

Metal-Gate Work-Function Variability in Emerging Device Technologies

ECE 225
Lecture 17

Kaustav Banerjee

Department of Electrical and Computer Engineering, University of
California, Santa Barbara, CA 93106

Main References

Grain-Orientation Induced Work-Function Variation in Nanoscale Metal-Gate Transistors—Part I: Modeling, Analysis, and Experimental Validation

Hamed F. Dadgour, Kazuhiko Endo, Vivek De, and Kaustav Banerjee
IEEE Transactions on Electron Devices, Vol. 57, No. 10, pp. 2504-2514, 2010.

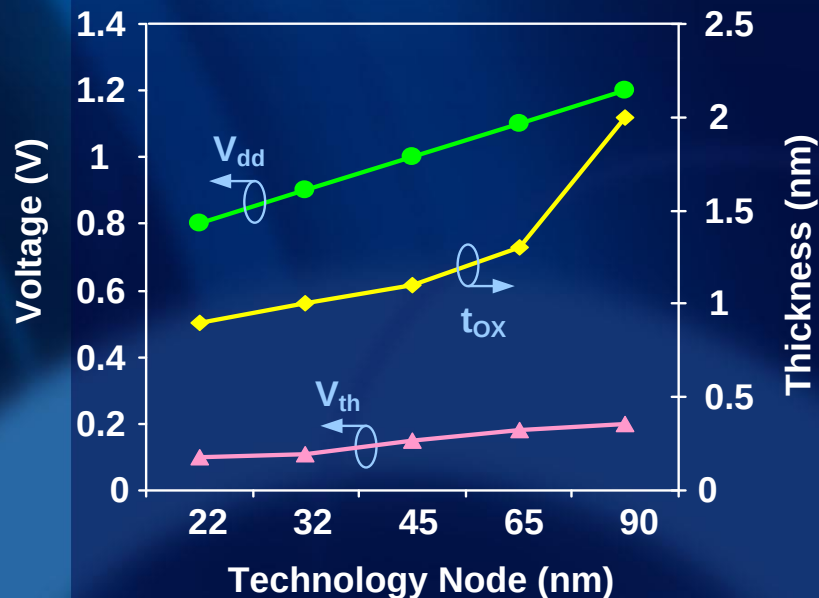
Grain-Orientation Induced Work-Function Variation in Nanoscale Metal-Gate Transistors—Part II: Implications for Process, Device, and Circuit Design

Hamed F. Dadgour, Kazuhiko Endo, Vivek De, and Kaustav Banerjee
IEEE Transactions on Electron Devices, Vol. 57, No. 10, pp. 2515-2525, 2010.

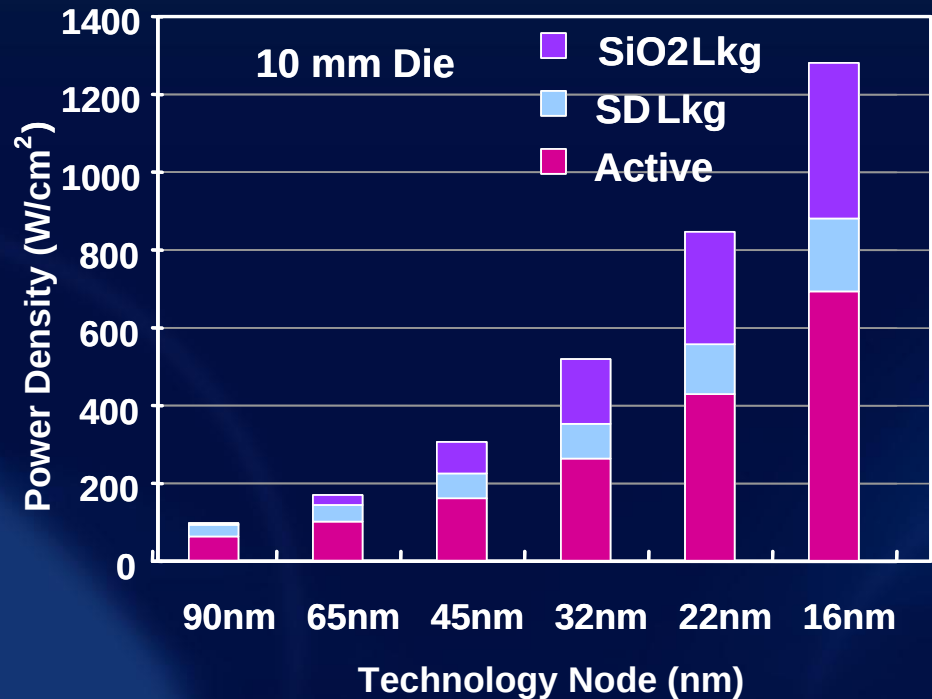
Outline

- Introduction
- Motivation
- Work Function Variation (WFV)
- Analytical modeling of WFV
- Implication of WFV for:
 - Fabrication process
 - Emerging devices
 - Circuit design
- Conclusions

CMOS Scaling Trends



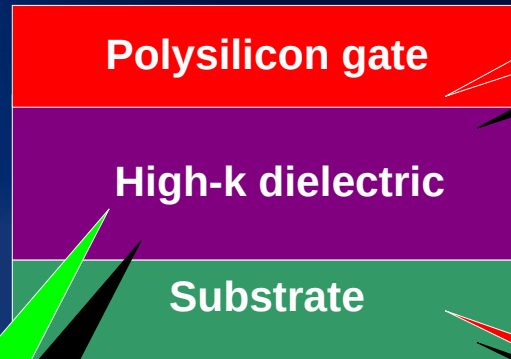
Source: ITRS 2005



Source: Intel

- Gate and subthreshold leakage increase dramatically with scaling device parameters

Introduction of High-k Dielectrics



Threshold
voltage
pinning



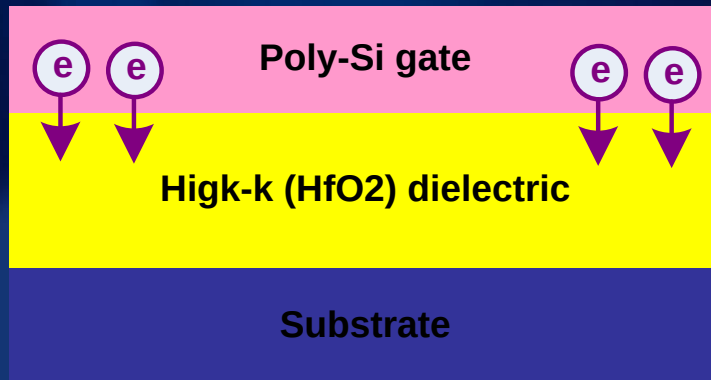
Low gate
leakage

Phonon
scattering

- Polysilicon gate on high-k dielectric results in:
 1. Threshold voltage pinning: difficulty in adjusting V_{th}
 2. Phonon scattering: reduced carrier mobility
- Solution is to replace polysilicon with metal gate

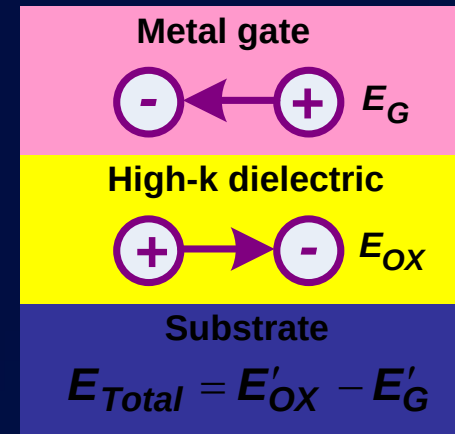
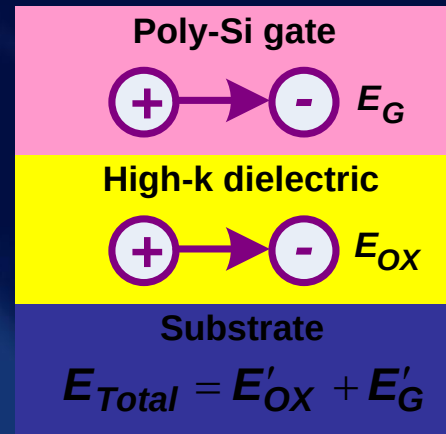
Why Metal Gates?

