Legend and Folklore:
A bestiary of electronics

Mark Rodwell,
University of California, Santa Barbara

~~Legend and Folklore:~~
~~A bestiary of~~ electronics

"Must-Haves"
for a Useful Electron Device

Mark Rodwell,
University of California, Santa Barbara

With apologies to my co-authors.

C.-Y. Huang, J. Rode, S. Lee, V. Chobpattanna,
P. Choudhary, A.C. Gossard, S. Stemmer
ECE and Materials Departments,
University of California, Santa Barbara

P. Long, E. Wilson, M. Povolotskyi, G. Klimeck
Network for Computational Nanotechnology, Purdue University

M. Urteaga, J. Hacker, M. Seo*, Z. Griffith, M. Fields, B. Brar
Teledyne Scientific and Imaging
**Now Sunkyunkwan University.*

Legend and Folklore: a bestiary of electronics

Let us now go hunting …

Terra incognita: beyond the transistor

"Moore's law,
transistor scaling, is over."

"We need more than Moore"

"We need new paradigms beyond charge as a state variable"

(Why do paradigms always shift?)

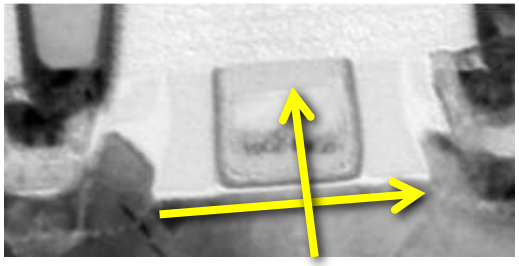
Moore or Dennard ?

Moore: predicted doubling of performance every ~18 months

Dennard: Proposed FET scaling laws (these now broken)

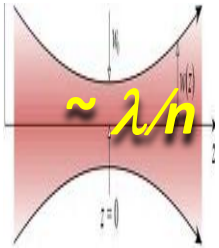
Robert and Gordon: What's your problem ?

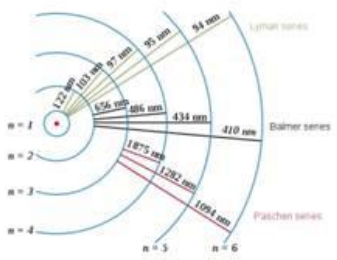
FETs will stop working when we make them much smaller.



Oxide tunneling
Source-drain tunneling

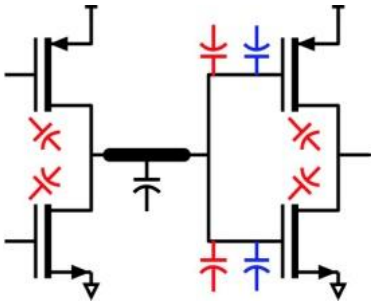
Lithography is getting hard.

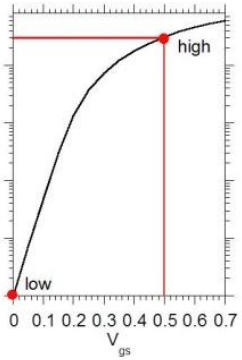




minimum focus spot size
deep UV absorbtion

Power density is becoming excessive





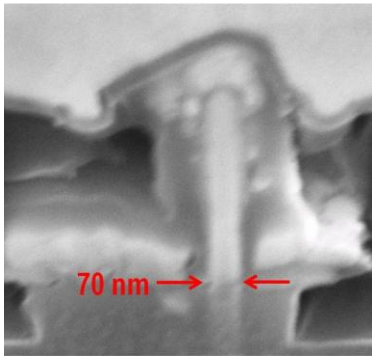
$C_{wire}V_{DD}^2$ *interconnect energy*

Static leakage $> I_{on}exp(-qV_{DD}/kT)$

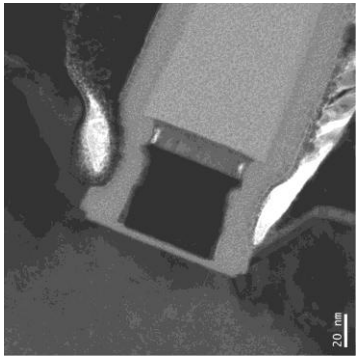
How electronics works today (same as in 1912)

Switching using charge-control

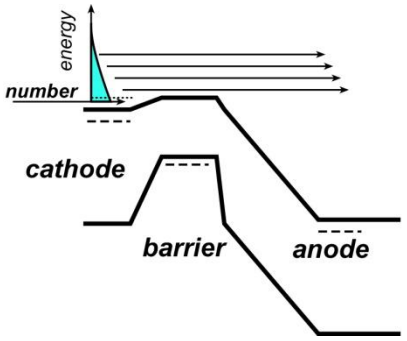
tubes



bipolar transistors

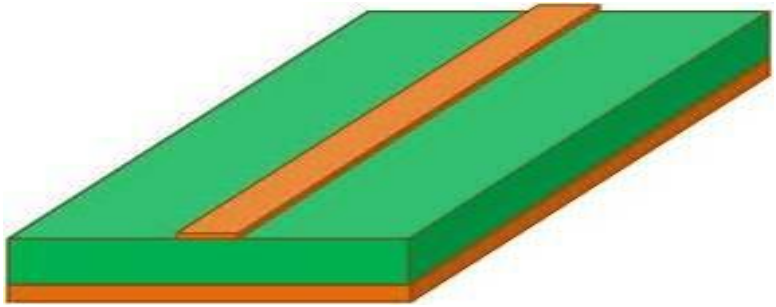


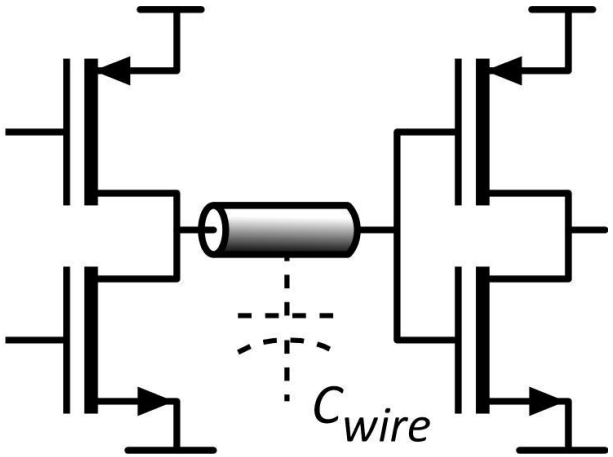
field-effect transistors



electrostatic barrier

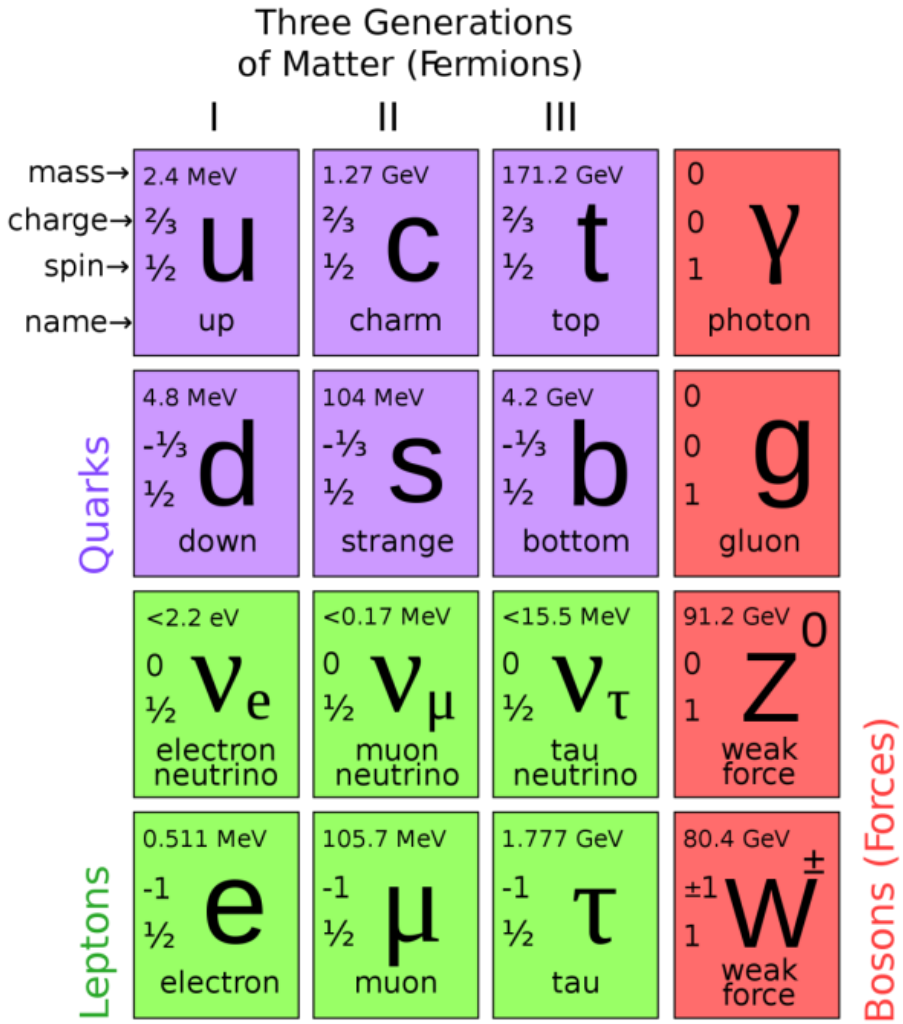
Communicating using... ...wires.





Time for new state variables ?

our tool-kit:



Gravitonics ?

Quarkonics ?

gluonics?

Neutrinoonics ?

Mechanics ?

millions of heavy nuclei ...

Magnetics ?

$F \sim v_1 v_2 / c^2 \rightarrow$ weak!

Our forebears chose wisely

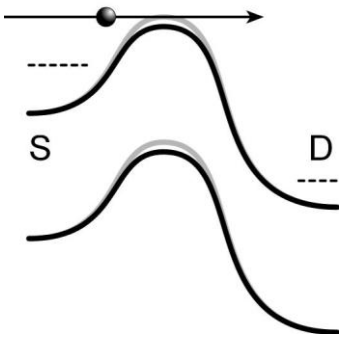
Charge-control devices (switches) seem best

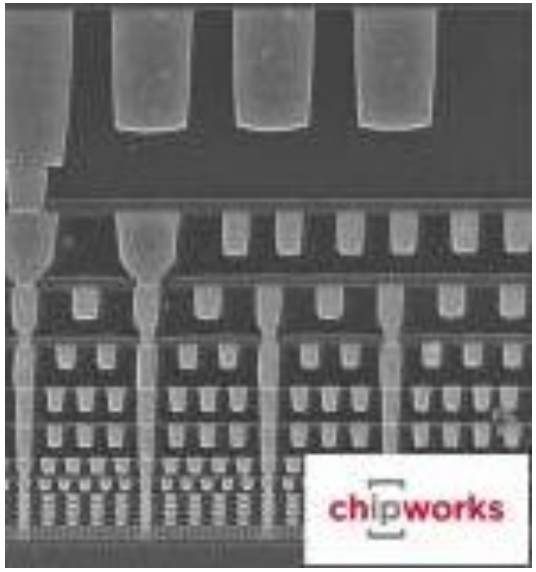
confine & release charge using electrostatic barrier.
electrons are light
electrostatic force is strong over moderate* range.

*range ~ source size, given $\nabla^2\phi = 0$

Communicating using wires seems best

signal using E&M waves
guide them using wires
E&M waves are strong & long-range.
wires can be very narrow

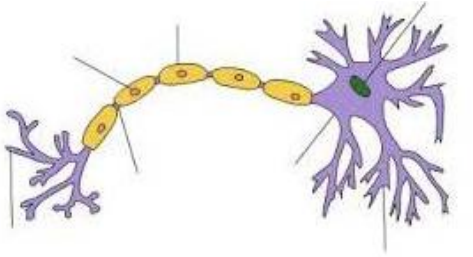




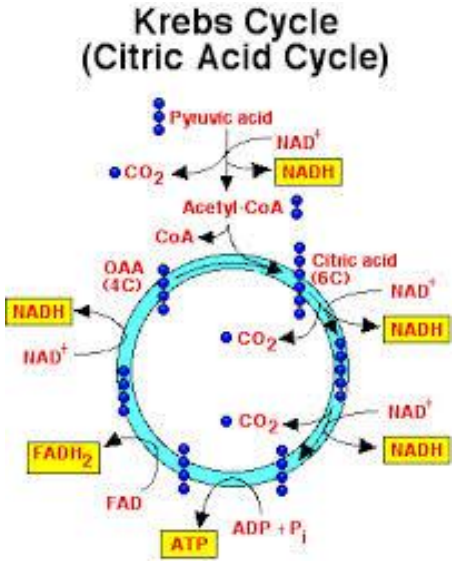
http://semimd.com/chipworks/2014/10/27/intels-14nm-parts-are-finally-here/ca

Two older computing technologies

Nerve Cells



Cellular chemistry



cascaded, inter-regulating
chemical reactions = computing machine

Both are charge-control
dense
slow
long development (~10^9 years)
large installed base

"$C_{wire}V^2$ dissipation constrains VLSI; optical interconnects will fix this."

Optical interconnects are better ?

Aren't they both E&M waves ?

So: what's the difference ?

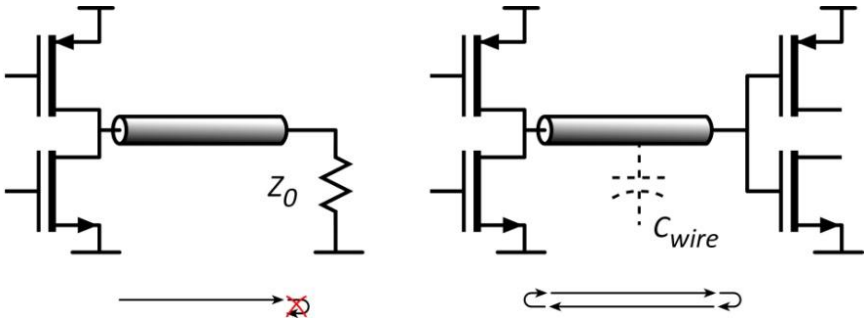
Both store energy $\varepsilon E^2/2 + \mu H^2/2$, a.k.a. $CV^2/2 + LI^2/2$

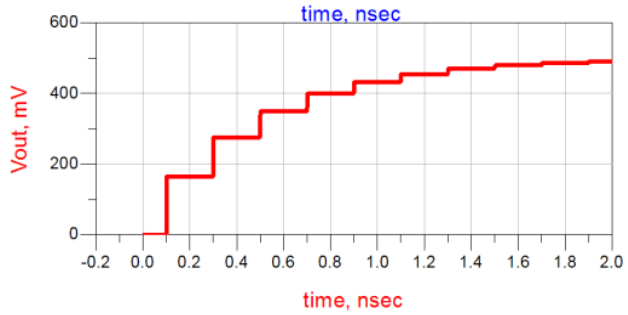
Wires can be either capacitors or transmission-lines

echoes or not

T-lines: static dissipation

Capacitors: $CV^2/2$ dissipation per transition





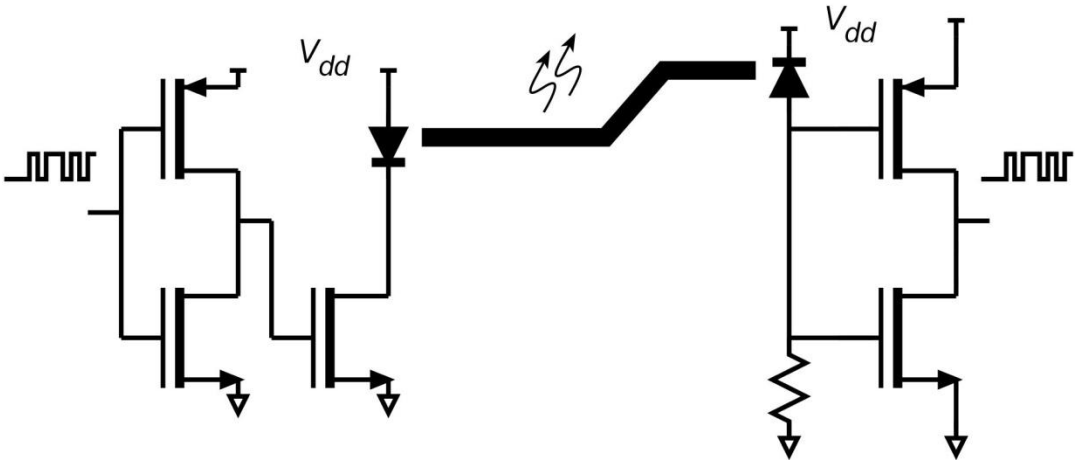
$$C_{wire} = l / vZ_0$$

Optical interconnects are better ?

If you shine a laser at a mirror, ...does it stop drawing current ?

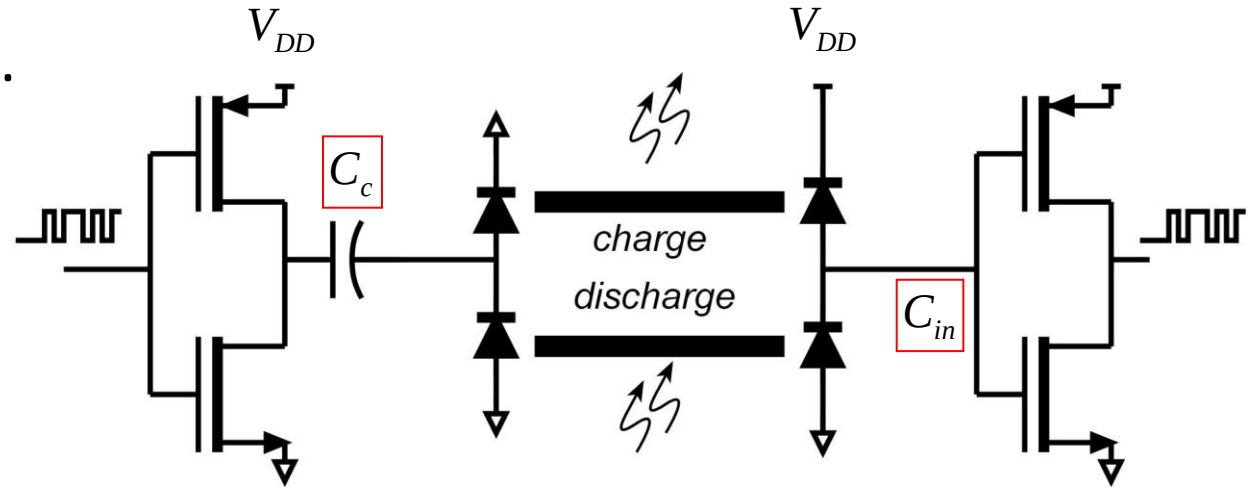
Optical interconnects have static dissipation

draws current
when "1" transmitted



No static dissipation....

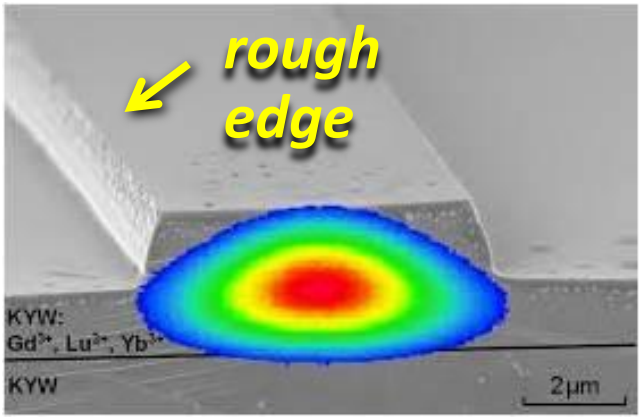
...looks tough.



$$\#\text{photons transmitted} = C_c V_{dd} / h\nu$$

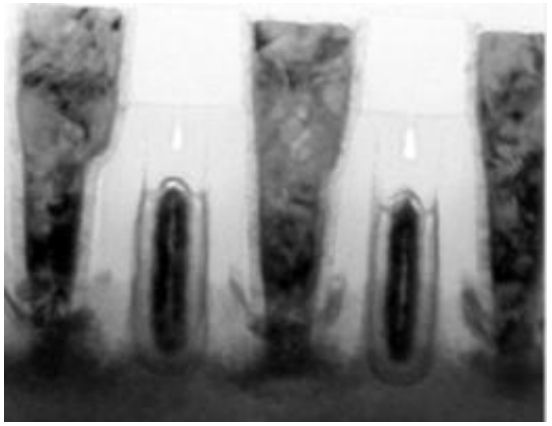
$$\#\text{photons needed} = C_{in} V_{dd} / h\nu$$

Other issues



optical losses

bend radius



20nm
contact pitch ?

optics benefit: lower loss in longer interconnects

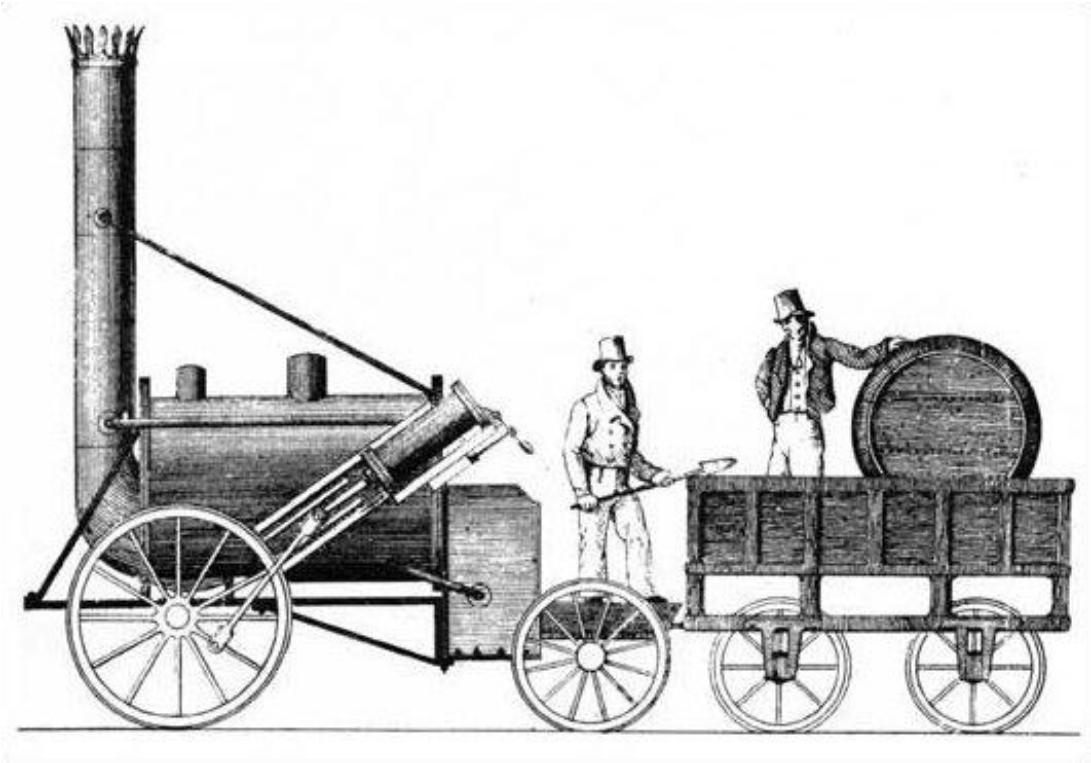
Where to use optics:

longer interconnect buses

where switching activity is high

Can man live at such speeds ?

"Can man live at such speeds ?"



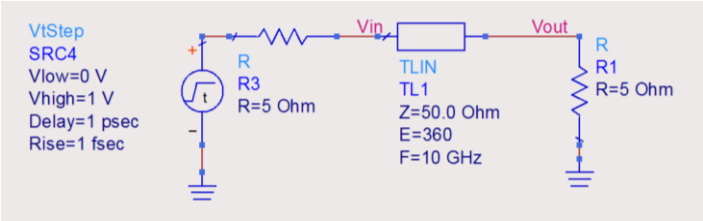
Stephenson's *Rocket*, 1829

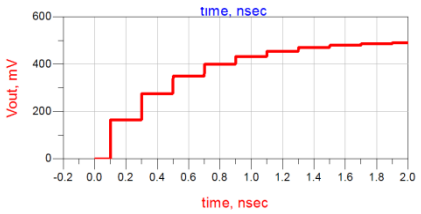
"Circuit theory doesn't work in the { *sub-mm-wave* / *THz* / *IR* / *etc."*

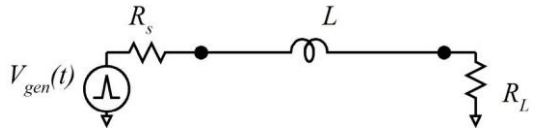
Circuit theory is just Maxwell's equations

Circuits: Maxwell's equations in 0-D limit
T-lines: Maxwell's equations in 1-D limit, etc.

