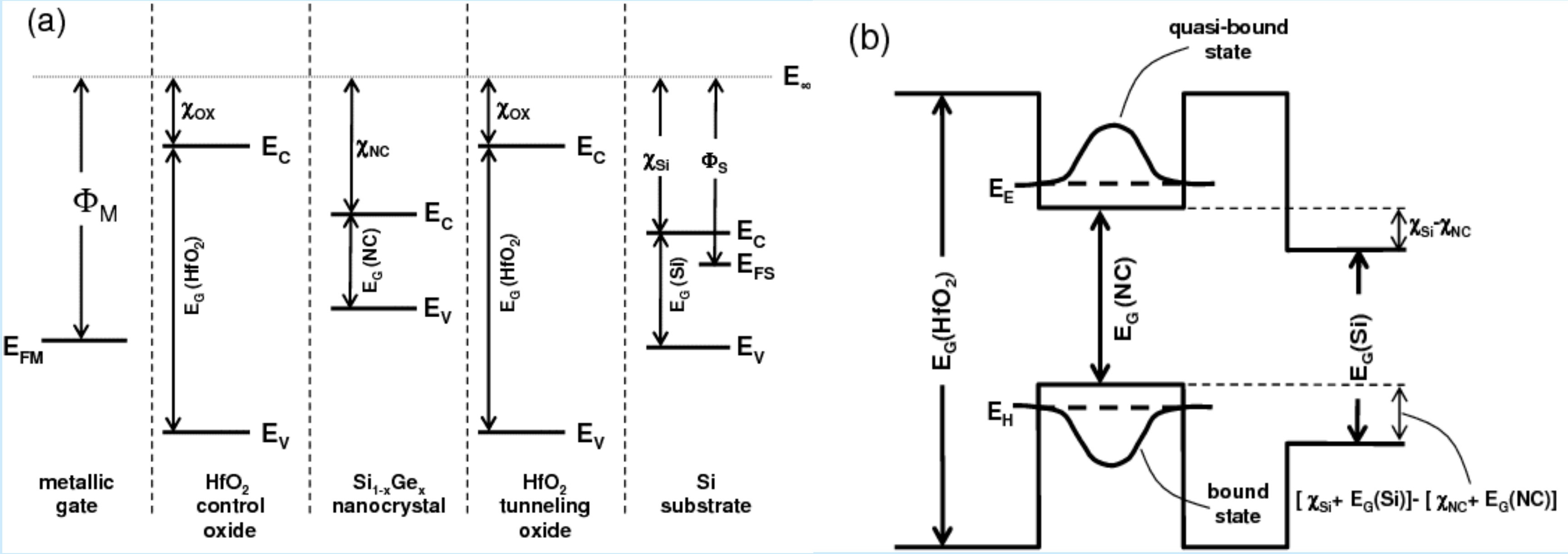
... the faster they are lost !

D= 70Å T_{OX} = 35Å

Shapes	Retention Time
Hemisphere	11 Days
Trunc. Sphere	3 Months
Sphere	10 Years

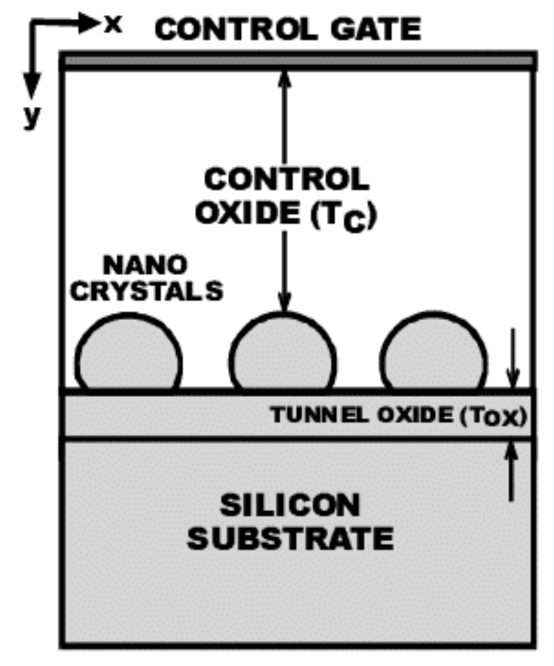
A. Thean et al., Proc. Nonvolatile Memory Technology Symp., 2000, pp.1621.