No threshold voltage pinning



- Metal, unlike Poly-si, does not react with high-k material to generate interface charges

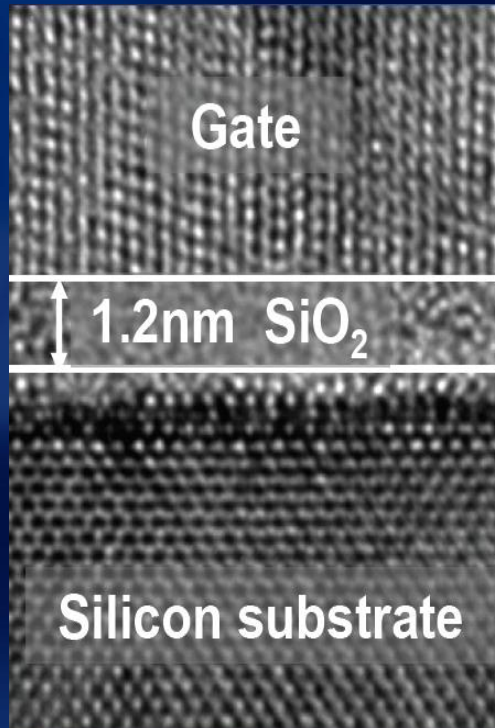
Reduced phonon scattering



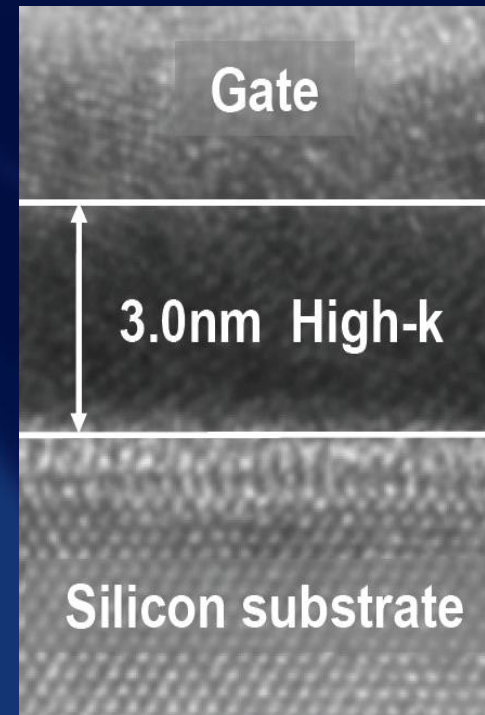
- Opposing electric dipoles in metal gate structure reduce the coupling of high-k phonons with carriers in the substrate

High-k/Metal Gate Technology

Gate leakage $\times 1$
Gate capacitance $\times 1$



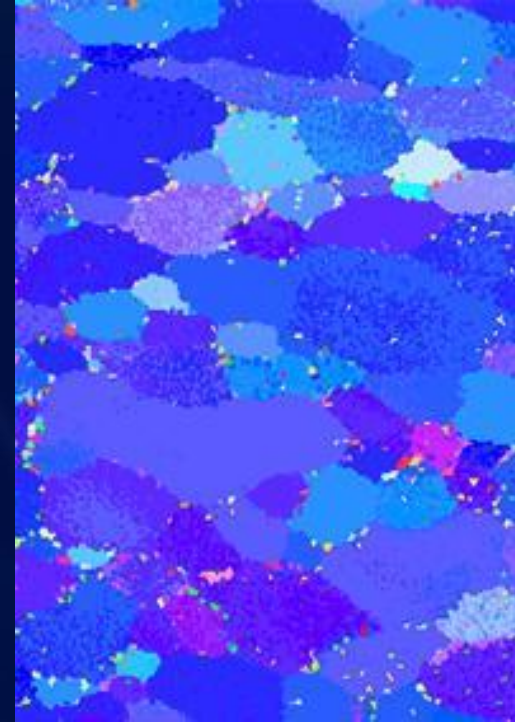
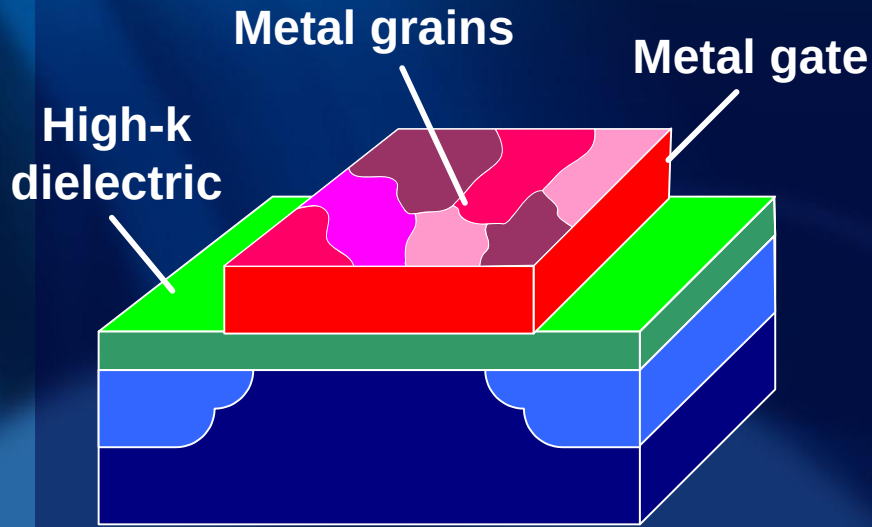
Gate leakage $\times 0.01$
Gate capacitance $\times 1.6$



Source: Intel

- High-k/metal gate devices offers lower gate leakage with small increase in gate capacitance

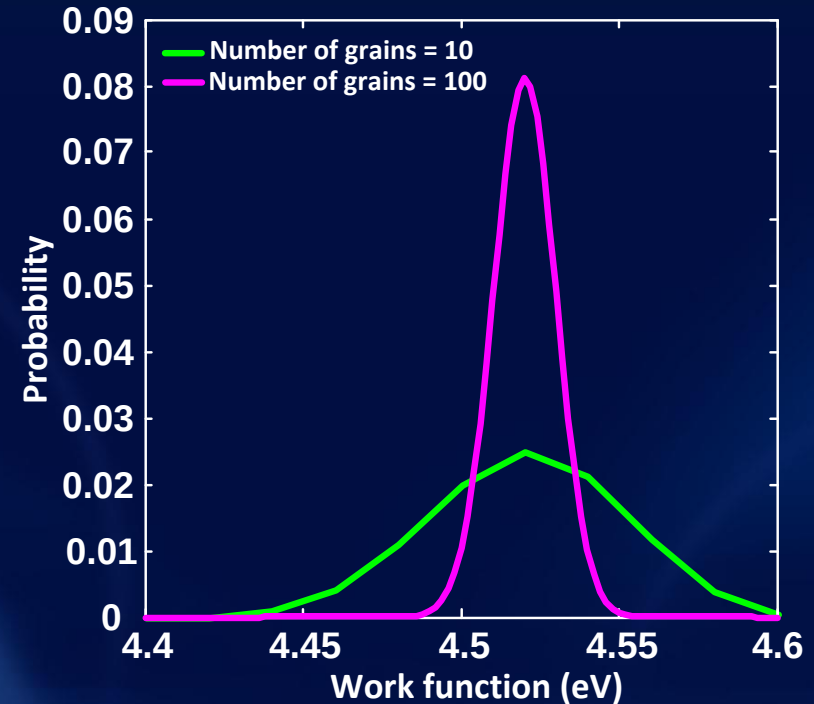
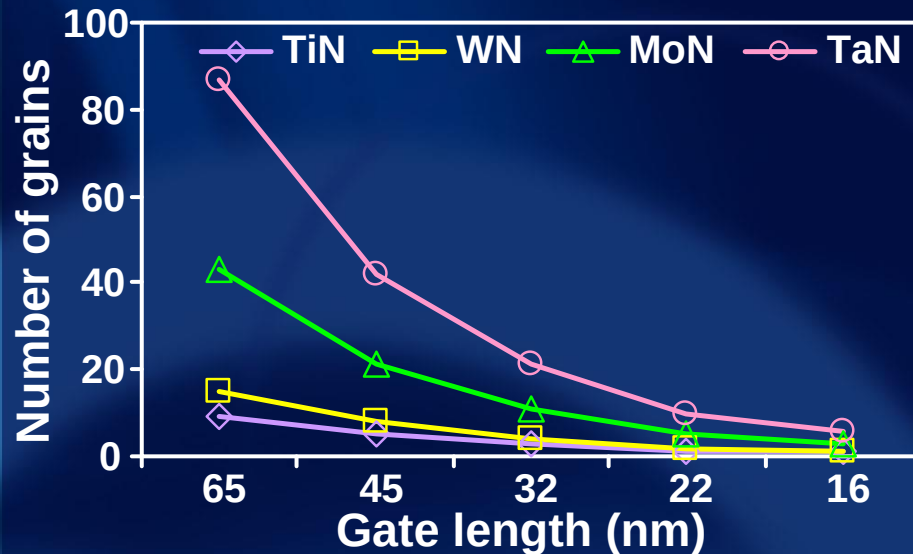
Work Function Variation (WFO)



