Transistors for VLSI, for Wireless: A View Forwards Through Fog

Mark Rodwell, UCSB

Low-voltage devices
P. Long, E. Wilson, S. Mehrotra, M. Povolotskyi, G. Klimeck: ***Purdue***

III-V MOS
C.-Y. Huang, S. Lee, A.C. Gossard,*
V. Chobpattanna, S. Stemmer, B. Thibeault, W. Mitchell : ***UCSB***

InP HBT:
*J. Rode**, P. Choudhary, A.C. Gossard, B. Thibeault, W. Mitchell:* ***UCSB***
M. Urteaga, B. Brar: ***Teledyne Scientific and Imaging***

*Now with: *IBM, **Intel*

Co-authors

In(Ga)As MOS

Cheng-Ying Huang

Sanghoon Lee

Varista Chobpattanna

Prof. Susanne Stemmer

THz InP HBT

Johann Rode

Prateek Choudhary

Steep FET design

Pengyu Long

Evan Wilson

Prof. Michael Povolotski

Prof. Gerhard Klimeck

III-V EPI

Prof. Art Gossard

Process

Brian Thibeault

Bill Mitchell

..and at Teledyne (HBT): ***Miguel Urteaga, Bobby Brar***

What's a Professor to do ?

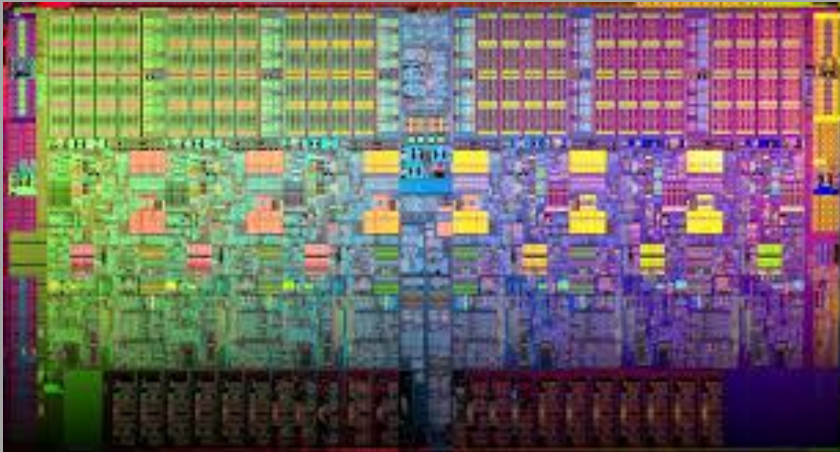
Transistors approaching scaling limits

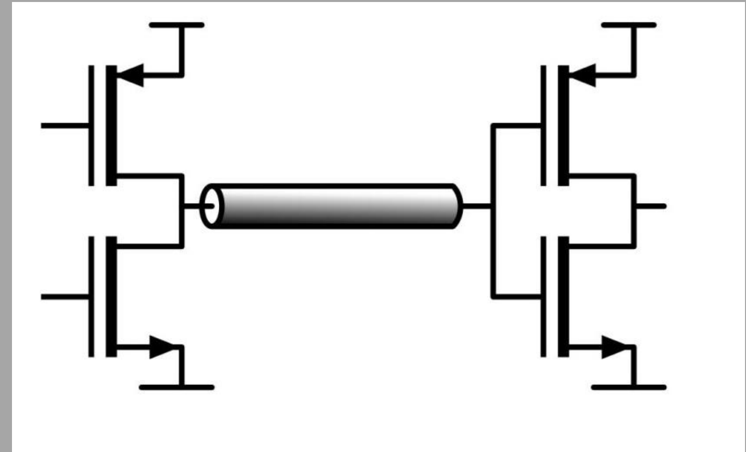
Process technology: it's getting hard.
extreme resolution, complex process, many steps
exhausted students

How can we steer the future of VLSI, of wireless ?

Beyond yet another new semiconductor ***(be it 3D or 2...)***
let's explore other options.

VLSI





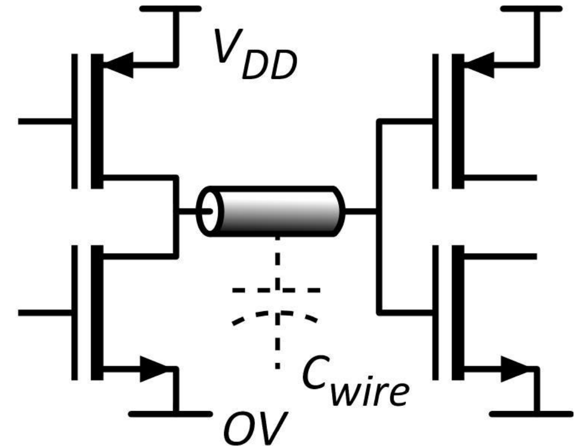
What does VLSI need ?

Small transistors: plentiful, cheap

Small transistors→ short wires

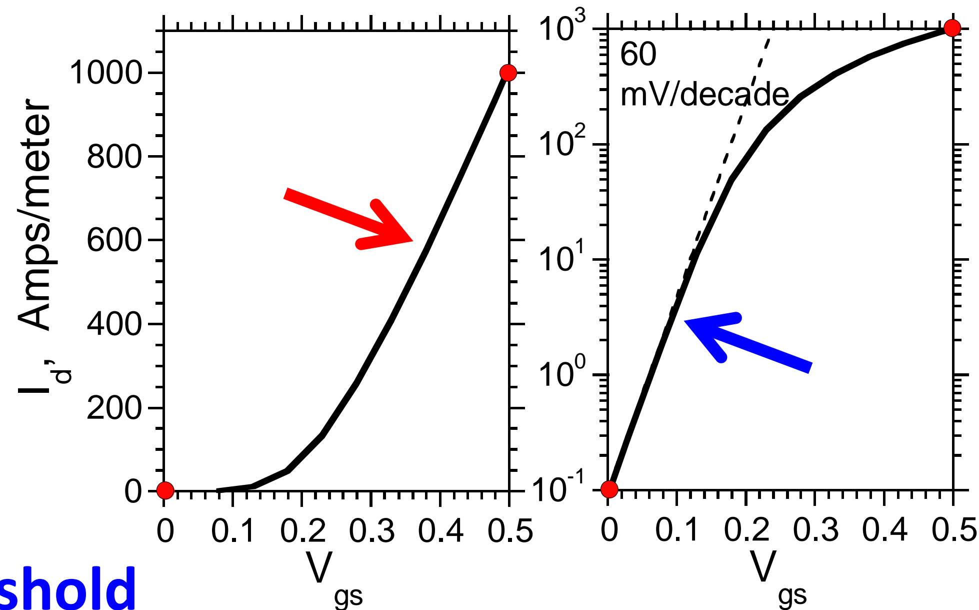
small delay CV_{DD}/I
low energy $CV_{DD}^2/2$

Small-area electronics is key



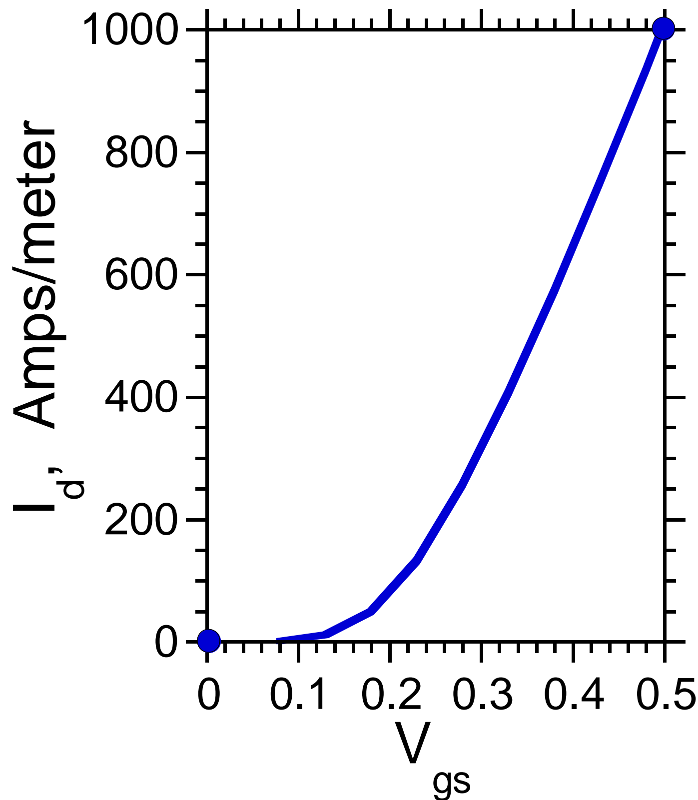
Low leakage current

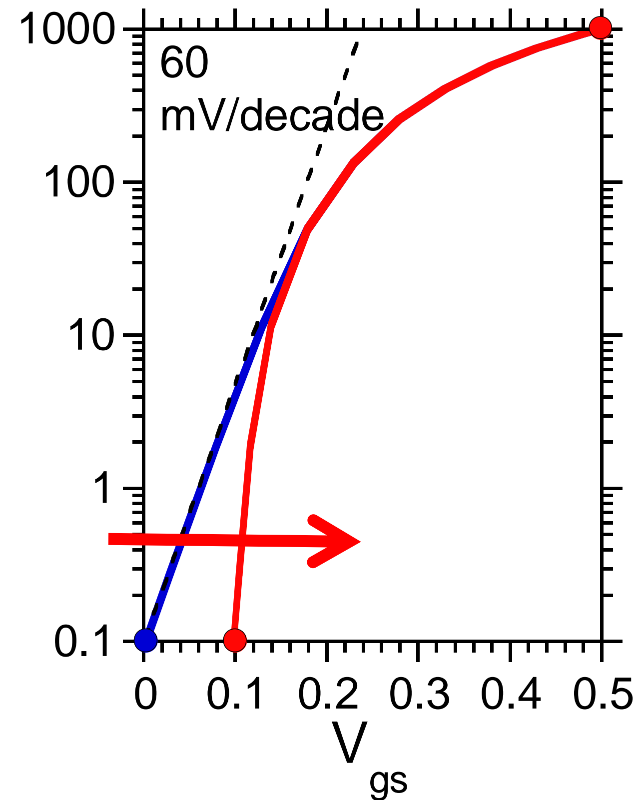
thermal: $I_{off} > I_{on}*\exp(-qV_{DD}/kT)$
want low V_{DD} yet low I_{off}.



Want:
Large *dI/dV* above threshold
Steeper than thermal below threshold

First: Steep-subthreshold-swing transistors

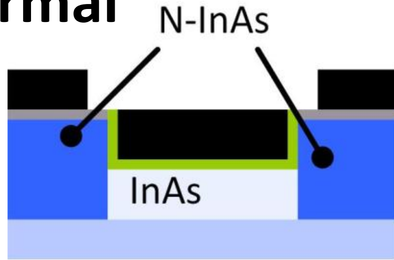


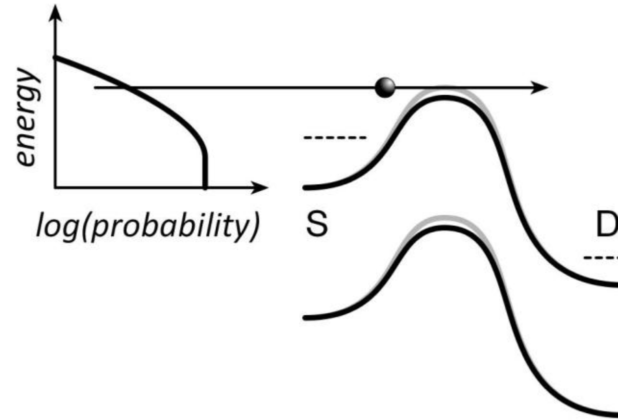


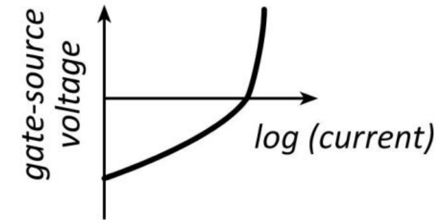
Characteristics steeper than thermal→ lower supply voltage

Tunnel FETs: truncating the thermal distribution

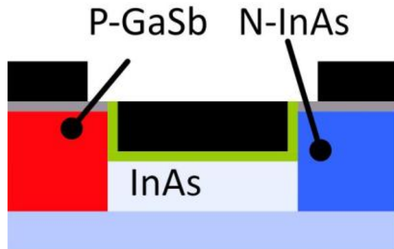
Normal

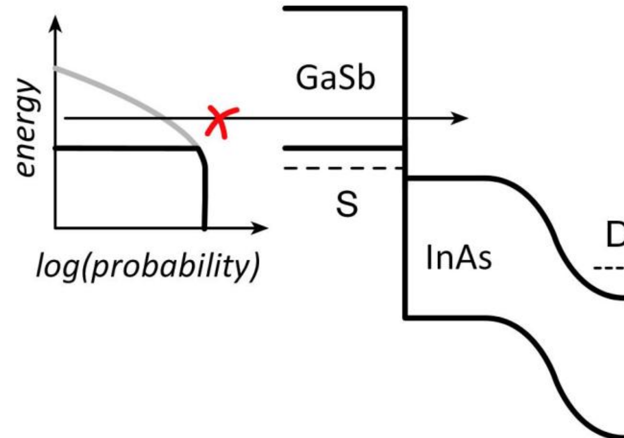


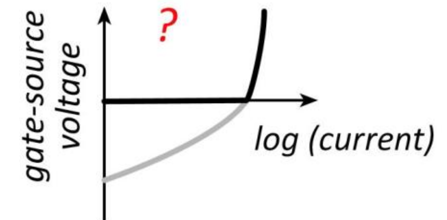




T-FET







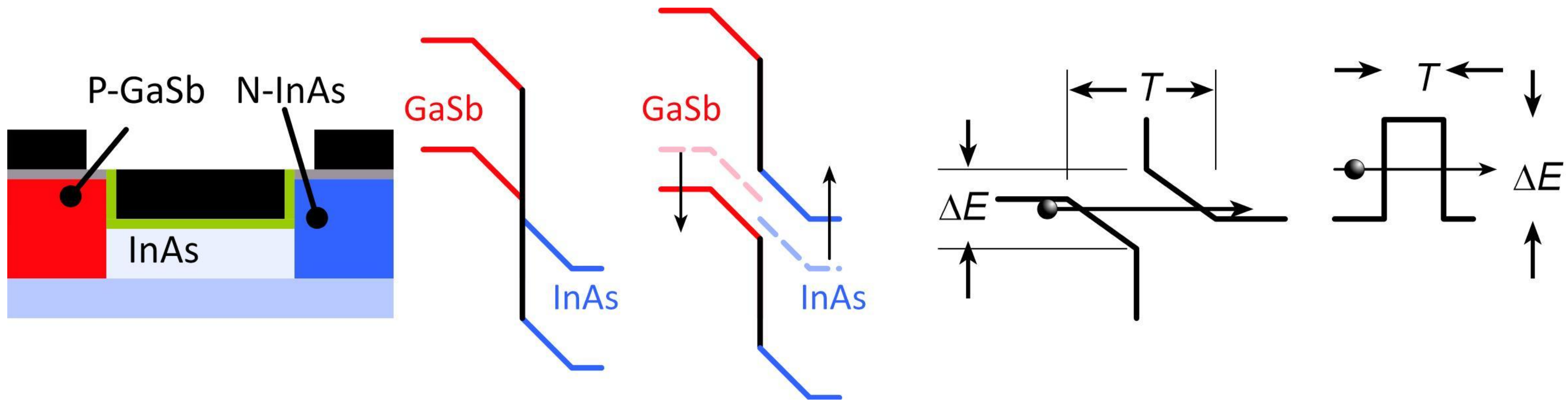
J. Appenzeller *et al.*, IEEE TED, Dec. 2005

Source bandgap truncates thermal distribution

Must cross bandgap: tunneling

Fix (?): broken-gap heterojunction

Tunnel FETs: are prospects good ?



Useful devices must be small

Quantization shifts band edges→ tunnel barrier

Band nonparabolicity increases carrier masses

Electrostatics: bands bend in source & channel

What actual on-current might we expect ?

Tunneling Probability

Transmission Probability (WKB, square barrier)

$$P \cong \exp(-2\alpha T_{\text{barrier}}),\ \text{where}\ \alpha \cong \frac{\sqrt{2m^* E_{\text{barrier}}}}{\hbar}$$

$$\text{Assume}: m^* = 0.06 \cdot m_0,\ E_b = 0.2\,\text{eV}$$

Then :

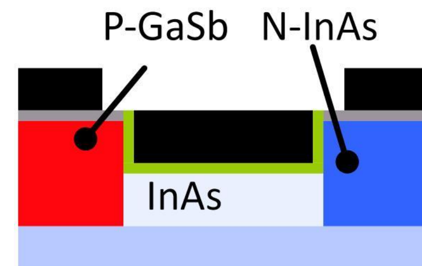
$P \cong 33\%$ for a 1nm thick barrier

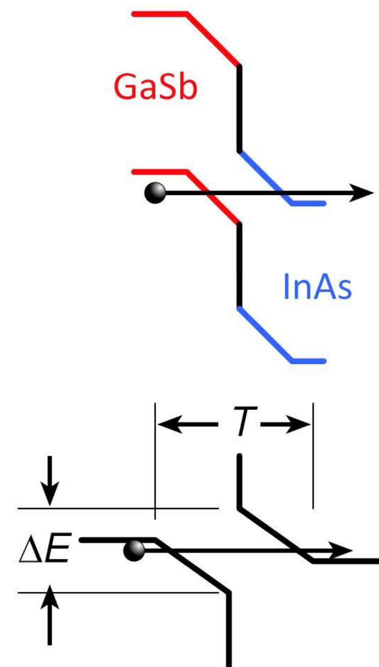
$\cong 10\%$ for a 2nm thick barrier

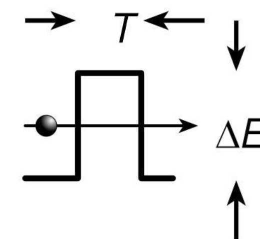
$\cong 1\%$ for a 4nm thick barrier

For high I_{on}, tunnel barrier must be *very* thin.