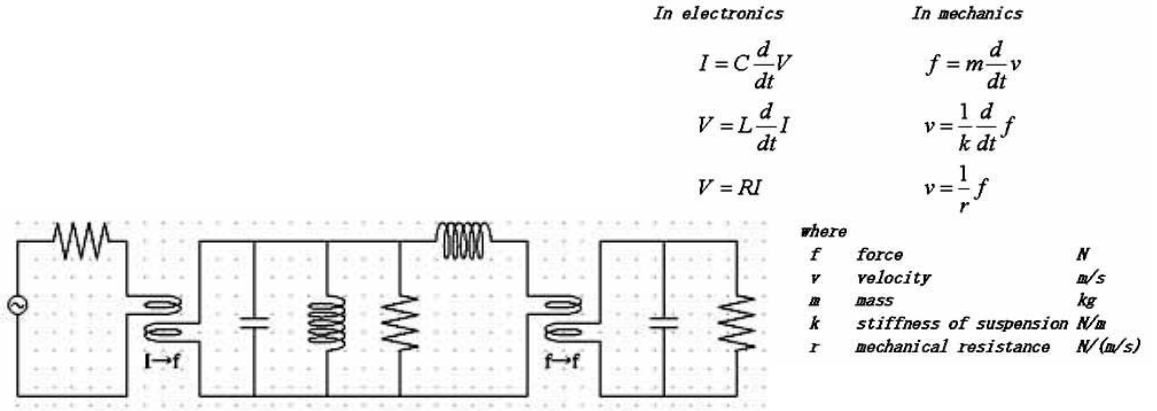
Example: short T-line approximating an inductor







Any system of PDEs→ mesh finely into ODEs→ ***equivalent circuit***

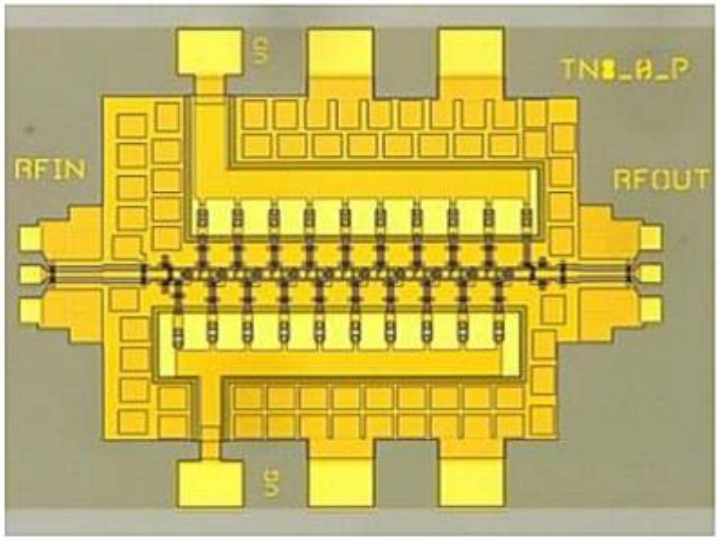


http://www5.ocn.ne.jp/

Circuit theory: alive & well at 1.0 THz

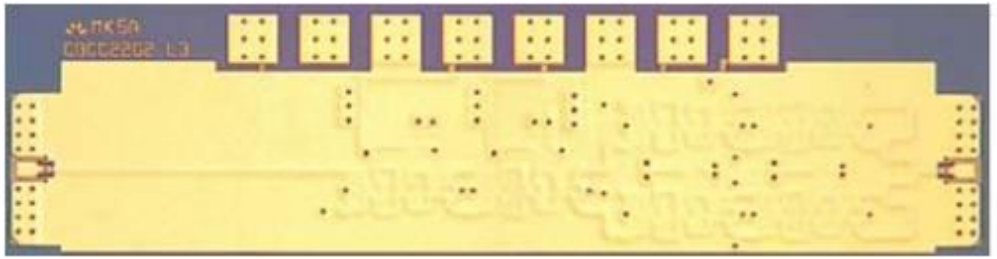
First Demonstration of Amplification at 1 THz Using 25-nm InP High Electron Mobility Transistor Process

Xiaobing Mei, et al, IEEE EDL, April 2015 doi: 10.1109/LED.2015.2407193



620 GHz, 20 dB gain HBT amplifier
M Seo, Teledyne, IMS 2013

Not shown: 670 GHz HBT amplifier:
J. Hacker, Teledyne: IMS 2013

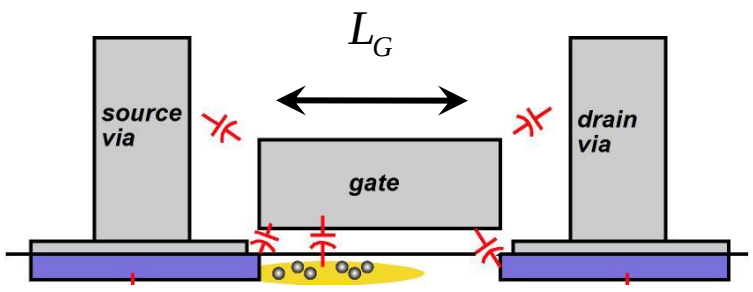


No question that 1 THz interconnects are challenging, but they work...

"Charge control doesn't work in the { *sub-mm-wave* / *THz* / *IR* / *etc."*

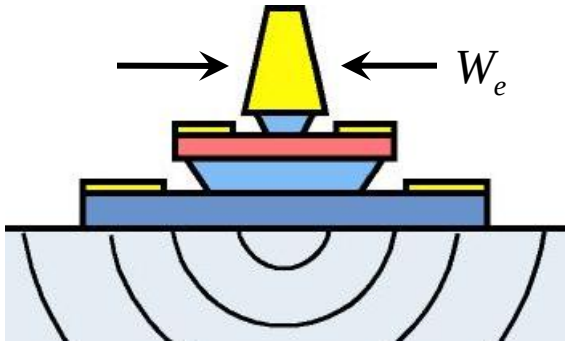
...we must use quantum transitions"

To double transistor bandwidth...



(gate width W_G)

FET parameter	change
gate length	decrease 2:1
current density (mA/μm), g_m (mS/μm)	increase 2:1
transport effective mass	constant
channel 2DEG electron density	increase 2:1
gate-channel capacitance density	increase 2:1
dielectric equivalent thickness	decrease 2:1
channel thickness	decrease 2:1
channel state density	increase 2:1
contact resistivities	decrease 4:1



(emitter length L_E)

HBT parameter	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
contact resistivities	decrease 4:1

Electron device scaling & frequency limits

$\tau \propto \text{thickness}$

$C \propto \text{area} / \text{thickness}$

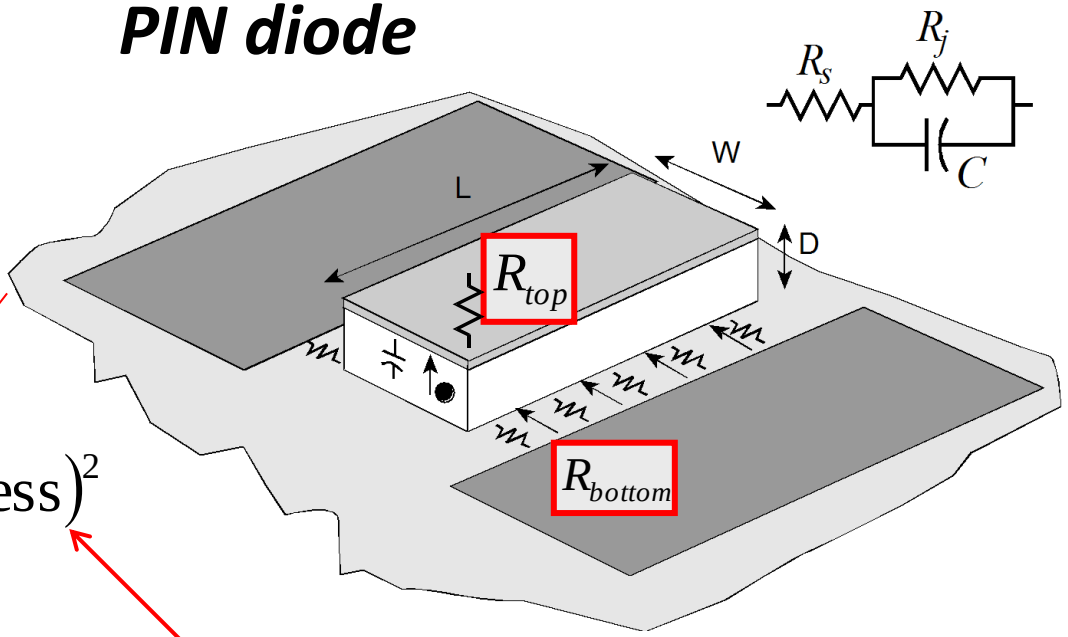
$R_{top} \propto \rho_{contact} / \text{area}$

$$R_{bottom} \propto \frac{\rho_{contact}}{\text{area}} + \frac{\rho_{sheet}}{4} \cdot \frac{\text{width}}{\text{length}}$$

$$I_{\text{max, space-charge-limit}} \propto \text{area} / (\text{thickness})^2$$

$$\Delta T \propto \frac{\text{power}}{\text{length}} \times \log\left(\frac{\text{length}}{\text{width}}\right)$$

PIN diode

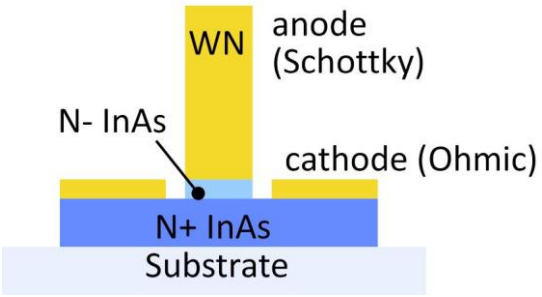


from $\partial^2\phi / \partial x^2 = \rho / \varepsilon = -J / \varepsilon v$

To double bandwidth:

- reduce thicknesses 2:1
- Improve contacts 4:1
- reduce width 4:1,
- keep constant length
- increase current density 4:1

How fast might it be ? 5nm diameter Schottky



$$\text{Assume: } T_{depl} \approx 3.5\,\text{nm, 5nm diameter, } 0.3\,\text{A}/\mu\text{m}^2$$

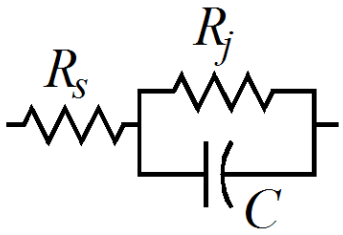
$$\text{Average velocity: } \bar{v} = (v_{Fermi}/2) = 3.5\cdot 10^7\,\text{cm/s}$$

$$\text{Transit time: } \tau_{transit} = T_{depl}/(v_{Fermi}/2) = 9.6\,\text{fs}$$

$$\text{Capacitance: } C = \varepsilon A/D = 0.43\,\text{aF}$$

$$\text{Series Resistance } (N^+, \text{contacts}): R_S = 148\,\Omega$$

$$\rightarrow R_S C = 65\,\text{as}$$

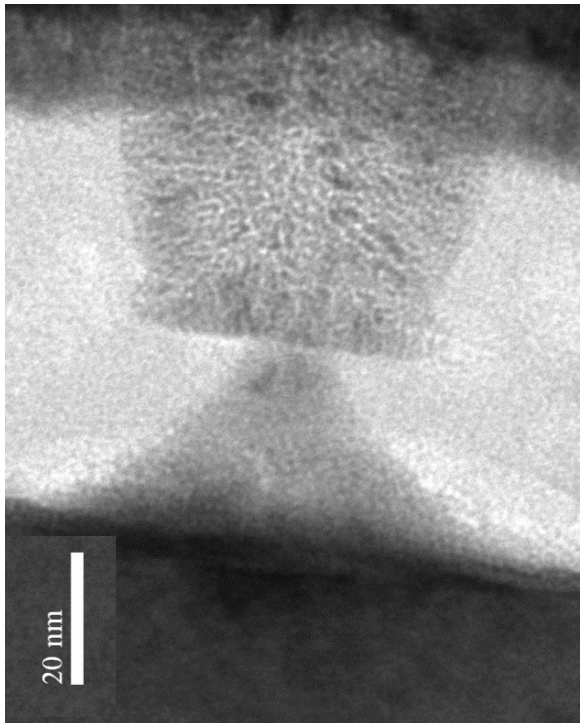


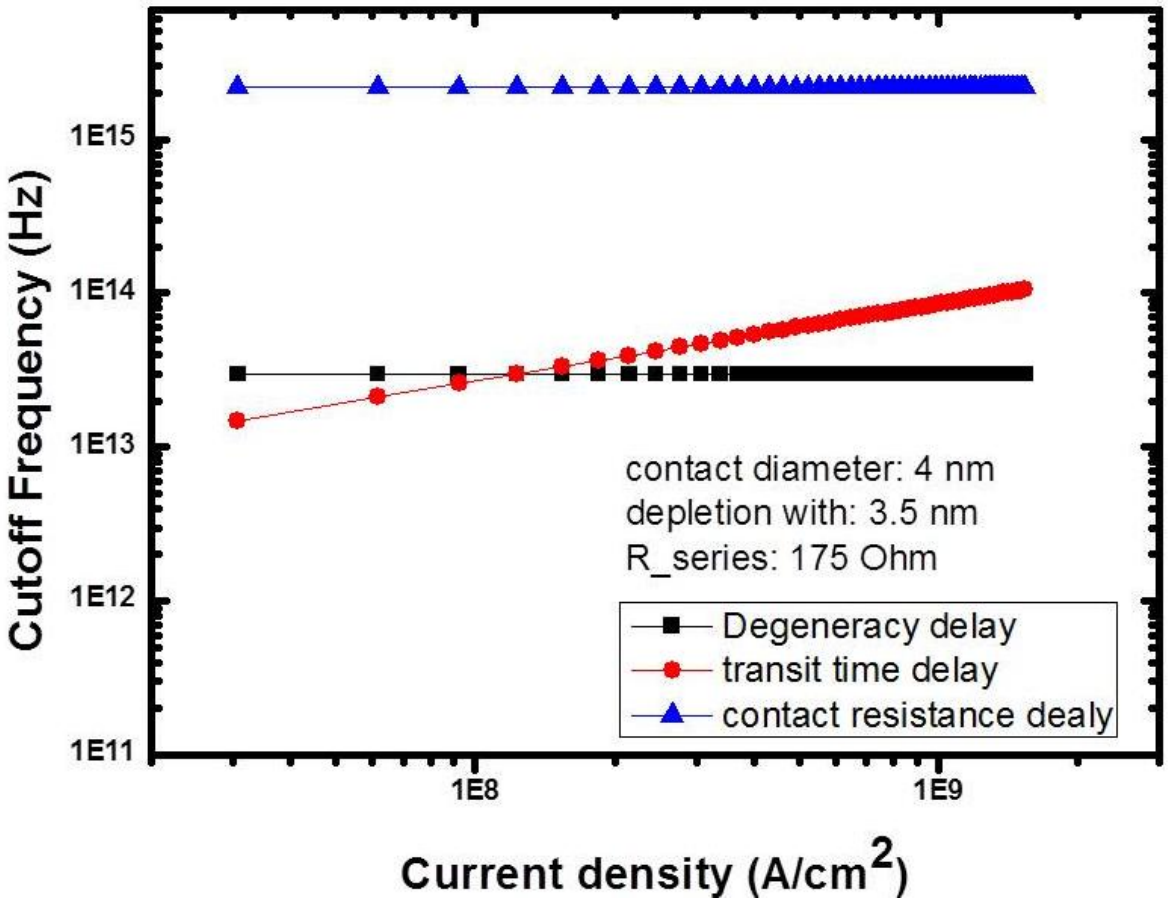
$$\text{Junction Resistance (degenerate 1D transport): } R_j \approx h/2q^2 + kT/qI = 17\text{k}\Omega$$

$$\rightarrow R_j C = 7.5\,\text{fs}$$

20-30 THz diode cutoff frequencies ?

How fast might it be ? 5nm diameter diode

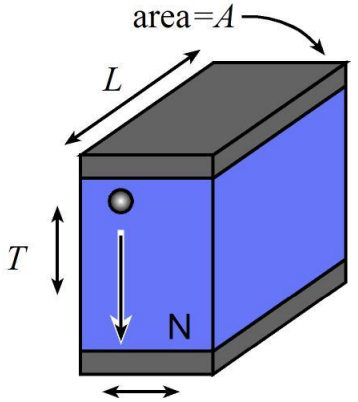




Rob Maurer, Cheng-Ying Huang, Unpublished

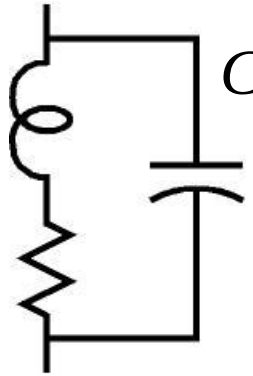
...plasma resonance sets an upper frequency limit "

Plasma resonance ? No worries !



$$L_{\text{kinetic}} = \frac{T}{A}\frac{1}{q^2 n m^*}$$

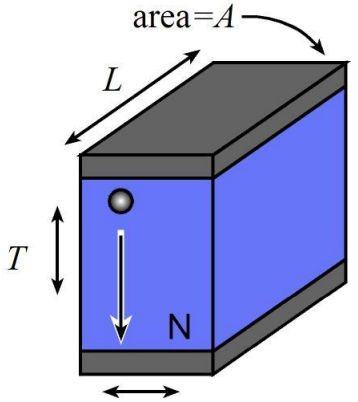
$$R_{\text{bulk}} = \frac{T}{A}\frac{1}{q^2 n m^* \tau_m}$$



$$C_{\text{displacement}} = \frac{\varepsilon A}{T}$$

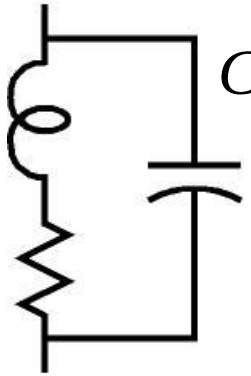
	dielectric relaxation frequency $f_{dielectric} = \frac{1/2\pi}{C_{\text{displacement}} R_{\text{bulk}}} = \frac{1}{2\pi}\frac{\sigma}{\varepsilon}$	scattering frequency $f_{scattering} = \frac{1}{2\pi}\frac{R_{\text{bulk}}}{L_{\text{kinetic}}} = \frac{1}{2\pi\tau_m}$	plasma frequency $f_{plasma} = \frac{1/2\pi}{\sqrt{L_{\text{kinetic}} C_{\text{displacement}}}}$
n - InGaAs $3.5\cdot10^{19}/\text{cm}^3$	800 THz	7 THz	74 THz
p - InGaAs $7\cdot10^{19}/\text{cm}^3$	80 THz	12 THz	31 THz

Plasma resonance ? No worries !



$$L_{\text{kinetic}} = \frac{T}{A}\frac{1}{q^2 n m^*}$$

$$R_{\text{bulk}} = \frac{T}{A}\frac{1}{q^2 n m^* \tau_m}$$



$$C_{\text{displacement}} = \frac{\varepsilon A}{T}$$

$$\text{scattering frequency} \quad f_{scattering} = \frac{1}{2\pi}\frac{R_{\text{bulk}}}{L_{\text{kinetic}}} = 7\text{THz}$$

Above 7 THz, kinetic inductance increases N+/P+ layer impedances.

But: contact resistances >> (N+/P+) resistances.

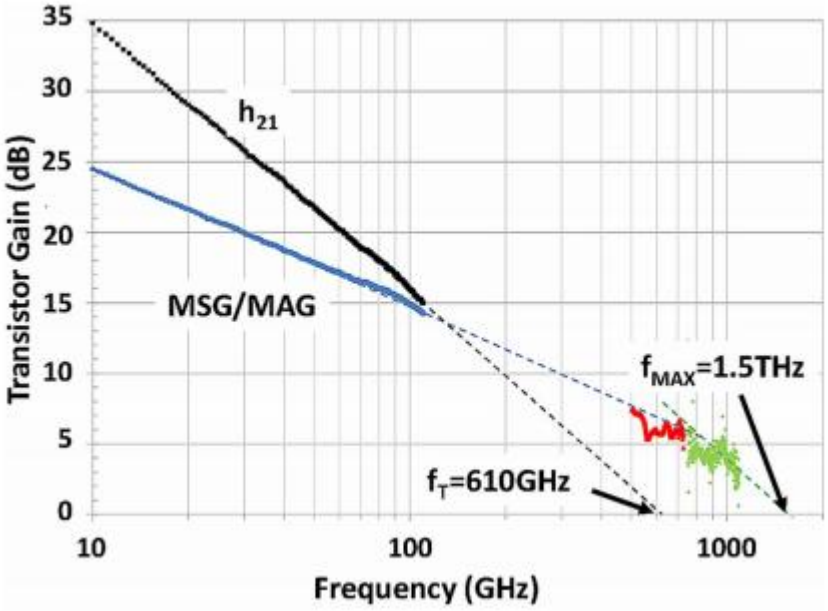
A non-dominant resistance is increasing with frequency.

Not a serious concern until ~30THz.

...optics is fast,
electronics is slow"

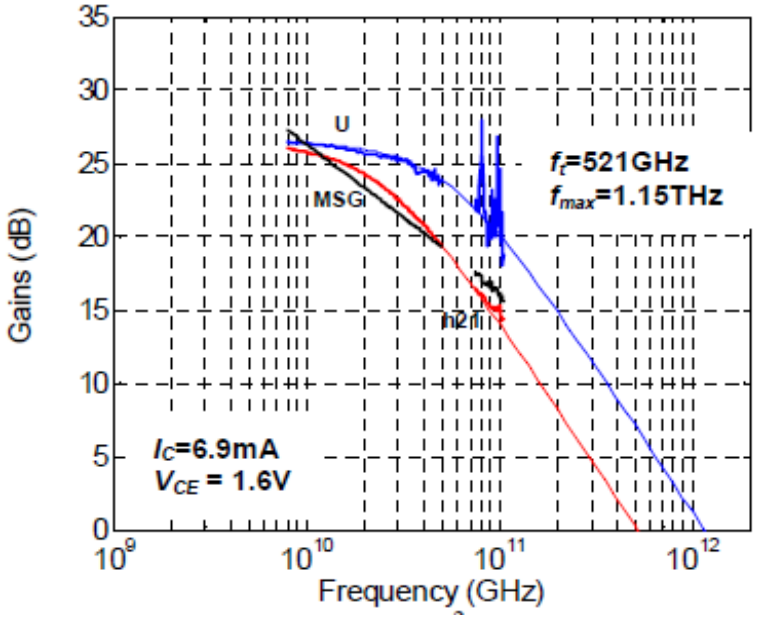
Transistors today

InP HEMTs



NGST
Xiaobing Mei *et al*, IEEE EDL, April 2015

InP HBTs



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

These made fast by ***scaling***	FET	Bipolar
lithographic dimensions:	25nm	130nm
epitaxial dimensions	~7nm	20nm base, 100nm collector
contact resistivities	~2-4Ω-μm² (both)	
current densities	~1mA/μm	~2 mA/μm
	100 mA/μm²	20-30 mA/μm²

Feasible to make these transistors (somewhat) faster.

Optics is fast ? Really ?

Optics has high *carrier* frequencies: 1.3 μm→ 230 THz

But, the per-channel *modulation* bandwidths are low.*

20~30 GHz modulation bandwidths for lasers

20~40 GHz for electro-absorbtion modulators

a few ~100 GHz traveling-wave EO modulators; these are big

Terabit optical fiber systems aggregate many channels.