TEM micrograph of an Al thin film
Source: www.mtalabs.com

- The grain orientation is uncontrollable during growth period
- Work function (and hence, threshold voltage) has random fluctuations

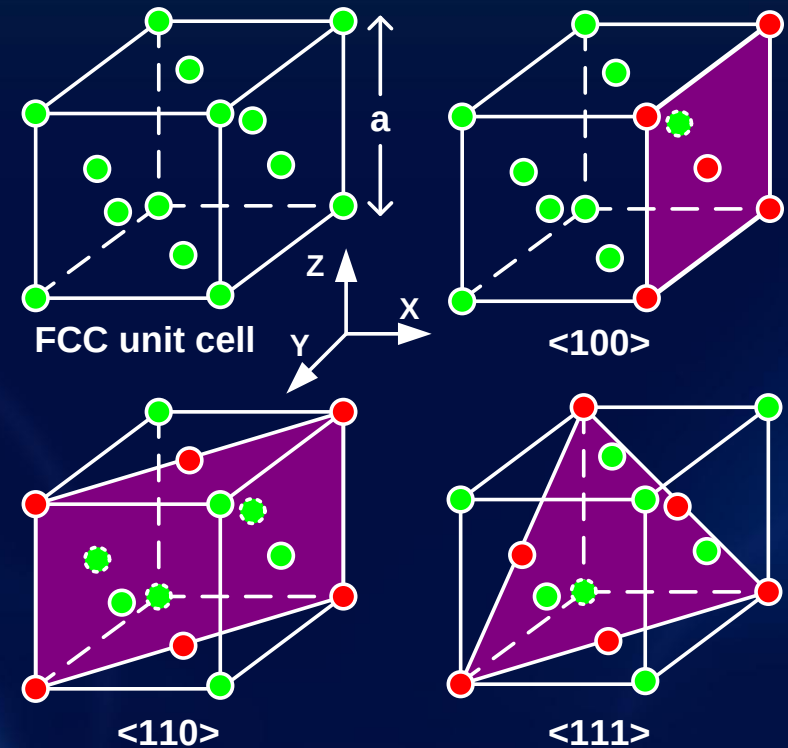
Work Function Variation (WFOV)



■ Work Function Variation (WFOV) increases with scaling

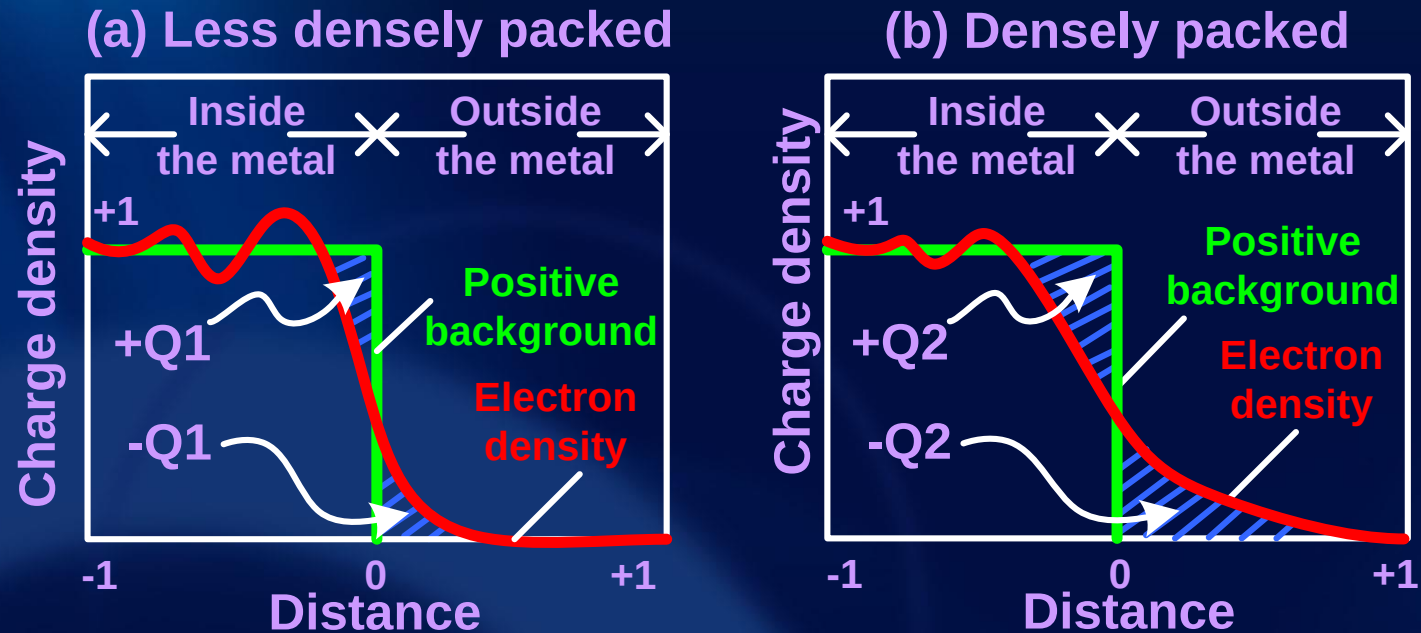
Surface Density

Orientation	Number of atoms	Area	Surface density
$\langle 111 \rangle$	2	$\frac{\sqrt{3}}{2} a^2$	$\frac{2.31}{a^2}$
$\langle 100 \rangle$ BCC	2	a^2	$\frac{2}{a^2}$
$\langle 100 \rangle$ FCC	2	a^2	$\frac{2}{a^2}$
$\langle 110 \rangle$	2	$\sqrt{2} a^2$	$\frac{\sqrt{2}}{a^2} \approx \frac{1.41}{a^2}$



- $\langle 111 \rangle$ orientation has the highest surface density followed by $\langle 100 \rangle$ and $\langle 110 \rangle$