~3-4nm minimum barrier thickness: P+ doping, body & dielectric thicknesses



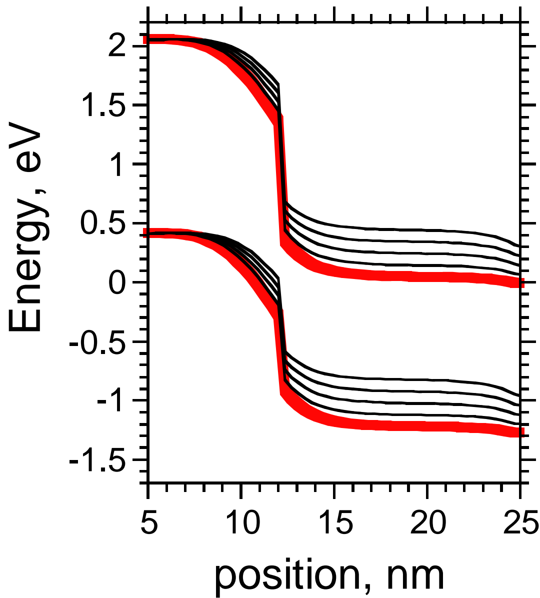


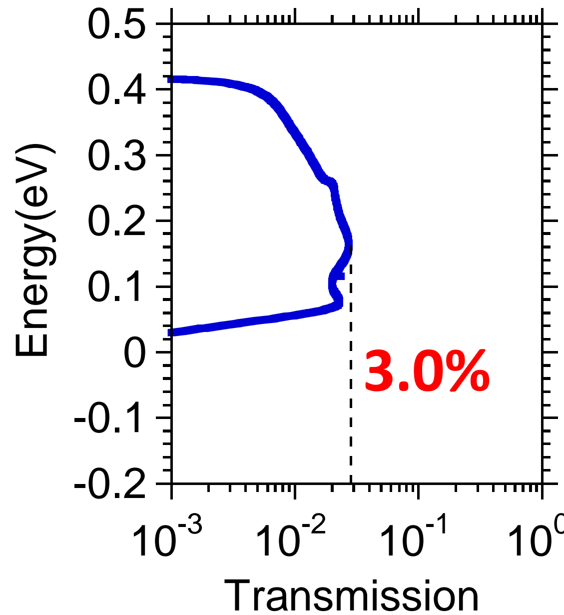


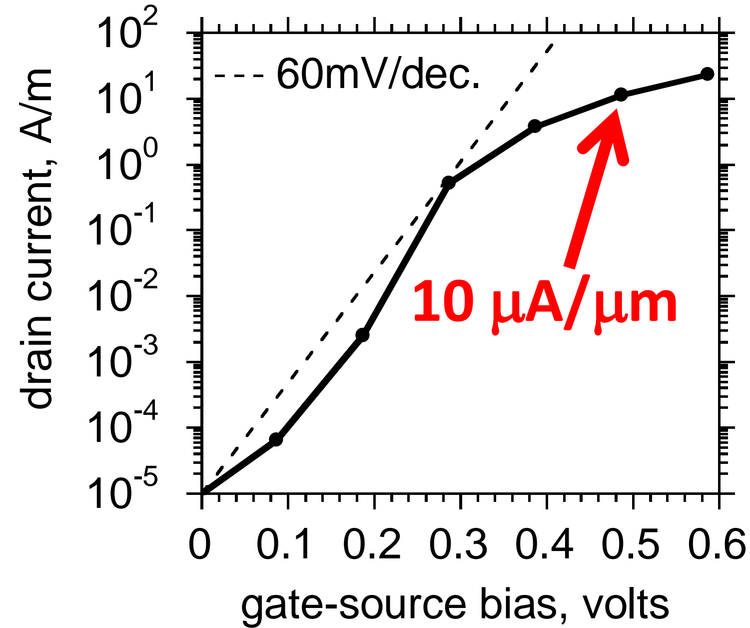
T-FET on-currents are low, T-FET logic is slow

NEMO simulation:

GaSb/InAs tunnel finFET: 2nm thick body, 1nm thick dielectric @ ε_r=12, 12nm L_g

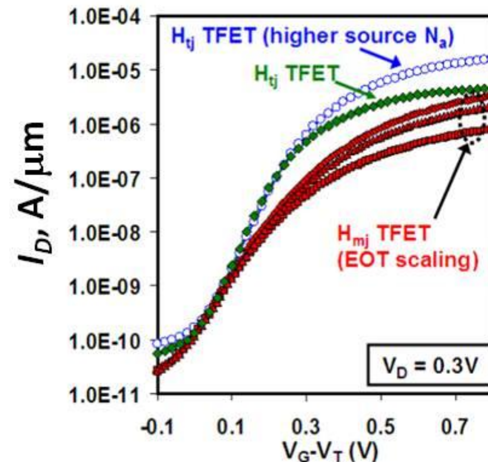






Experimental:

InGaAs heterojunction HFET,
Dewey et al,
2011 IEDM,
2012 VLSI Symp.



~15 µA/µm @0.7V

Low current → slow logic

Resonant-enhanced tunnel FET **Avci & Young, (Intel) 2013 IEDM**

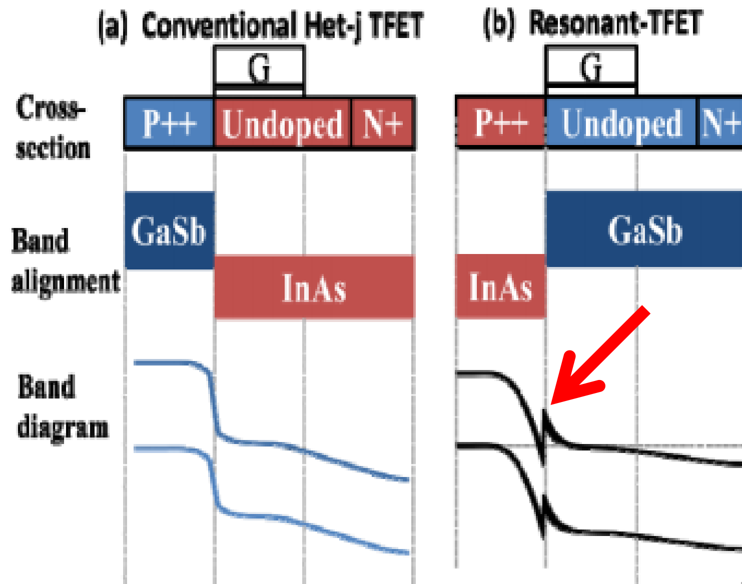


Figure 11 Device cross-section, band-alignments and band diagrams for (a) conventional Het-j TFET and (b) R-TFET. R-TFET uses the same materials but with reverse order.

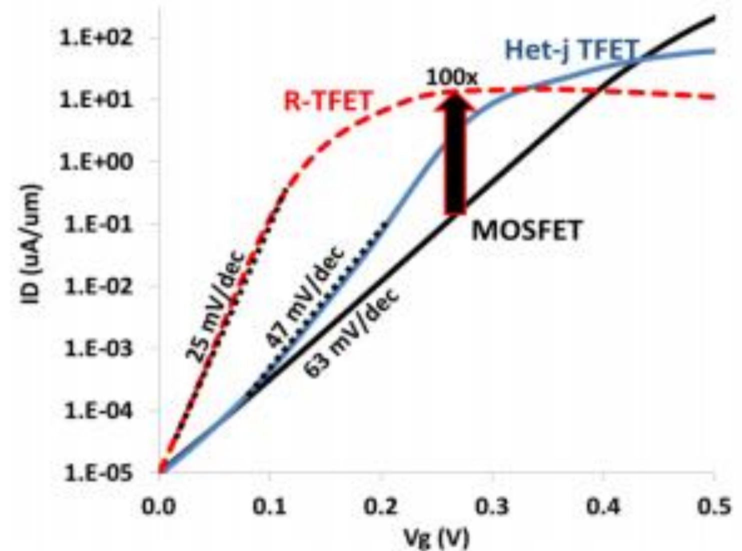


Figure 16 I-V curves for Lg=9nm NW R-TFET, Het-j TFET and MOSFET. R-TFET has 100x higher I_{dsat} than MOSFET at V_{DD}=0.27V. (I_{off}=10pA/um, V_{ds}=0.3V)

2nd barrier: bound state

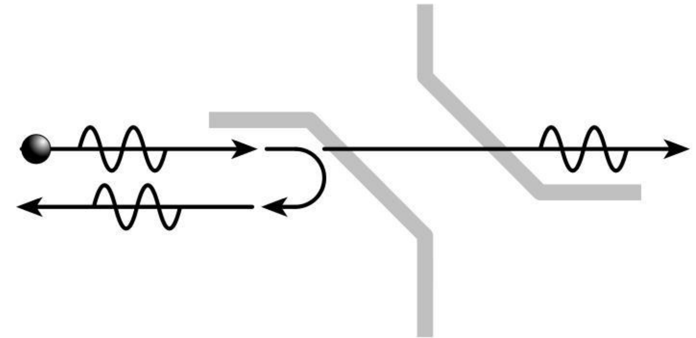
dI/dV **peaks as state aligns with source**

improved subthreshold swing.

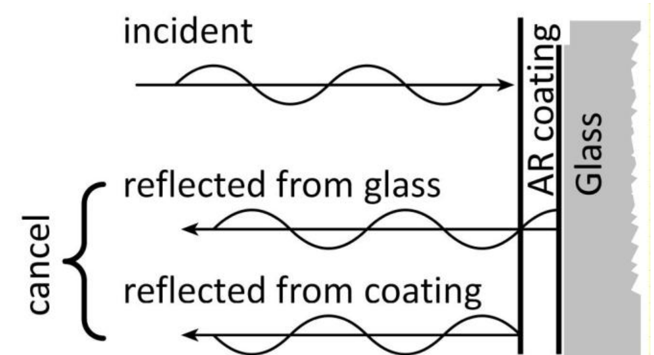
Can we also increase the on-current ?

Electron anti-reflection coatings

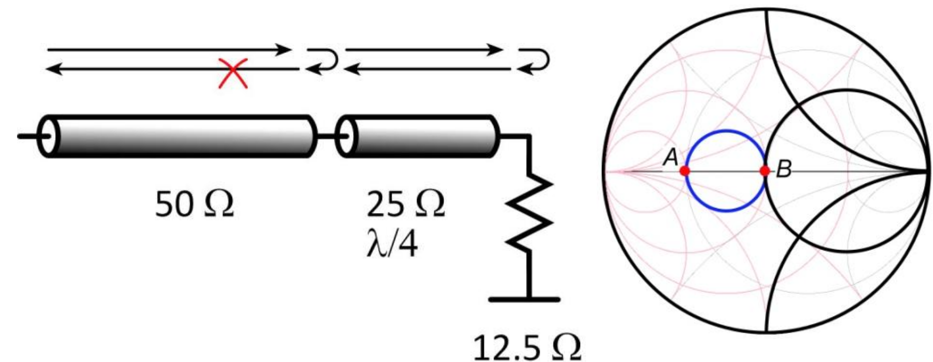
Tunnel barrier:
transmission coefficient < 100%
reflection coefficient > 0%
want: 100% transmission, zero reflection
familiar problem



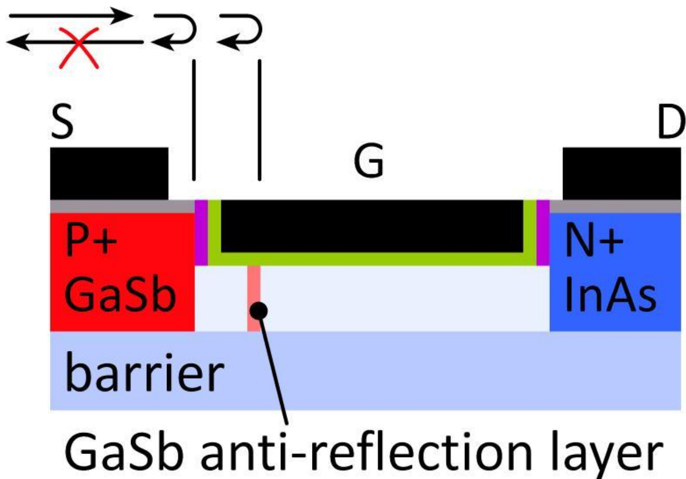
Optical coatings
reflection from lens surface
quarter-wave coating, appropriate *n*
reflections cancel

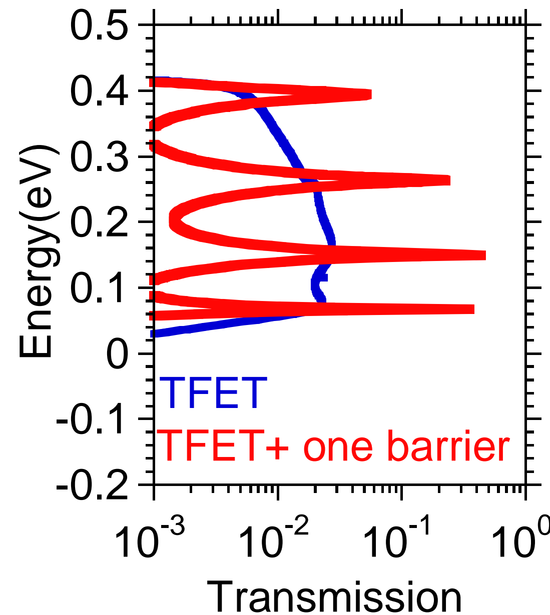


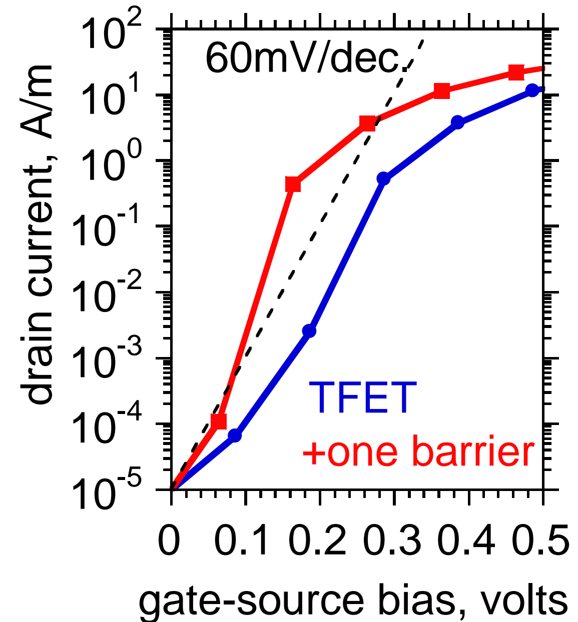
Microwave impedance-matching
reflection from load
quarter-wave impedance-match
no reflection
Smith chart.



T-FET: single-reflector AR coating







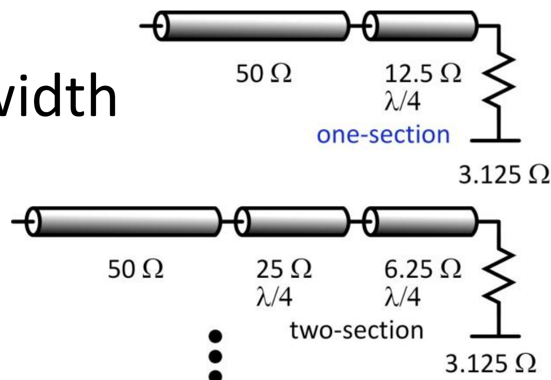
Peak transmission approaches 100%

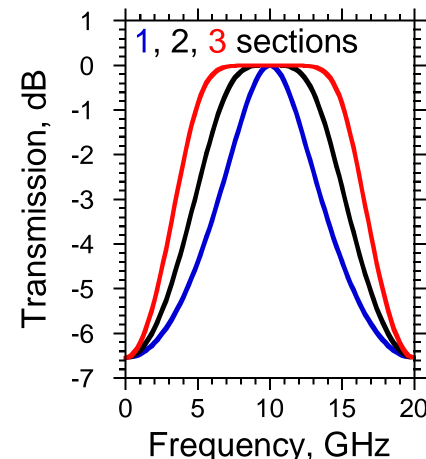
Narrow transmission peak; limits on-current

Can we do better ?

Limits to impedance-matching bandwidth

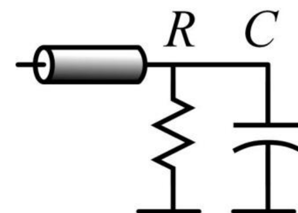
Microwave matching:
More sections→ more bandwidth
Is there a limit ?





Bode-Fano limits
R. M. Fano, J. Franklin Inst., Jan. 1960
Bound bandwidth for high transmission
example: bound for RC parallel load→

$$\int_0^\infty \ln\left(\frac{1}{\|\Gamma\|^2}\right) d\omega \le \frac{\pi}{RC}$$

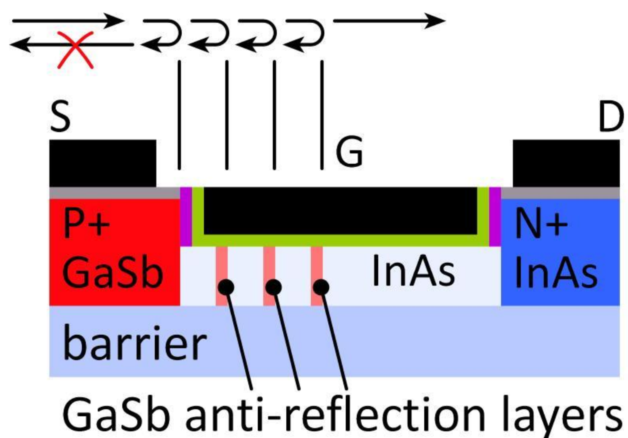


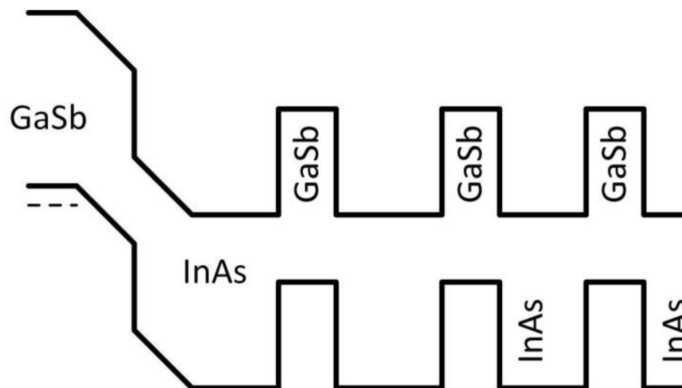
Do electron waves have similar limits ?