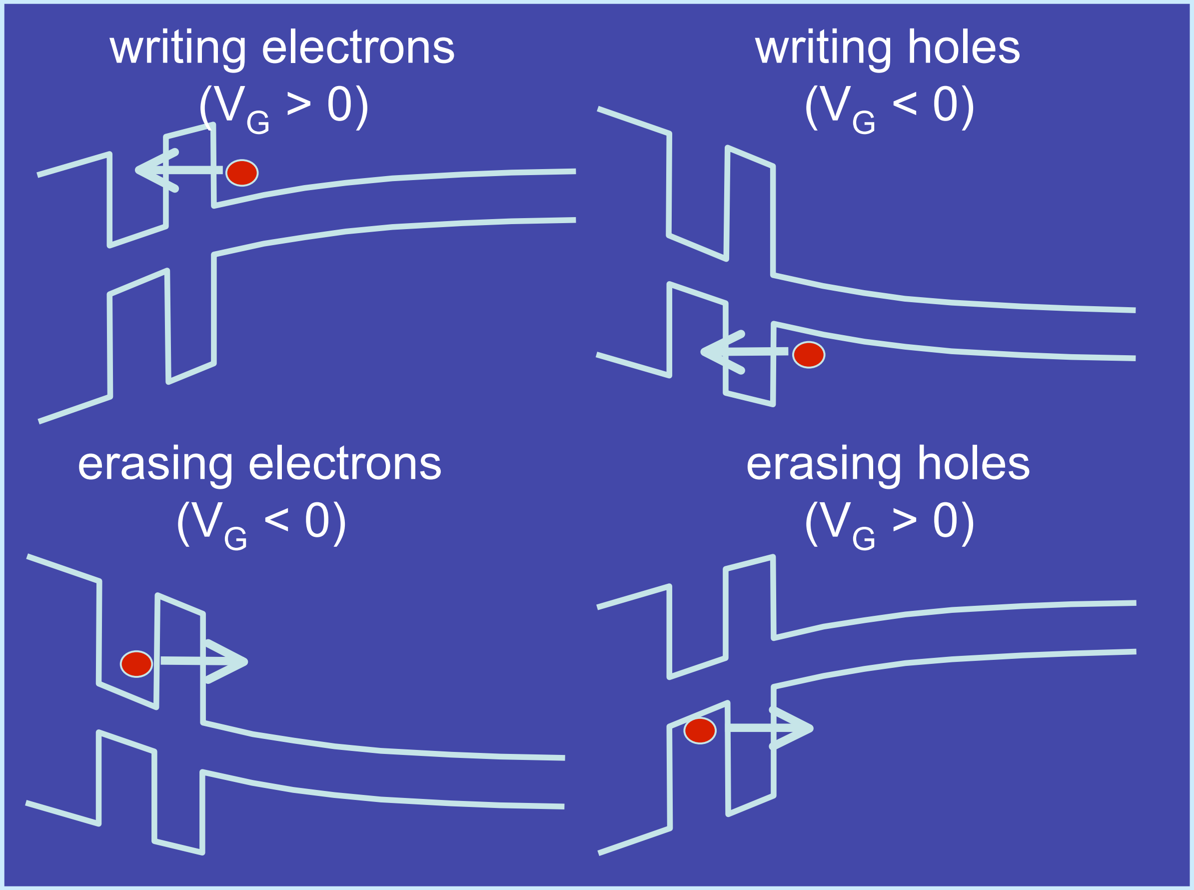
$Si_{1-x}Ge_x$ NC’s: Advantages



Due to the misaligment between the NC and substrate valence band edges, hole-based operations appear to be appropriate for simultaneous good programming performances and reliable data retention !