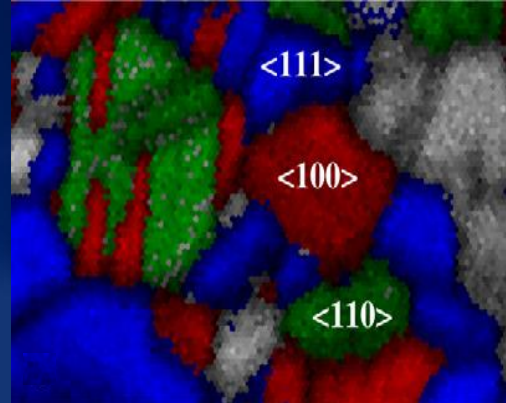
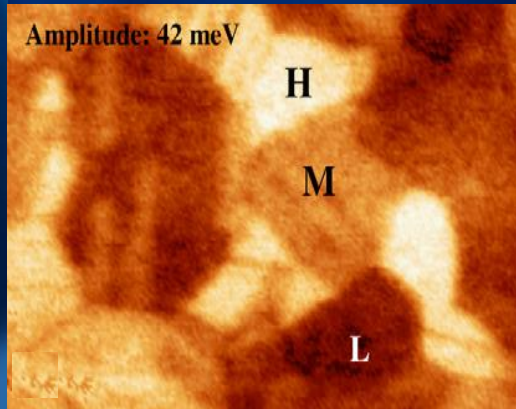
Dependency of WF on the Grain Orientation



N. Lang and W. Kohn, *Physical Review B* (1970)

- Densely packed orientation gives rise to stronger dipoles and hence, higher work function

Experimental Validation of WF Dependency on Grain Orientation



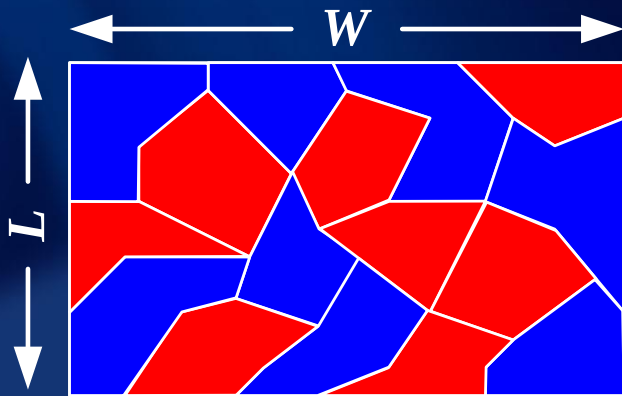
TEM picture of a deca-nanometer Copper wire
N. Gaillard et al., *Microelectronic Engineering* (2006)

Orientation	Surface density
<111>	$\frac{4}{\sqrt{3}a^2} \approx \frac{2.31}{a^2}$
<100>	$\frac{2}{a^2}$
<110>	$\frac{\sqrt{2}}{a^2} \approx \frac{1.41}{a^2}$

- Experimental measurements validates Lang and Kohn's model

Statistical Modeling of WFV

Known Parameters:



■ Φ_1, P_1
■ Φ_2, P_2

G = Average grain size

N = Total number of grains
 $= (W \times L) / G^2$

Unknown Variables:

X_i = Random variable representing number of grains with $WF = \Phi_i$

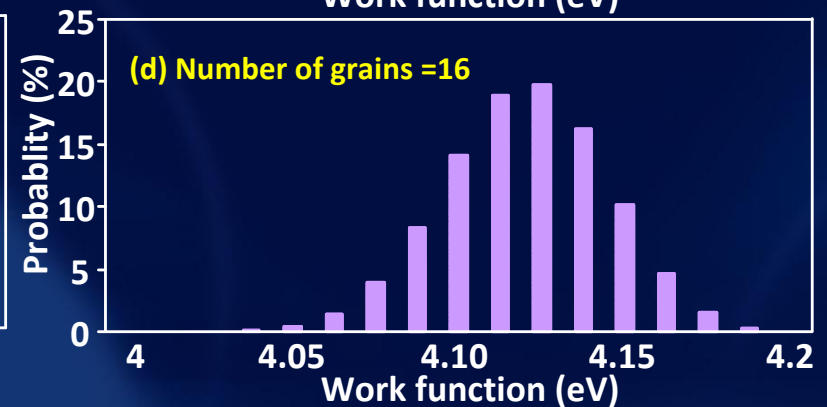
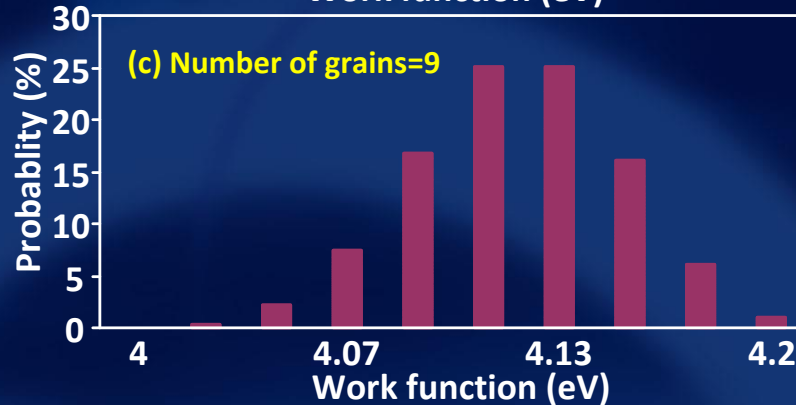
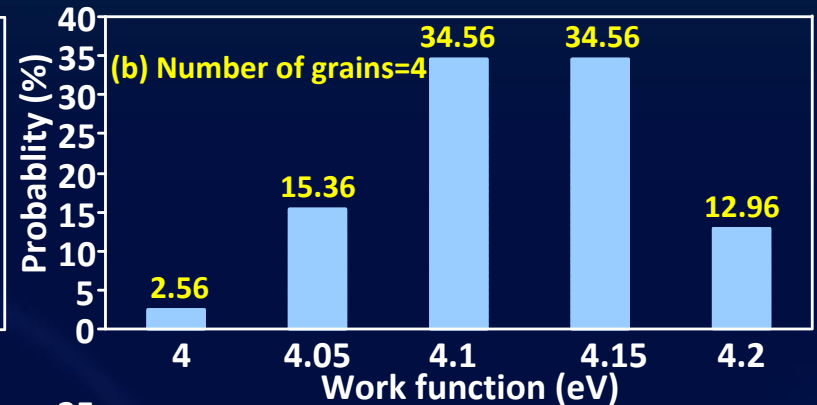
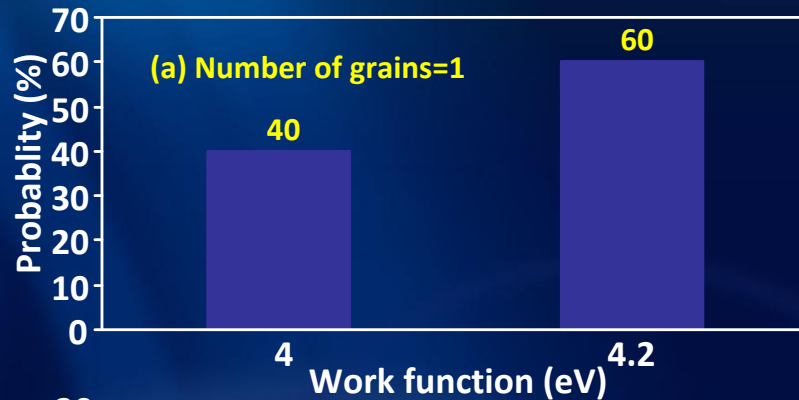
Φ_M = WF value of the metal gate

$$f_{X_1}(k) = \binom{N}{k} P_1^k (1 - P_1)^{N-k} \quad \binom{N}{k} = \frac{N!}{k!(N-k)!}$$

$$\Phi_M = \left(\frac{X_1}{N} \right) \Phi_1 + \left(\frac{X_2}{N} \right) \Phi_2 = \left(\frac{X_1}{N} \right) \Phi_1 + \left(\frac{N - X_1}{N} \right) \Phi_2$$

- Probability density function ($f_{X_1}(k)$) of getting exactly $X_1 = k$ of grains with WF Φ_1 can be calculated by a “binomial distribution”