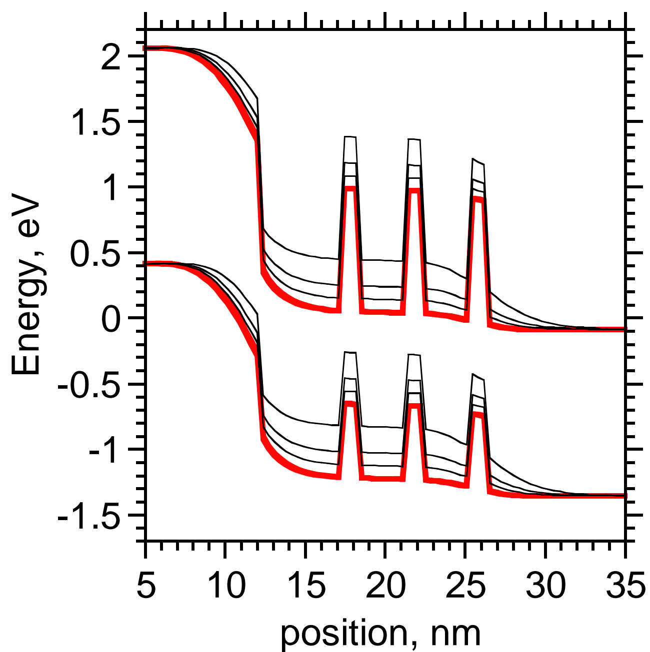
Yes ! Schrödinger's equation is isomorphic to E&M plane wave.
Khondker, Khan, Anwar, JAP, May 1988
T-FET design→ microwave impedance-matching problem
Fano: limits energy range of high transmission
Design T-FETs using Smith chart, optimize using filter theory
Working on this: for now design by random search

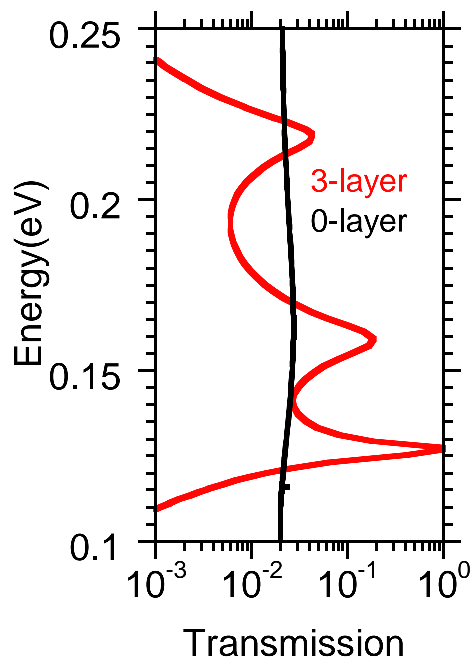
*$E/h \leftrightarrow f$, $\phi \leftrightarrow V$, $\psi \leftrightarrow I$, probability current $\leftrightarrow$ power, where $\phi(x) = (\hbar / jm^*)(\partial\psi/\partial x)$

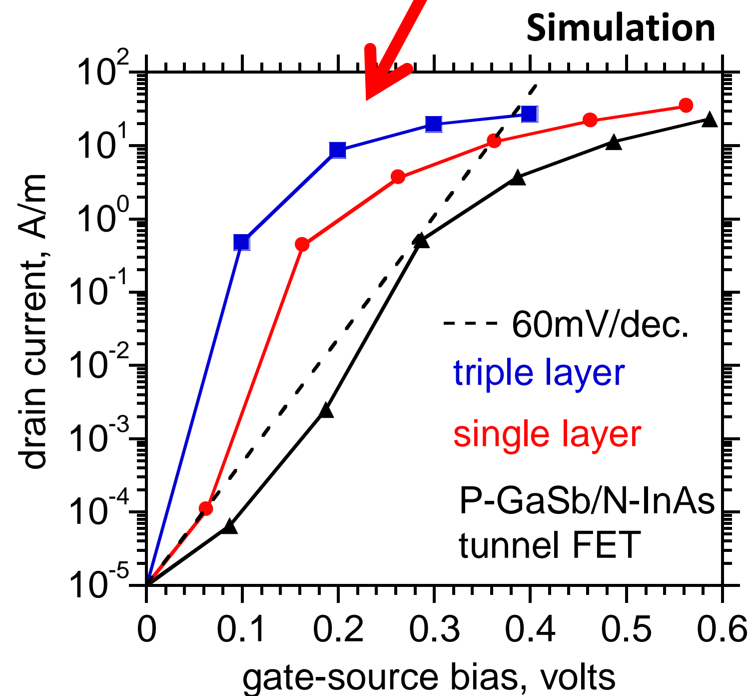
T-FET with 3-layer antireflection coating





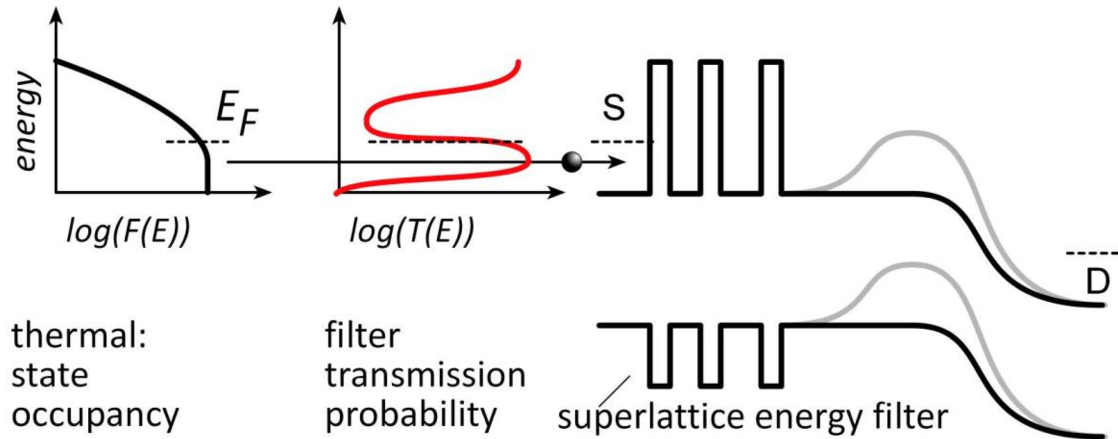






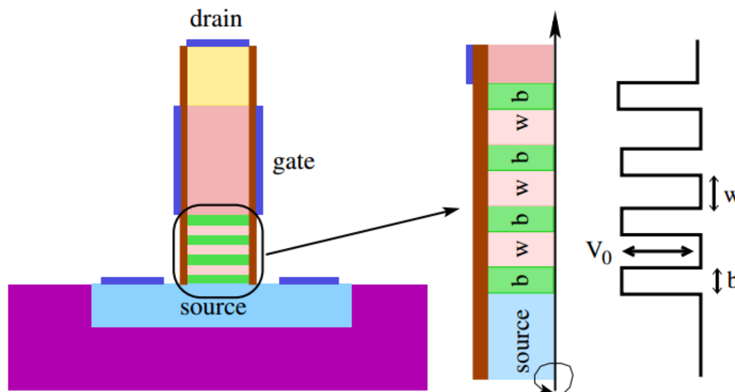
Interim result; still working on design

Source superlattice: truncates thermal distribution

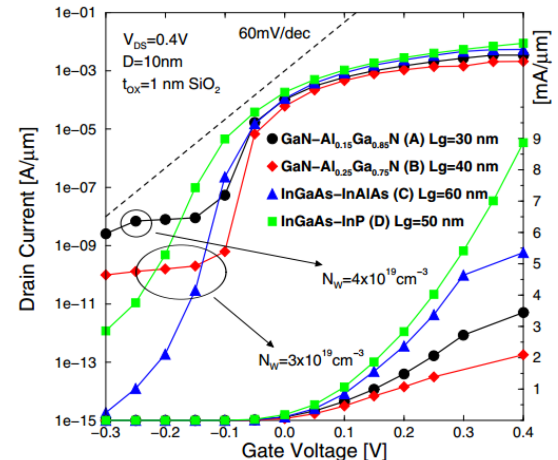


Proposed 1D/nanowire device:

M. Bjoerk *et al*., U.S. Patent 8,129,763, 2012. E. Gnani *et al*., 2010 ESSDERC



Gnani, 2010 ESSDERC



Gnani, 2010 ESSDERC:

simulation

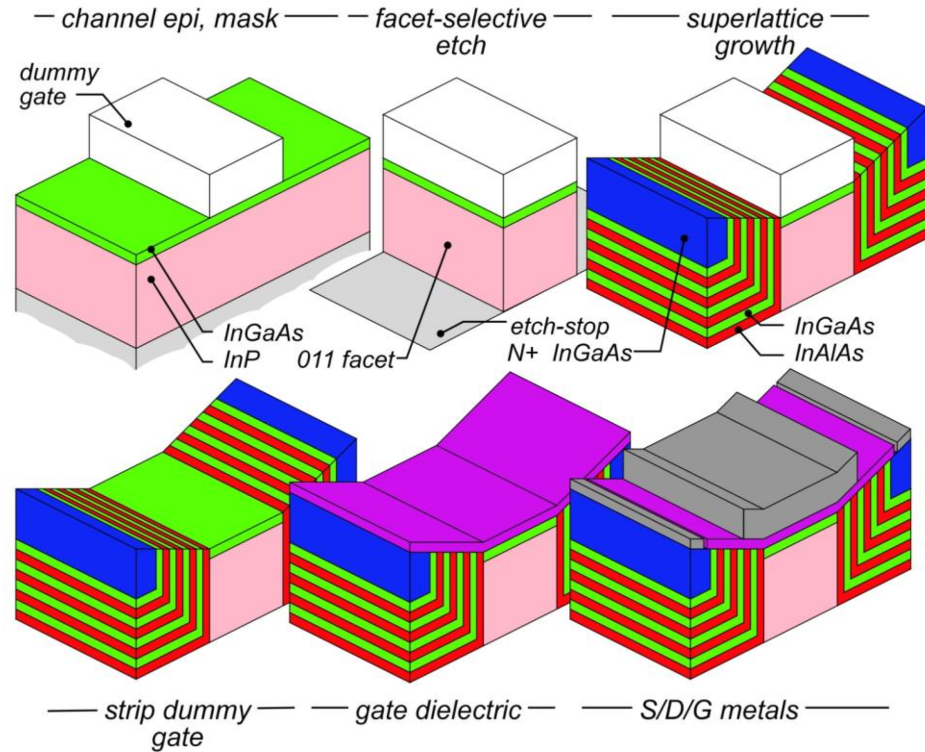
Long *et al.*, EDL, Dec. 2014

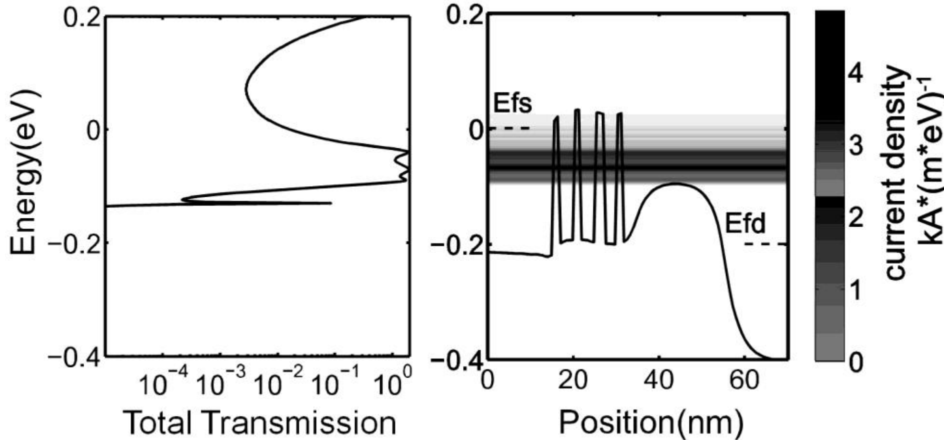
Planar (vs. nanowire) superlattice steep FET

Planar superlattice FET
superlattice by ALE regrowth
easier to build than nanowire (?)

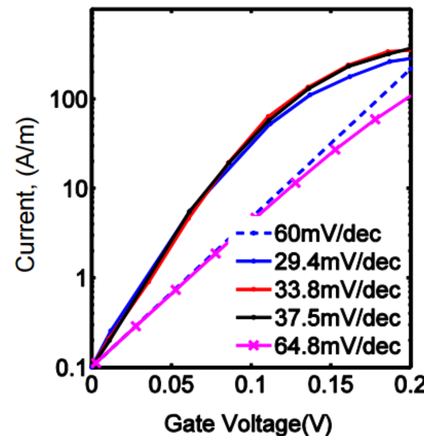
Performance (simulations):
~100% transmission in miniband.
0.4 mA/μm I_{on} , 0.1μA/μm I_{off} ,0.2V

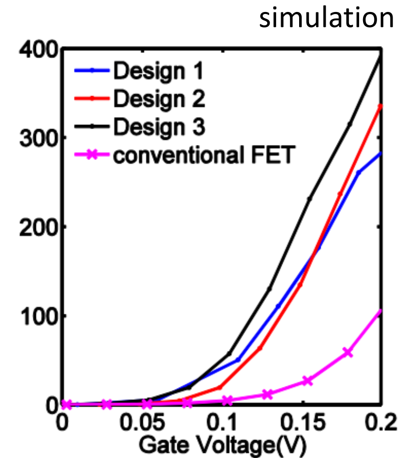
Ease of fabrication ?
Tolerances in SL growth ?
Effect of scattering ?





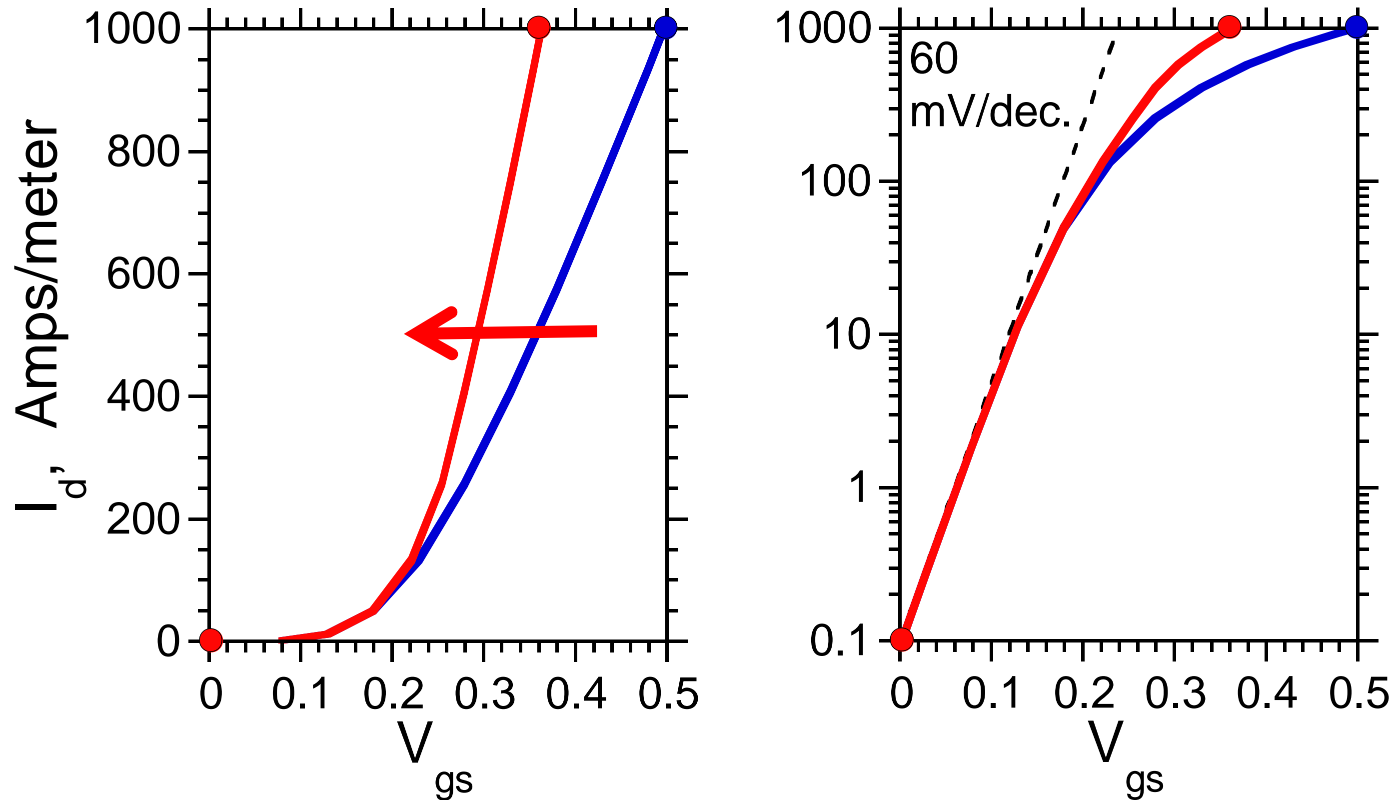
simulation





What if steep FETs prove not viable ?

Steep FETs will not be easy.



Instead, increase *dI/dV* above threshold.

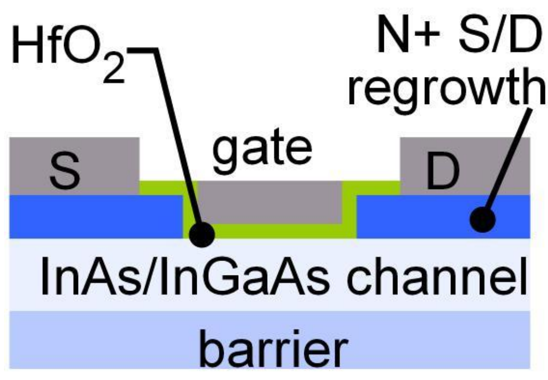
dI/dV: a.k.a. transconductance, g_m.

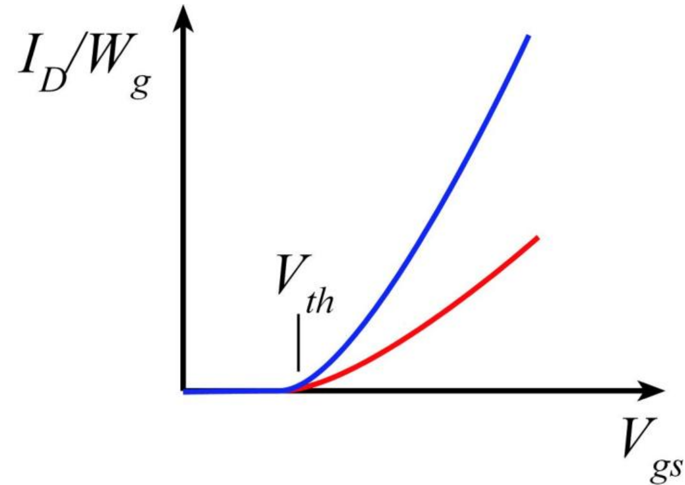
Reduced voltage, reduced CV^2

First: III-V MOS as (potential) high-(dI/dV) device

Why III-V MOS ?