Electron & Hole Operations: Schematics




writing electrons ($V_G > 0$)

writing holes ($V_G < 0$)

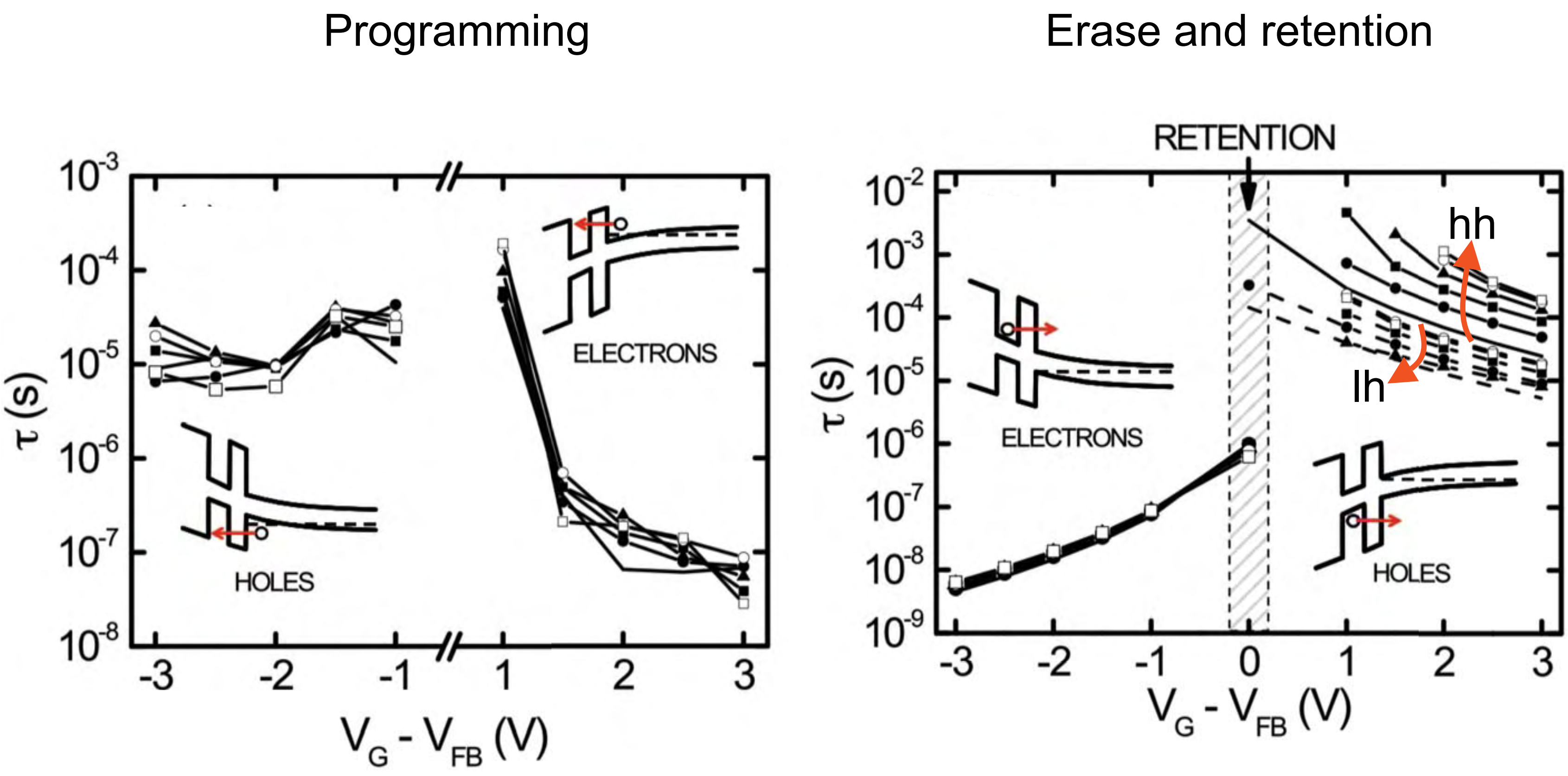
erasing electrons ($V_G < 0$)

erasing holes ($V_G > 0$)

	$V_G < 0$	$V_G = 0$	$V_G > 0$	V_{FB} *
e-based operation	Erase	Off	Write	~ -1.0 V
h-based operation	write	Off	Erase	~ - 0.2 V

$N_A = 10^{17}$ cm^{-3} for p-type Si, $N_D = 10^{17}$ cm^{-3} for n type Si and Al metallic gate.