Statistical Modeling of WFV



- For $N > 9$, the ϕ_M distribution resembles a Gaussian distribution

Statistical Modeling of WFFV

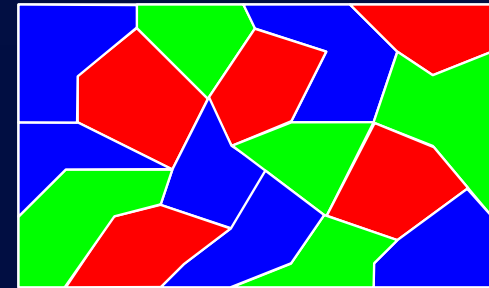
The special case with three grain orientations:

$$f_X(k_1, k_2, \dots, k_n) = \frac{N!}{k_1! k_2! \dots k_n!} P_1^{k_1} P_2^{k_2} \dots P_3^{k_3}$$

$$\Phi_M \sim N(\mu_{\Phi_M}, \sigma_{\Phi_M})$$

$$\mu_{\Phi_M} = P_1 \Phi_1 + P_2 \Phi_2 + P_3 \Phi_3$$

$$\sigma_{\Phi_M} = \frac{1}{N} (P_1 \Phi_1^2 + P_2 \Phi_2^2 + P_3 \Phi_3^2) - \frac{1}{N} (P_1 \Phi_1 + P_2 \Phi_2 + P_3 \Phi_3)^2$$



■ Φ_1, P_1
■ Φ_2, P_2
■ Φ_3, P_3

General case with r grain orientations:

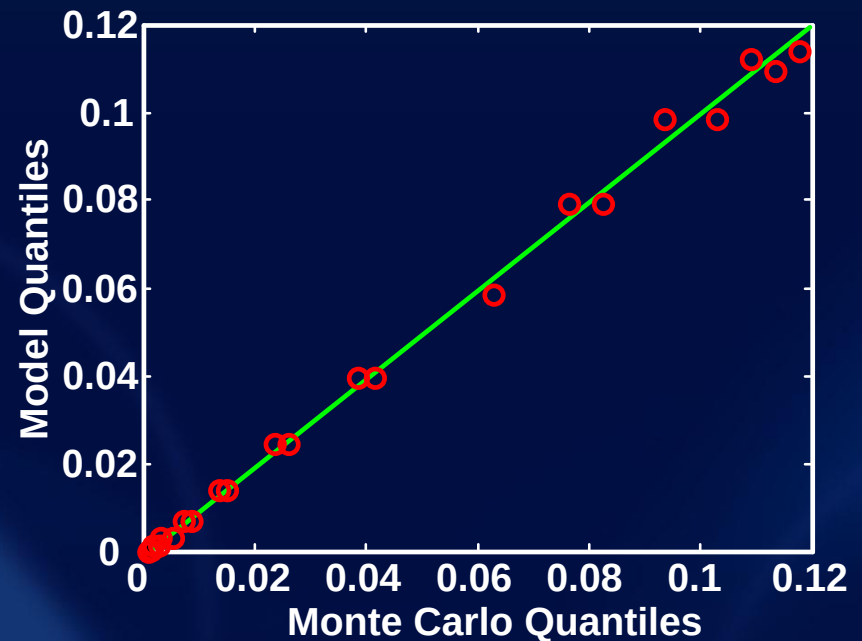
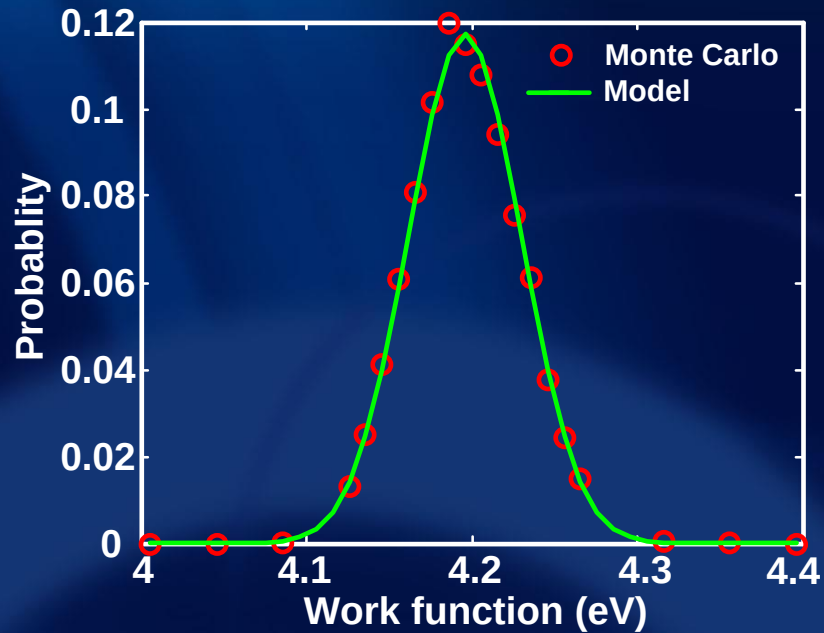
$$\Phi_M \sim N(\mu_{\Phi_M}, \sigma_{\Phi_M})$$

$$\mu_{\Phi_M} = \sum_{i=1}^r P_i \Phi_i$$

$$\sigma_{\Phi_M} = \frac{1}{N} \left[\sum_{i=1}^r P_i \Phi_i^2 - \left(\sum_{i=1}^r P_i \Phi_i \right)^2 \right]$$

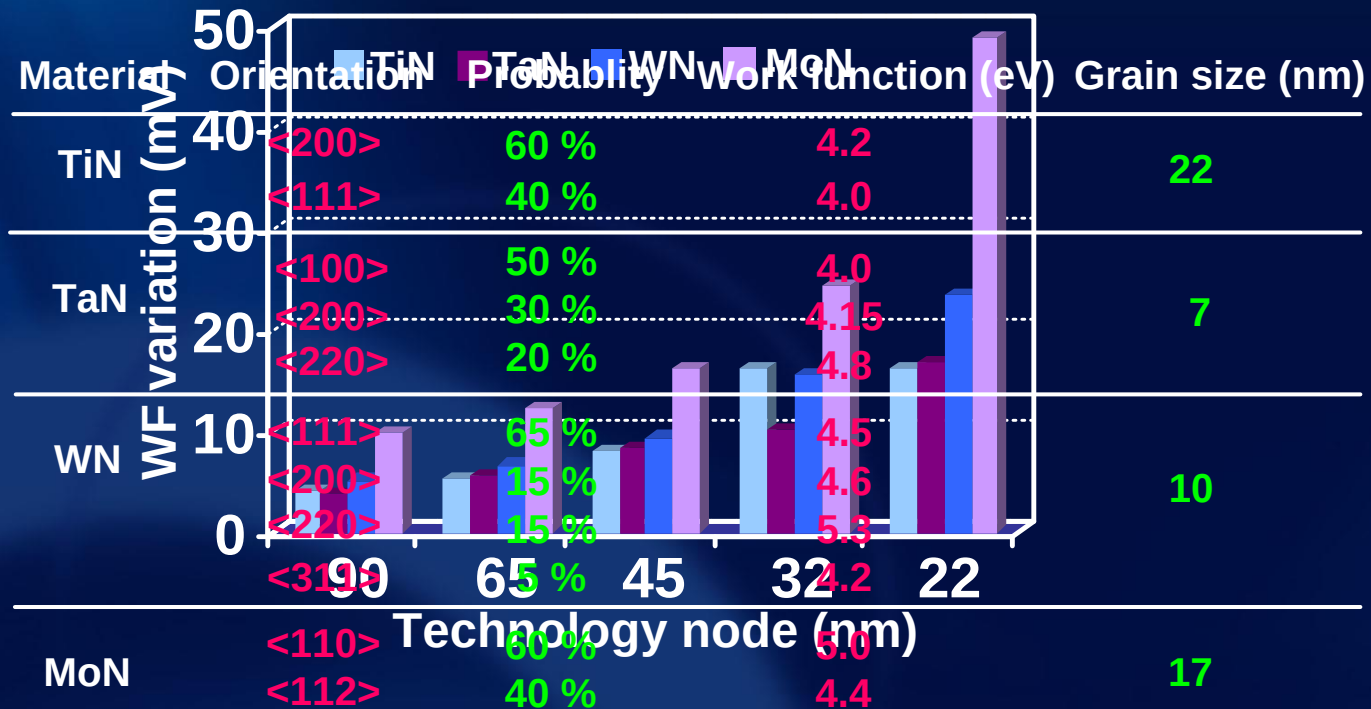
- For $N > 10$, the distribution of Φ_M can be approximated by a Gaussian distribution