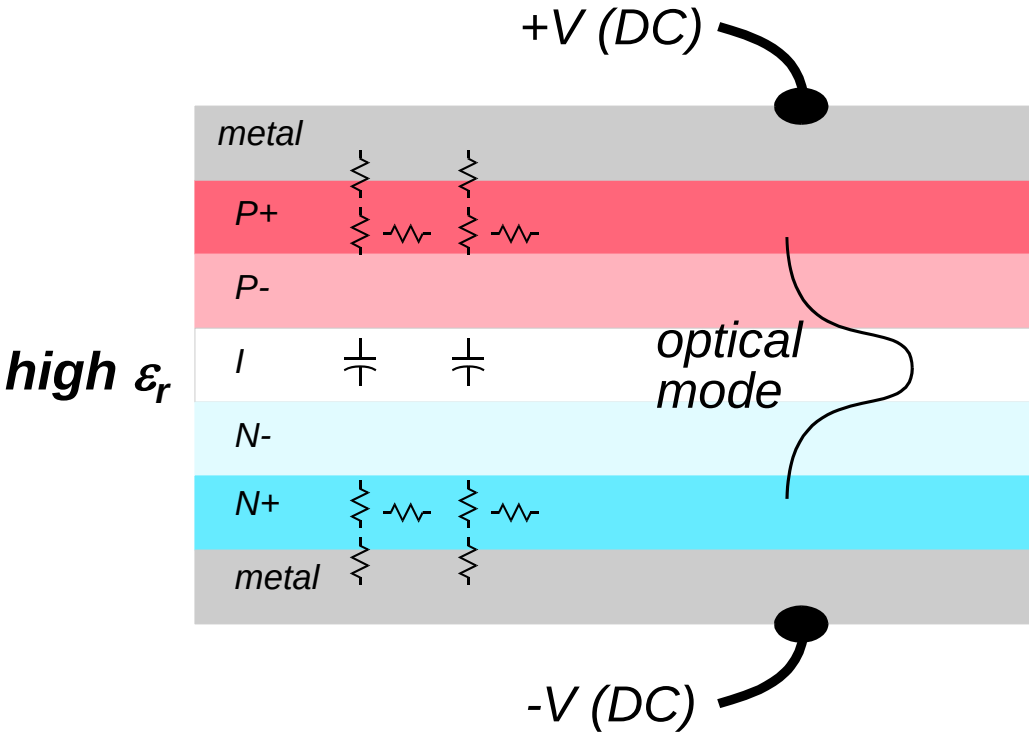
WDM, polarization, etc

Need lots of channels→ cost

Why are optical devices slow ?

*sure, a mode-locked-laser might have a 0.5fs pulse width,
but how rapidly can you impose a signal (information) on this pulse?

Optical devices are hard to scale



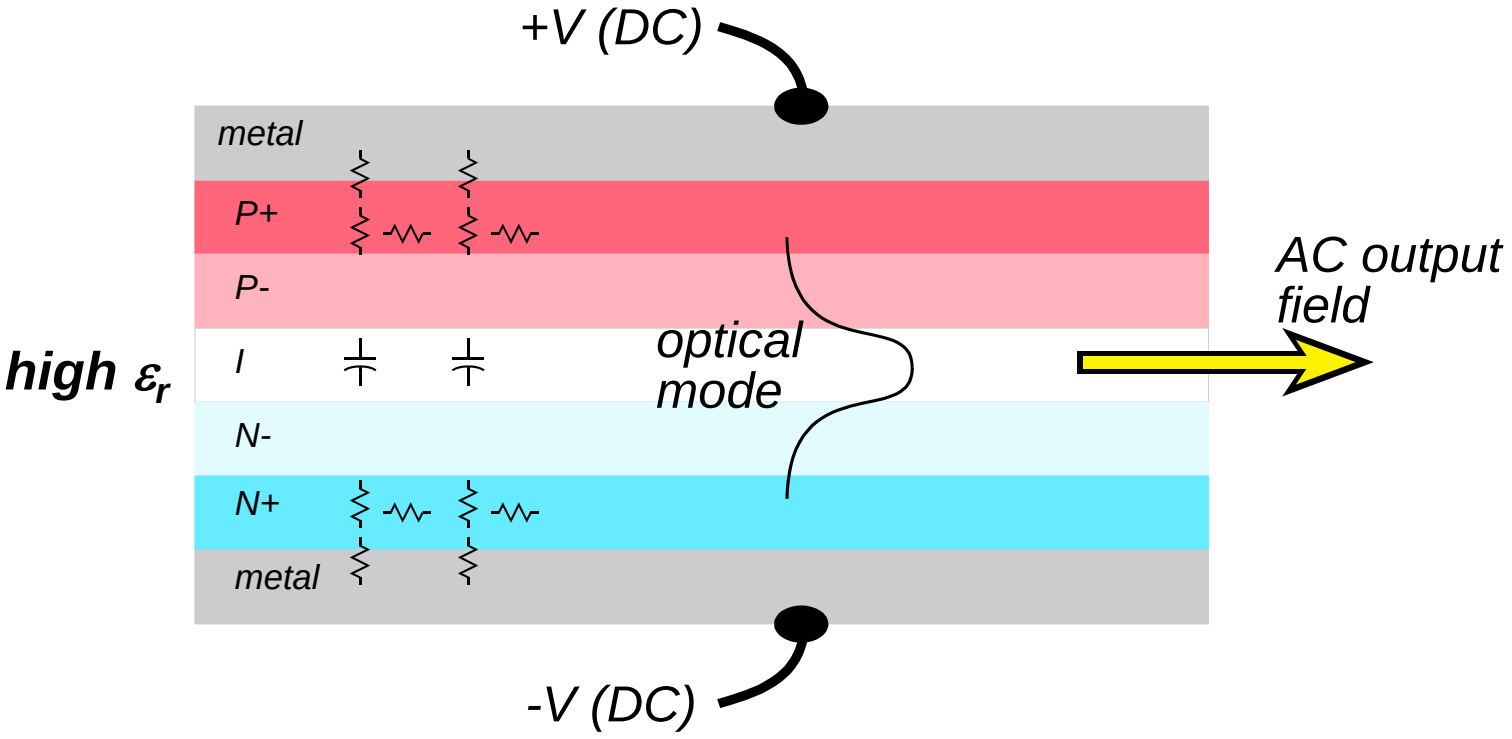
Optical mode size prevents scaling:

minimum I-layer thickness

minimum lateral junction width

maximum (P-/N-) doping (free-carrier losses)

But their carrier frequencies are high



Transistor R/C/τ limits don't apply to laser (etc) carrier frequency

carrier optical field guided by dielectric waveguide

AC field kept away from resistive bulk and contact regions.

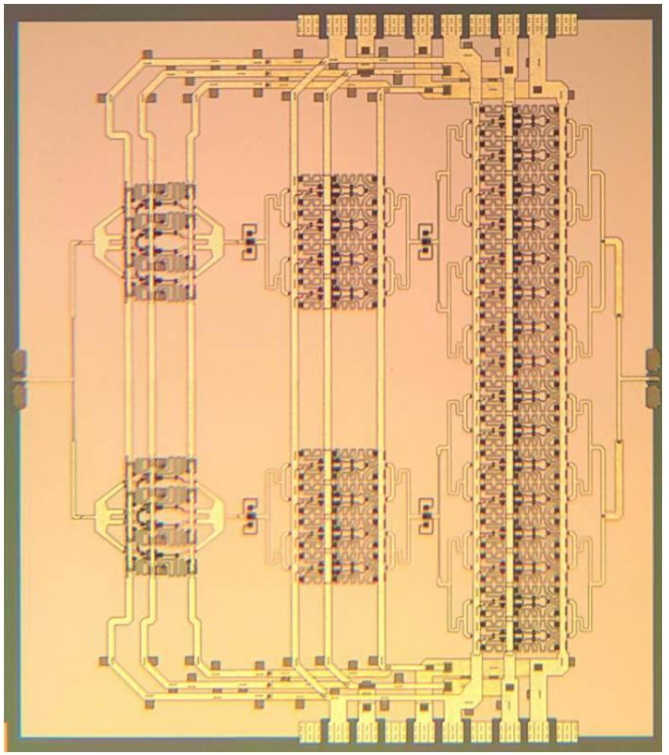
AC signal not coupled through electrical contacts

such dielectric mode confinement hard at low frequencies

"Tubes are ancient history"

Tubes are ancient ? They still beat transistors !

We developed a
180mW , 220 GHz
HBT power amplifier...



T Reed et al, CSICS 2014

... as part of a *driver* for a 20W, 220 GHz traveling-wave tube...
... for DARPA's 220 GHz radar

We develop solid-state mm-wave sources,
seeking to *drive* or *replace* existing high-power tubes.

Myths about III-Vs

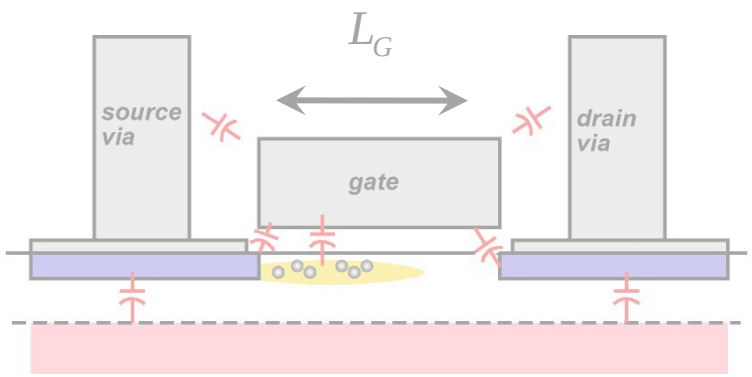
Heard often at IEDM *(even said by some III-V MOS folk)*
(never at DRC) :

"III-V contacts much poorer than Si"

"...can't be doped above $10^{18}/cm^3$."

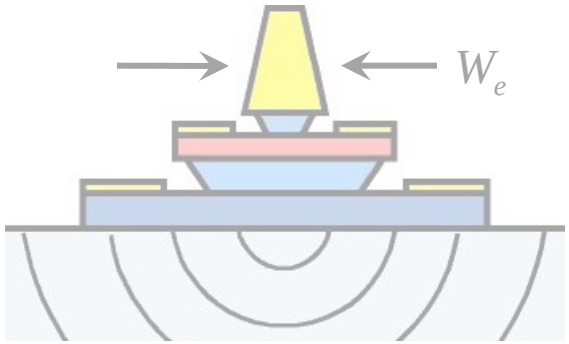
"...need to unpin the Fermi level under the contacts."

To double transistor bandwidth...



(gate width W_G)

FET parameter	change
gate length	decrease 2:1
current density (mA/μm), g_m (mS/μm)	increase 2:1
transport effective mass	constant
channel 2DEG electron density	increase 2:1
gate-channel capacitance density	increase 2:1
dielectric equivalent thickness	decrease 2:1
channel thickness	decrease 2:1
channel state density	increase 2:1
contact resistivities	decrease 4:1



(emitter length L_E)

HBT parameter	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
contact resistivities	decrease 4:1

III-V contacts much poorer than Si ???

For N-type, Si seems to be just a bit better

Si at ~0.3 Ω-μm^2 , InAs at ~0.5 Ω-μm^2

For P-type, Si significantly better

P-SiGe:~0.15 Ω-μm^2 (ZHANG et al, EDL, June 2013)

P-InGaAs at ~0.5 Ω-μm^2

III-V's can't be N-doped above ~10^{18}/cm^3 ???

Very strange, very persistent myth.

Easy fix: read any of 100's of papers in the literature

N-type InGaAs to ~$8*10^{19}$/cm^3, InAs to 10^{20}/cm^3,
P-type InGaAs to ~$2*10^{20}$/cm^3

Myth seems to arise from low conduction-band state density

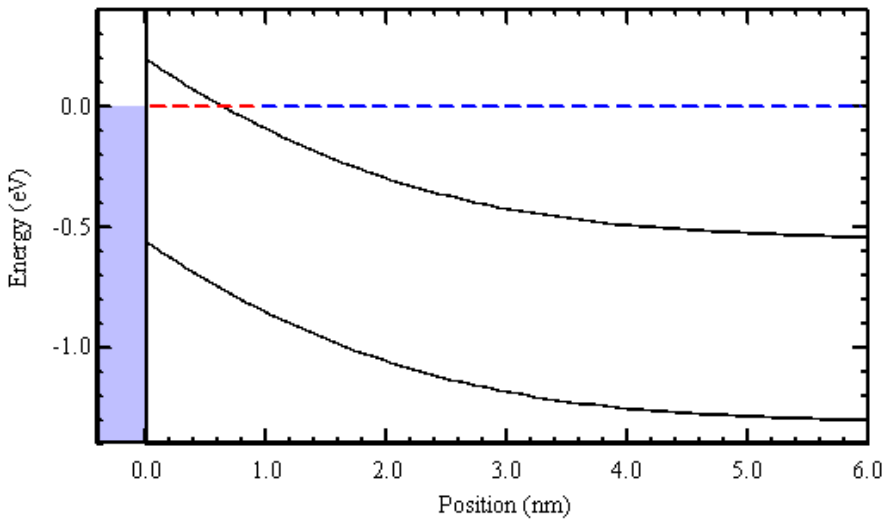
InGaAs effective state density Ns~$4*10^{17}$/cm^3,
Surely it is not possible to dope higher than that ? ;-)

Need to unpin the contact Fermi level ???

N-InGaAs seems to have a ~0.2 eV Schottky barrier

~0.6nm depletion depth

~one lattice constant



N-InAs, or course, has a *negative* ~0.2 eV Schottky barrier

with no barrier, is contact resistivity zero ??? No→ Landauer !

$$\rho_c = \left(\frac{\hbar}{q^2}\right)\cdot\left(\frac{8\pi}{3}\right)^{2/3}\cdot\frac{1}{\|T\|^2}\cdot\frac{1}{n^{2/3}}$$

$n = \text{carrier concentration},\ T = \text{transmission coefficient}$

Wavefunction reflects due to mass, energy change → |T|<1

(over)simplified theory, heavily-doped InAs : $|T|^2$ ~0.3,

Experimental: $|T|^2$ ~0.1

TLM Resistance, at zero spacing, is the contacts ?

That's how we all learned to characterize contacts.

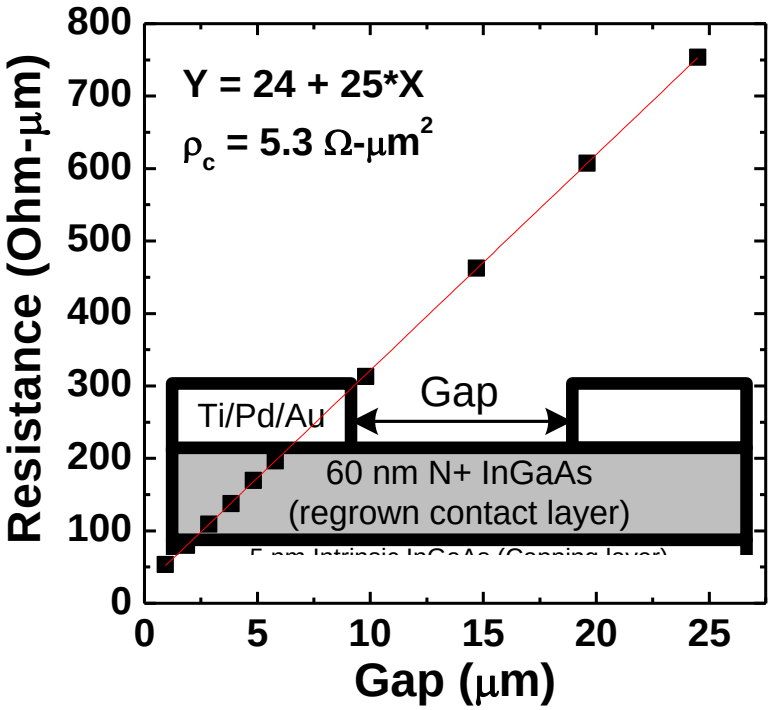
But the zero-gap resistance also contains a Landauer term:

$$R_{\text{Landauer}} = \left(\frac{\hbar}{q^2}\right)\cdot\left(\frac{8\pi}{3}\right)^{2/3}\cdot\frac{1}{HW}\cdot\frac{1}{n^{2/3}}$$

$n = \text{carrier concentration},$

$H = \text{epi-layer thickness}$

$W = \text{TLM width}$



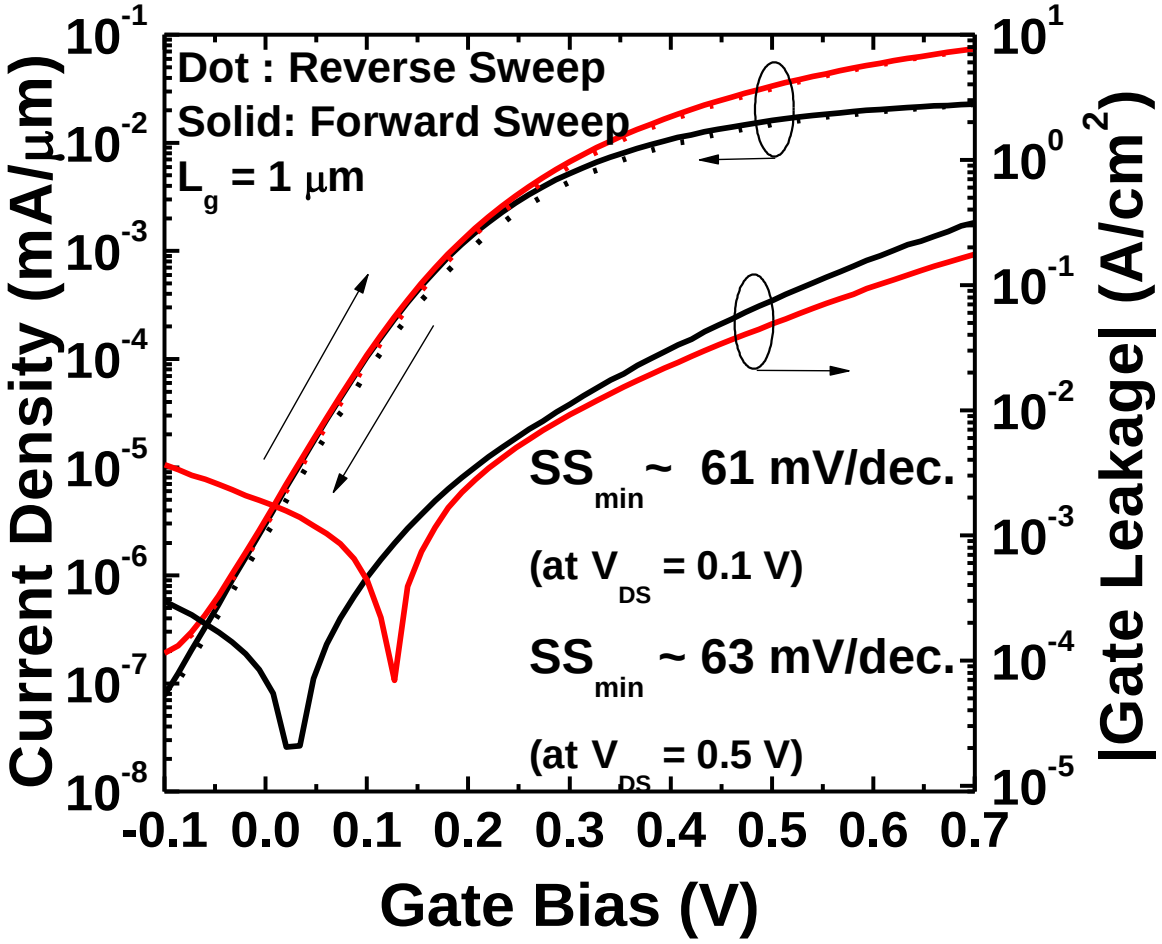
Lee et al, 2015 VLSI symposium

Correction can be significant
Contacts are better than we think
UCSB's published N-InAs contact data is pessimistic

"III-V dielectrics are still very poor"

Are III-V dielectrics still very poor ?

Maybe not perfect. But perhaps better than one might think.
Here are Susanne Stemmer's dielectrics in our FETs.



Lee et al, 2015 VLSI symposium

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

BJT Myths

"...bipolar transistors
are current controlled,

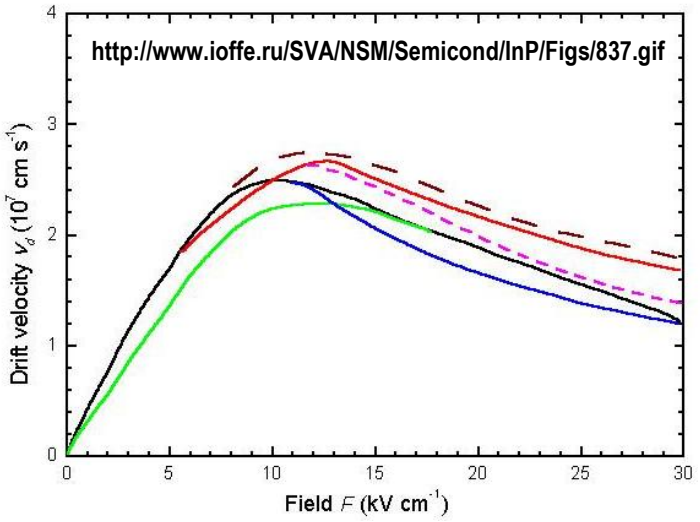
FETs are voltage-controlled"

Given some finite, nonzero input impedance,
there's a 1-1 relationship between voltage & current...

"InP is poor for power devices:
*v_{sat} is only ~1.5*10^7 cm/s"*

Is the velocity low in InP ??

This is the bulk velocity-field curve

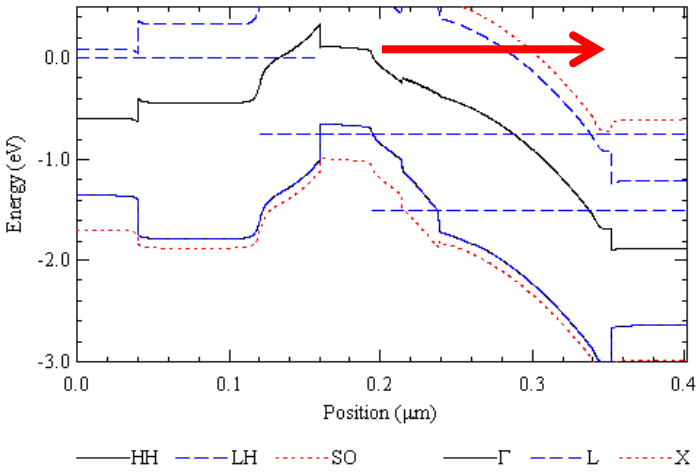


But, in InP collector, no Γ-L scattering until band energies allow it.

Typical velocities are ~3E7 cm/s.

"velocity overshoot"

Drops at higher voltages...

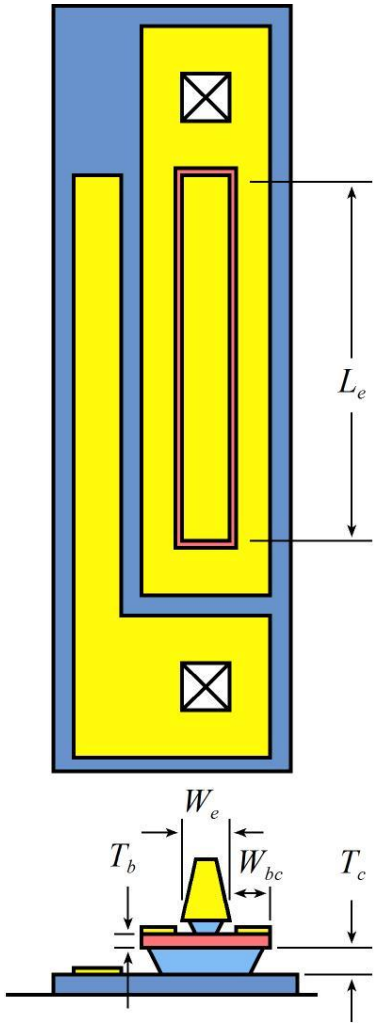


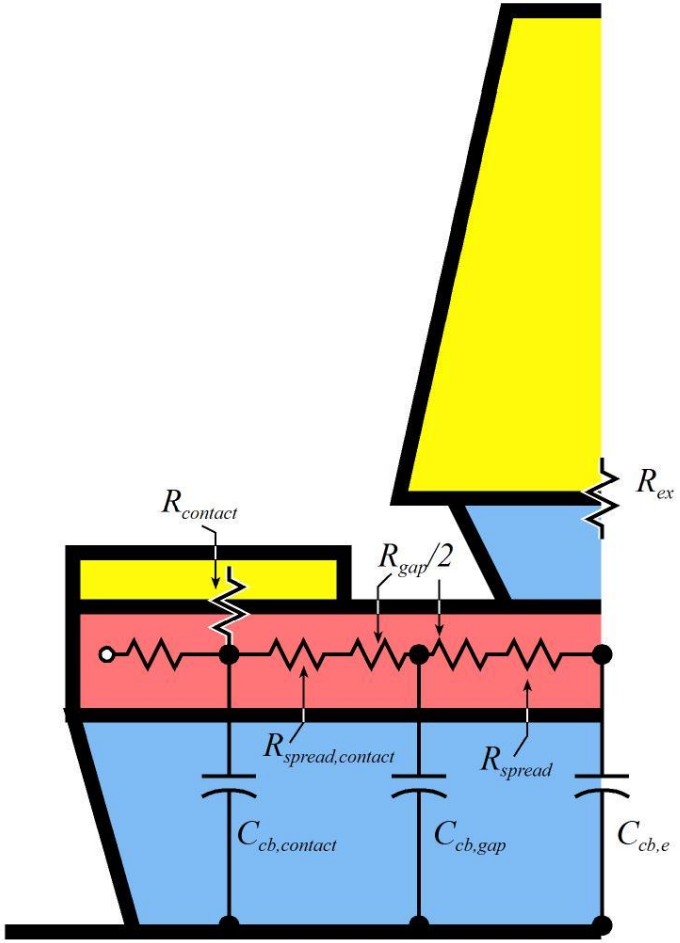
R_{bb} is not "base spreading resistance"

It's mostly from the contacts

..and a bit from the gap

..and quite a lot from the metal !