Dynamical Performances*



Programming

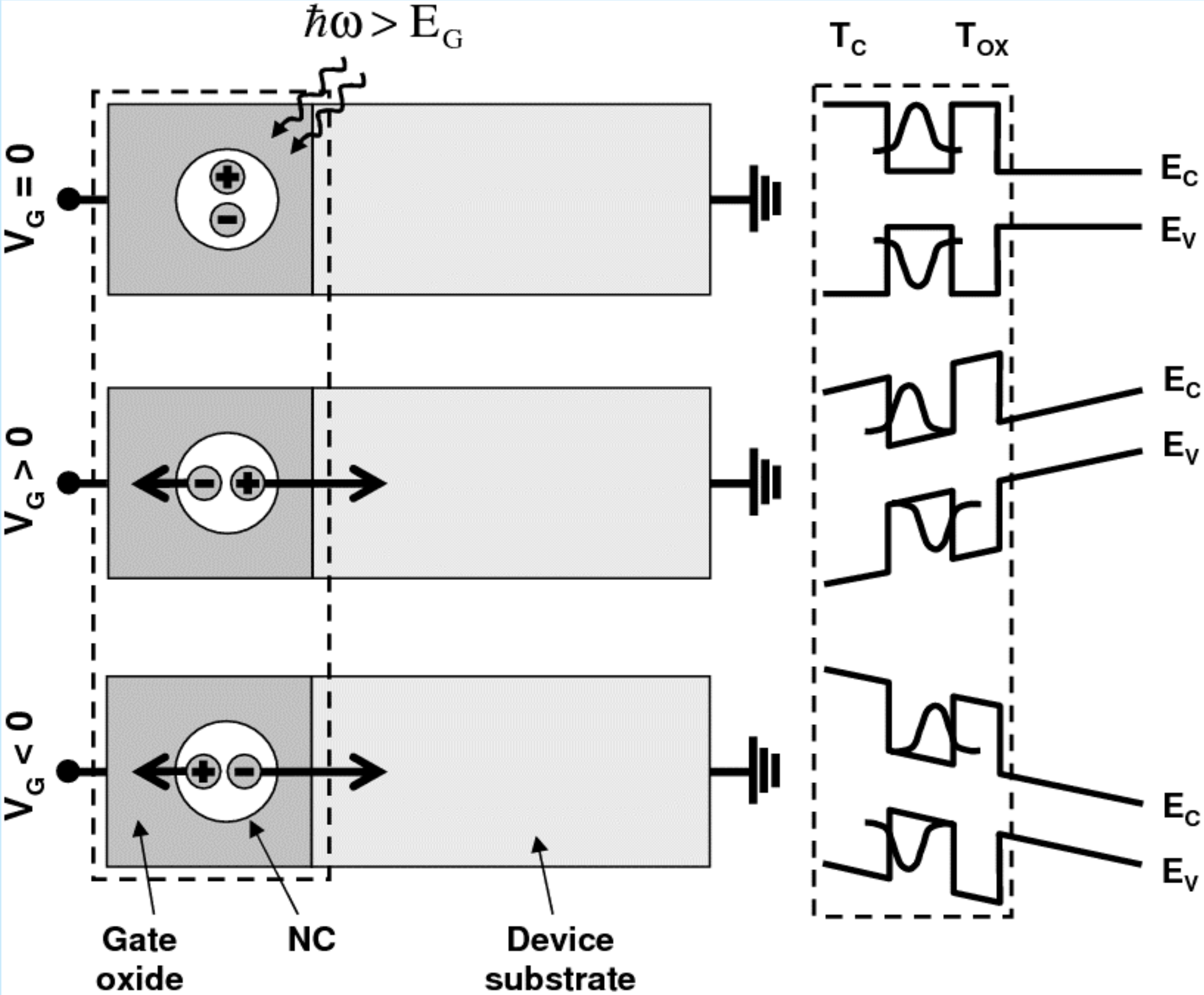
Erase and retention

x=0.0 (solid line)
x=0.2 (circle)
x=0.4 (square)

x=0.6 (triangle)
x=0.8 (open circle)
x=1.0 (open square)