Validation of the WFV Model



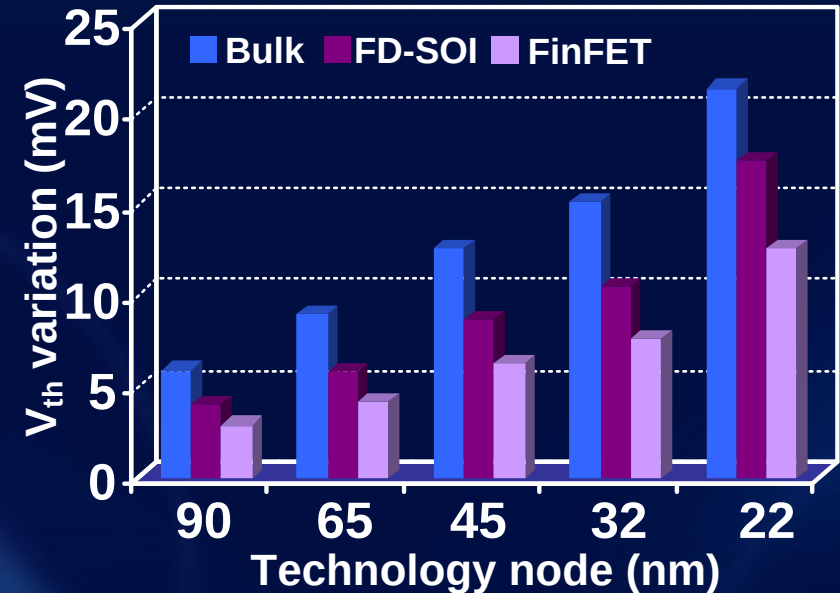
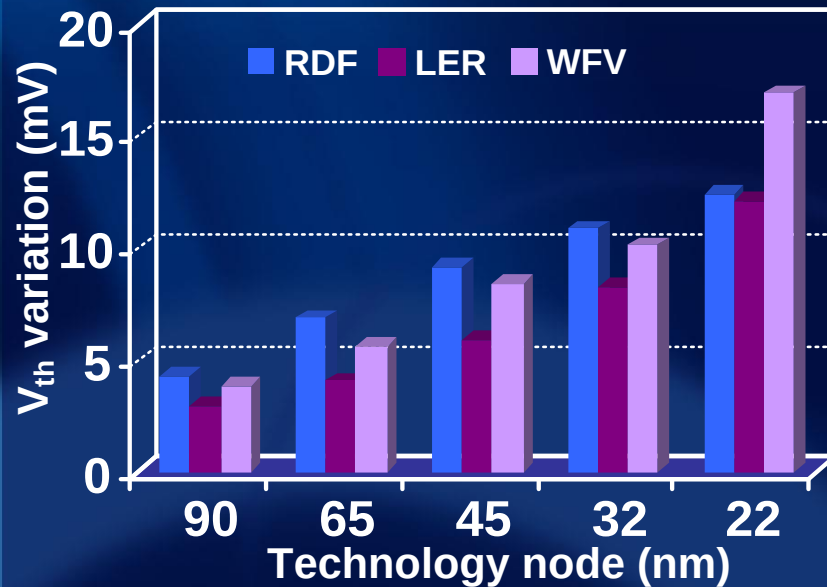
- The proposed model shows excellent agreement with Monte Carlo simulations

WFV for Different Gate Materials



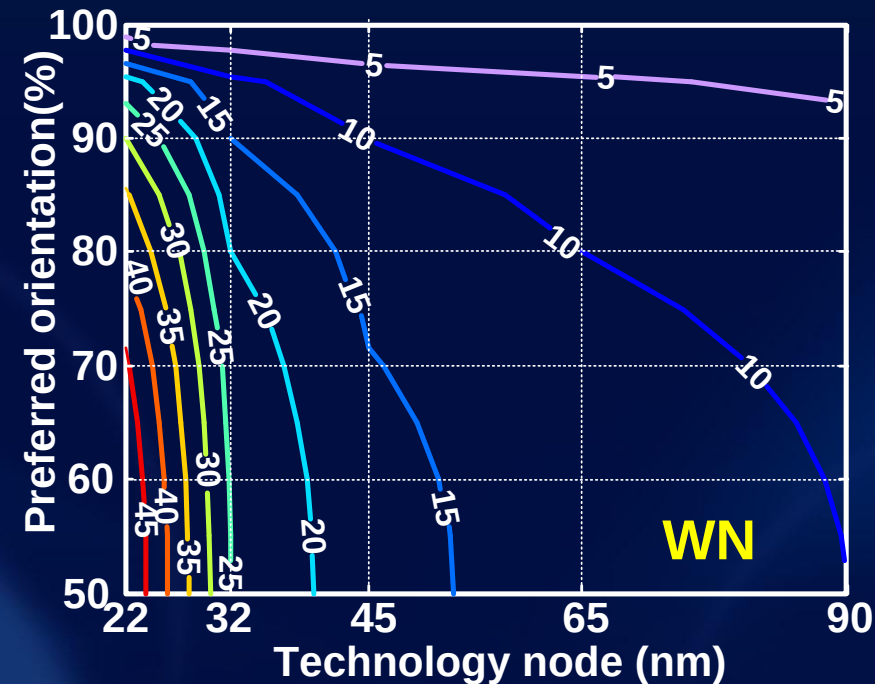
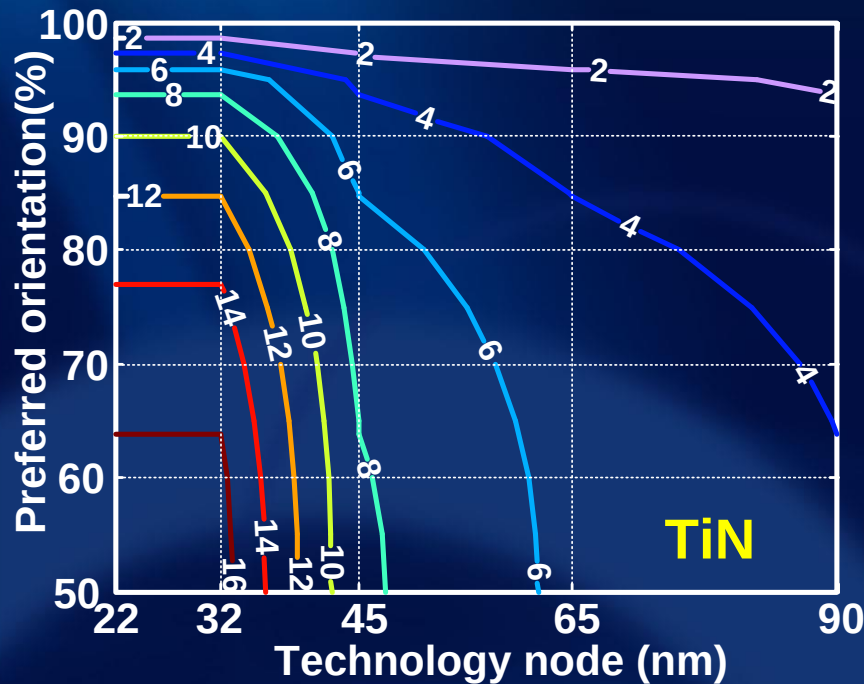
- V_{th} fluctuations in NMOS (TaN and TiN) and PMOS (MoN and WN) devices predicted by the proposed model