Are base-emitter heterojunctions important ?

InGaAs emitter & base

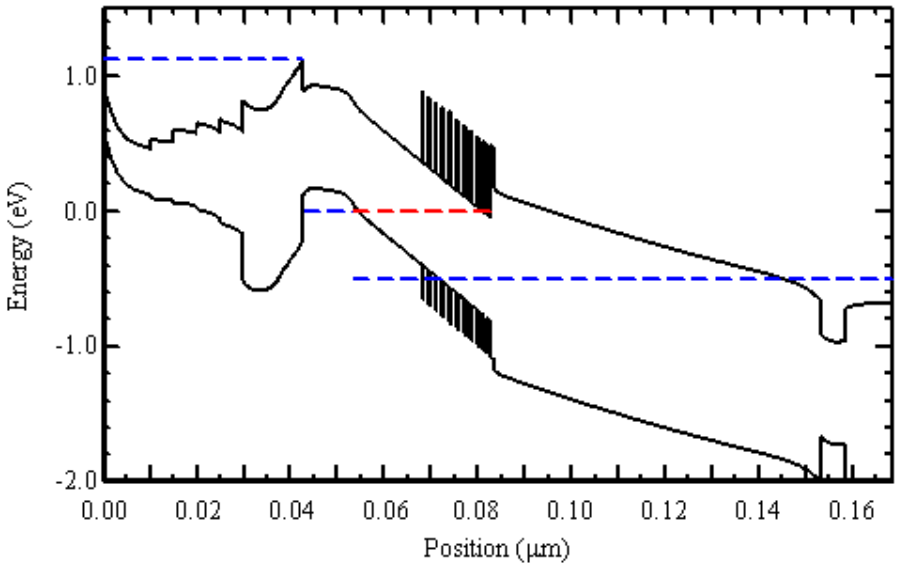
no EB heterojunction
large electron degeneracy
don't need heterojunction !

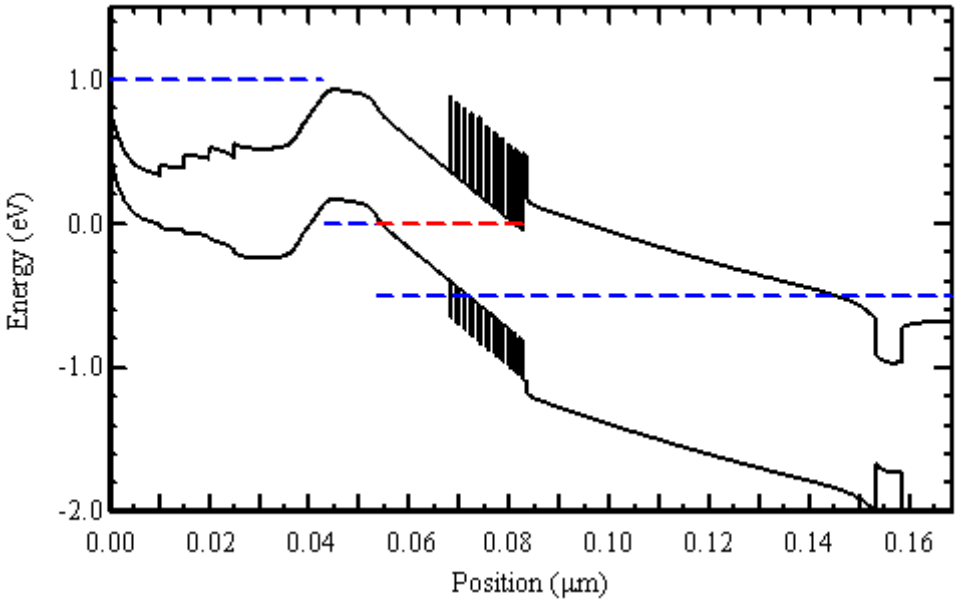
History

Woodall pointed it out.
Ritter did the experiment.
Verified !

Why do we keep the InP ?

InP/InGaAs selective etch
precision placement
of base contact.





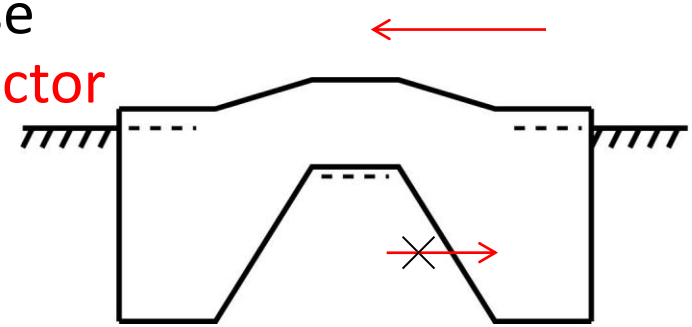
Are DHBTs slow when saturated ?

Classic bipolar transistor saturation:

moderate minority carrier storage in base

large minority carrier storage in subcollector

subcollector stored charge dominates



DHBTs don't store holes in the subcollector:

base-collector junction blocks hole injection into collector

much less saturation charge.

Circuit implications:

base-collector diodes are Schottky like

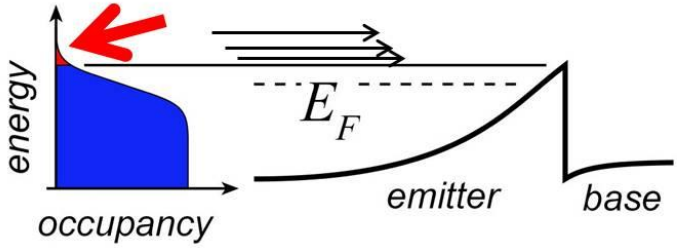
Daneshgar 2014: use in fast sample-hold gates

CMOS-like saturating-HBT logic

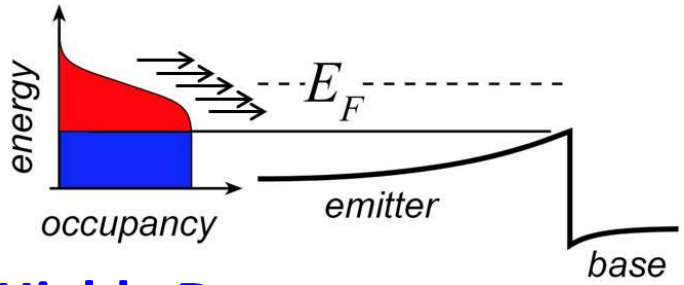
Taur et al 2015: proposed as CMOS logic replacement

HBTs have exponential I-V characteristics ????

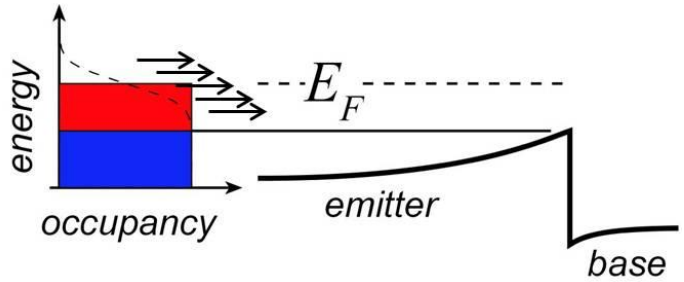
Boltzmann

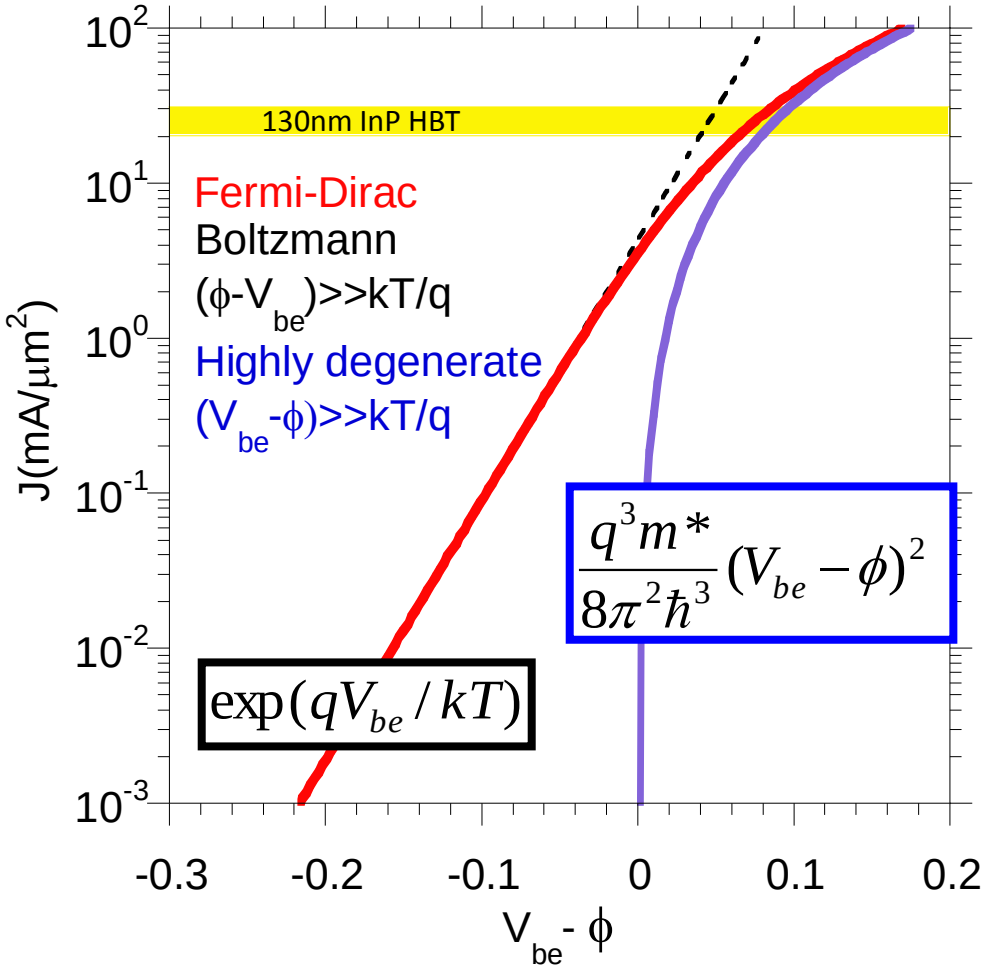


Fermi-Dirac



Highly Degenerate





drops g_m → hurts bandwidth

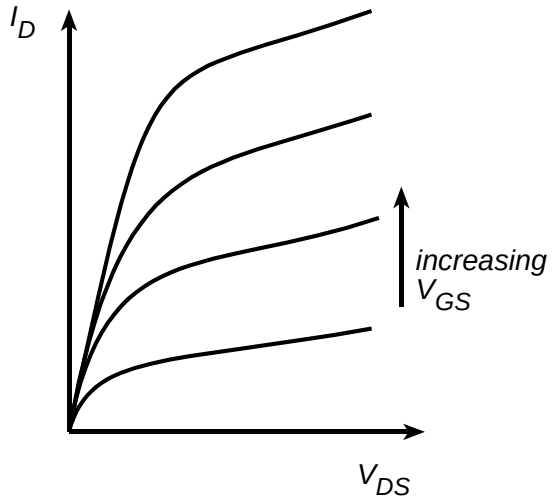
FET Myths

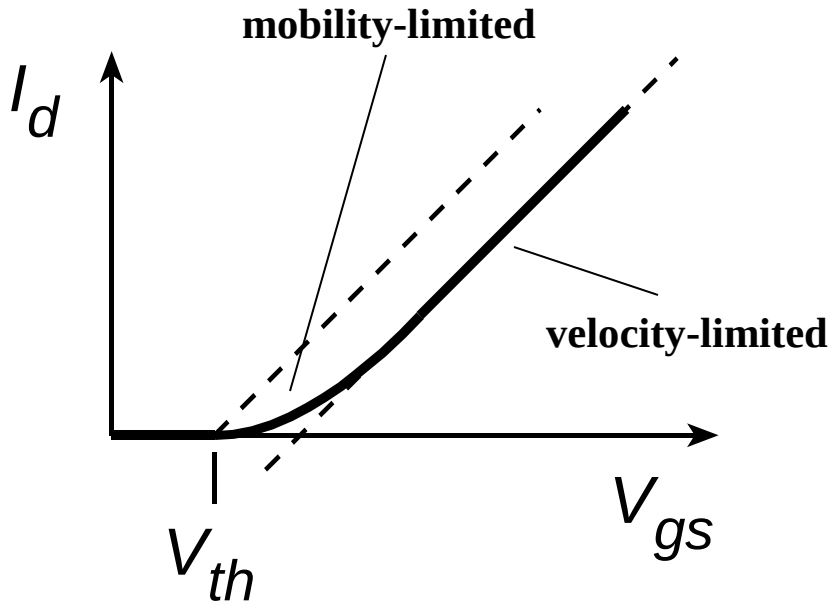
"Long-channel FETs
are limited by mobility...

...short-channel FETs
are limited by
saturation drift velocity".

Mobility/saturation velocity model ???

My undergraduate class notes





For drain voltages larger than the knee voltage :

mobility – limited current

$$I_{D,\mu} = \mu c_{ox} W_g (V_{gs} - V_{th})^2 / 2L_g$$

velocity – limited current

$$I_{D,v} = c_{ox} W_g v_{sat} (V_{gs} - V_{th})$$

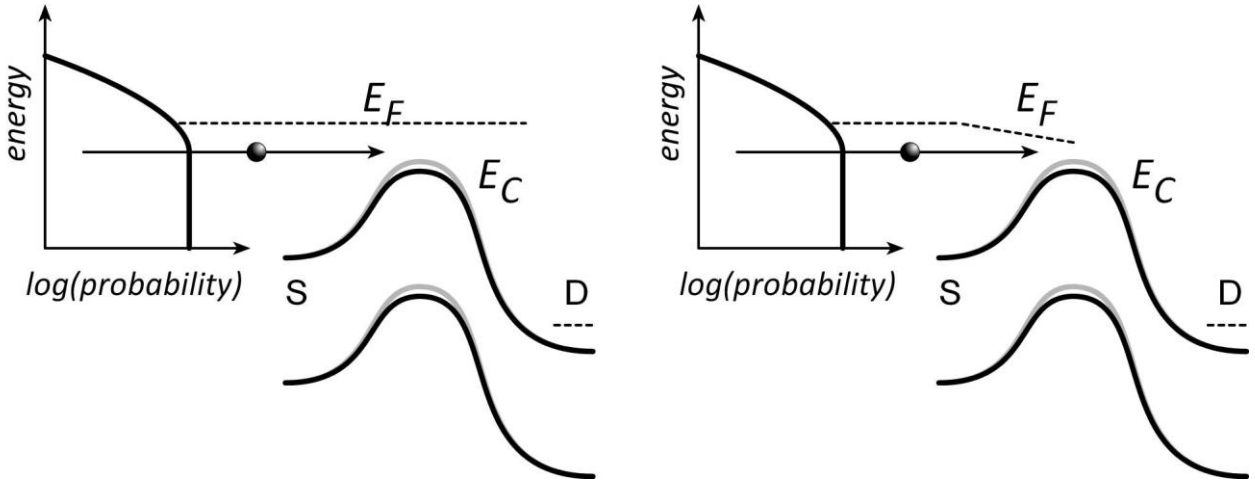
Generalized Expression

$$\left(\frac{I_D}{I_{D,v}}\right)^2 + \left(\frac{I_D}{I_{D,\mu}}\right) = 1$$

This is not correct !

Short-channel: Ballistic top-of-barrier model

Natori, Lundstrom, Antoniadis



If zero scattering between source and barrier:

velocity set by $mv^2 = E_{electron} - E_{c,barrier}$.

current= #states above barrier times velocity of each

scattering in drain region: reduces f_τ, doesn't reduce current

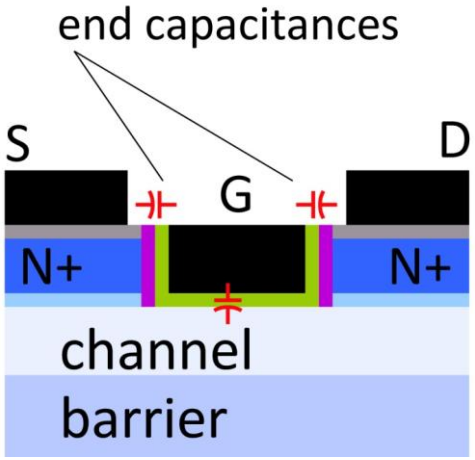
Scattering near source:

drops E_f near barrier

reduces current

Are FET f_τ's set by transit times ?

$$1/2\pi f_\tau = L_g/\bar{v} + (\text{end capacitances})/g_m + \ldots$$



Everyone knows this.

But, the significance is not always noted.

End capacitances

about 0.3 fF/μm total in HEMTs

about 1.0 fF/μm total in CMOS VLSI.

Impact on f_τ (assuming zero gate length):

g_m=2mS/μm, 0.3 fF/μm→f_τ =1.1 THz (InP HEMT)

g_m=1mS/μm, 0.3 fF/μm→f_τ =550 GHz (GaN HEMT)

g_m=2mS/μm, 1.0 fF/μm→f_τ =320 GHz (CMOS)

Yet, g_m is also hard to increase (gate dielectric scaling)

C_{end}/g_m time constant is major bandwidth limit.

But the roadmap says VLSI will give 1 THz f_τ...

They have fixed it now, but...

...one recent RF ITRS roadmap predicted:

f_τ increases as 1/(technology note)

Physics-free prediction

1) gate length no longer proportional to technology node

2) $1/2\pi f_\tau = L_g/v + C_{ends}/g_m + \ldots$ ...some terms no longer scale.

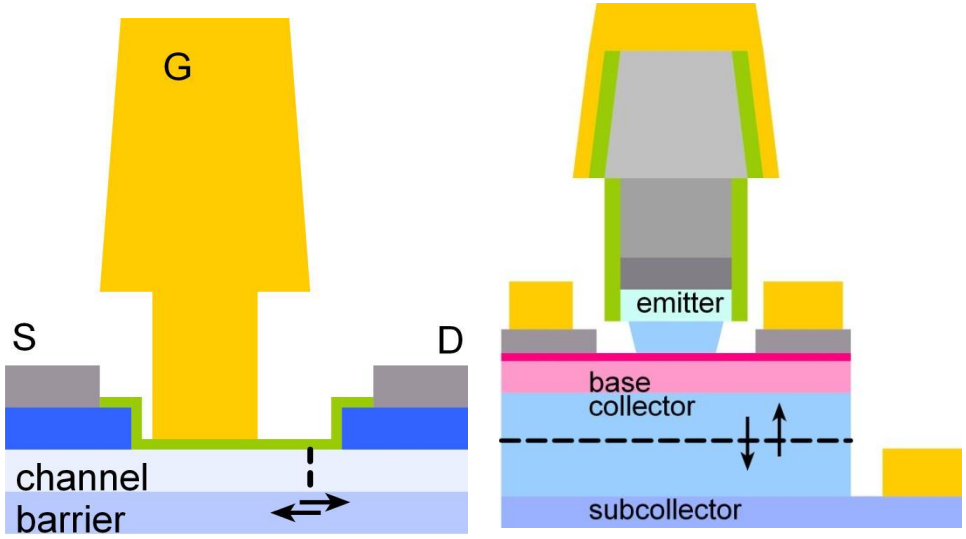
Embellishment, fantabulism, and balderdash.

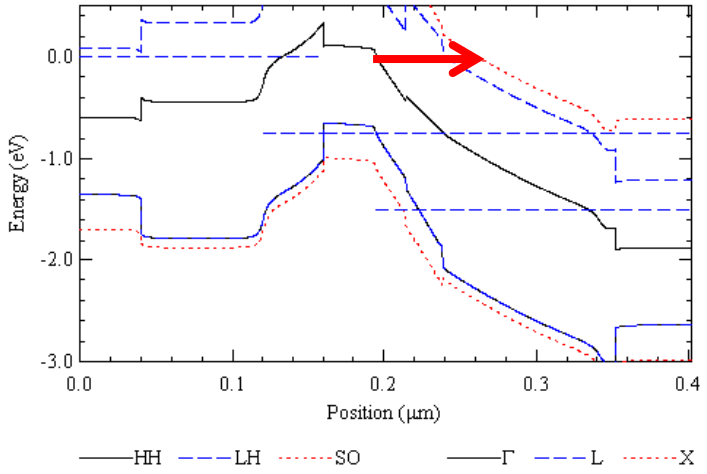
Power Transistor Gamesmanship

Quote bandwidth at a low voltage, breakdown at a high one.

For InP HBTs, f_τ drops with increased V_{ce}:
Movement of depletion edge
decrease in distance before Γ-L scattering..

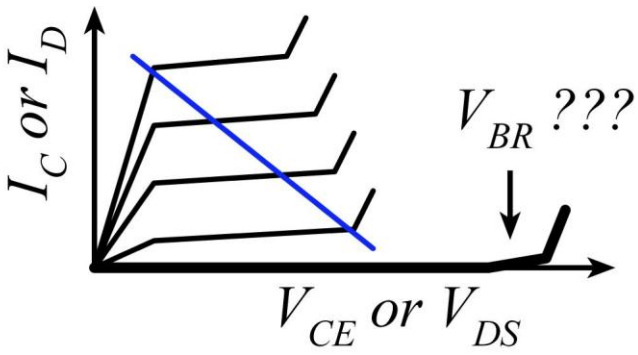
For HEMTs, f_τ drops with increased V_{ce}:
Lateral movement of gate-drain depletion edge



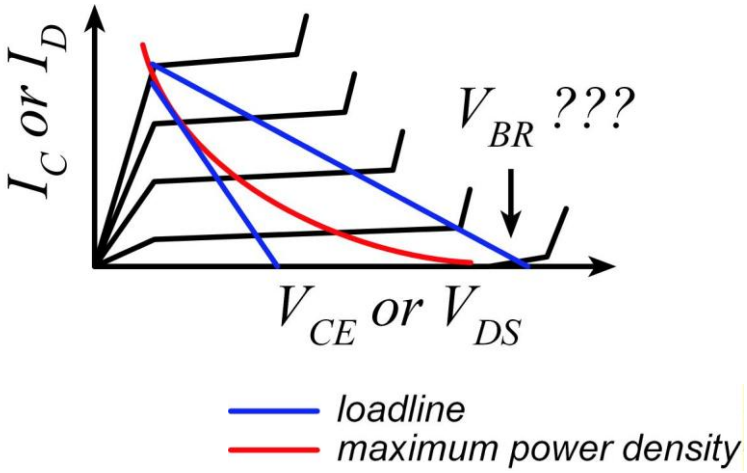


Gamesmanship with breakdown

Is this breakdown useful ?