III-V vs. Si: Low m*→ higher velocity. Fewer states→ less scattering → higher current. Then trade for lower voltage or smaller FETs.





Problems: Low m*→ less charge. Low m* → more S/D tunneling. Narrow bandgap→ more band-band tunneling, impact ionization.

In(Ga)As: low m*→ high velocity → high current (?)

Ballistic on-current:

Natori, Lundstrom, Antoniadis (Rodwell)

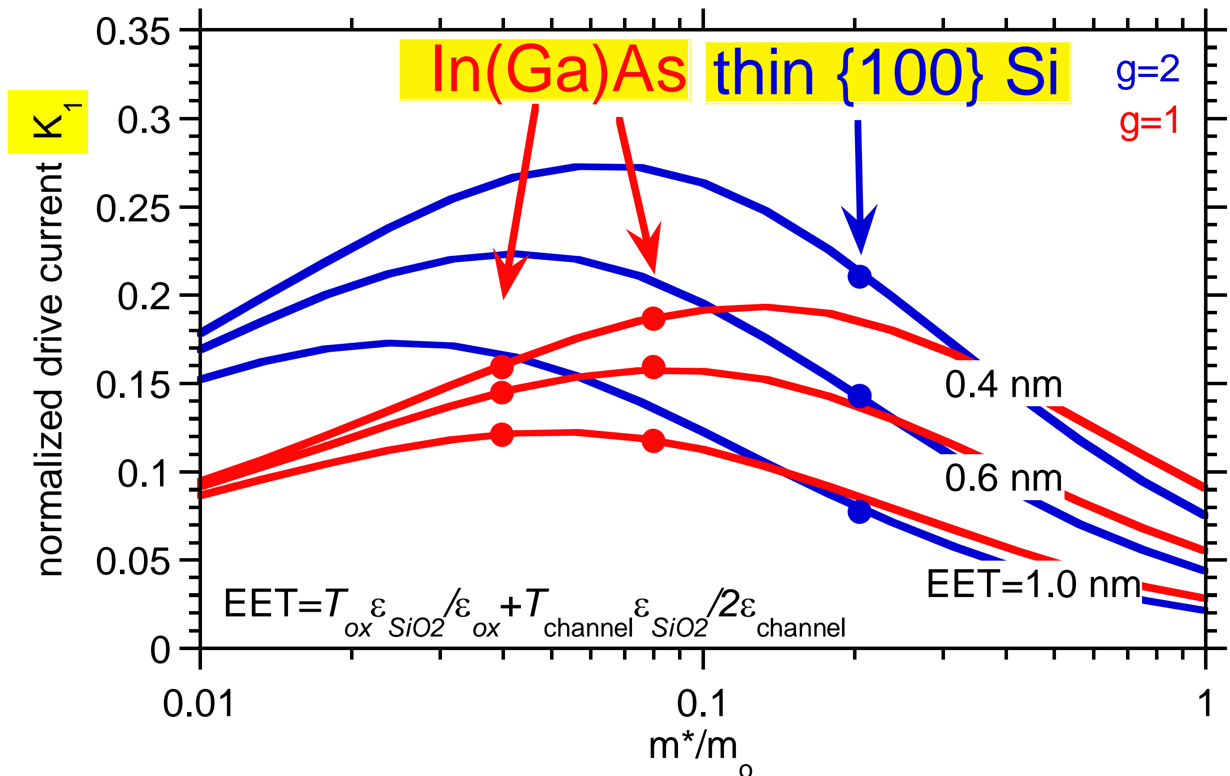
$$J = K_1 \cdot \left(84\,\frac{\text{mA}}{\mu\text{m}}\right)\cdot\left(\frac{V_{gs}-V_{th}}{1\,\text{V}}\right)^{3/2},$$

$$\frac{1}{c_{equiv}} = \frac{T_{ox}}{\varepsilon_{ox}} + \frac{T_{\text{channel}}}{2\varepsilon_{\text{semiconductor}}}$$

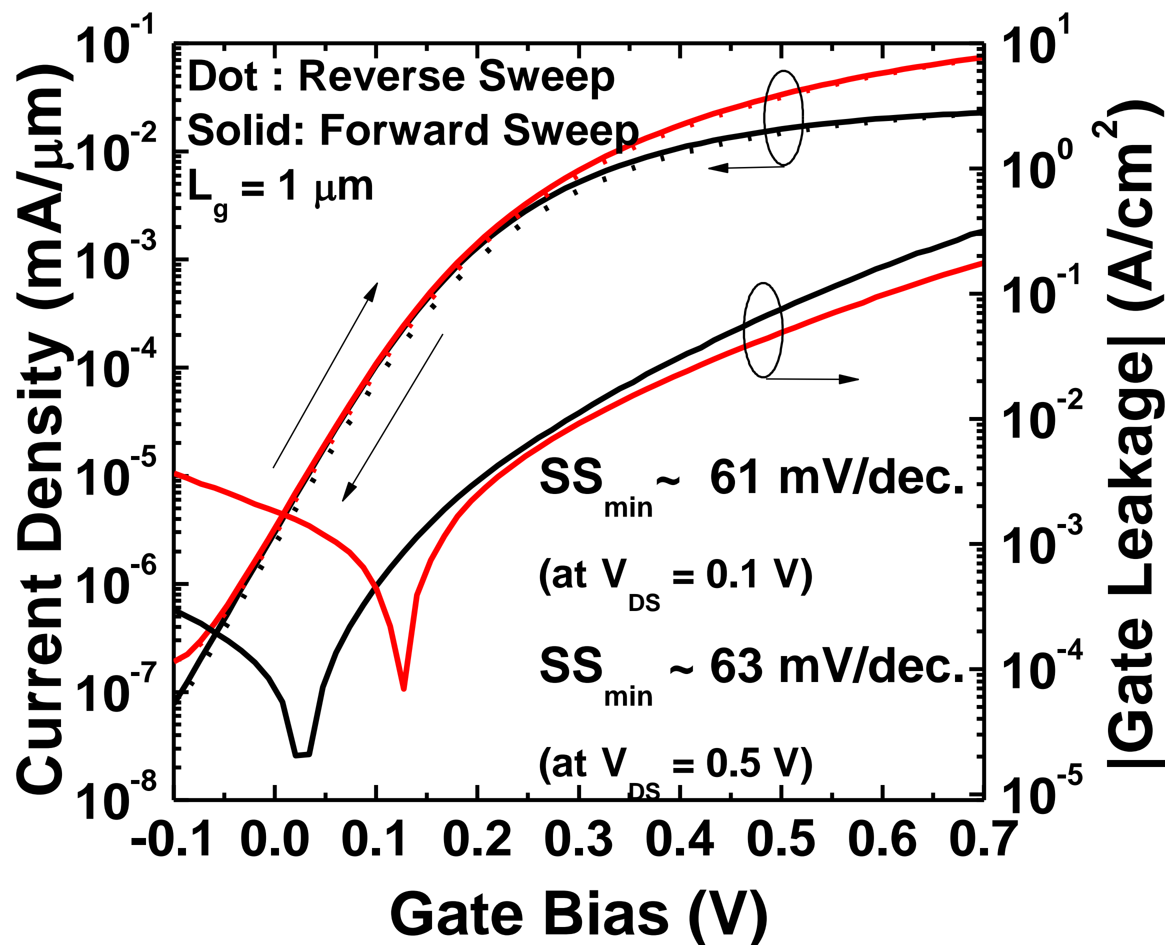
$$g = \#\,\text{valleys}$$

$$K_1 = \frac{g\cdot\left(m^*/m_o\right)^{1/2}}{\left(1+(c_{dos,o}/c_{equiv})\cdot g\cdot(m^*/m_o)\right)^{3/2}}$$

More current unless dielectric, and body, are extremely thin.



Excellent III-V gate dielectrics



2.5nm ZrO_2
1nm Al_2O_3
2.5nm InAs

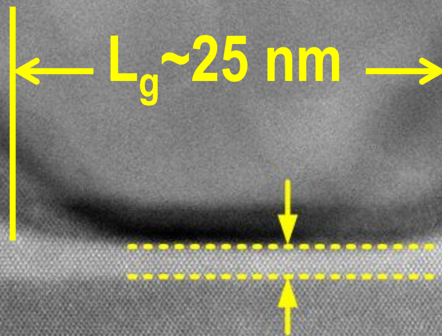
V. Chobpattanna, S. Stemmer

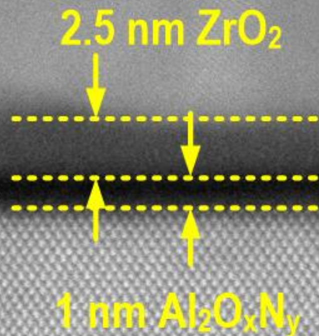
FET data: S Lee, 2014 VLSI Symp.

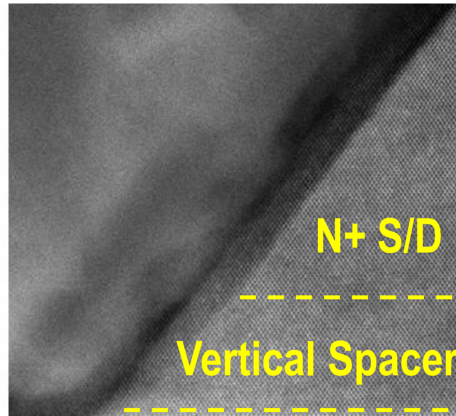
61 mV/dec Subthreshold swing at V_{DS}=0.1 V

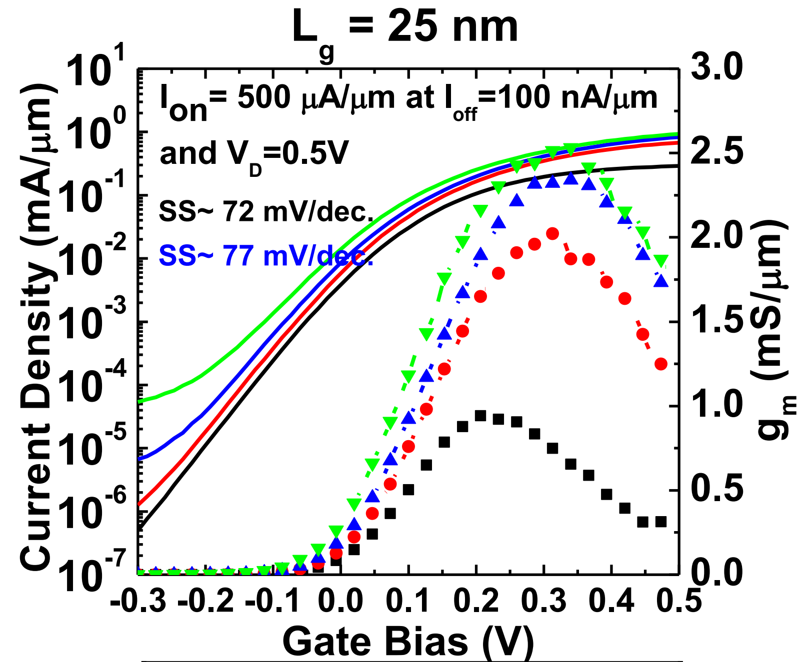
Negligible hysteresis

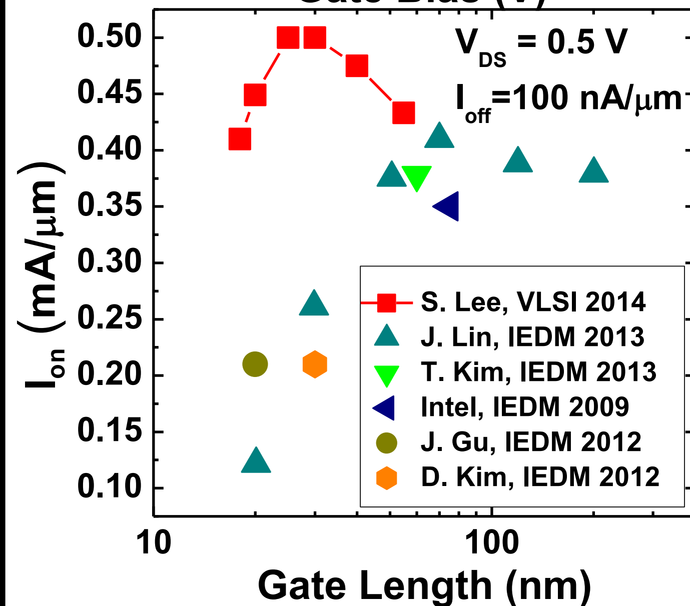
Record III-V MOS







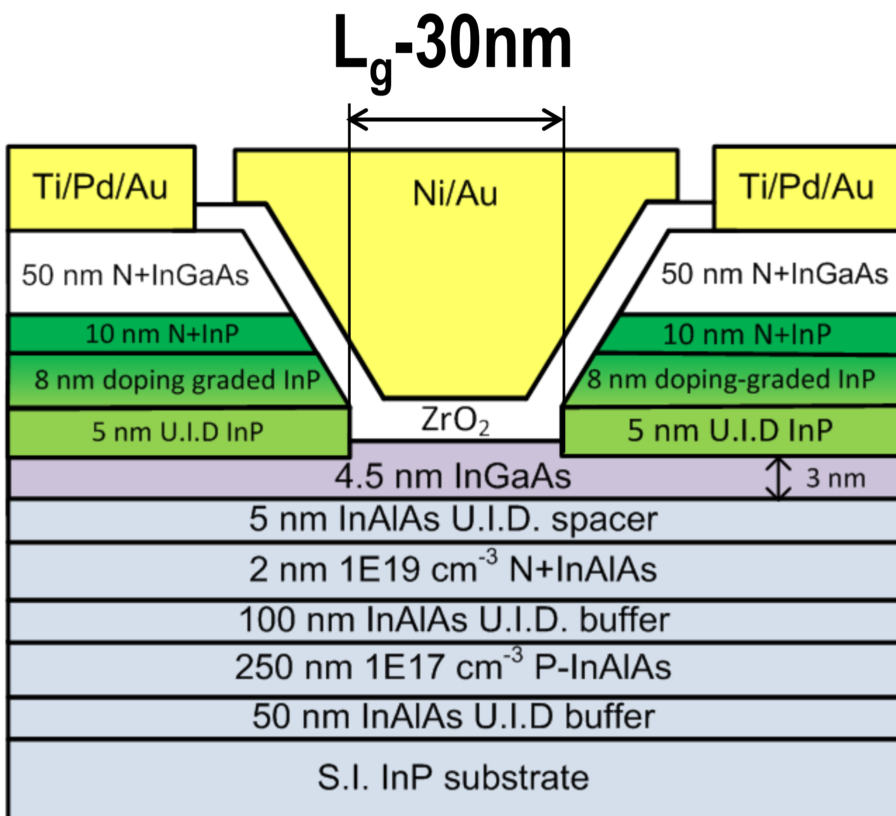




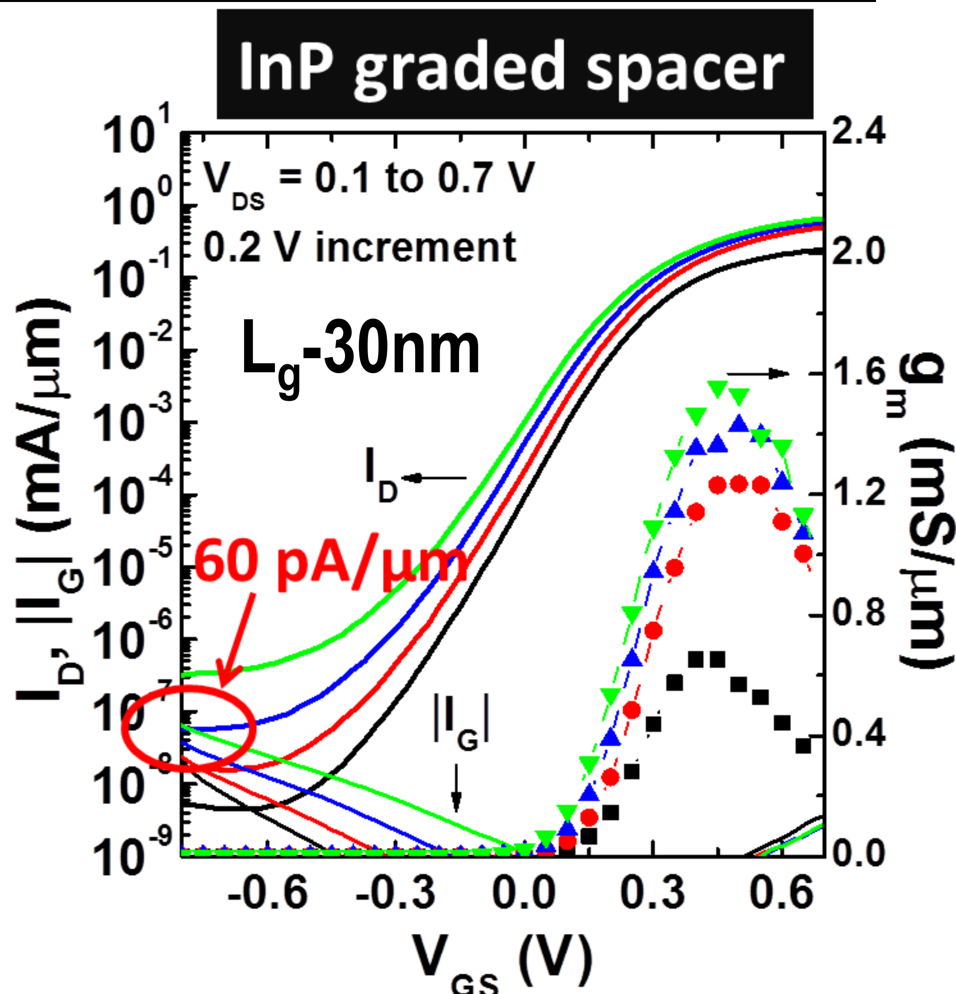
record for III-V

= best UBT SOI silicon

Double-heterojunction MOS: 60 pA/μm leakage



C. Y. Huang et al., IEDM 2014



- **Minimum I_{off}~ 60 pA/μm at V_D=0.5V for L_g-30 nm**
- **100:1 smaller I_{off} compared to InGaAs spacer**
- **BTBT leakage suppressed→isolation leakage dominates**

III-V MOS @ L_g = ???