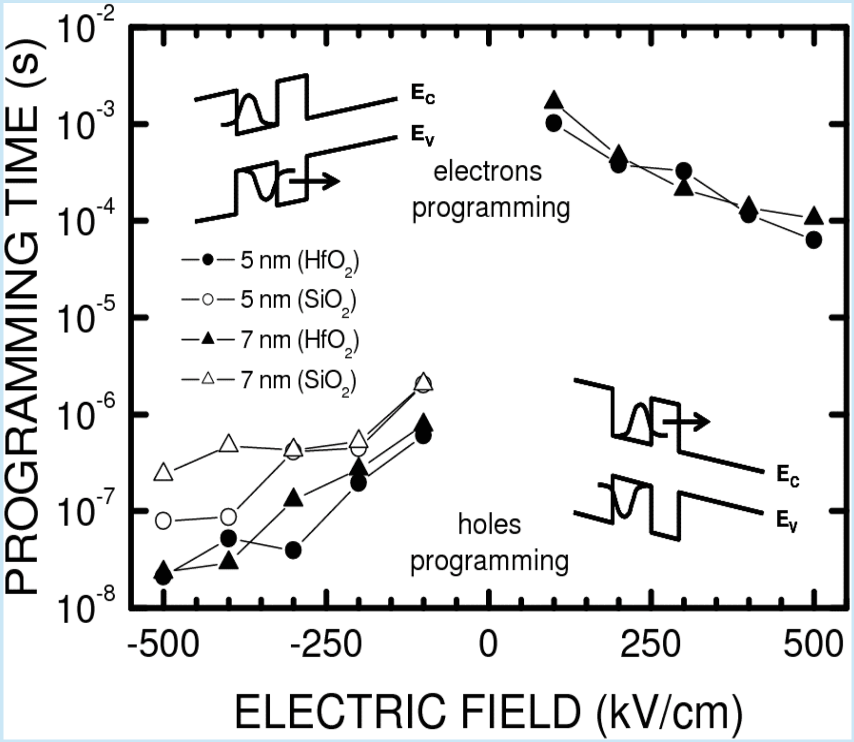
J. S. de Sousa et al., Appl. Phys. Lett. 90, 223504 (2007)

Optical Programming*



*J. S. de Sousa et al., Appl. Phys. Lett. 92, 103508 (2008)