Device-Level Implications of WFV



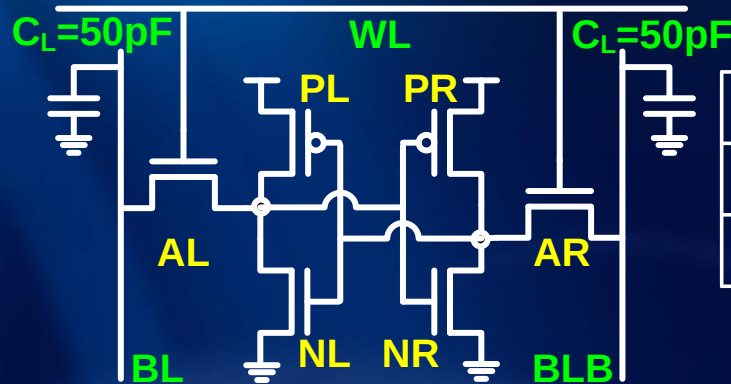
- For TiN based bulk devices, the impact of WFV can exceed those of RDF and LER in 22 nm node
- Compared to bulk and FD-SOI devices, FinFET transistors are less influenced by WFV due to larger gate area

Implications of WFV for Process Design



- Composition of the metal gates (ratio of the preferred orientation) to limit WFV to a certain upper-bound
- Each contour corresponds to a specific level of V_{th} variation
- To limit the V_{th} variation to 10 mV in 32-nm technology, the ratio of the preferred grain must be at least 90%, whereas for the MoN device, this ratio has to be no less than 95%.

The Impact of WFV on SRAM Reliability



$V_{dd} = 1.1 \text{ V}$ $L_{min} = 65 \text{ nm}$

Transistor	AR	AL	NR	NL	PR	PL
L	$2L_{min}$	$2L_{min}$	$2L_{min}$	$2L_{min}$	$2L_{min}$	$2L_{min}$
W/L	1.5	1.5	2	2	1.5	1.5

$$\beta = \left(\frac{(W/L)_{NR}}{(W/L)_{AR}} \right) = \left(\frac{(W/L)_{NL}}{(W/L)_{AL}} \right) \approx 1.33$$

$$SNM = SNM_0 + \frac{\partial SNM}{\partial V_{th,AR}} \Delta V_{th,AR} + \frac{\partial SNM}{\partial V_{th,AL}} \Delta V_{th,AL} + \dots + \frac{\partial SNM}{\partial V_{th,PL}} \Delta V_{th,PL}$$

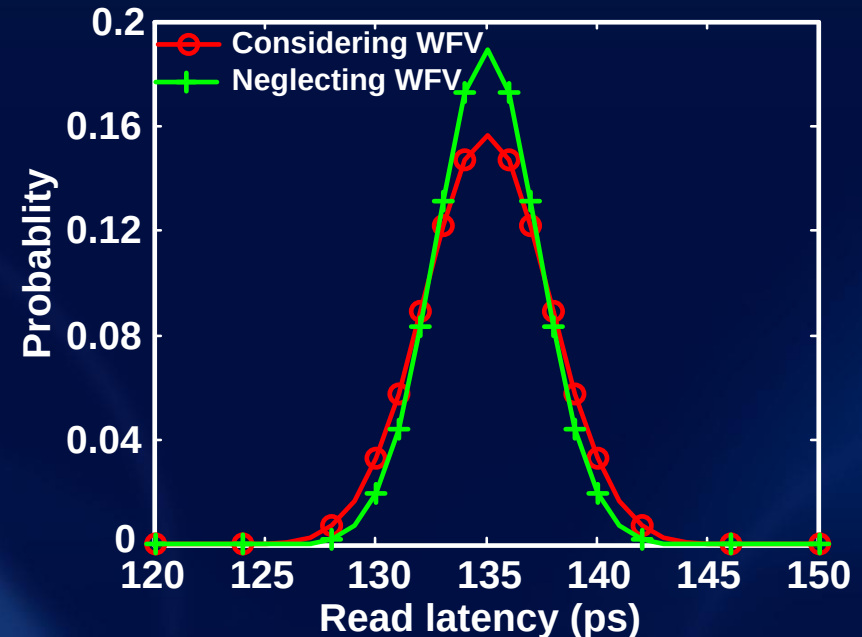
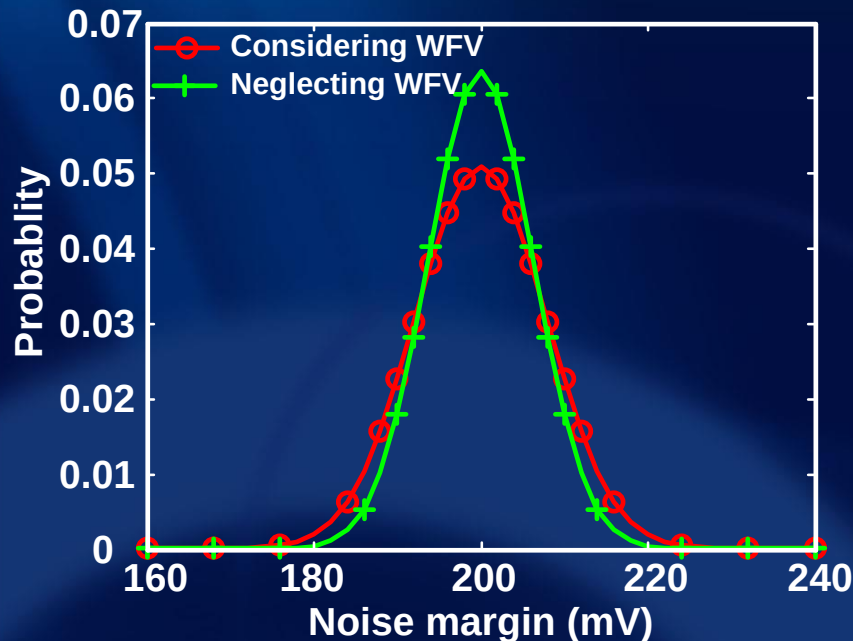
$$SNM \sim N(\mu_{SNM}, \sigma_{SNM})$$

$$\mu_{SNM} = SNM_0$$

$$\sigma_{SNM} = \sqrt{\left(\frac{\partial SNM}{\partial V_{th,AR}} \sigma V_{th,AR} \right)^2 + \left(\frac{\partial SNM}{\partial V_{th,AL}} \sigma V_{th,AL} \right)^2 + \dots + \left(\frac{\partial SNM}{\partial V_{th,PL}} \sigma V_{th,PL} \right)^2}$$