N+ InGaAs

N+ InP

InP spacer

InAlAs Barrier

L_g~12nm

~ 8nm

t_{ch}~ 2.5 nm

5 nm

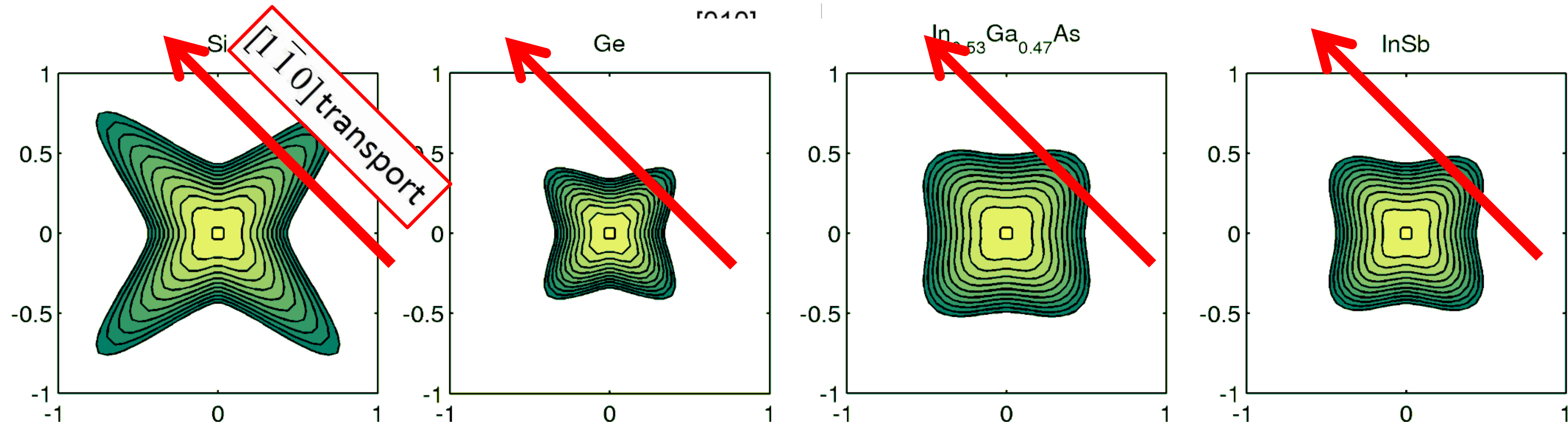
Huang *et al.*, this conference

Courtesy of S. Kraemer (UCSB)

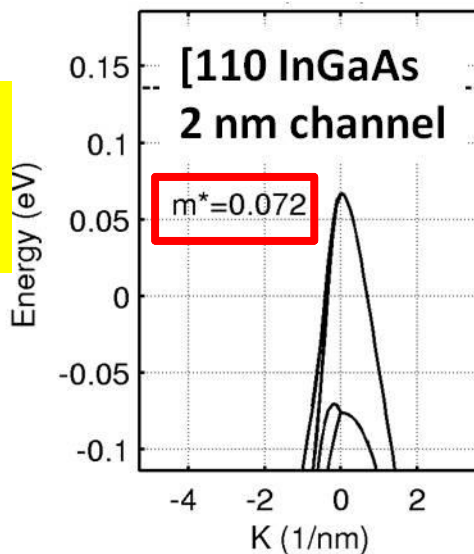
High-current III-V PMOS

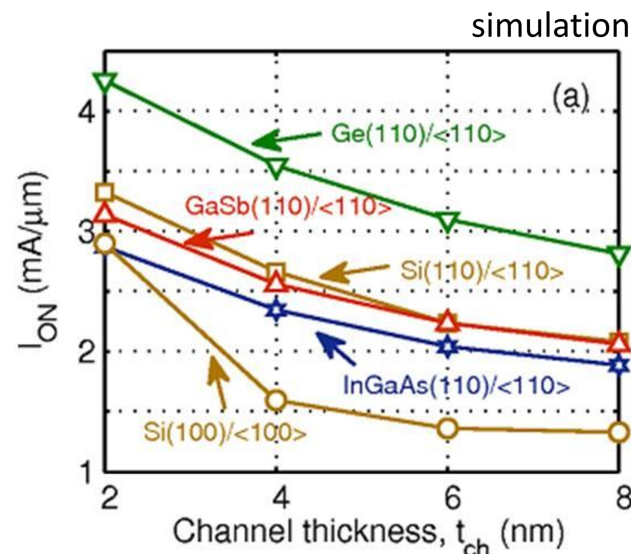
Silicon PMOS: Wang *et al.*, IEEE TED 2006 (Intel)
III-V: S. Mehrotra (Purdue), unpublished

nm thickness [110]-oriented PMOS channels→ low transport mass



Very low m*
Current approaching NMOS
finFETs are naturally [110]

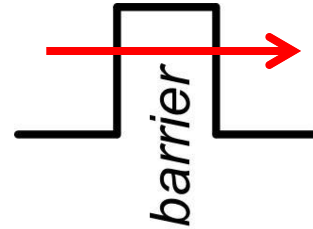




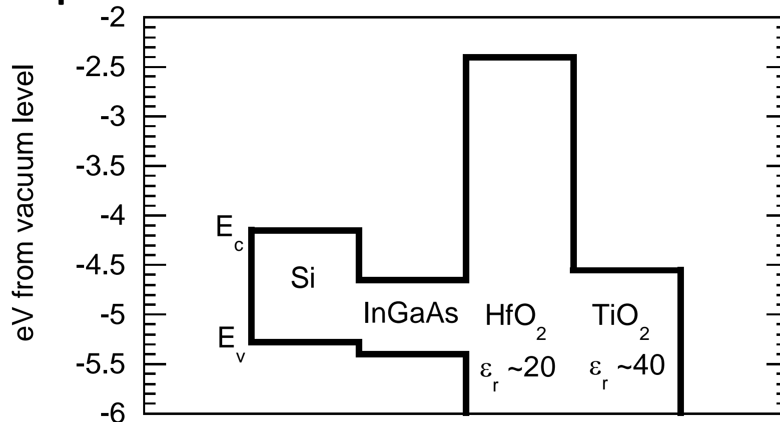
Minimum Dielectric Thickness & Gate Leakage

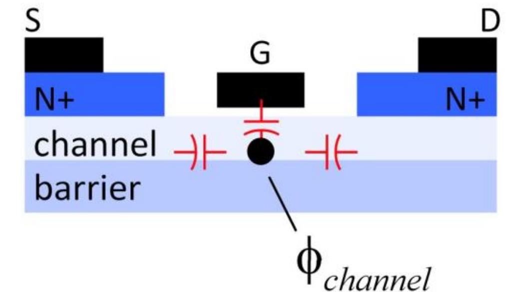
Thin dielectrics are leaky

Transmission Probability $P \cong \exp(-2\alpha T_{\text{barrier}})$, where $\alpha \cong \hbar^{-1}\sqrt{2m^* E_{\text{barrier}}}$



High-ε_r materials have lower barriers



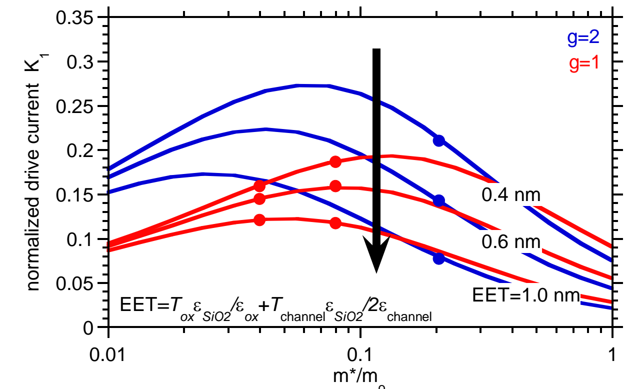


→ **0.5-0.7nm minimum EOT**

constrains on-current

electrostatics degrades with scaling

→ **fins, nanowires**

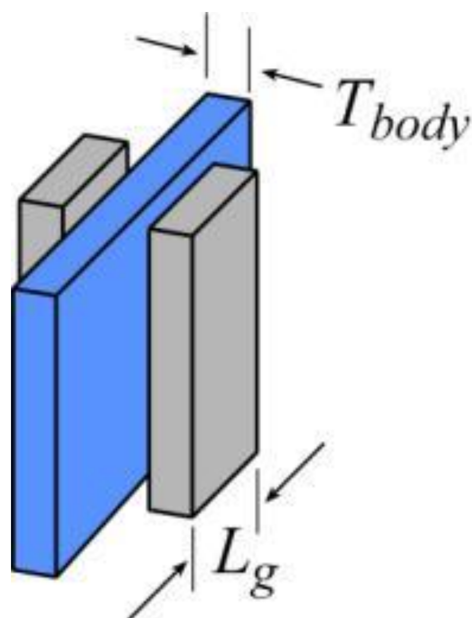


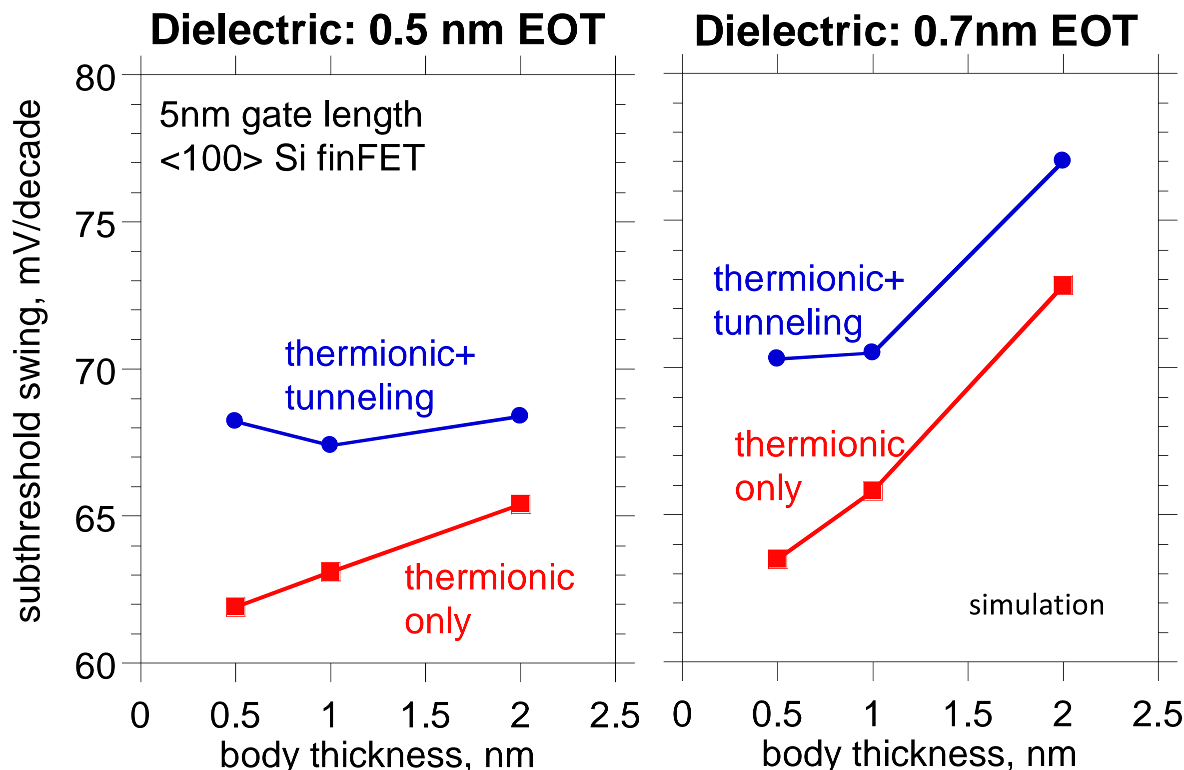
Quick check: scaling limits

NEMO ballistic simulations

finFET: 5 nm physical gate length.

Channel: <100> Si, 0.5, 1, or 2nm thick **dielectric:** ε_r=12.7, 0.5 or 0.7 nm EOT





Given EOT limits, ~1.5-2nm body is acceptable.

Source-drain tunneling often dominates leakage.

Do 2-D semiconductors help ?

3D: Is body thickness a scaling limit ?

recall the previous slide

If oxides won't scale, we must make fins

with 2D, can we make fins ?

later, will need to make nanowires...

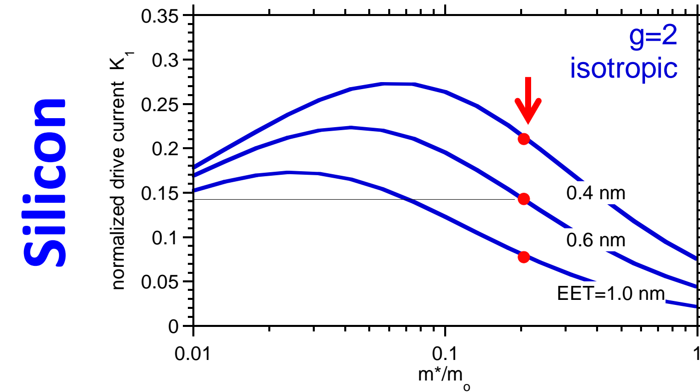
Ballistic drive currents don't win either

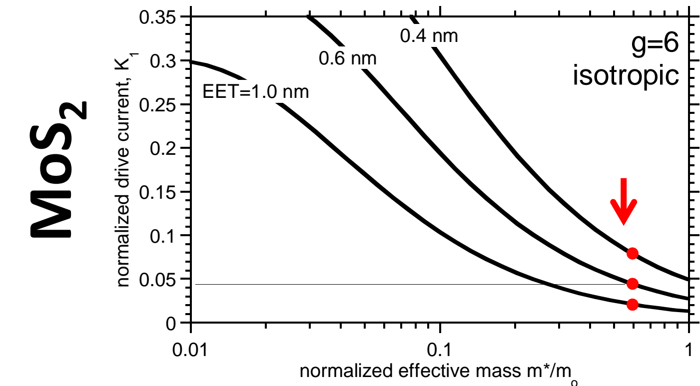
high m*, and/or high DOS

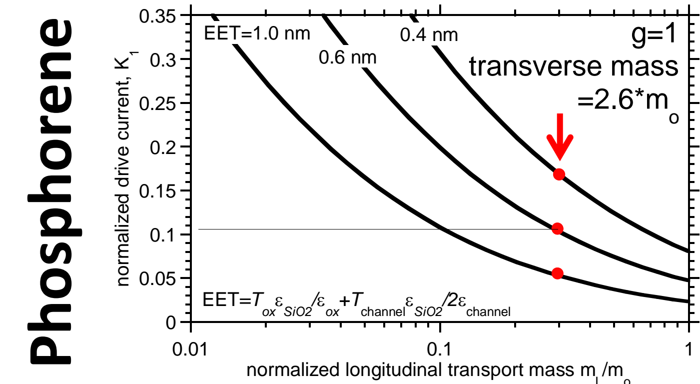
mobility sufficient for ballistic ?

$$J = K \cdot \left(84\,\frac{\text{mA}}{\mu\text{m}}\right) \cdot \left(\frac{V_{gs} - V_{th}}{1\,\text{V}}\right)^{3/2},$$

$$\text{where } K = \frac{g \cdot (m_{\perp}^{1/2} / m_o^{1/2})}{\left(1 + (c_{dos,o} / c_{equiv}) \cdot g \cdot (m_{\perp}^{1/2} m_{\parallel}^{1/2} / m_o)\right)^{3/2}}$$







When it gets crowded, build vertically

Los Angeles: sprawl

Manhattan: dense

2-D integration:
wire length $\propto$ # gates$^{1/2}$

3-D integration:
wire length $\propto$ #gates$^{1/3}$

LA is interconnect-limited

1) Chip stacking (skip)
2) 3D transistor integration

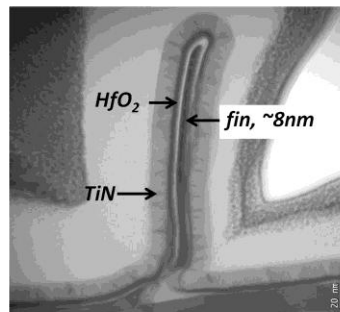
Corrugated surface→ more surface per die area

Corrugated surface→ more current per unit area

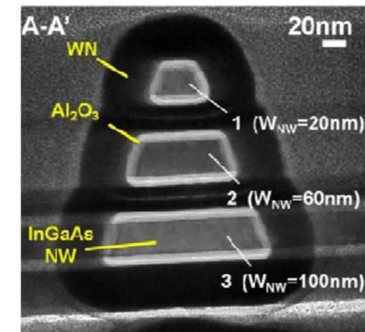
channel
N+ S/D
gate

planar   fin   tall fin   "4-D"