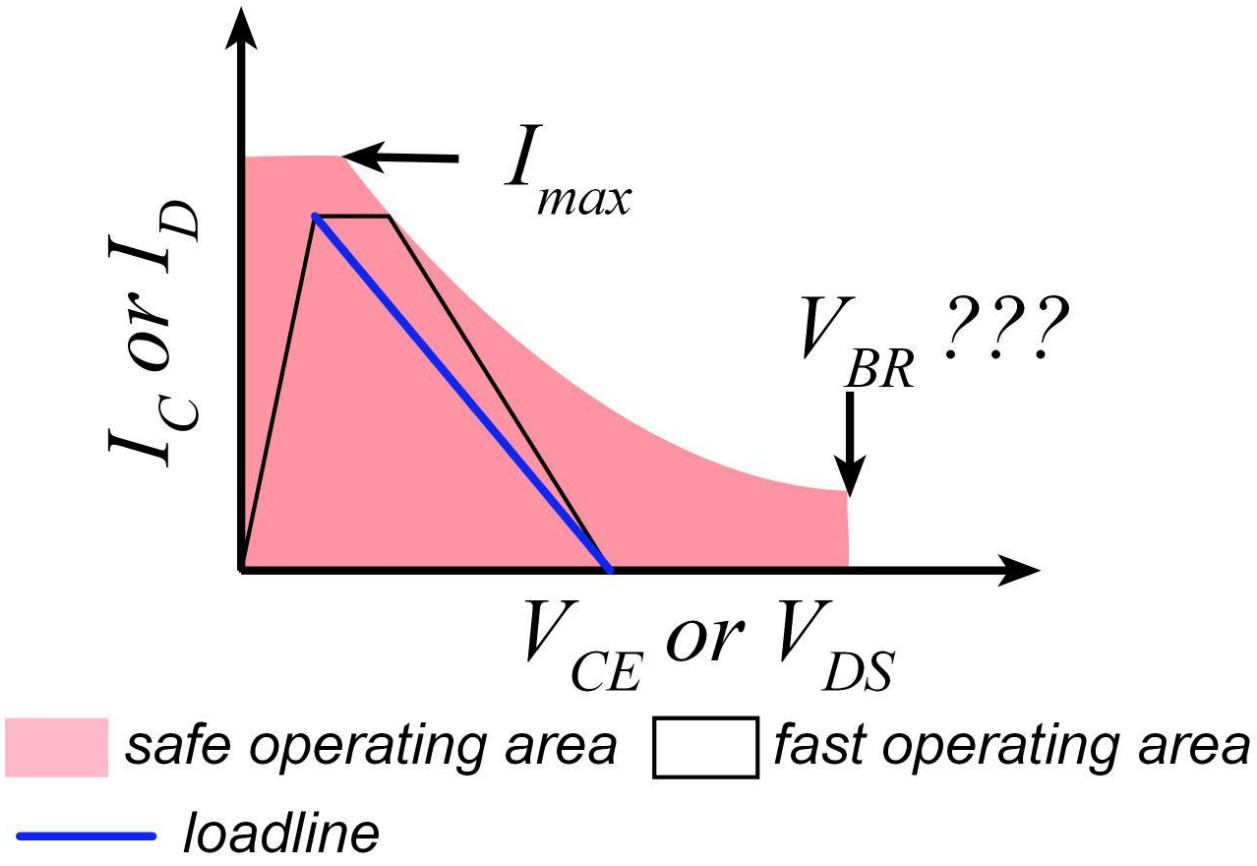


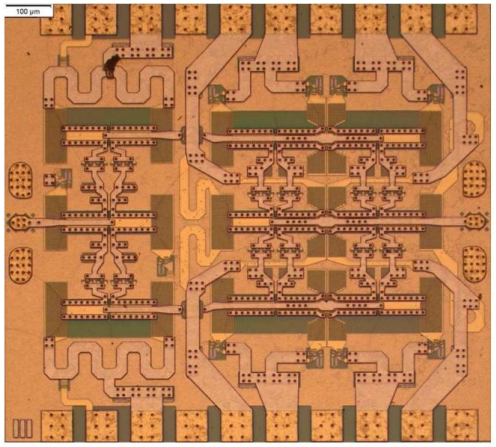
How about this ?

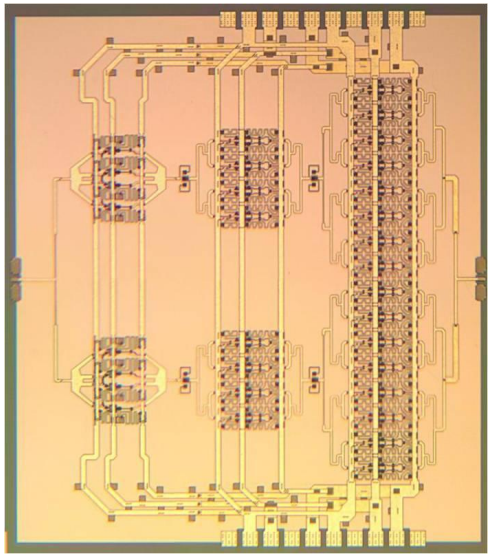


In designing PAs, such games would kill the IC.

This is how we specify our HBTs internally…

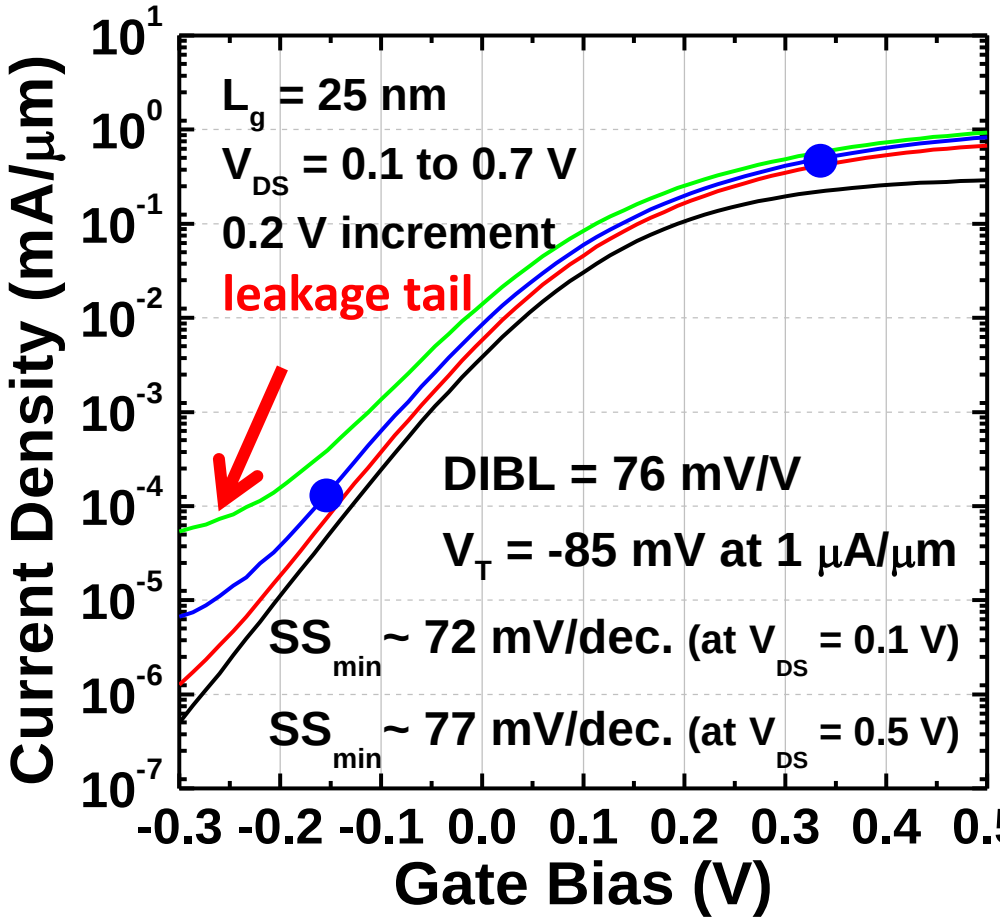






VLSI specsmanship

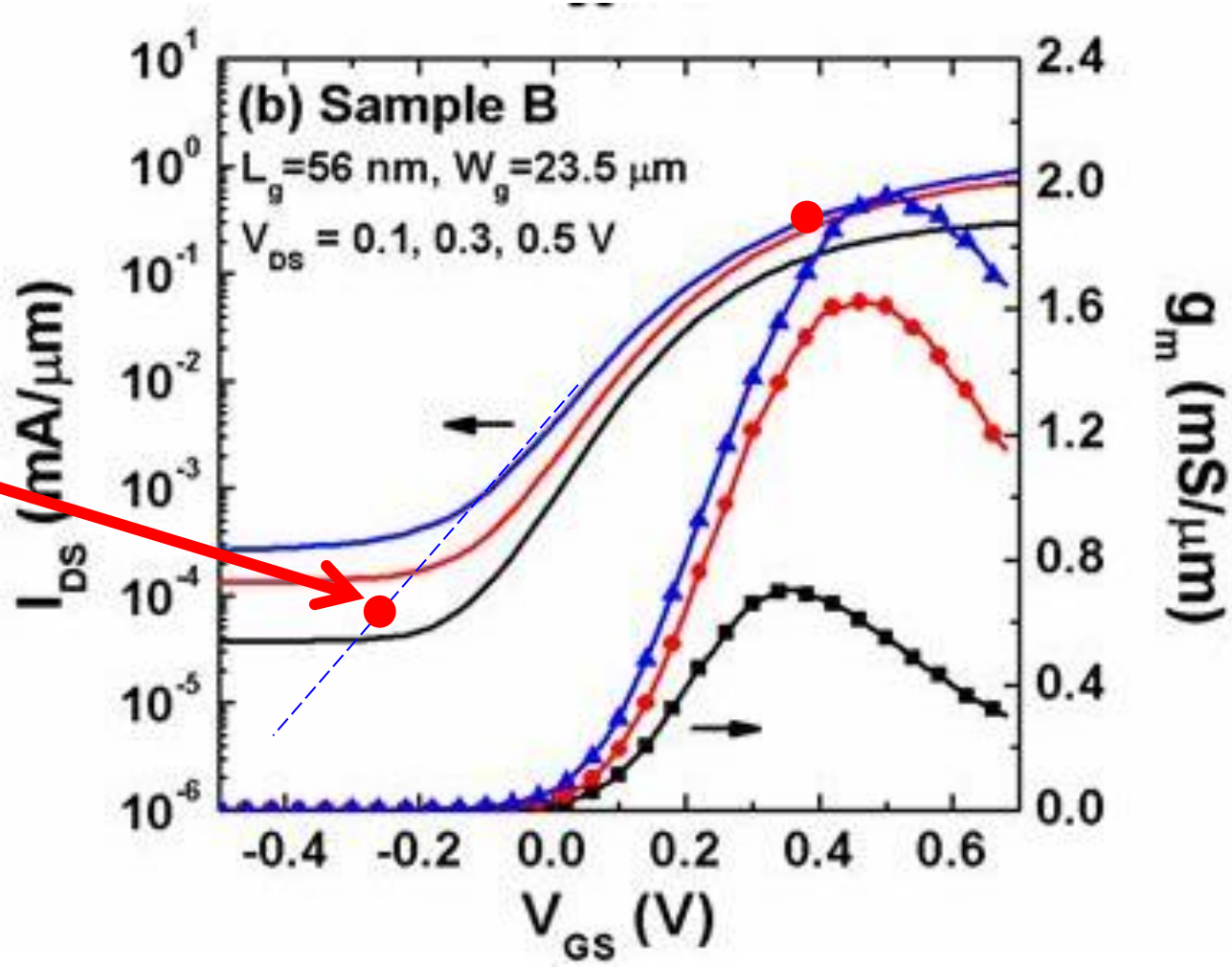
It's nice to have high g_m, low SS, but (I_{on}, I_{off}, V_{DD}) is better
you can have low SS, but a bad leakage tail will increase I_{off}.



VLSI specmanship

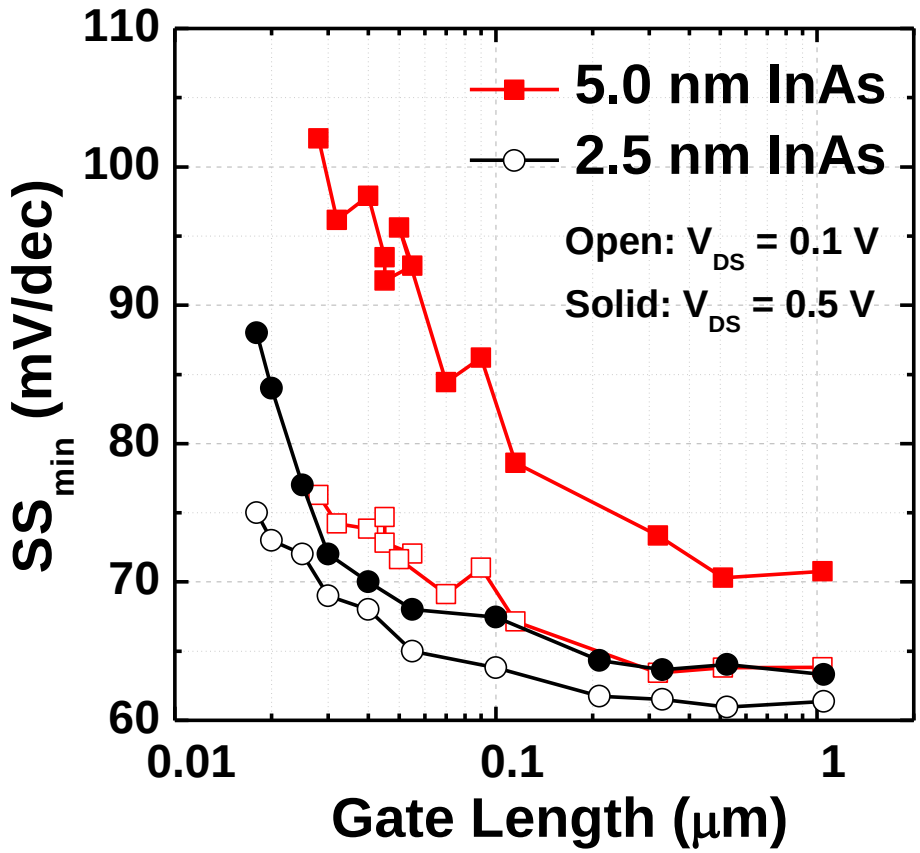
A few recent papers have even done this...

stated off-current using extrapolation



VLSI specmanship

Or quote SS, gm, DIBL etc on a long-channel device
Nice to look at...but the VLSI IC will use short-channel FETs...



Transistor benchmarks for circuits

Gain

How much gain can we get from a transistor ?

A CMOS 210-GHz Fundamental Transceiver With OOK Modulation

Zheng Wang, *Student Member, IEEE*, Pei-Yuan Chiang, *Student Member, IEEE*, Peyman Nazari, *Student Member, IEEE*, Chun-Cheng Wang, *Member, IEEE*, Zhiming Chen, *Member, IEEE*, and Payam Heydari, *Senior Member, IEEE*

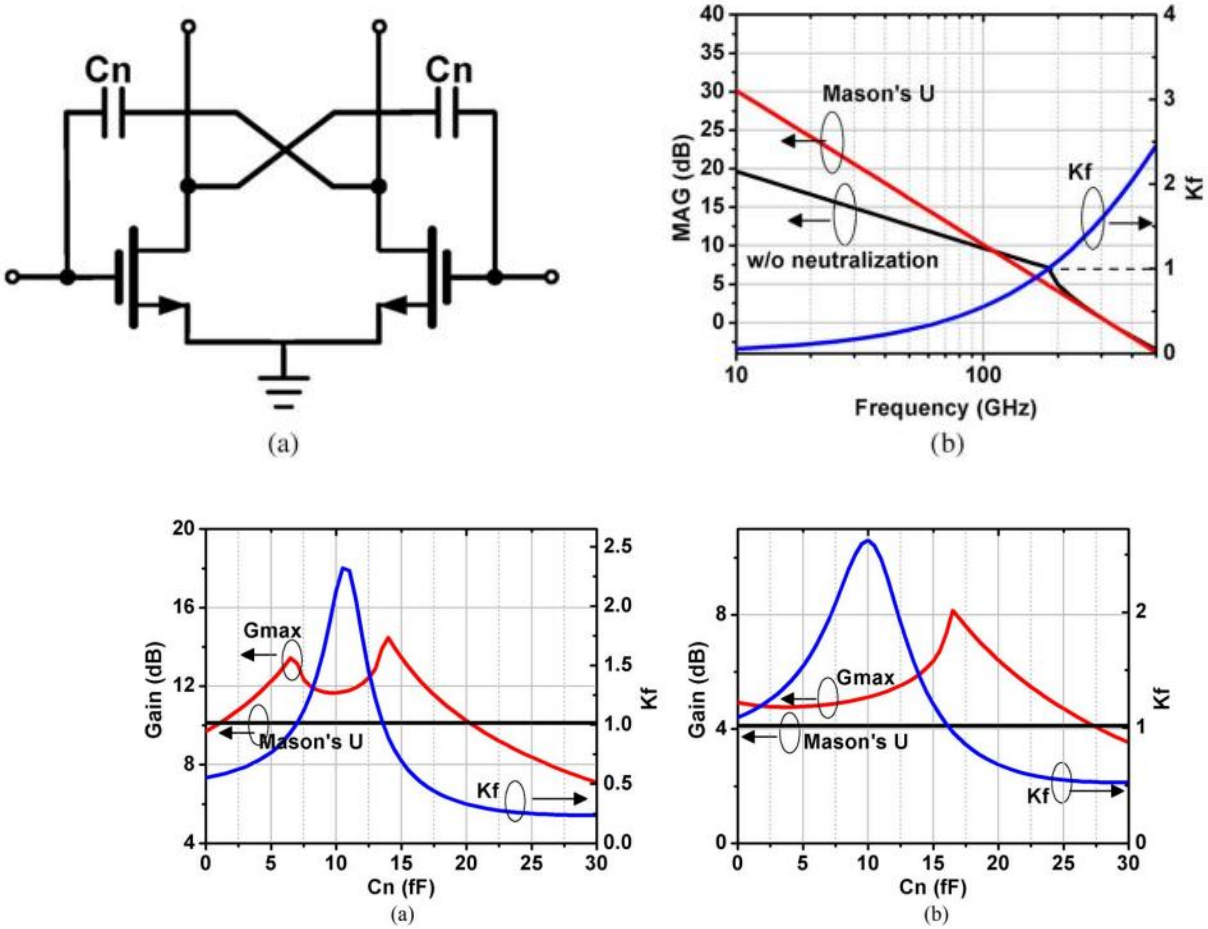
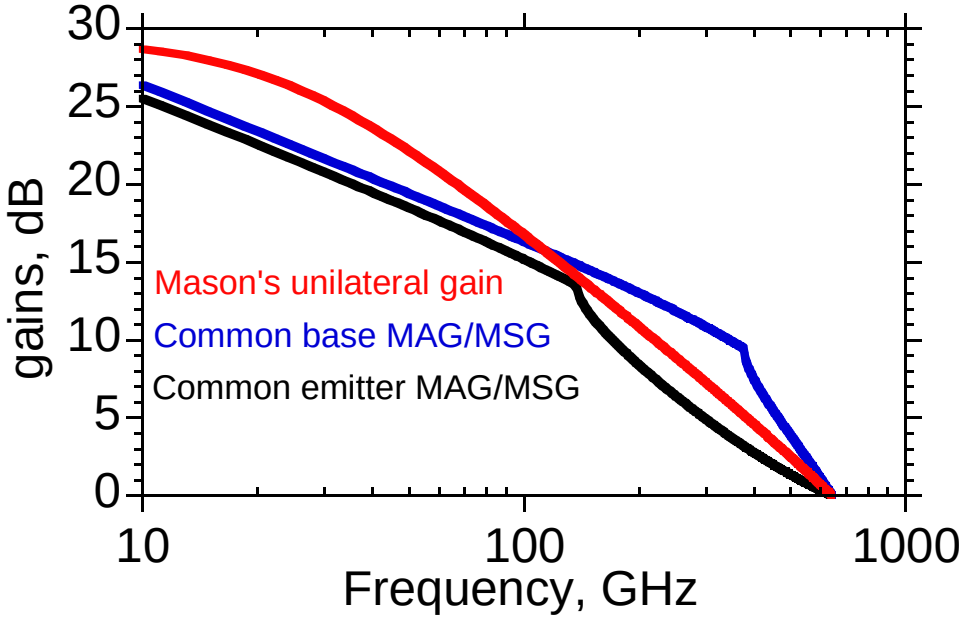


Fig. 11. Maximum power gain with respect to neutralization capacitance at (a) 100 GHz and (b) 200 GHz.

How much gain can we get from a transistor ?



MSG: maximum obtainable gain, with no added feedback,
If we add sufficient stabilizing resistance to ensure that no change in the generator or load can cause oscillation

U: gain if we unilateralize (add lossless reciprocal feedback) and then match.

Note that the common-base MSG is larger than U !

How much gain can we get from a transistor ?

The common-base MSG is larger than U.

What does this tell us ?

Adding feedback, and then re-stabilizing can increase gain.

Even if we ensure unconditional stability.

Is there any upper bound to the gain so obtained ?

No !

Mason showed that U is

(1) an invariant (with respect to lossless reciprocal embedding)

(2) the only invariant calculable from the network parameters.

→ The only limit to such design tricks is component tolerance.

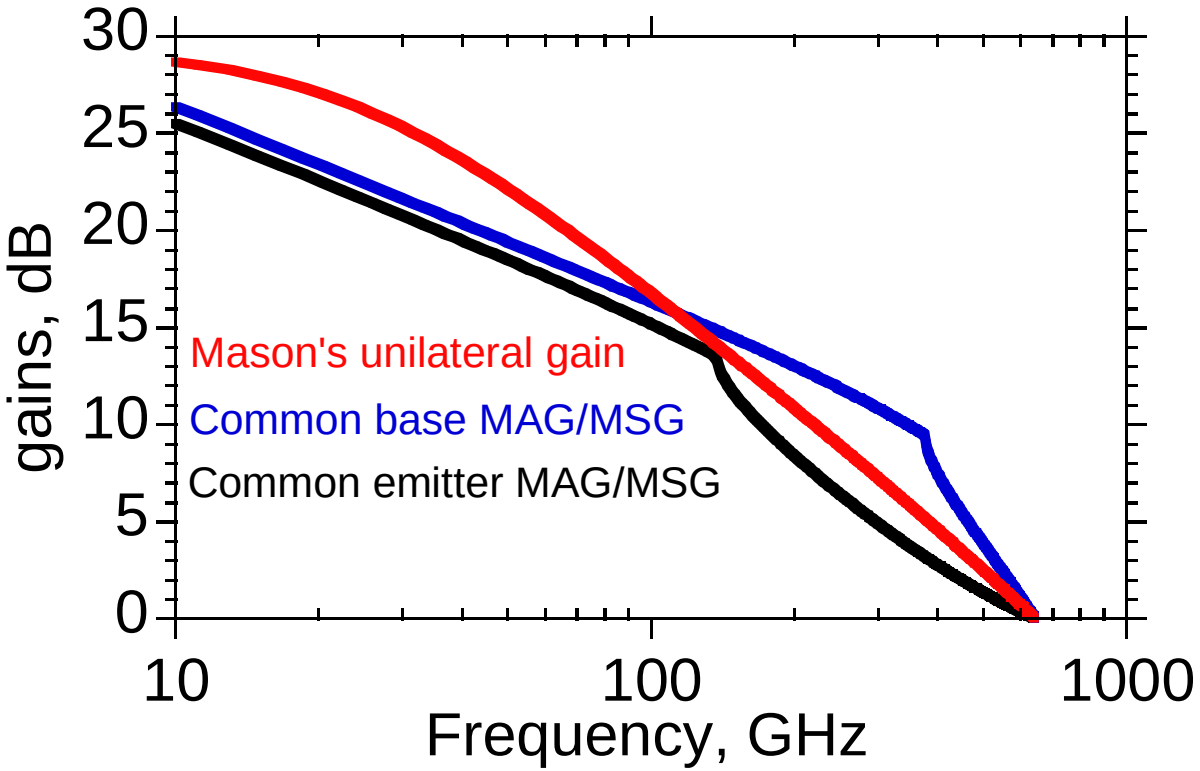
Will such techniques be widely used ?

Radios need LNAs and PAs

These have quite different design constraints

There are few applications for RF "A's" (i.e., not LNA, not PA)

Common-base stages are less stable ????



MSG is the maximum obtainable gain

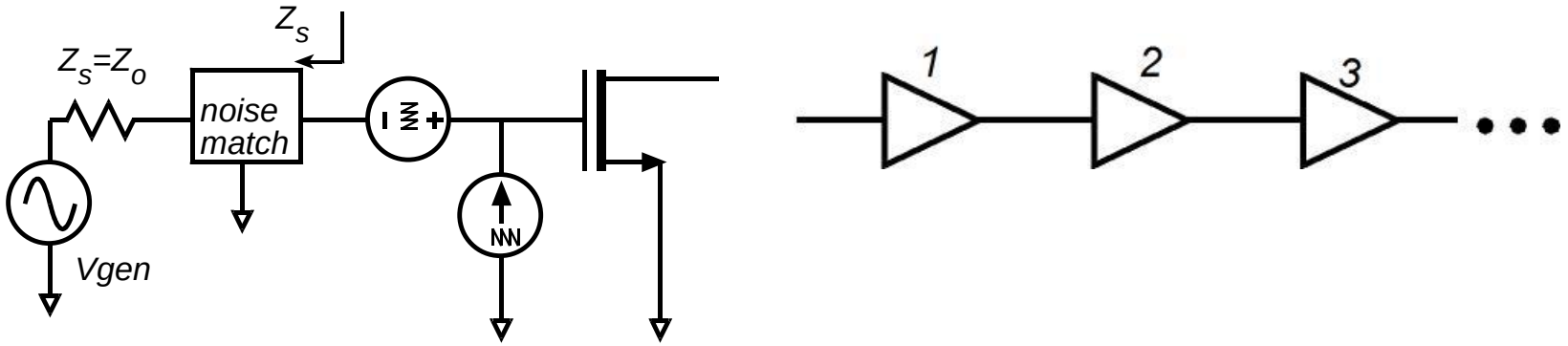
If you don't add feedback

and if you ensure that termination changes won't cause oscillation

...so, in what sense is the CB stage less stable ?