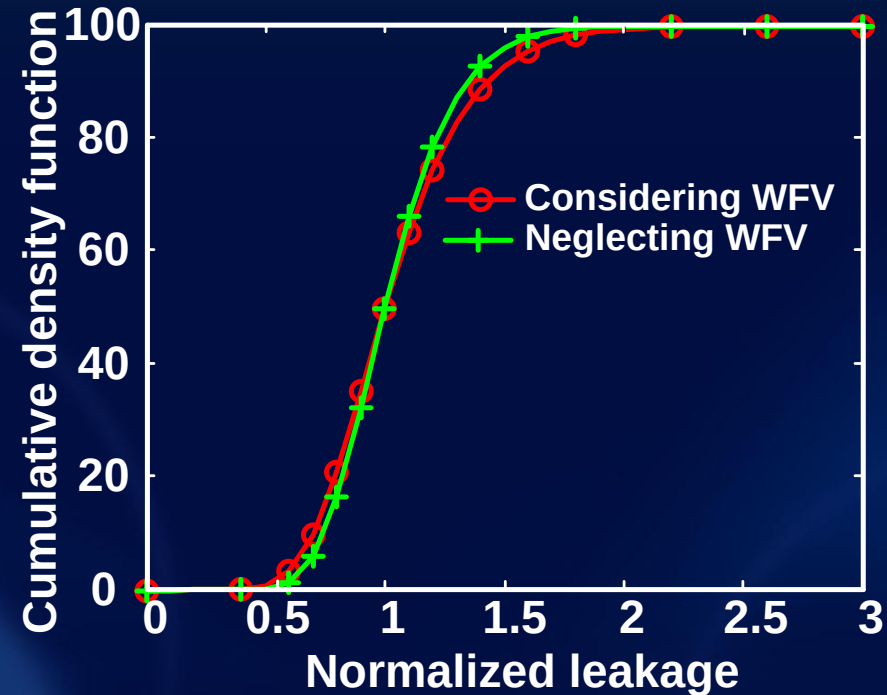
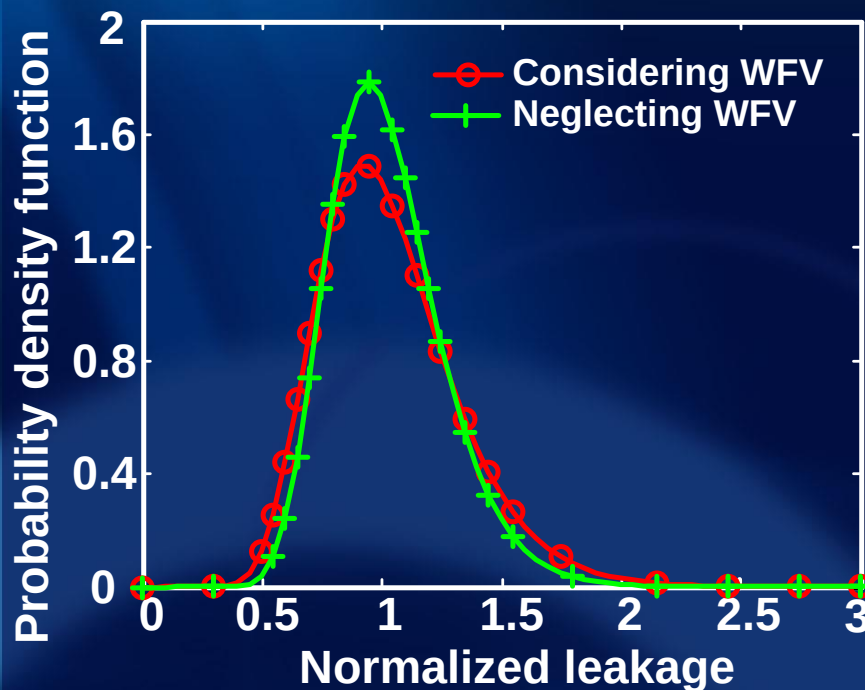
- Using Taylor expansion and the proposed framework, main characteristics of an SRAM cell can be statistically modeled

Analysis of SRAM Cells under WFV



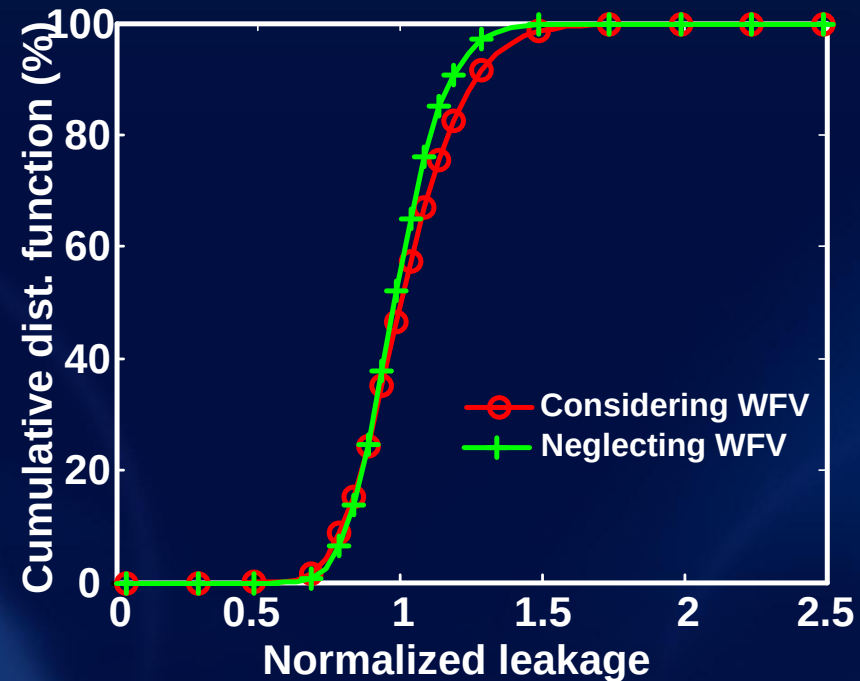
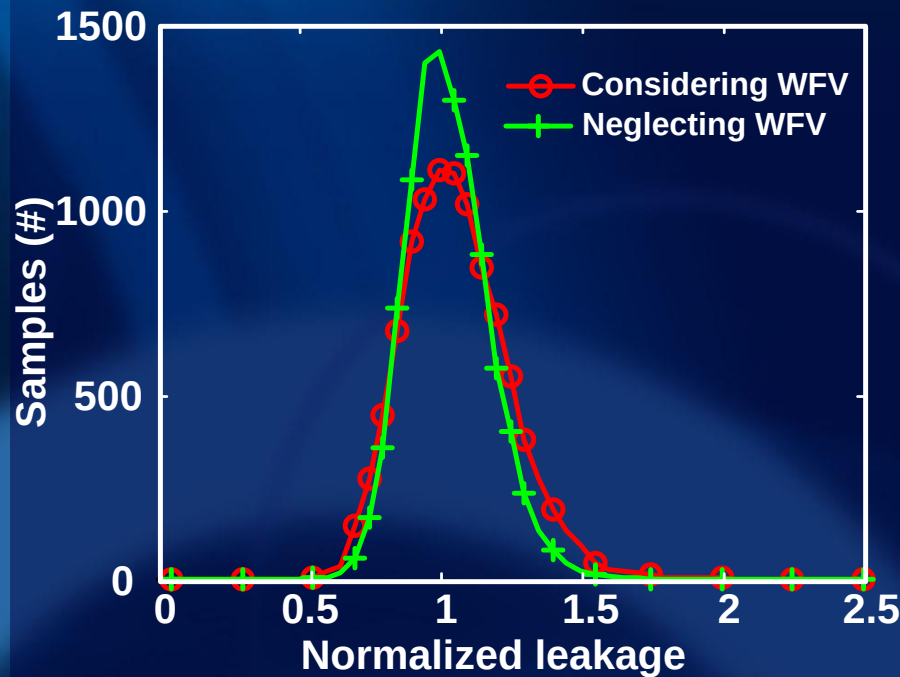
- Taking WFV into account increases the likelihood of the Noise margin and Read latency falling short of the design requirements

Analysis of Device Leakage under WFV



- If WFV is neglected, $\approx 96\%$ of devices will have less than $1.5X$ of the normalized leakage. However, considering WFV reduces the percentage of those devices to $\approx 93\%$

Analysis of SRAM Cells under WFV



- Ignoring WFV, 98% of cells will have less than $1.5\times$ the normalized leakage, whereas with WFV, the percentage of those devices decreases to $\approx 96\%$

Conclusions

- WFV is caused by the dependency of metal work-function on the orientation of its grains
- A statistical framework is developed, which enables estimation of the key parameters of work-function distribution

The proposed model can be used in:

1. Identifying suitable materials and fabrication processes (TiN / WN cause lower WFV for NMOS / PMOS metal gates)
2. Evaluating the WFV level of various types of transistors (Compared to bulk and FD-SOI devices, FinFET transistors are less influenced by WFV)
3. Investigating the impact of WFV on circuit reliability and performance (considering WFV reduces the percentage of cells with $1.5\times$ the nominal leakage from 98% to $\approx 96\%$)