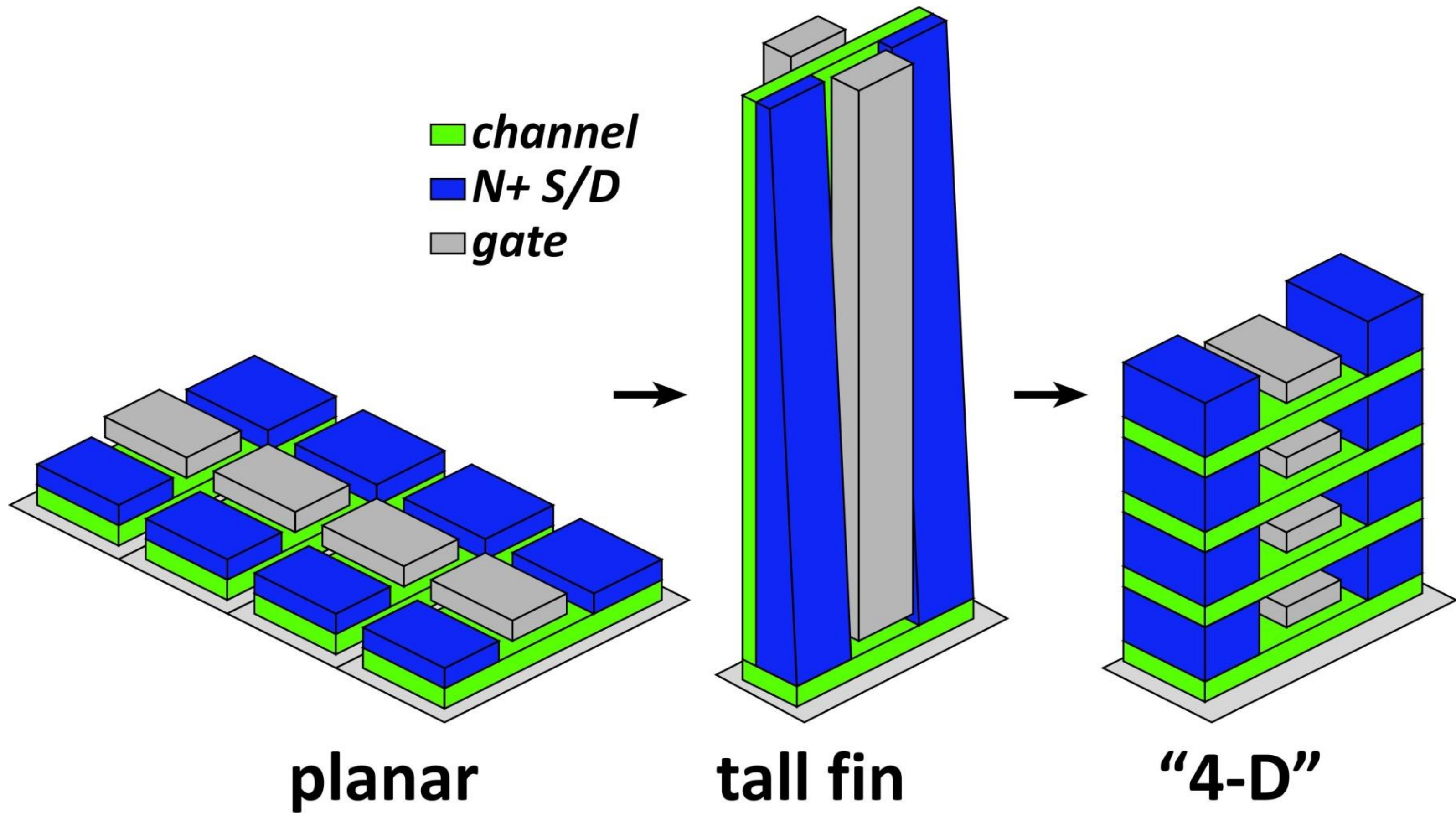
Cohen-Elias *et al.*,
UCSB
2013 DRC



J .J. Gu *et al.*, 2012 DRC,
Purdue
2012 IEDM



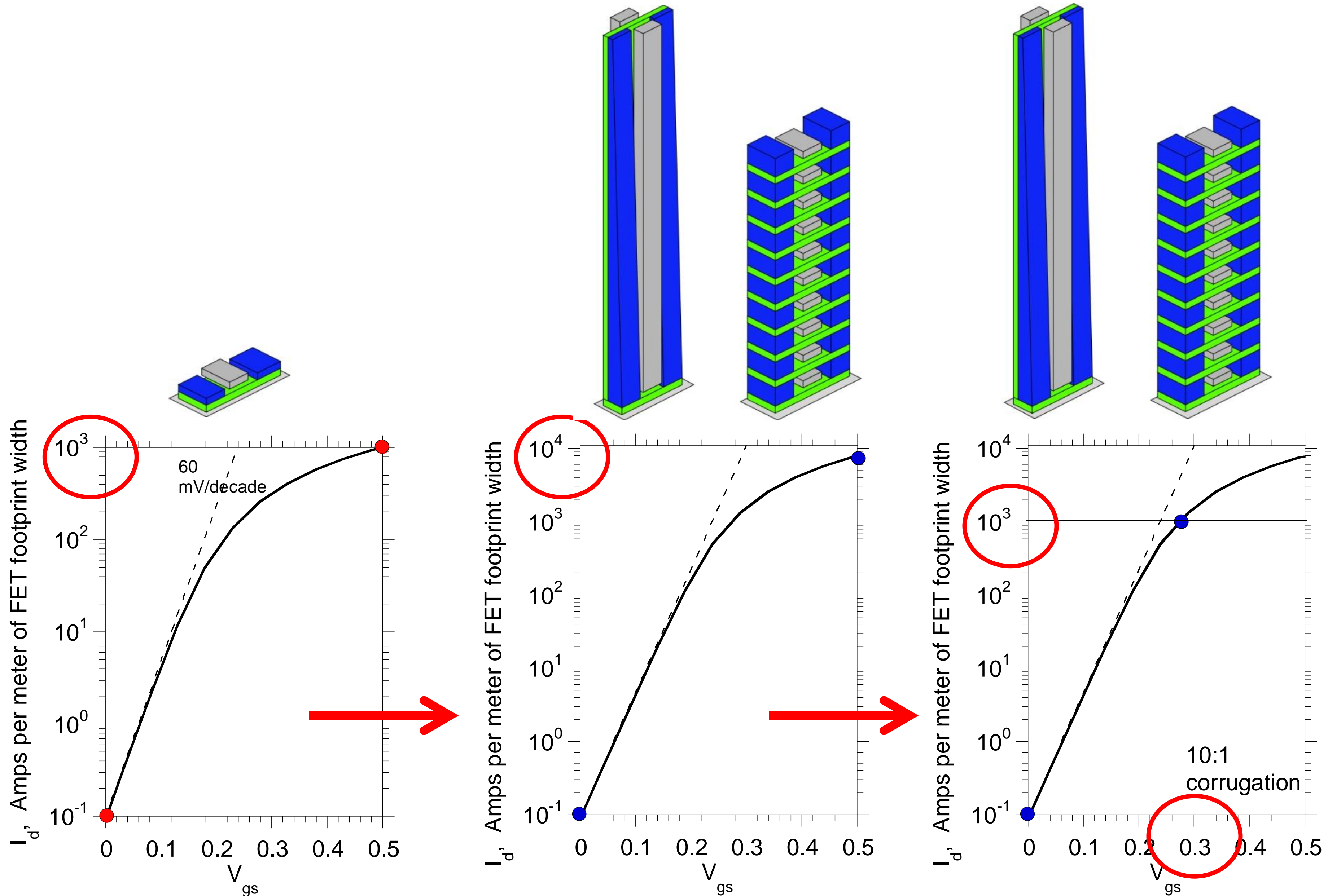
3D→shorter wires→less capacitance→less CV^2



All three have same drive current, same gate width

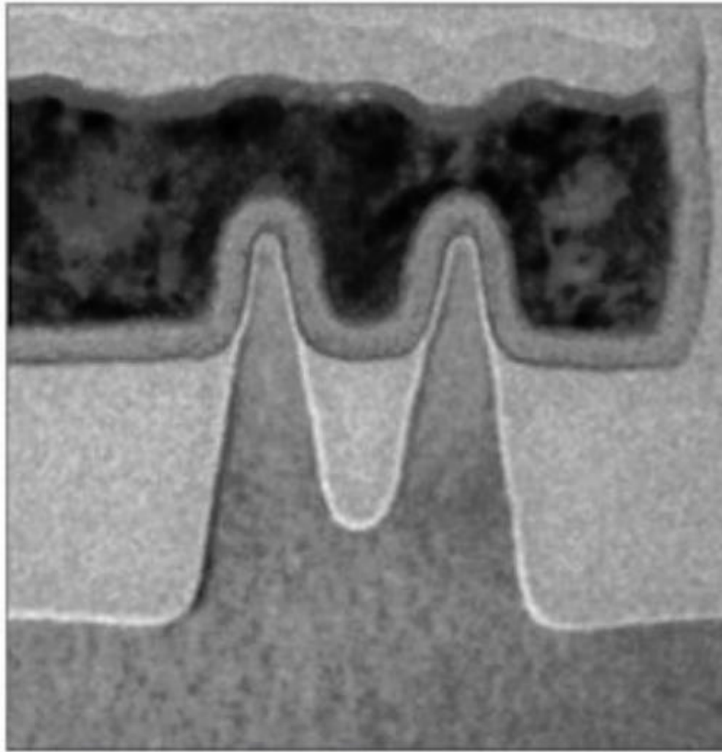
Tall fin, "4-D": smaller footprint→ shorter wires

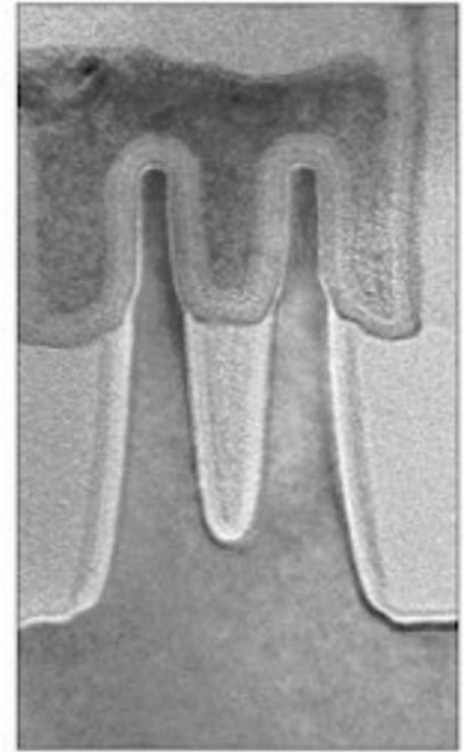
Corrugation: same current, less voltage, less CV^2



Industry is moving to taller fins.

Transistor Fin Improvement



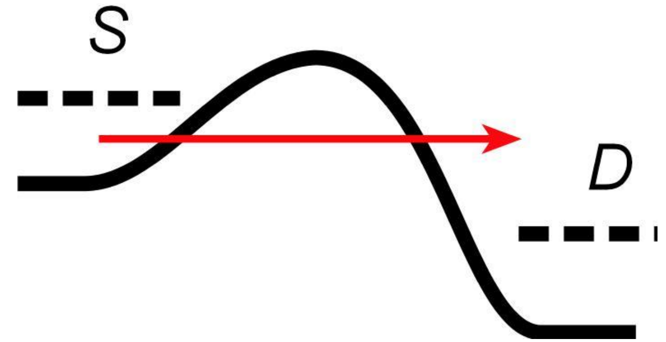


http://techreport.com/review/26896/intel-broadwell-processor-revealed

Fixing source-drain tunneling by increasing mass ?

Source-drain tunneling leakage:

$$I_{off} \cong \exp(-2\alpha L_g), \text{ where } \alpha \cong \hbar^{-1}\sqrt{2m^*(qV_{th})}$$

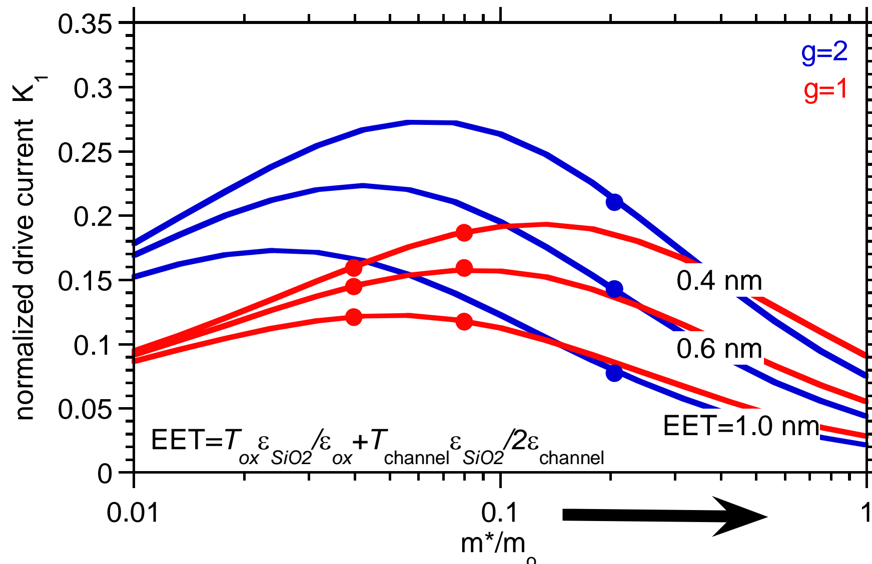


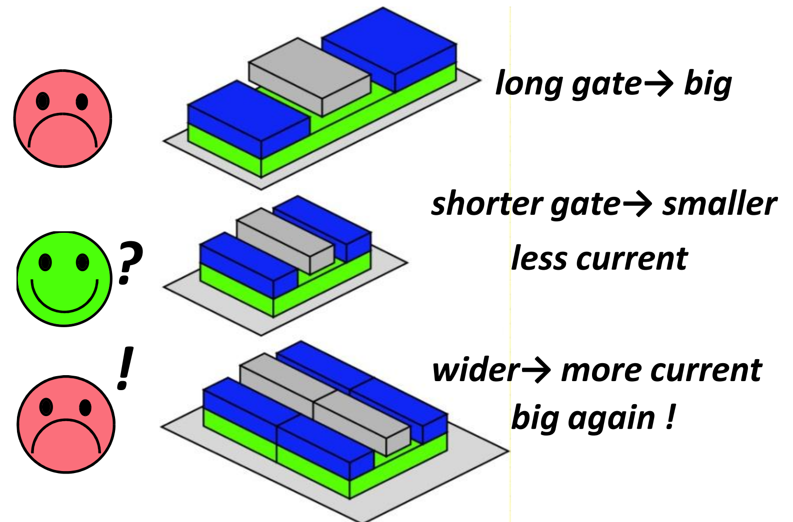
Fix by increasing effective mass ?

$$\alpha L_g = \text{constant} \rightarrow m^* \propto 1/L_g^2$$

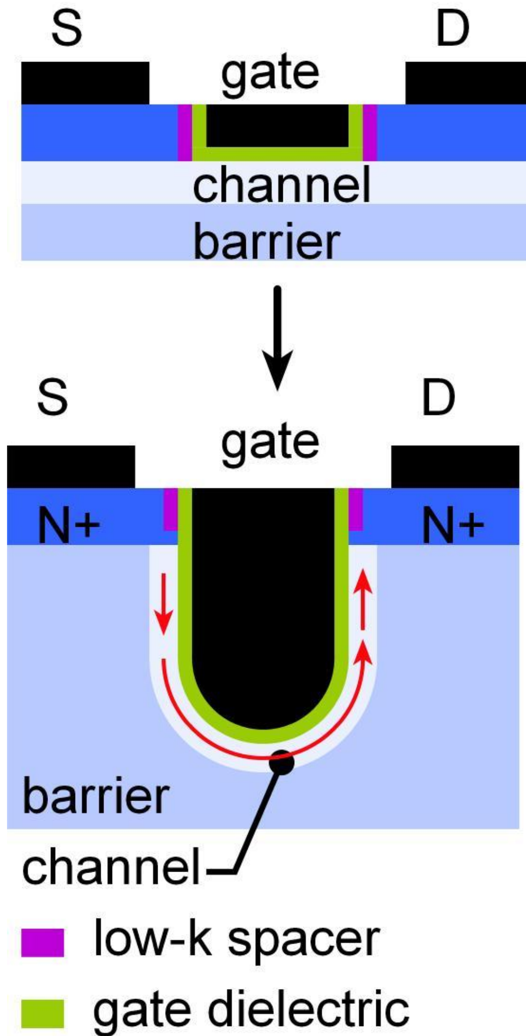
This will decrease the on-current:

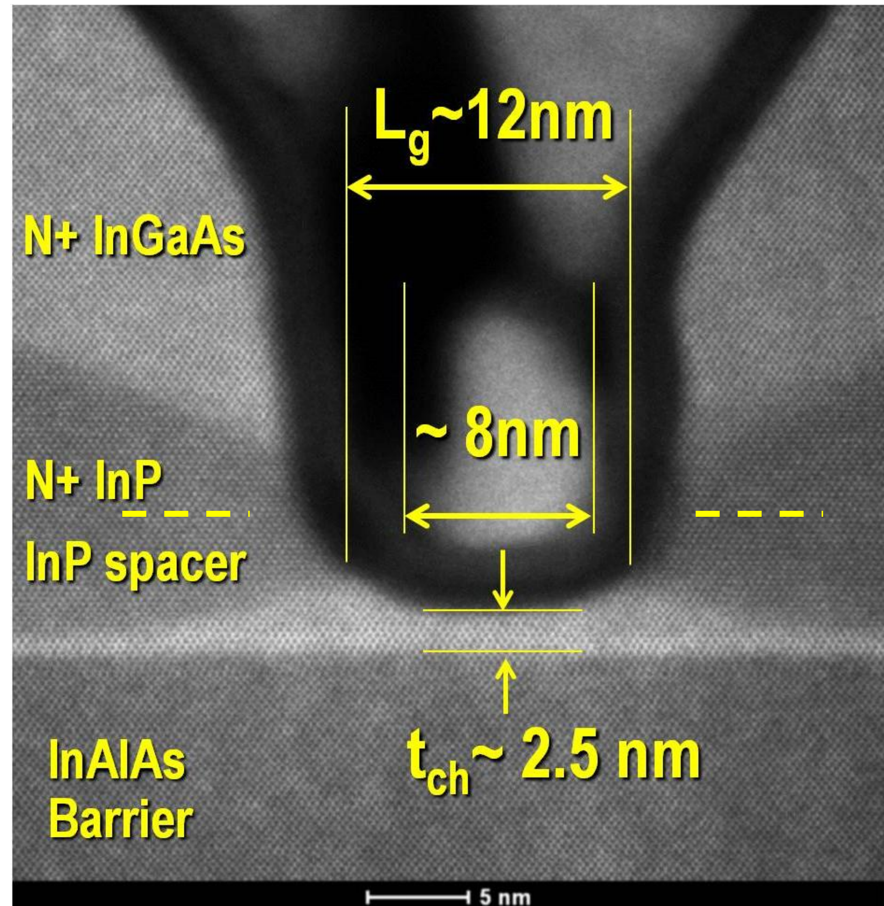
(also increases transit time)





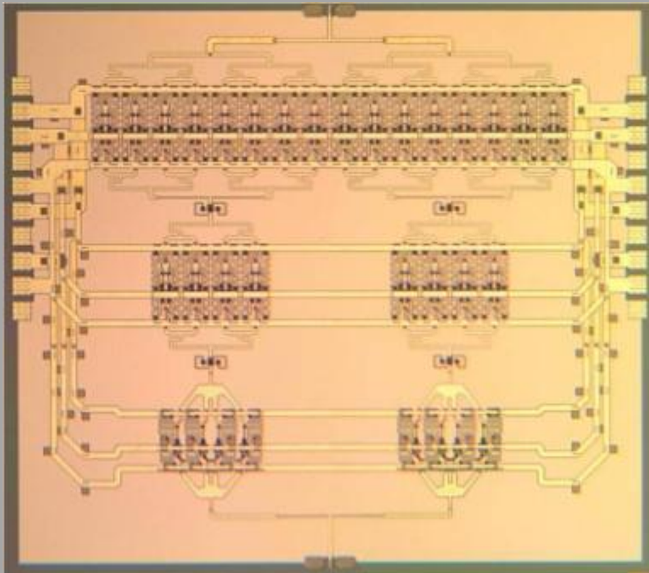
Fixing source-drain tunneling by corrugation