Noise

We have a clean measure of noise performance



Define F_∞ as the noise figure of an infinite cascade

$$F_\infty = F + \frac{F-1}{G_A} + \frac{F-1}{G_A^2} + \frac{F-1}{G_A^3} + \ldots$$

Define M, the noise measure, as $M = F_\infty - 1$

Mason proves that M is invariant w.r.t lossless reciprocal embedding.

Note that $F_{\min}$ is not such an invariant.

So, the best low-noise transistor is the one with the smallest M.

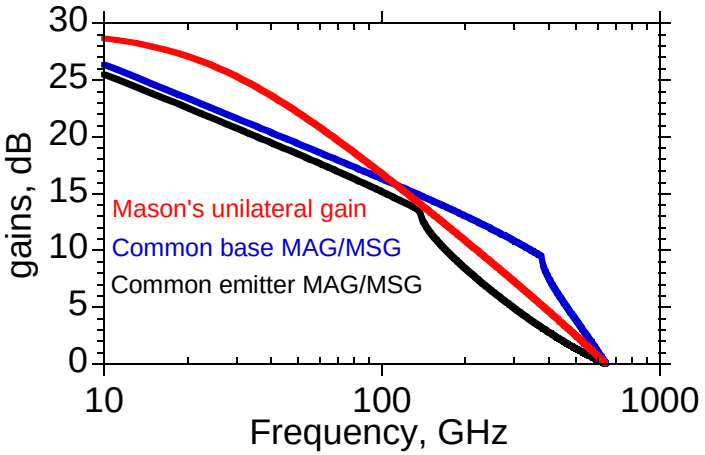
Power

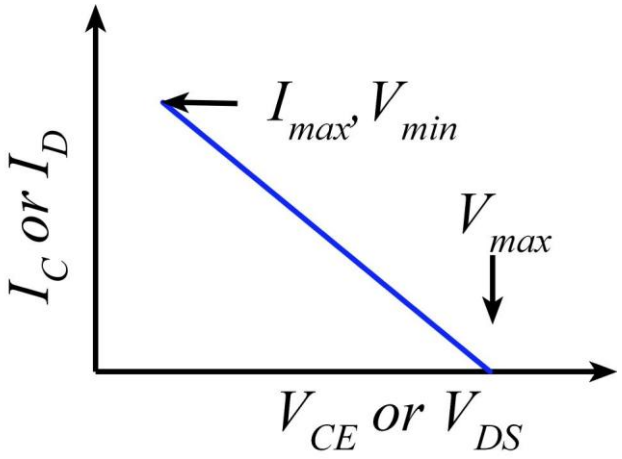
Does f_{max} determine PA gain ? No !

Load is different for gain and for power

MAG/MAG/U....obtained with load giving optimum gain

load for maximum power is $R_{Lopt} = (V_{max}-V_{min})/I_{max}$, plus parallel L





Power gain with the optimum load match

Can calculate this from transistor model (or S_{ij}) and $R_{L,opt}$.

Load-pull measures this gain.

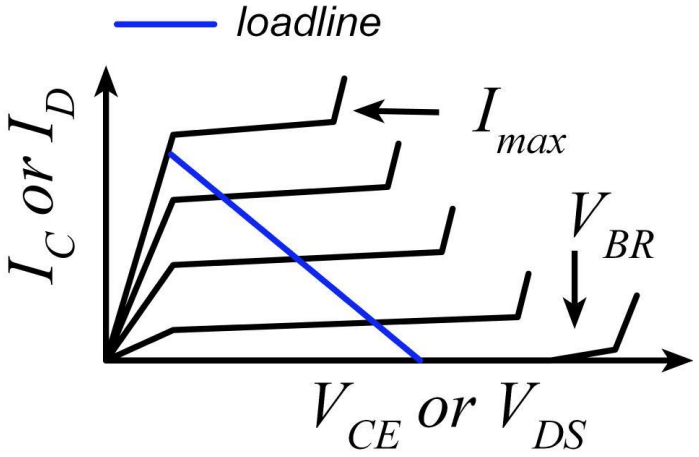
Are super-high breakdown voltages useful ?

Suppose we have transistor with

(1) $J_{\max} = 1\,\text{mA}/\mu\text{m}$ (2) $V_{br} = 100\,\text{V}$.

Is this useful for a 100GHz PA ?

Note: $R_L > R_{\max} \approx 100\Omega$ can't be realized



Design A: Pick $W_g = 100\,\mu\text{m}$:

$I_{\max} = 100\text{mA}, V_{br} = 100\,\text{V} \rightarrow R_{L,opt} = 1\text{k}\Omega \rightarrow$ can't match

Design B: pick $W_g = 1\text{mm} \rightarrow$ huge device

$I_{\max} = 1\text{A}, V_{br} = 100\,\text{V} \rightarrow R_{L,opt} = 100\Omega$

Are super-high breakdown voltages useful ?

Design B: pick $W_g = 1\text{mm} \rightarrow$ huge device

$I_{\max} = 1\text{A}, V_{br} = 100\text{ V} \rightarrow R_{L,opt} = 100\Omega$

Suppose $W_{finger} = 10\mu\text{m}, D_{gg} = 10\mu\text{m}$

$\rightarrow$ need 100 fingers @10μm spacing

$\rightarrow$ Wide FET cell : $D_{cell} = 1\text{mm}$

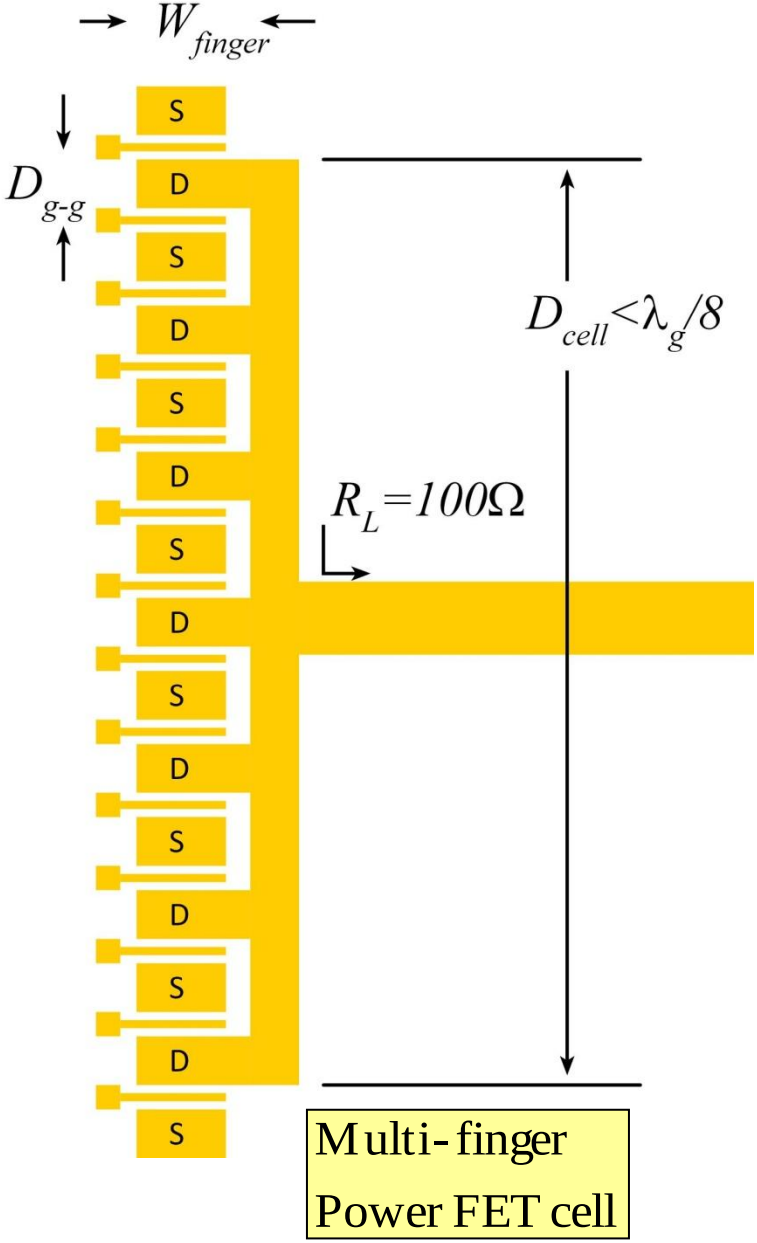
But $\lambda_g \approx \lambda_0 / \varepsilon_r^{1/2} = 1.0\text{mm} @ 94\text{GHz}$

$\lambda_g / 8 \approx 126\mu\text{m}$

Our FET cell is ~ 1 wavelength wide.

$\rightarrow$ Huge impedance change in feed

Won't work!



Maximum useful breakdown voltage

If $D_{cell} > \lambda_g / 8$, then R_L will drop.

Maximum current per cell :

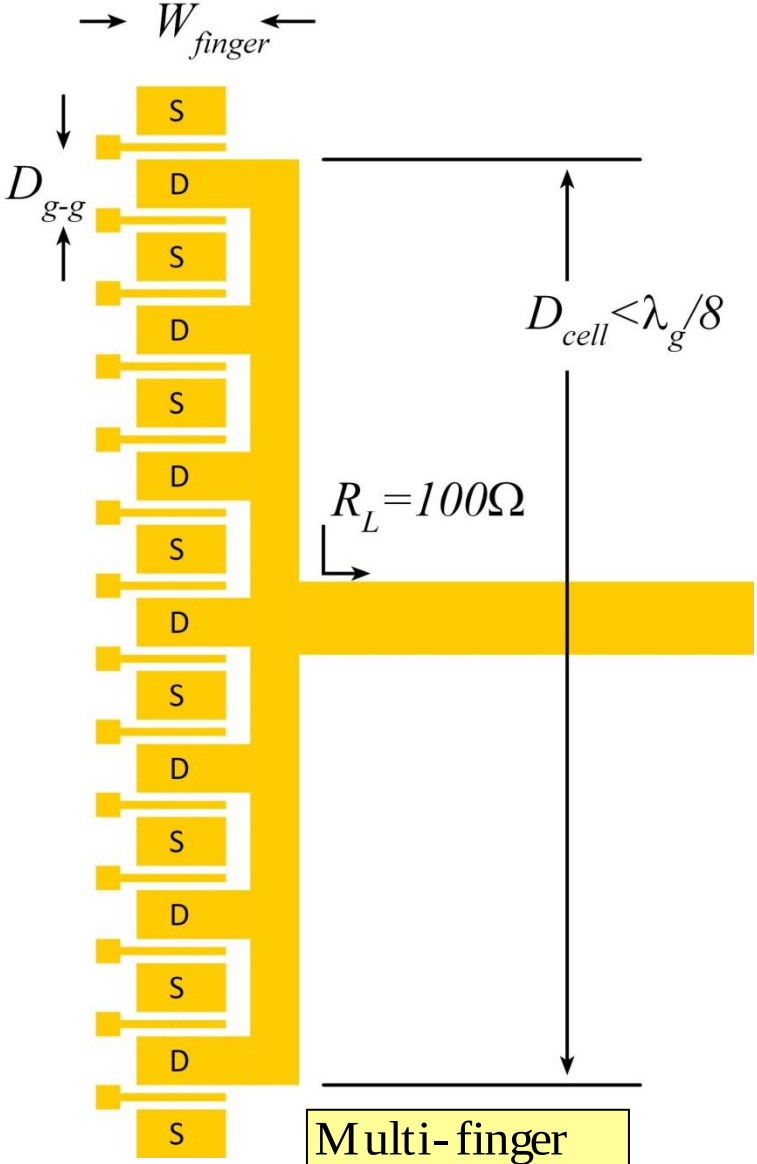
$$I_{\max} = \frac{W_{finger}}{D_{gg}} \left(J_{\max} \cdot \frac{\lambda_0}{8\varepsilon_{r,eff}^{1/2}} \right)$$

Maximum useful voltage

$$V_{\max,useful} - V_{\min} = I_{\max} R_{L,\max}$$

$$= \frac{W_{finger}}{D_{gg}} \left(J_{\max} \cdot \frac{\lambda_0}{8\varepsilon_{r,eff}^{1/2}} \right) R_{L,\max}$$

$$= 12.6 \text{Volts}$$



Multi-finger Power FET cell

Maximum power per cell

If $D_{cell} > \lambda_g / 8$, then R_L will drop.

Maximum current per cell :

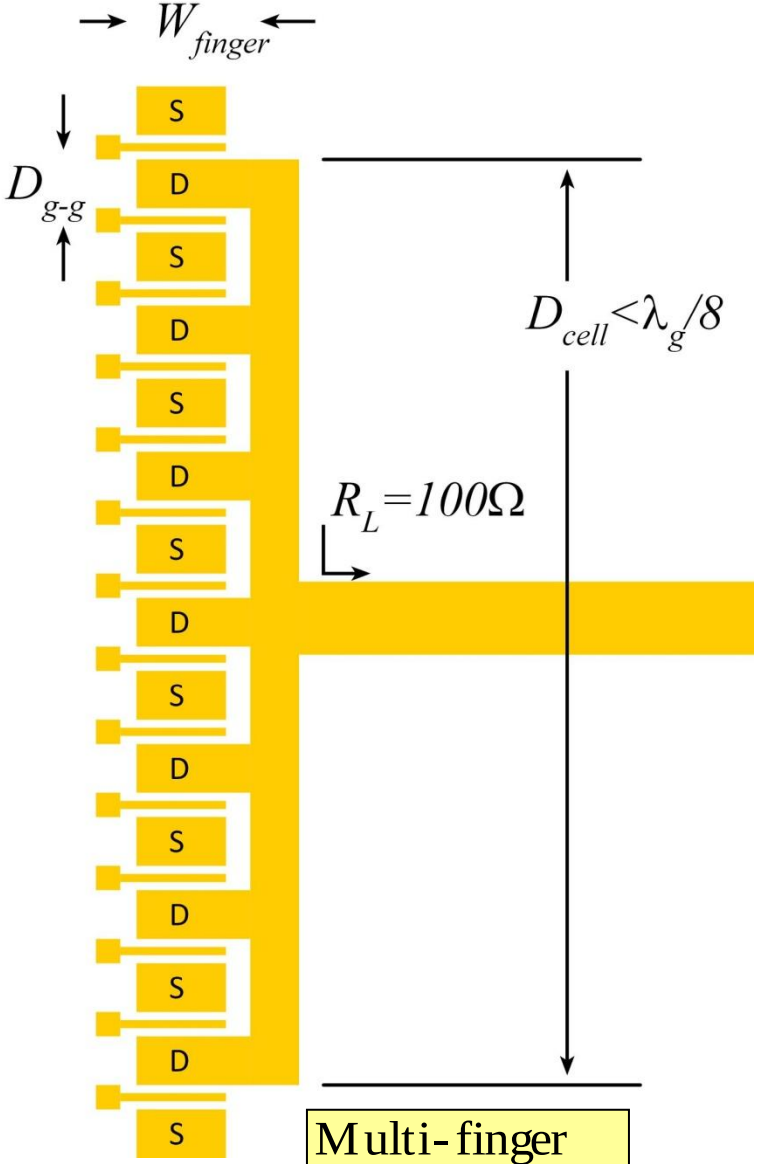
$$I_{\max} = \frac{W_{finger}}{D_{gg}} \left(J_{\max} \cdot \frac{\lambda_0}{8\varepsilon_{r,eff}^{1/2}} \right)$$

Maximum power per cell

$$P_{\max,cell} = I_{\max}^2 R_{L,\max} / 8$$

$$= \left(\frac{W_{finger}}{D_{gg}} J_{\max} \cdot \frac{\lambda_0}{8\varepsilon_{r,eff}^{1/2}} \right)^2 \frac{R_{L,\max}}{8}$$

$$= 0.2 \text{ Watts}$$



A current-density-limit on mm-wave power

Established breakdown limit on power

(Johnson figure of merit)

$$P_{\max,cell} = V_{\max}^2 / 8R_{L,\min}$$

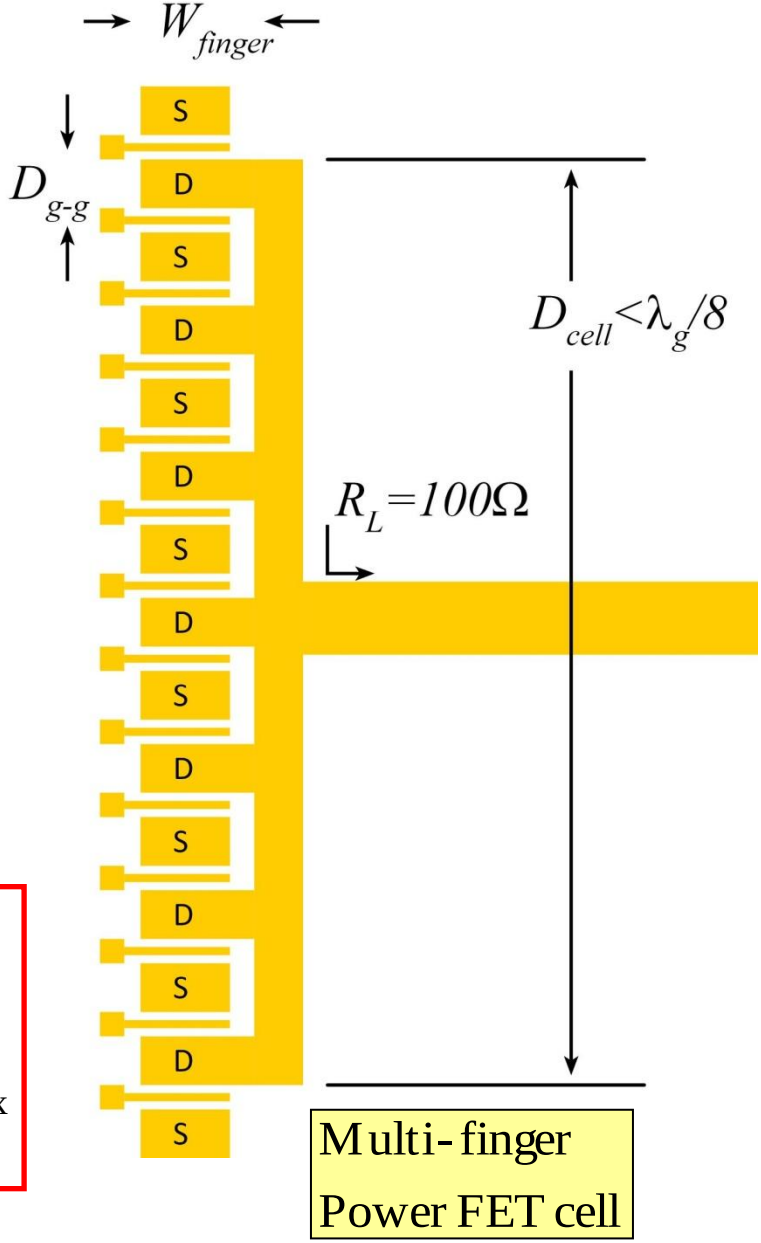
Current limit on power

$$P_{\max,cell} = I_{\max}^2 R_{L,\max} / 8$$

$$= \left(\frac{W_{finger}}{D_{gg}} J_{\max} \cdot \frac{\lambda_0}{8\varepsilon_{r,eff}^{1/2}} \right)^2 \frac{R_{L,\max}}{8}$$

Maximum useful voltage

$$V_{\max,useful} - V_{\min} = \left(\frac{W_{finger}}{D_{gg}} J_{\max} \cdot \frac{\lambda_0}{8\varepsilon_{r,eff}^{1/2}} \right) R_{L,\max}$$



Must-Haves for Electronics

RF/microwave/mmwave devices must have...

For general RF amplification, the answer's unclear

Unilateral gain & f_{max} seem reasonable metrics.

Fortunately, we always want PAs or LNAs

For low-noise RF transistors, the real metric is noise measure.

For high-power RF transistor, the metrics are

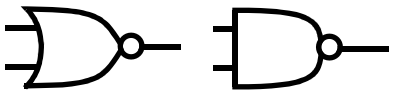
maximum output power, and associated gain, into realizable load between ~10 and 100 Ω.

this involves both the discussed current & voltage metrics

load-pull measures this.

Logic : some desirables

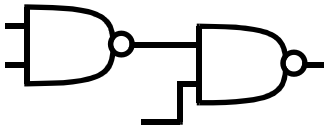
Logic should perform Boolean functions



NAND alone is sufficient, as is NOR

AND and OR are not

Gates should Cascade



input and output in the form of the same physical variable

same values defining "1" and "0"

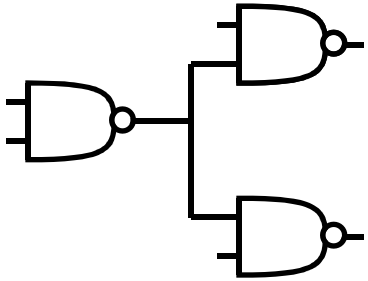
include translation devices in speed/size/power analysis

e.g. , problem if input is DC H-field and output is 50 GHz spin wave amplitude

e.g. , problem if input is DC current and output is DC B-field

e.g. , problem if input is at 2 GHz, and output is at 25 GHz (parametric gain)

Fan-out is probably needed, too



Logic Elements Should Communicate

...over significant distances.

The bigger the gates, the longer the distances.

gates should be very small (Boolean complexity /area)

bigger devices needed to drive longer interconnects--how big ?

Rent's rule: useful chips have many long wires.

Power / Energy / Delay analyses must explicitly include interconnects.

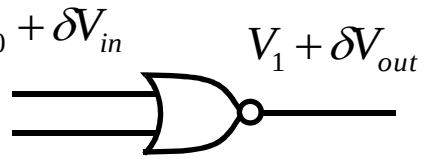
ion / reagent concentration in solution (biology)
wires
gears (adding machines)
optical waveguides

Logic Should Be Robust

Jan Rabaey says "Digital ICs scale, analog ICs don't"

Analog: errors accumulate

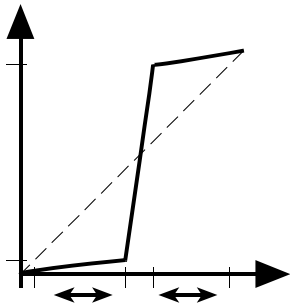
Digital: levels are restored



$dV_{out} / dV_{in} < 1$

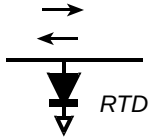
This seems to imply nonlinearities.

This seems to imply gain.



Logic Elements should be (nearly) uni-directional

output → input reverse propagation is a headache



e.g., nondegenerate parametric gain ---bilateral

e.g. clockless tunnel diode logic---bilateral

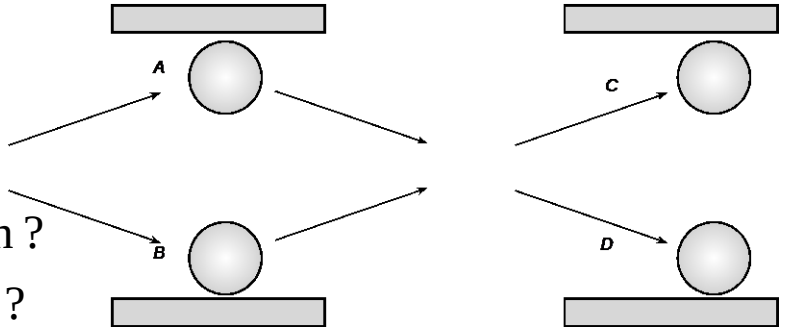
Zero power logic worries me----despite Landauer

zero dissipation → reversible

reversible: $f(\text{time}) \Leftrightarrow f(-1 \cdot \text{time})$

a) do we lose input/output cause / effect distinction ?

b) do small deviations accumulate --exponentially ?



Second Part

Battling to make good electron devices

Mark Rodwell,
University of California, Santa Barbara

With apologies to my co-authors.