Optical Programming*



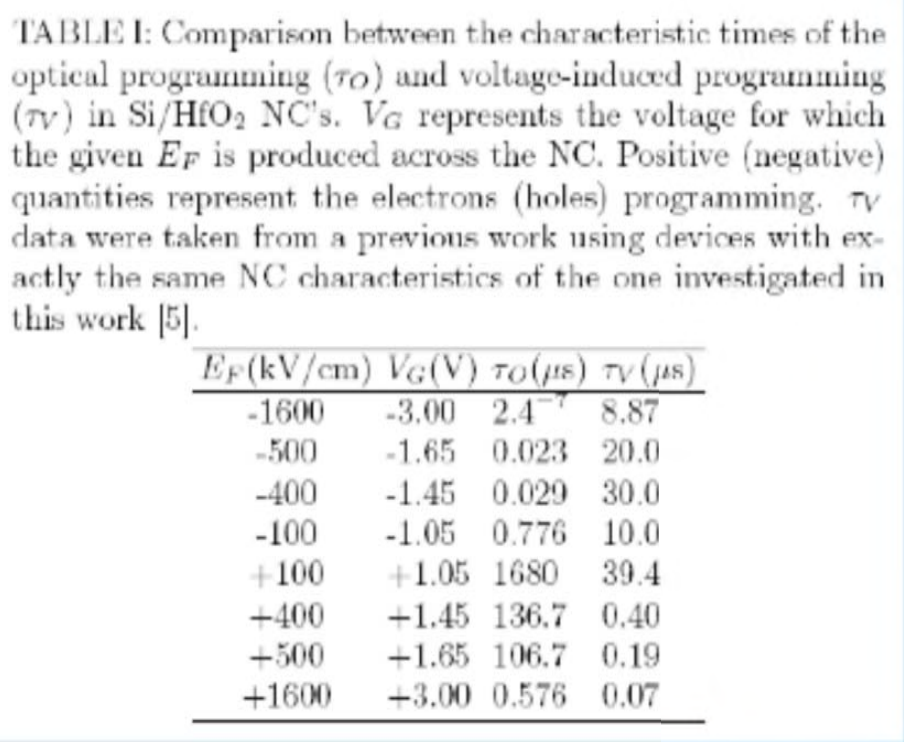


TABLE I: Comparison between the characteristic times of the optical programming (τ_O) and voltage-induced programming (τ_V) in Si/HfO_2 NC's. V_G represents the voltage for which the given E_F is produced across the NC. Positive (negative) quantities represent the electrons (holes) programming. τ_V data were taken from a previous work using devices with exactly the same NC characteristics of the one investigated in this work [5].

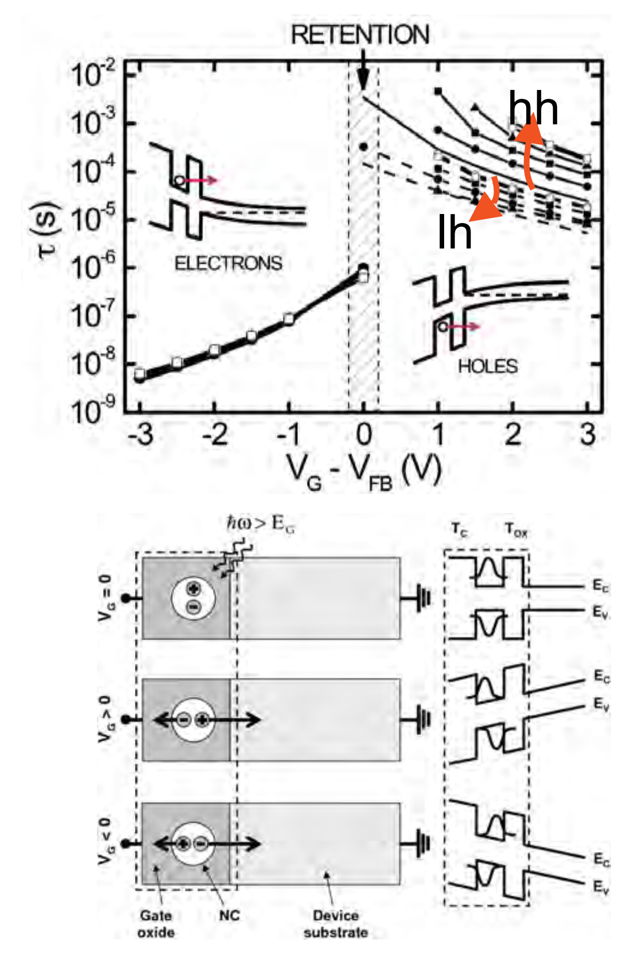
E_F(kV/cm)	V_G(V)	τ_O(μs)	τ_V(μs)
-1600	-3.00	2.4^{-7}	8.87
-500	-1.65	0.023	20.0
-400	-1.45	0.029	30.0
-100	-1.05	0.776	10.0
+100	+1.05	1680	39.4
+400	+1.45	136.7	0.40
+500	+1.65	106.7	0.19
+1600	+3.00	0.576	0.07

*J. S. de Sousa et al., Appl. Phys. Lett. 92, 103508 (2008)

Conclusions*

Nanocrystal flash memories

- Many device features can be used to optimize the trade-off between retention and preformance: NC shape, oxide thickness, high-k oxides.
- Although slower, hole-based device operation is suitable for long retention.
- Optical programming may lead to extremely fast memory operation without compromising data retention



*J. S. de Sousa et al., Appl. Phys. Lett. 92, 103508 (2008)