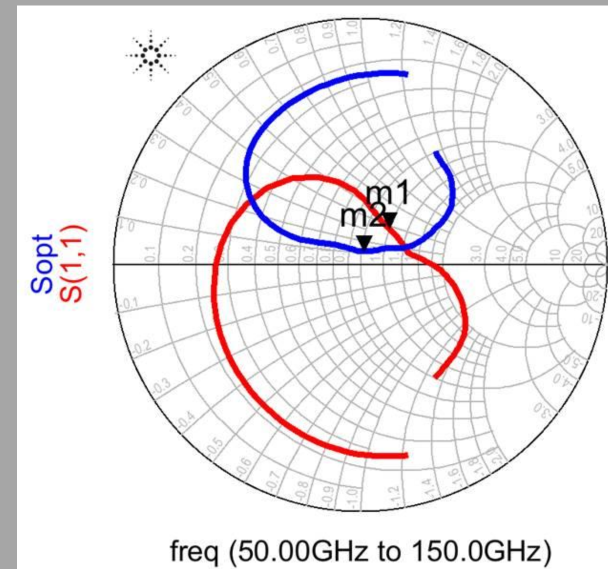




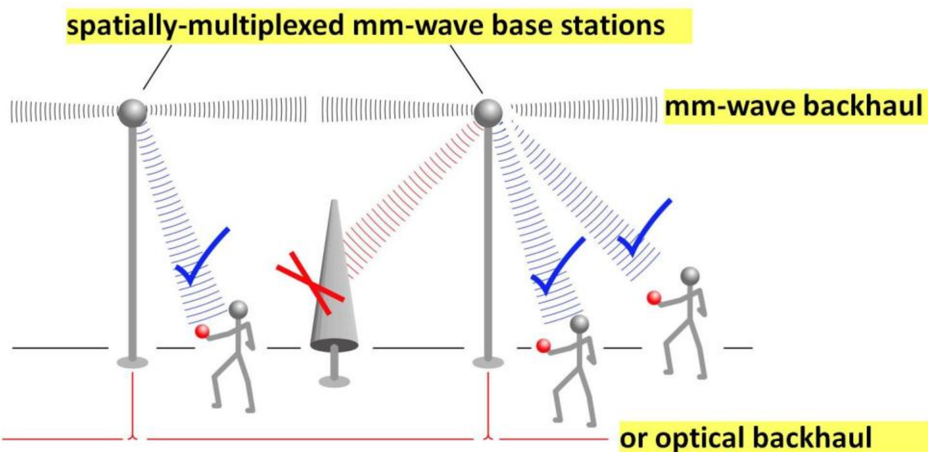
Transport distance > gate footprint length
Only small capacitance increase

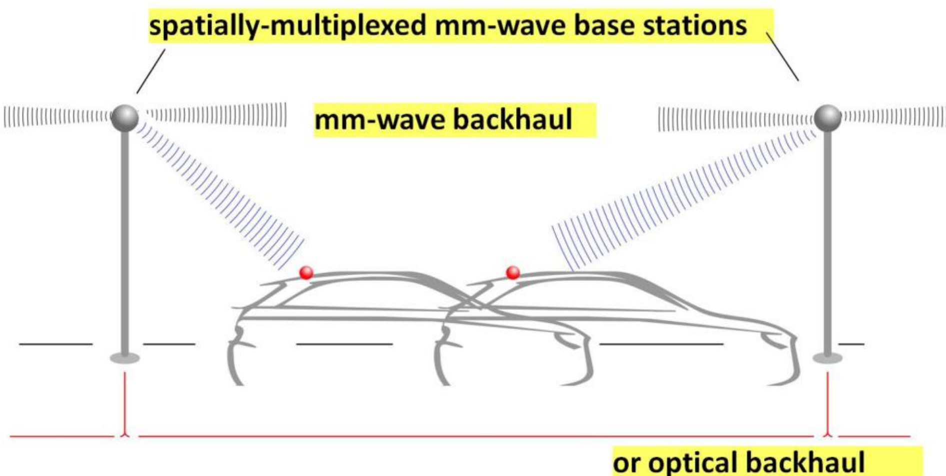
RF/Wireless



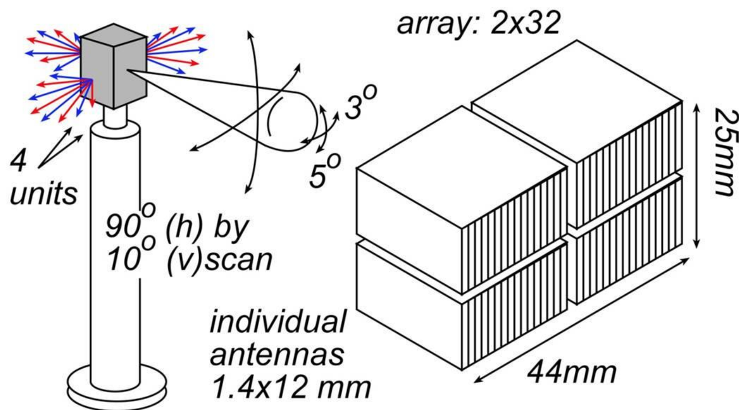


mm-Waves: high-capacity mobile communications

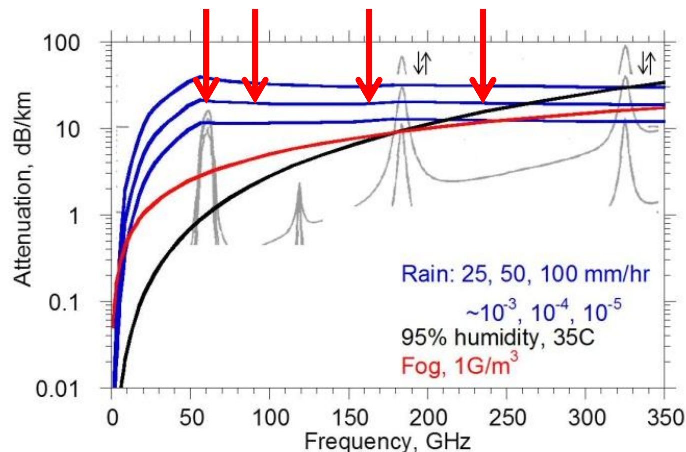




140 GHz, 10 Gb/s Adaptive Picocell Backhaul



wide, useful bandwidths from 60 to ~300 GHz



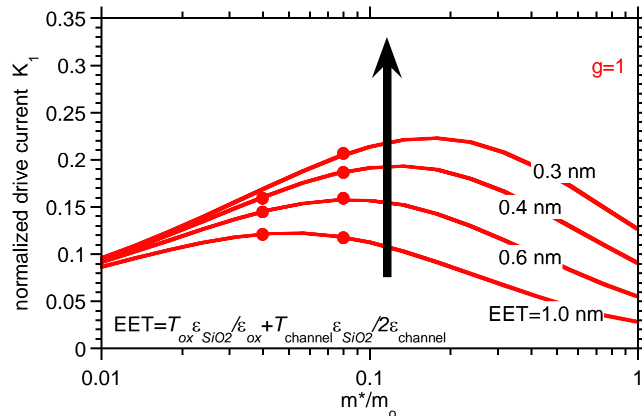
Needs→ research:

<u>RF front end:</u> phased array ICs, high-power transmitters, low-noise receivers

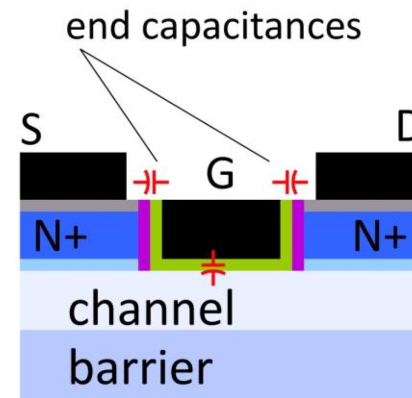
<u>IF/baseband:</u> ICs for multi-beam beamforming, for ISI/multipath suppression, ...

mm-Wave CMOS won't scale much further

Gate dielectric can't be thinned
→ on-current, g_m can't increase



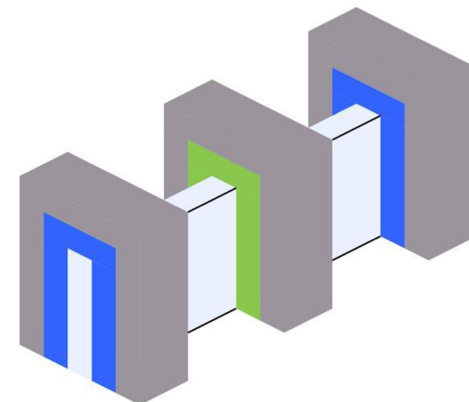
Shorter gates give no less capacitance
dominated by ends; ~1fF/μm total



Maximum g_m, minimum C→ upper limit on f_τ.
about 350-400 GHz.

Tungsten via resistances reduce the gain
Inac et al, CSICS 2011

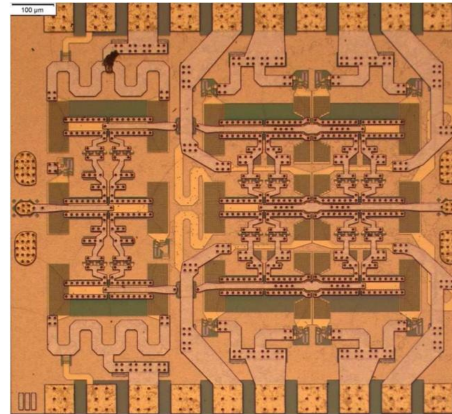
Present finFETs have yet larger end capacitances



III-V high-power transmitters, low-noise receivers

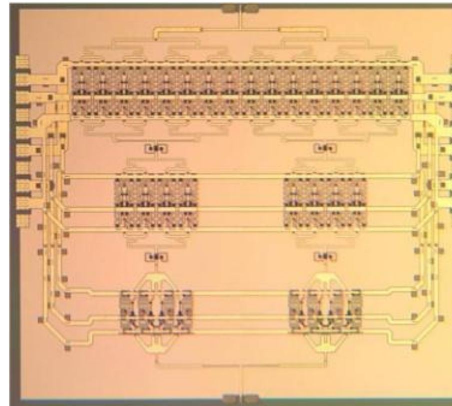
Cell phones & WiFi:
GaAs PAs, LNAs

mm-wave links need
high transmit power,
low receiver noise



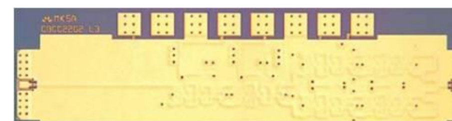
0.47 W @86GHz

H Park, UCSB, IMS 2014



0.18 W @220GHz

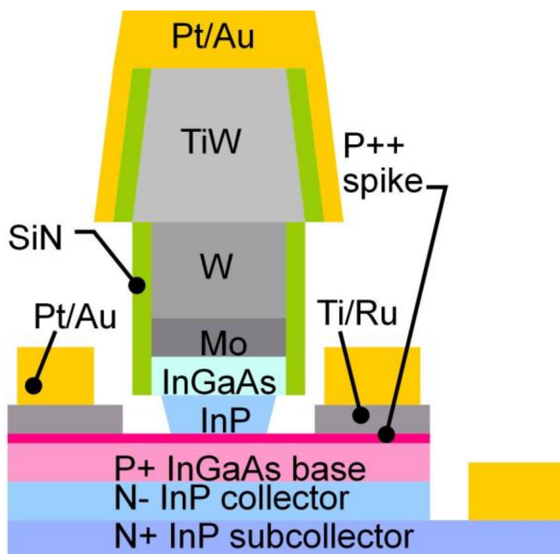
T Reed, UCSB, CSICS 2013



1.9mW @585GHz

M Seo, TSC, IMS 2013

Making faster bipolar transistors



to double the bandwidth:	change
emitter & collector junction widths	decrease 4:1
current density (mA/μm^2)	increase 4:1
current density (mA/μm)	constant
collector depletion thickness	decrease 2:1
base thickness	decrease 1.4:1
emitter & base contact resistivities	decrease 4:1

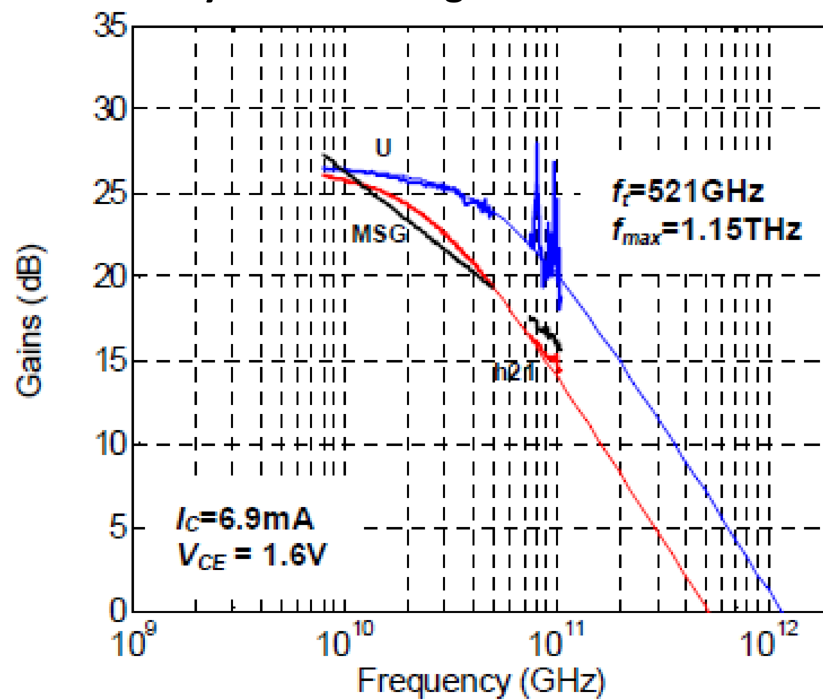
Narrow junctions.

Thin layers

High current density

Ultra low resistivity contacts

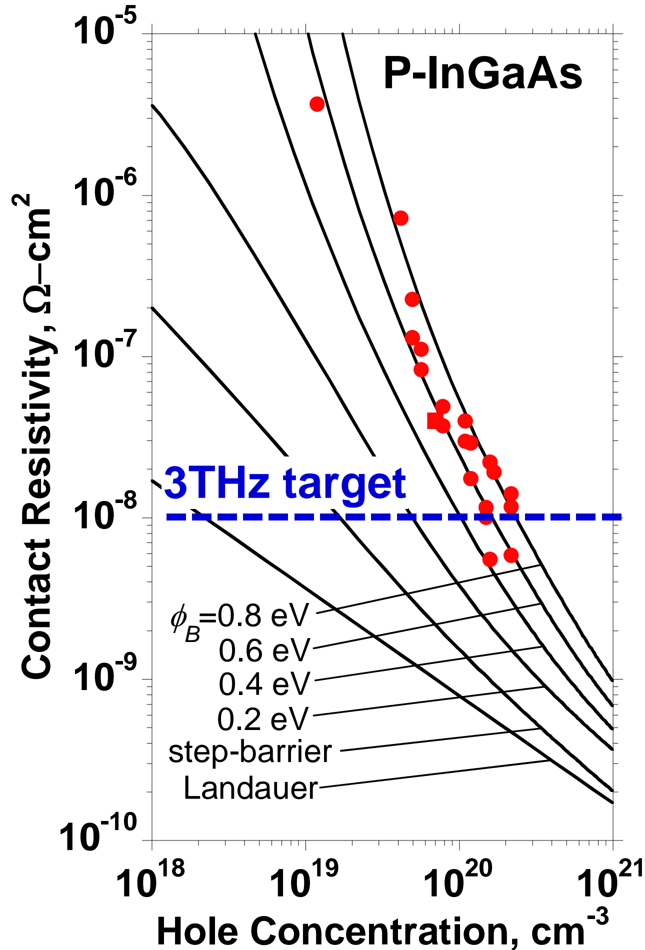
Teledyne: M. Urteaga *et al*: 2011 DRC



THz HBTs: The key challenges

Obtaining good base contacts

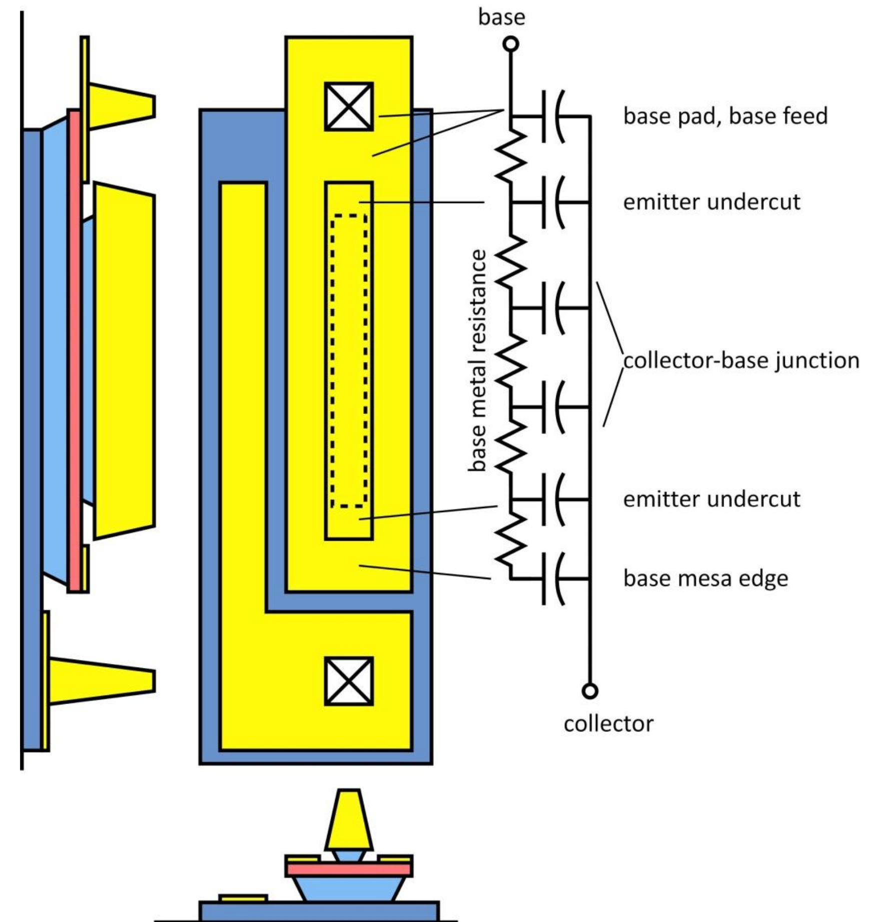
in full HBT process flow
(vs. in TLM structure)