C.-Y. Huang, J. Rode, S. Lee, V. Chobpattanna,
P. Choudhary, A.C. Gossard, S. Stemmer
ECE and Materials Departments,
University of California, Santa Barbara

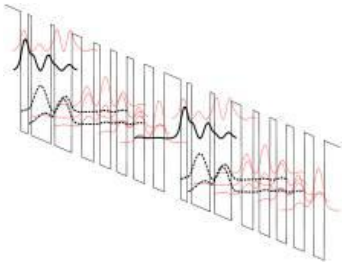
P. Long, E. Wilson, M. Povolotskyi, G. Klimeck
Network for Computational Nanotechnology, Purdue University

M. Urteaga, J. Hacker, M. Seo*, Z. Griffith, M. Fields, B. Brar
Teledyne Scientific and Imaging
**Now Sunkyunkwan University.*

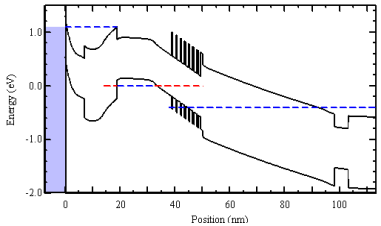
RF, Fast Digital Performance Figures of Merit

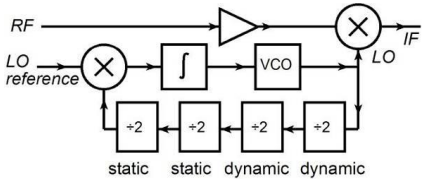
nm Transistors, Far-Infrared Integrated Circuits

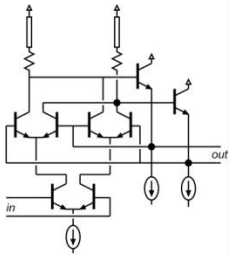
IR today→ lasers & bolometers → generate & detect

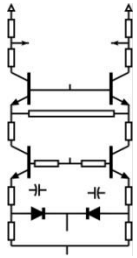


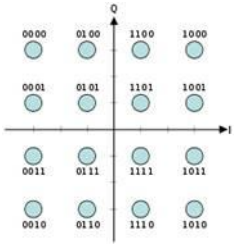
Far-infrared ICs: classic device physics, classic circuit design









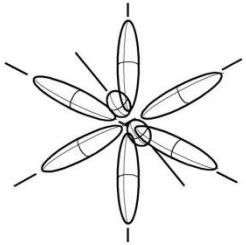


It's all about the interfaces: contact and gate dielectrics...

...wire resistance,...

...heat,...

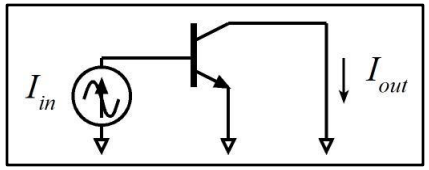
...& charge density.



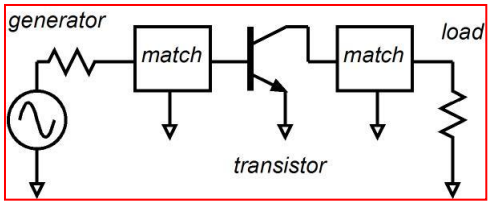
band structure and density of states !

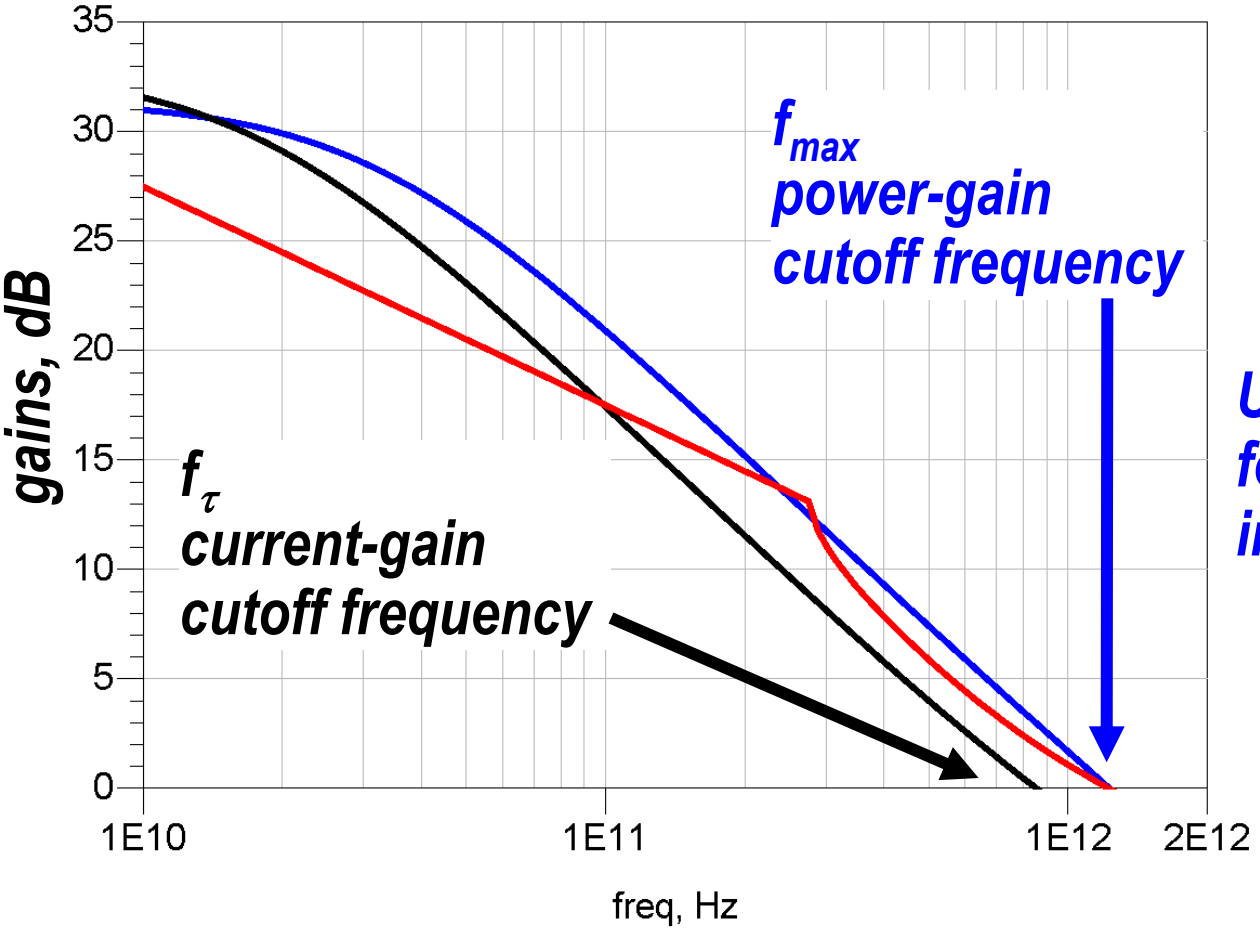
Transistor figures of Merit / Cutoff Frequencies

H_{21}=short-circuit current gain

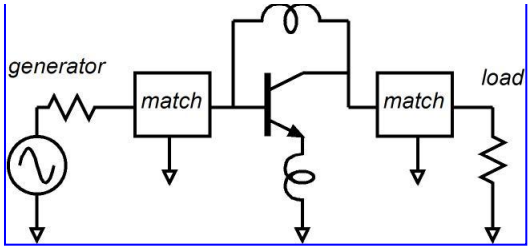


MAG = maximum available power gain: impedance-matched





U= unilateral power gain: feedback nulled, impedance-matched



What Determines Gate Delay ?

Gate Delay Determined by :

Depletion capacitance charging through the logic swing

$$\left(\frac{\Delta V_{LOGIC}}{I_C}\right)\left(C_{cb}+C_{be,\text{depletion}}\right)$$

Depletion capacitance charging through the base resistance

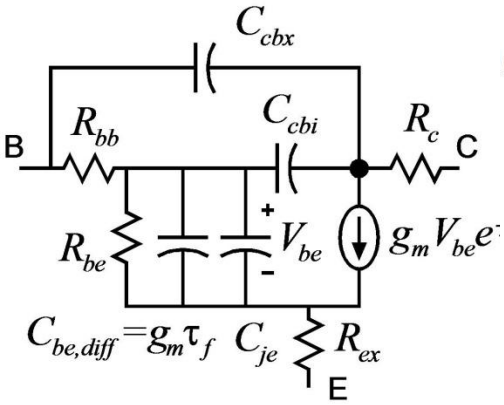
$$R_{\text{bb}}\left(C_{cbi}+C_{be,\text{depletion}}\right)$$

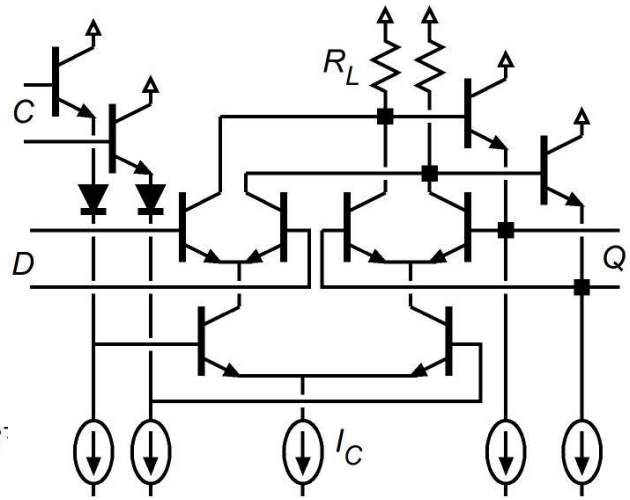
Supplying base + collector stored charge through the base resistance

$$R_{\text{bb}}\left(\tau_b+\tau_c\right)\left(\frac{I_C}{\Delta V_{LOGIC}}\right)$$

The logic swing must be at least

$$\Delta V_{LOGIC} > 4\cdot\left(\frac{kT}{q}+R_{ex}I_c\right)$$





$(\tau_b+\tau_c)$ typically 10 - 25% of total delay;
Delay not well correlated with f_τ

$(\Delta V_{LOGIC}/I_C)(C_{cb}+C_{be,\text{depl}})$ is 55% - 80% of total.

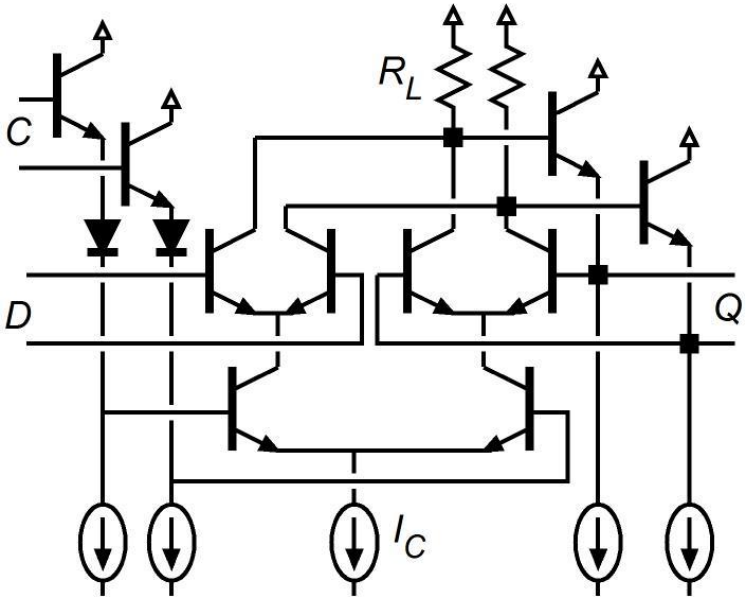
High (I_C/C_{cb}) is a key HBT design objective.

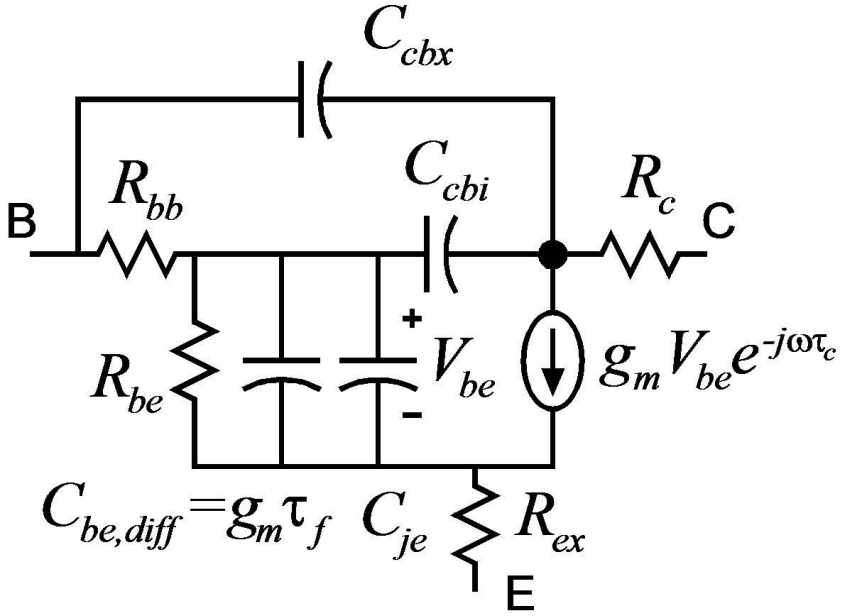
$$J_{\max,Kirk}=2\varepsilon\bar{v}_{electron}(V_{ce,\text{operating}}+V_{ce,\text{full depletion}})/T_c^2$$

$$\Rightarrow \frac{C_{cb}\Delta V_{LOGIC}}{I_C}=\frac{\Delta V_{LOGIC}}{2V_{CE,\min}}\left(\frac{A_{\text{collector}}}{A_{\text{emitter}}}\right)\left(\frac{T_C}{2\bar{v}_{electron}}\right)$$

R_{ex} must be very low for low ΔV_{logic} at high J

HBT Design For Digital & Mixed-Signal Performance





from charge-control analysis:

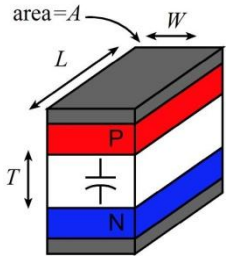
$$T_{gate} \approx (\Delta V_L / I_C)(C_{je} + 6C_{cbx} + 6C_{cbi}) + \tau_f$$
$$+ (kT / qI_C)(0.5C_{je} + C_{cbx} + C_{cbi} + 0.5\tau_f I_C / \Delta V_L)$$
$$+ R_{ex}(0.5C_{cbx} + 0.5C_{cbi} + 0.5\tau_f I_C / \Delta V_L)$$
$$+ R_{bb}(0.5C_{je} + C_{cbi} + 0.5\tau_f I_C / \Delta V_L).$$

analog ICs have similar bandwidth constraints...

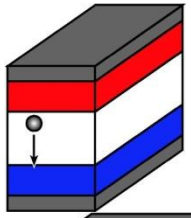
Electron Device Design

Transistor scaling laws: (V,I,R,C,t) vs. geometry

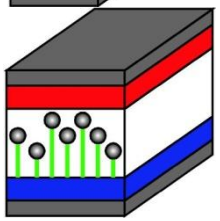
Depletion Layers



$$C = \varepsilon \cdot \frac{A}{T}$$

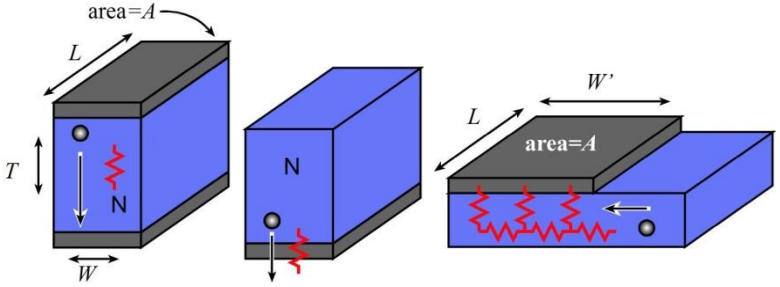


$$\tau = \frac{T}{2v}$$



$$\frac{I_{\max}}{A} = \frac{4\varepsilon v_{sat}(V_{appl} + \phi)}{T^2}$$

Bulk and Contact Resistances



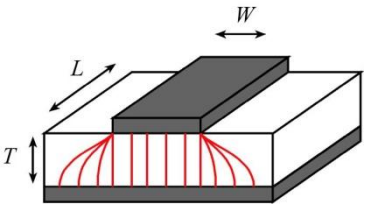
$$R \cong \rho_{contact} / A$$

contact terms dominate

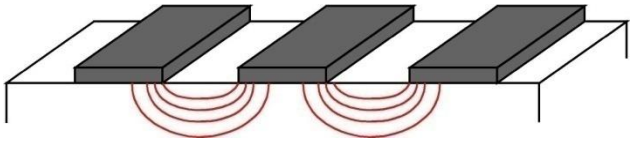
Fringing Capacitances

1) FET fringing capacitances

2) IC interconnect capacitances

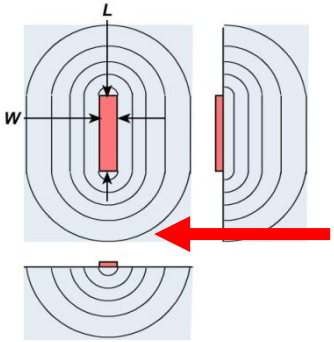


$$C_{finging} / L \sim \varepsilon$$

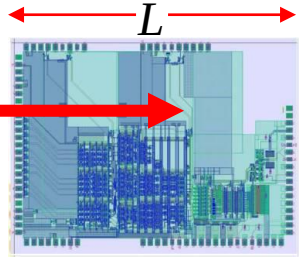


$$C_{finging} / L \sim \varepsilon$$

Thermal Resistance

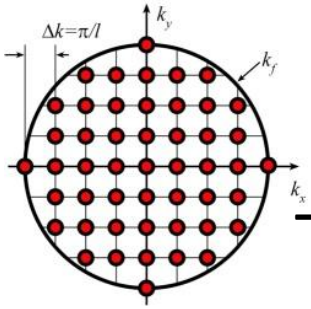


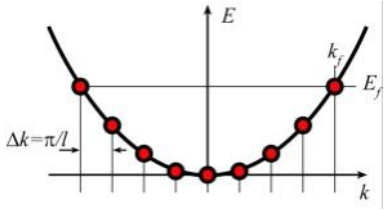
$$\Delta T_{IC} \propto \frac{P_{IC}}{K_{th} L}$$



$$\Delta T_{\text{transistor}} \sim \frac{P}{\pi K_{th} L} \ln\left(\frac{L}{W}\right)$$

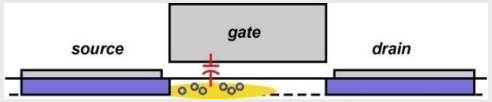
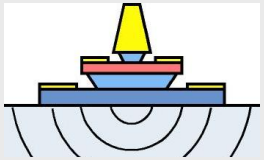
Available quantum states to carry current

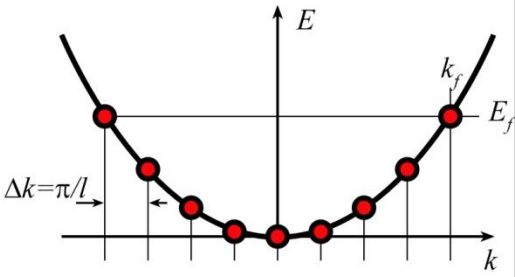


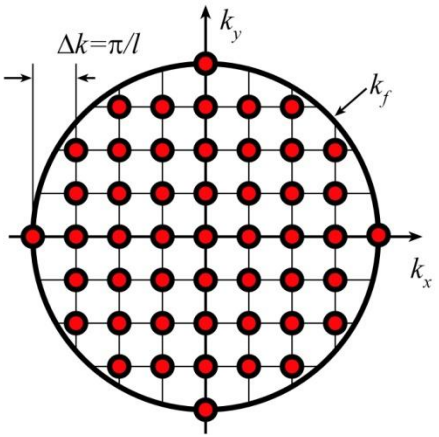


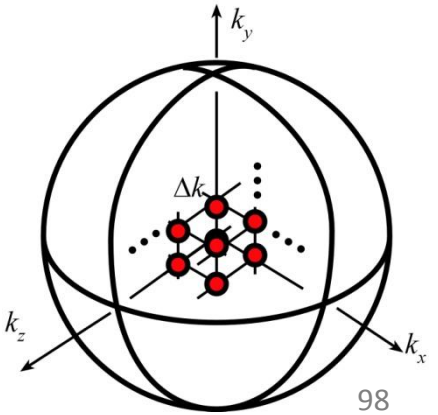
→ ***capacitance, transconductance contact resistance***

THz & nm Transistors: State Density Limits

	2-D: FET	3-D: BJT
		
capacitance	$C_{DOS} = \frac{q^2 m^*}{2\pi\hbar^2}$	
current	$J_{sheet} = \frac{2^{3/2} q^{5/2} (m^*)^{1/2} V^{3/2}}{3\pi^2\hbar^2}$	$J = \frac{q^3 m^* V^2}{4\pi^2\hbar^3}$
conductivity	$\sigma_c = \left(\frac{q^2}{\hbar}\right) \cdot \left(\frac{2}{\pi^3}\right)^{1/2} \cdot n^{1/2}$	$\sigma_c = \left(\frac{q^2}{\hbar}\right) \cdot \left(\frac{3}{8\pi}\right)^{2/3} \cdot n^{2/3}$



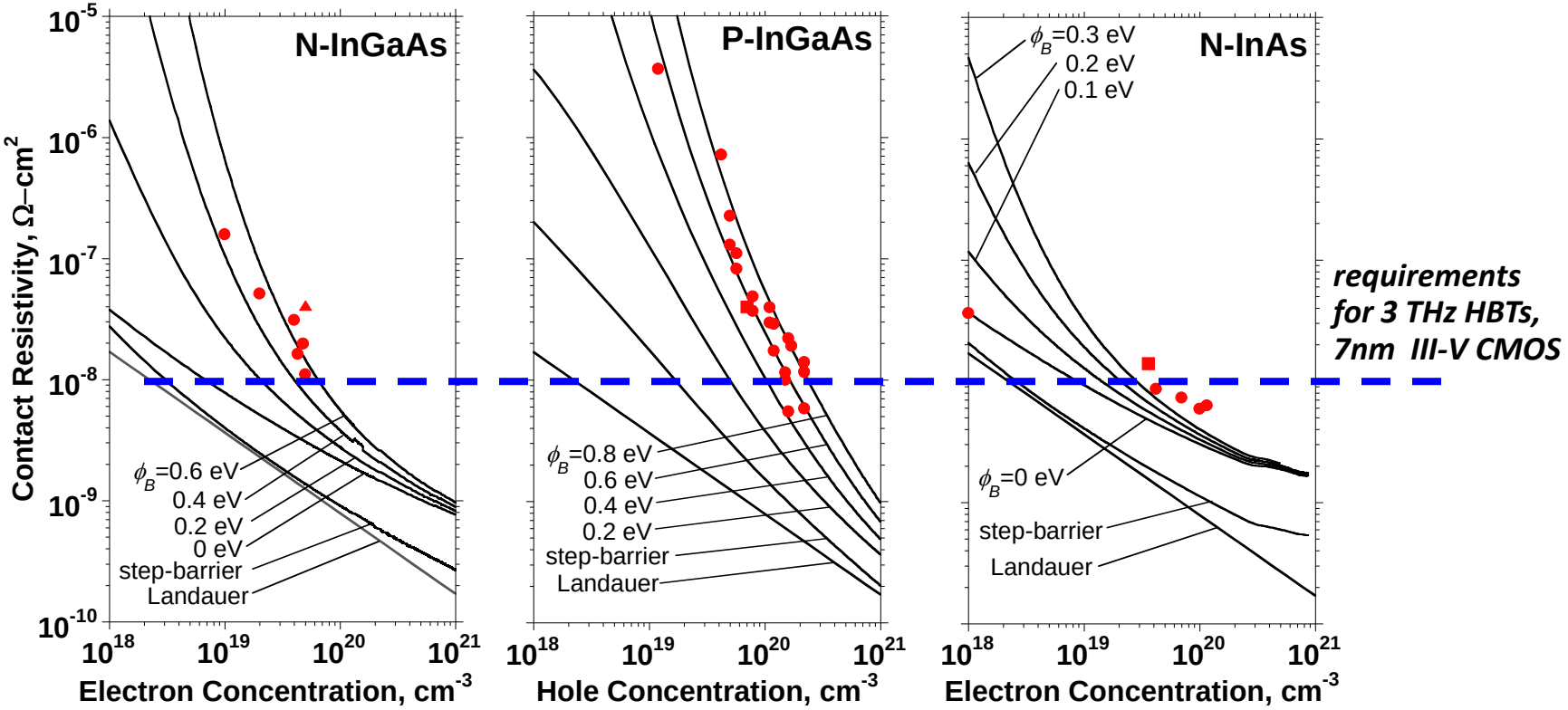




# of available quantum states / energy
- **determines FET channel capacitance**
- **FET and bipolar transistor current**
- **access resistance of Ohmic contact**

Refractory Contacts to In(Ga)As

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



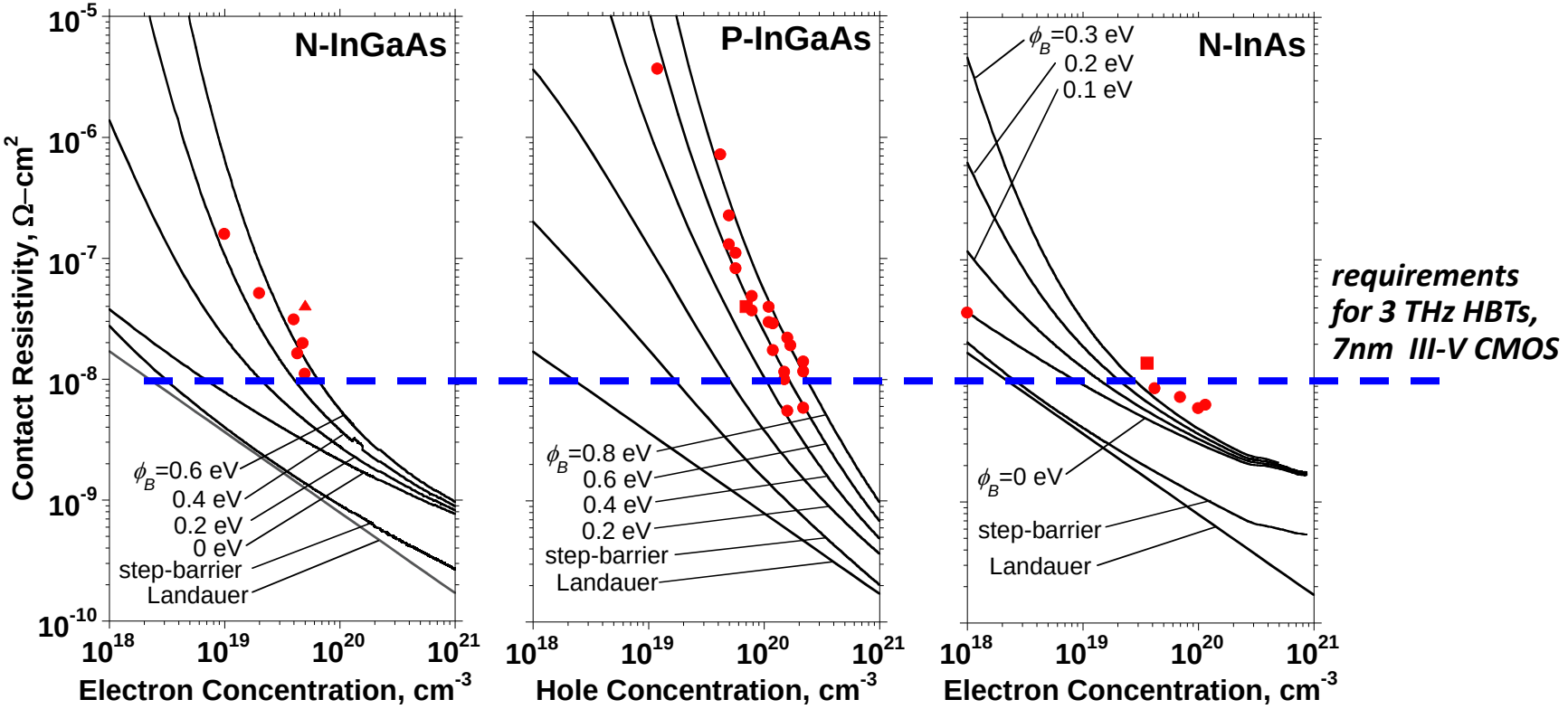
Refractory: robust under high-current operation / Low penetration depth: ~ 1 nm / Performance sufficient for 32 nm /2.8 THz node.