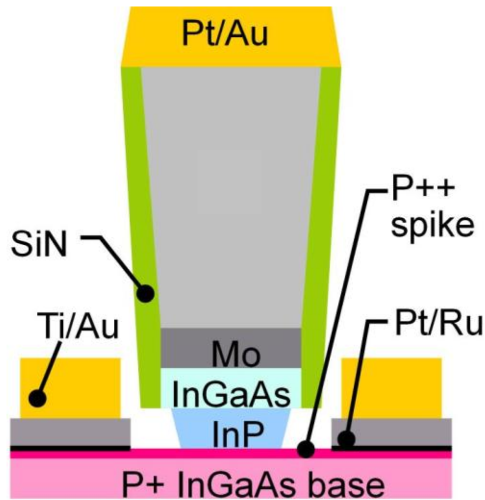
Baraskar *et al*, Journal of Applied Physics, 2013

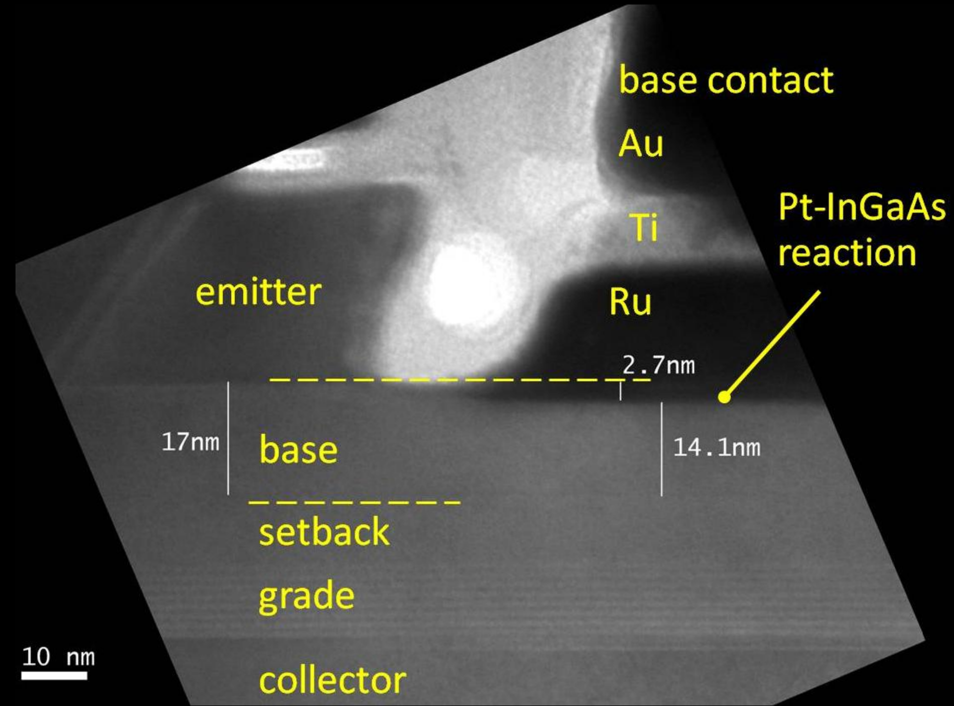
RC parasitics along finger length

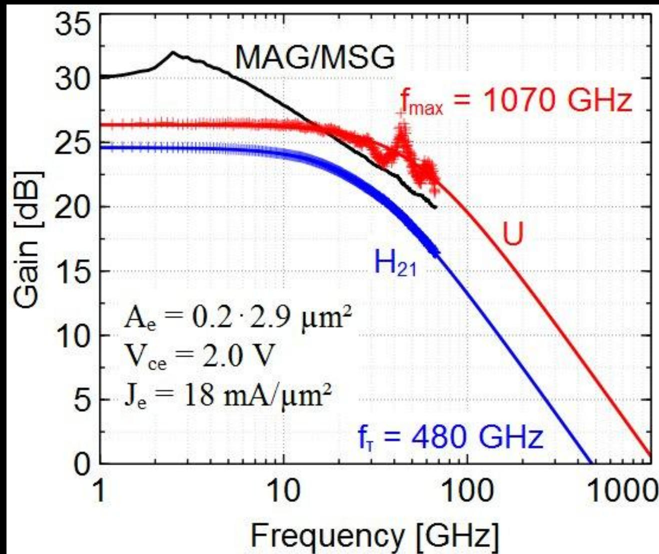
metal resistance, excess junction areas



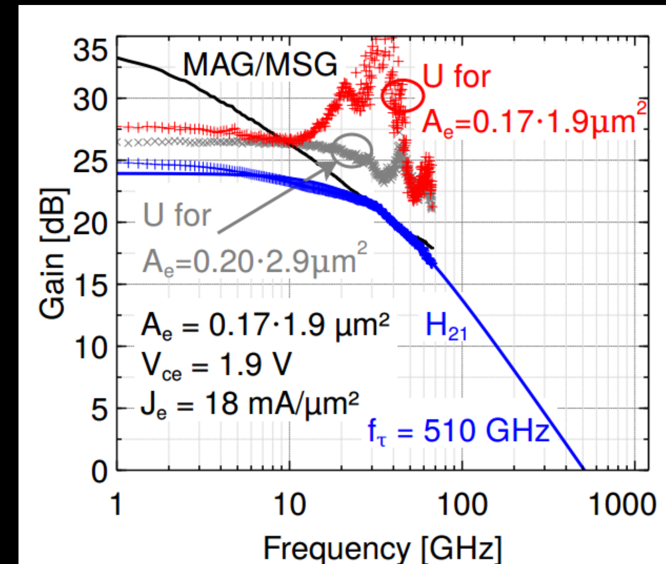
THz InP HBTs





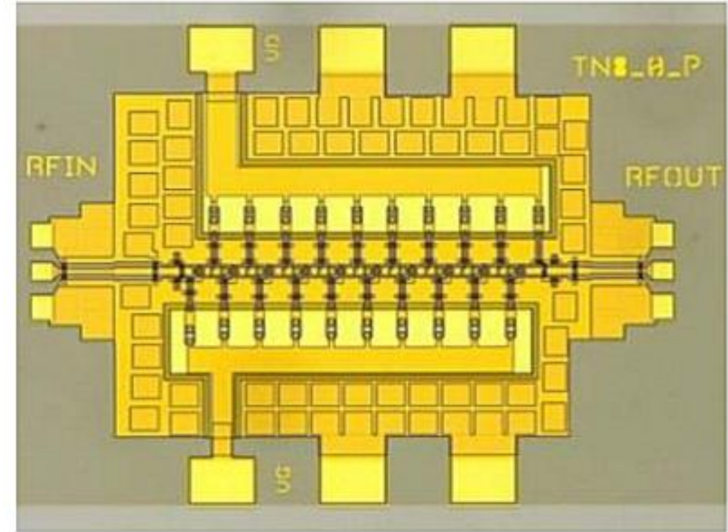


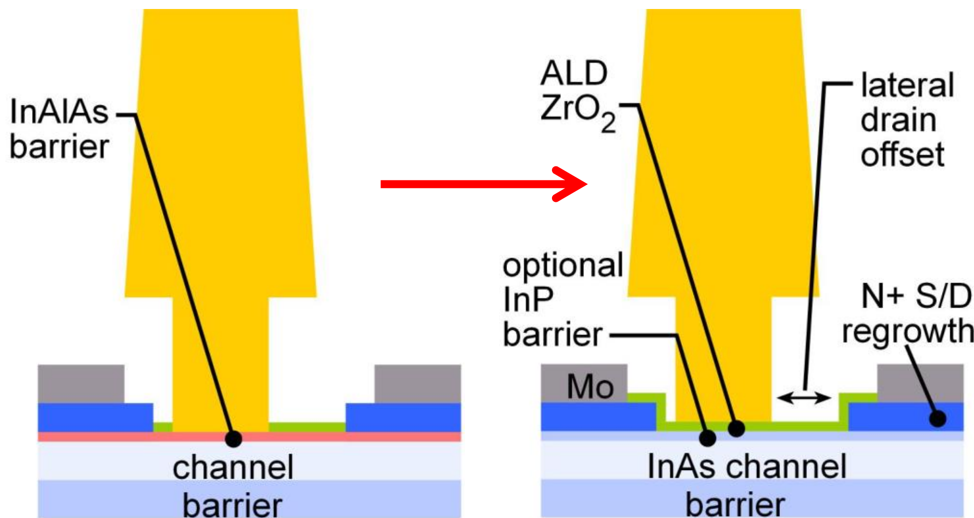
J. Rode, in review



THz HEMTs: one more scaling generation ?

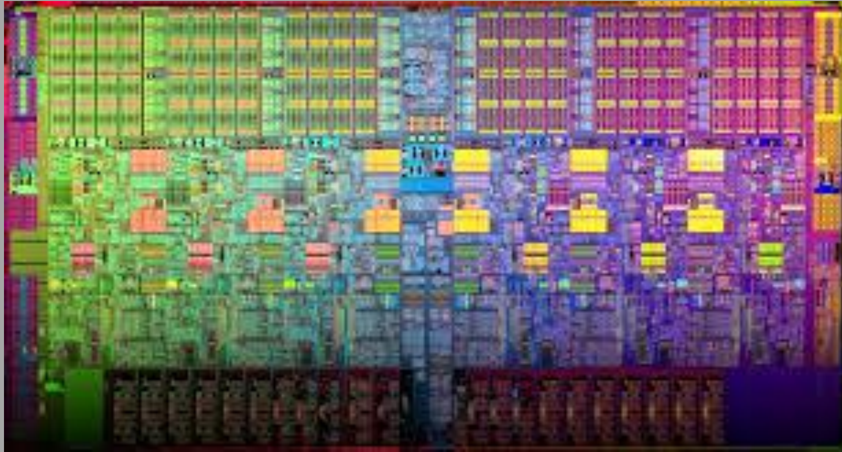
First Demonstration of Amplification at 1 THz Using 25-nm InP High Electron Mobility Transistor Process
Xiaobing Mei, et al, IEEE EDL, April 2015 doi: 10.1109/LED.2015.2407193

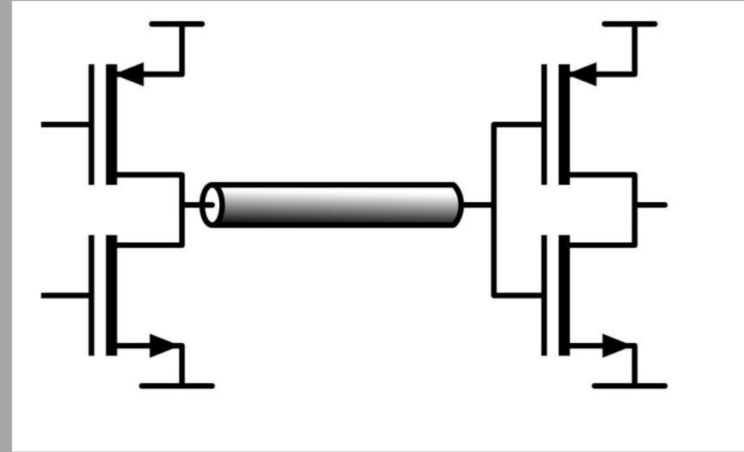


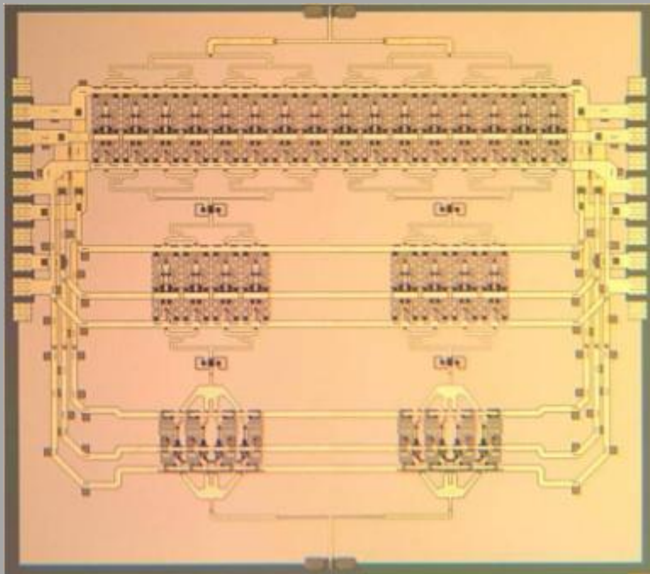


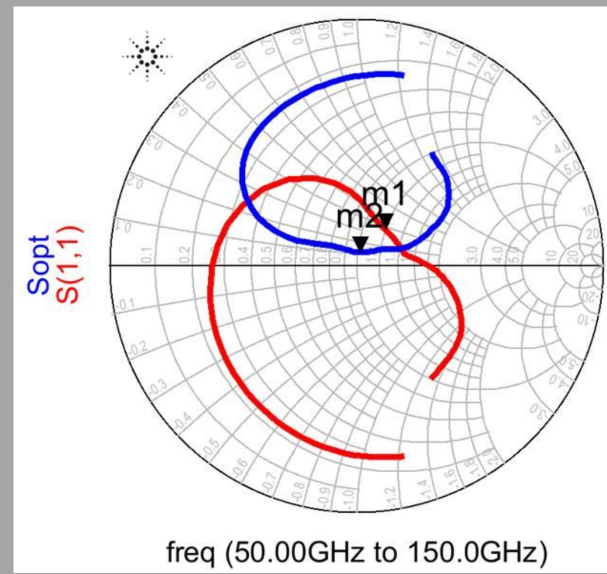
gate length	36	18	9	nm
EOT	0.8	0.4	0.2	nm
well thickness	5.6	2.8	1.4	nm
effective mass	0.05	0.08	0.08	times m_0
# bands	1	1	1	--
S/D resistivity	150	74	37	Ω-µm
extrinisic g_m	2.5	4.2	6.4	mS/µm
on-current	0.55	0.8	1.1	mA/µm
f_τ	0.70	1.2	2.0	THz
f_{max}	0.81	1.4	2.7	THz

nm & THz electronics





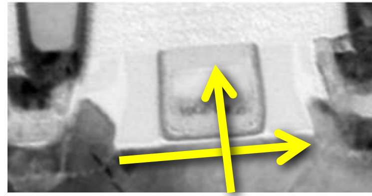




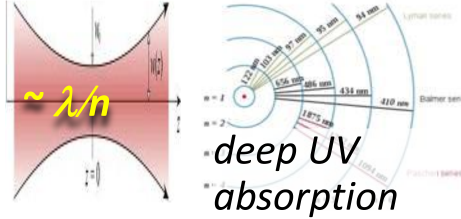
Electron devices: What's next ?

Problems:

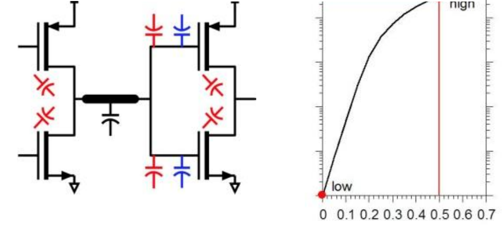
oxide, S/D tunneling



lithography



interconnect energy



& static dissipation

Why transistors are best:

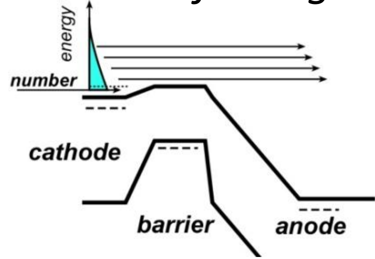
	I	II	III	
Quarks	2.4 MeV, ⅔, ½, **u** up	1.27 GeV, ⅔, ½, **c** charm	171.2 GeV, ⅔, ½, **t** top	0, 0, 1, **γ** photon
Quarks	4.8 MeV, -⅓, ½, **d** down	104 MeV, -⅓, ½, **s** strange	4.2 GeV, -⅓, ½, **b** bottom	0, 0, 1, **g** gluon
Leptons	<2.2 eV, 0, ½, ν_e electron neutrino	<0.17 MeV, 0, ½, ν_μ muon neutrino	<15.5 MeV, 0, ½, ν_τ tau neutrino	91.2 GeV, 0, 1, Z^0 weak force
Leptons	0.511 MeV, -1, ½, **e** electron	105.7 MeV, -1, ½, **μ** muon	1.777 GeV, -1, ½, **τ** tau	80.4 GeV, ±1, 1, $W^\pm$ weak force

(mass, charge, spin, name; Bosons (Forces))

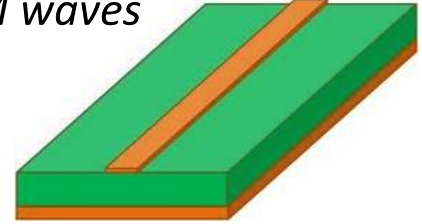
http://en.wikipedia.org/wiki/Standard_Model

our best tools are:

..electrostatic control of charge

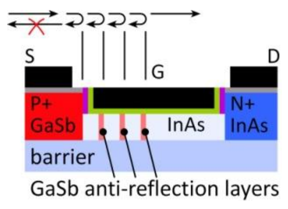


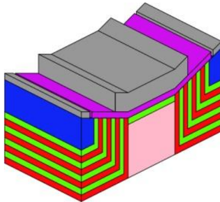
...and communicating by E&M waves

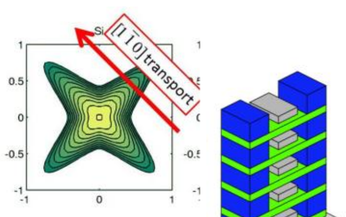


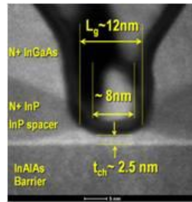
Opportunities:

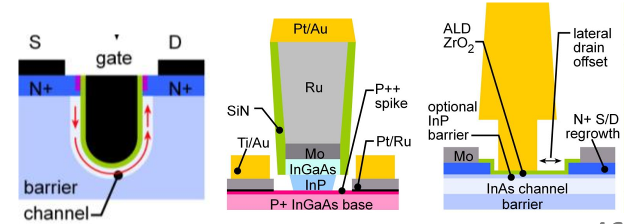
low voltages *high currents* *nm via 3D* *RF→ THz*











(backup slides follow)