Why no ~2THz HBTs today ?

Problem: reproducing these base contacts in full HBT process flow

Refractory Contacts to In(Ga)As

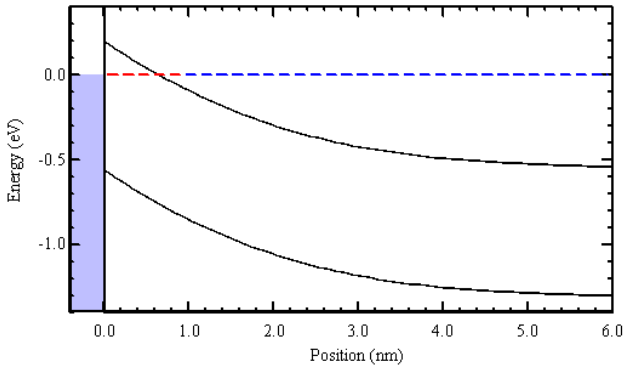
Baraskar *et al*, Journal of Applied Physics, 2013



Schottky Barrier is about one lattice constant

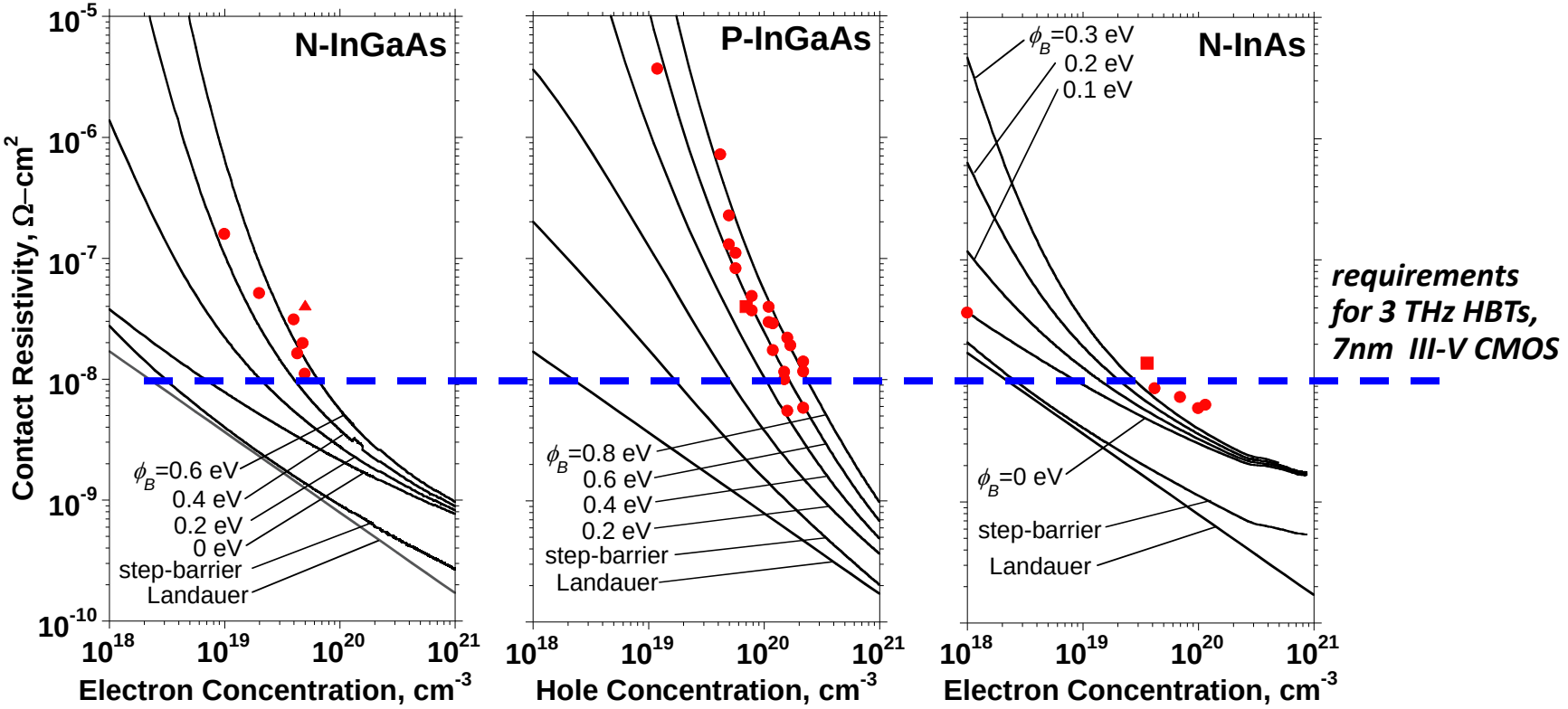
what is setting contact resistivity ?

what resistivity should we expect ?



Refractory Contacts to In(Ga)As

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



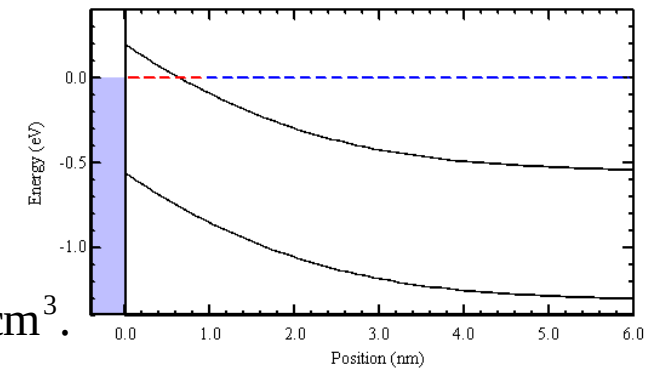
Zero-barrier contact resistivity:
(state density and
quantum-reflectivity limit)

$$\rho_c = \left(\frac{\hbar}{q^2}\right)\cdot\left(\frac{8\pi}{3}\right)^{2/3}\cdot\frac{1}{T^2}\cdot\frac{1}{n^{2/3}}$$

n = carrier concentration

T = transmission coefficient

$\rho_c \cong 0.1\,\Omega-\mu\text{m}^2$ at $n = 7\cdot10^{19}/\text{cm}^3$.



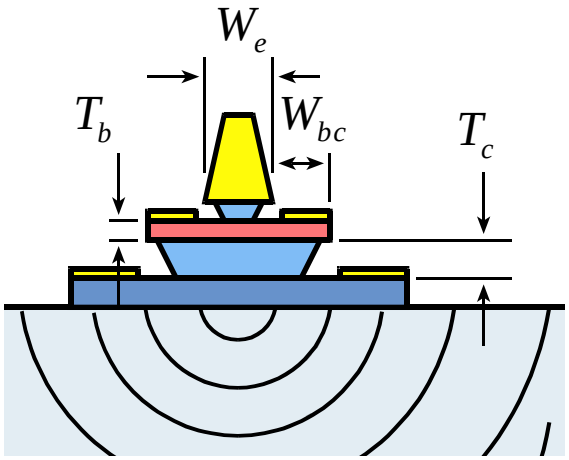
Bipolar Transistors

Bipolar Transistor Design

$$\tau_b \approx T_b^2 / 2D_n$$

$$\tau_c = T_c / 2v_{sat}$$

$$C_{cb} = \varepsilon A_c / T_c$$



(emitter length L_E)

$$I_{c,\max} \propto v_{sat} A_e (V_{\text{ce,operating}} + V_{\text{ce,punch-through}}) / T_c^2$$

$$\Delta T \propto \frac{P}{L_E}\left[1 + \ln\left(\frac{L_e}{W_e}\right)\right]$$

$$R_{ex} = \rho_{\text{contact}} / A_e$$

$$R_{bb} = \rho_{\text{sheet}}\left(\frac{W_e}{12L_e} + \frac{W_{bc}}{6L_e}\right) + \frac{\rho_{\text{contact}}}{A_{\text{contacts}}}$$

Bipolar Transistor Design: Scaling

$$\tau_b \approx T_b^2 / 2D_n$$

$$\tau_c = T_c / 2v_{sat}$$

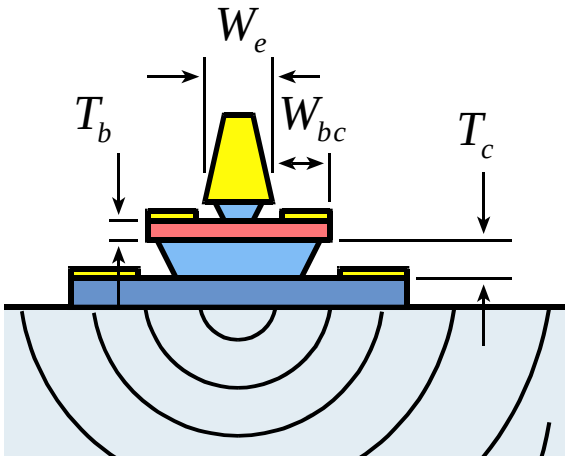
$$C_{cb} = \varepsilon A_c / T_c$$

$$I_{c,\max} \propto v_{sat} A_e (V_{\text{ce,operating}} + V_{\text{ce,punch-through}}) / T_c^2$$

$$\Delta T \propto \frac{P}{L_E}\left[1 + \ln\left(\frac{L_e}{W_e}\right)\right]$$

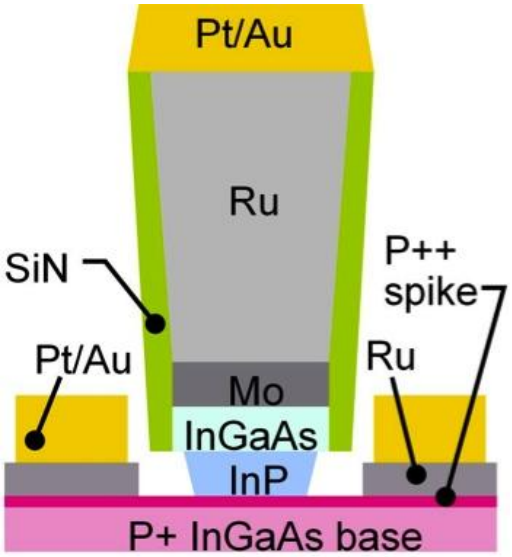
$$R_{ex} = \rho_{\text{contact}} / A_e$$

$$R_{bb} = \rho_{\text{sheet}}\left(\frac{W_e}{12L_e} + \frac{W_{bc}}{6L_e}\right) + \frac{\rho_{\text{contact}}}{A_{\text{contacts}}}$$



(emitter length L_E)

Scaling Laws, Scaling Roadmap



HBT parameter	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

Scaling Node	64	32	16	nm
Emitter Width	64	32	16	nm
Resistivity	2	1	0.5	Ω-μm^2
Base Thickness	18	15	13	nm
Contact width	60	30	15	nm
Contact ρ	2.5	1.25	0.63	Ω-μm^2
Collector Width	180	90	45	nm
Thickness	53	37.5	26	nm
Current Density	36	72	140	mA/μm^2
f_τ	1.0	1.4	2.0	THz
f_{max}	2.0	2.8	4.0	THz

Can we make a 2 THz SiGe Bipolar Transistor ?

Simple physics clearly drives scaling

transit times, C_{cb}/I_c

→ thinner layers, higher current density

high power density → narrow junctions

small junctions→ low resistance contacts

Key challenge: Breakdown

15 nm collector → very low breakdown

Also required:

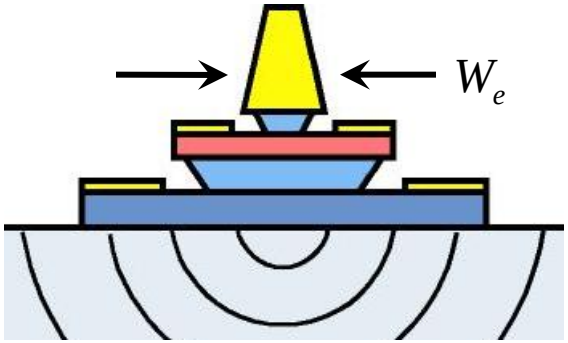
low resistivity Ohmic contacts to Si

very high current densities: heat

	InP	SiGe	
emitter	64	18	nm width
	2	**0.6**	$\Omega\cdot\mu m^2$ access ρ
base	64	18	nm contact width,
	2.5	**0.7**	$\Omega\cdot\mu m^2$ contact ρ
collector	53	**15**	nm thick
	36	125	$mA/\mu m^2$
	2.75	**1.3?**	V, breakdown
f_τ	**1000**	**1000**	GHz
f_{max}	**2000**	**2000**	GHz
PAs	1000	1000	GHz
digital	480	480	GHz
(2:1 static divider metric)			

Assumes collector junction 3:1 wider than emitter.
Assumes SiGe contacts no wider than junctions

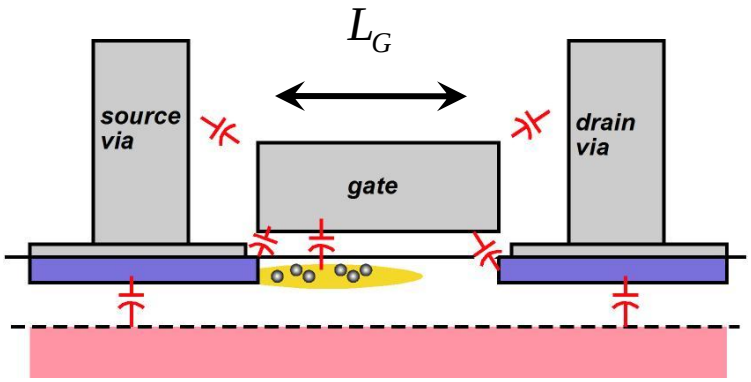
To double transistor bandwidth...



(emitter length L_E)

HBT parameter	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

nearly constant junction temperature → linewidths vary as (1 / bandwidth)²

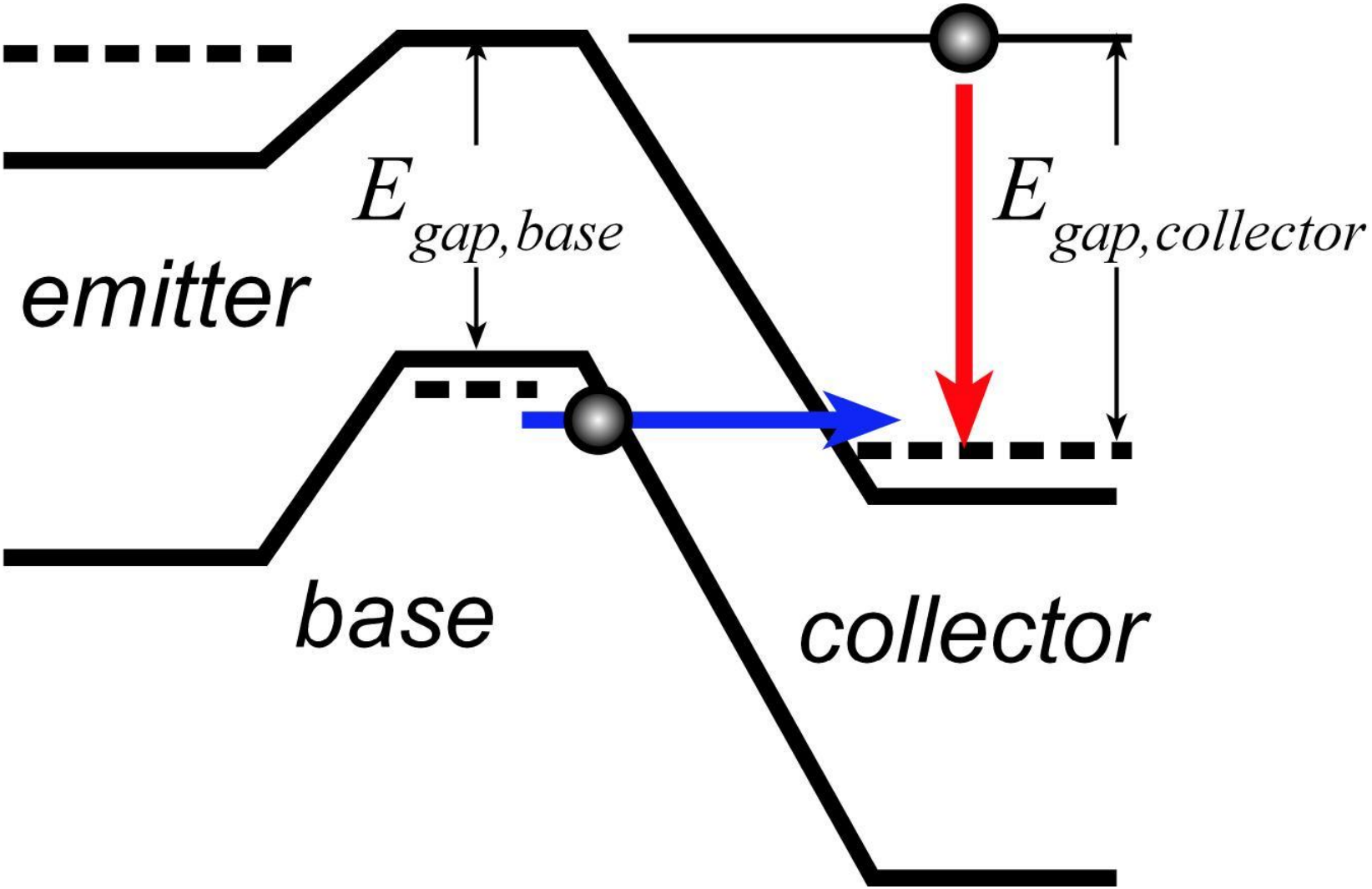


(gate width W_G)

FET parameter	change
gate length	decrease 2:1
current density (mA/μm), g_m (mS/μm)	increase 2:1
channel 2DEG electron density	increase 2:1
gate-channel capacitance density	increase 2:1
dielectric equivalent thickness	decrease 2:1
channel thickness	decrease 2:1
channel density of states	increase 2:1
source & drain contact resistivities	decrease 4:1

fringing capacitance does not scale → linewidths scale as (1 / bandwidth)

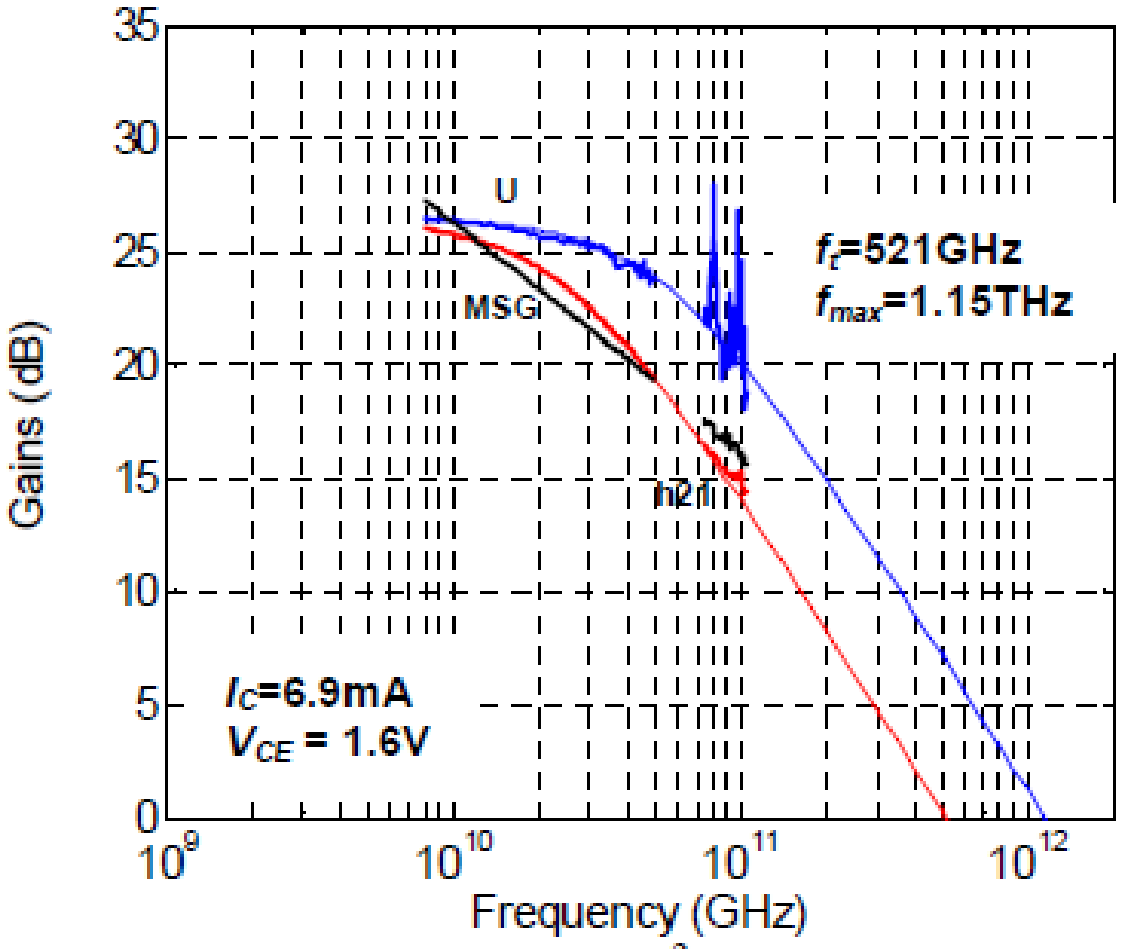
Energy-limited vs. field-limited breakdown



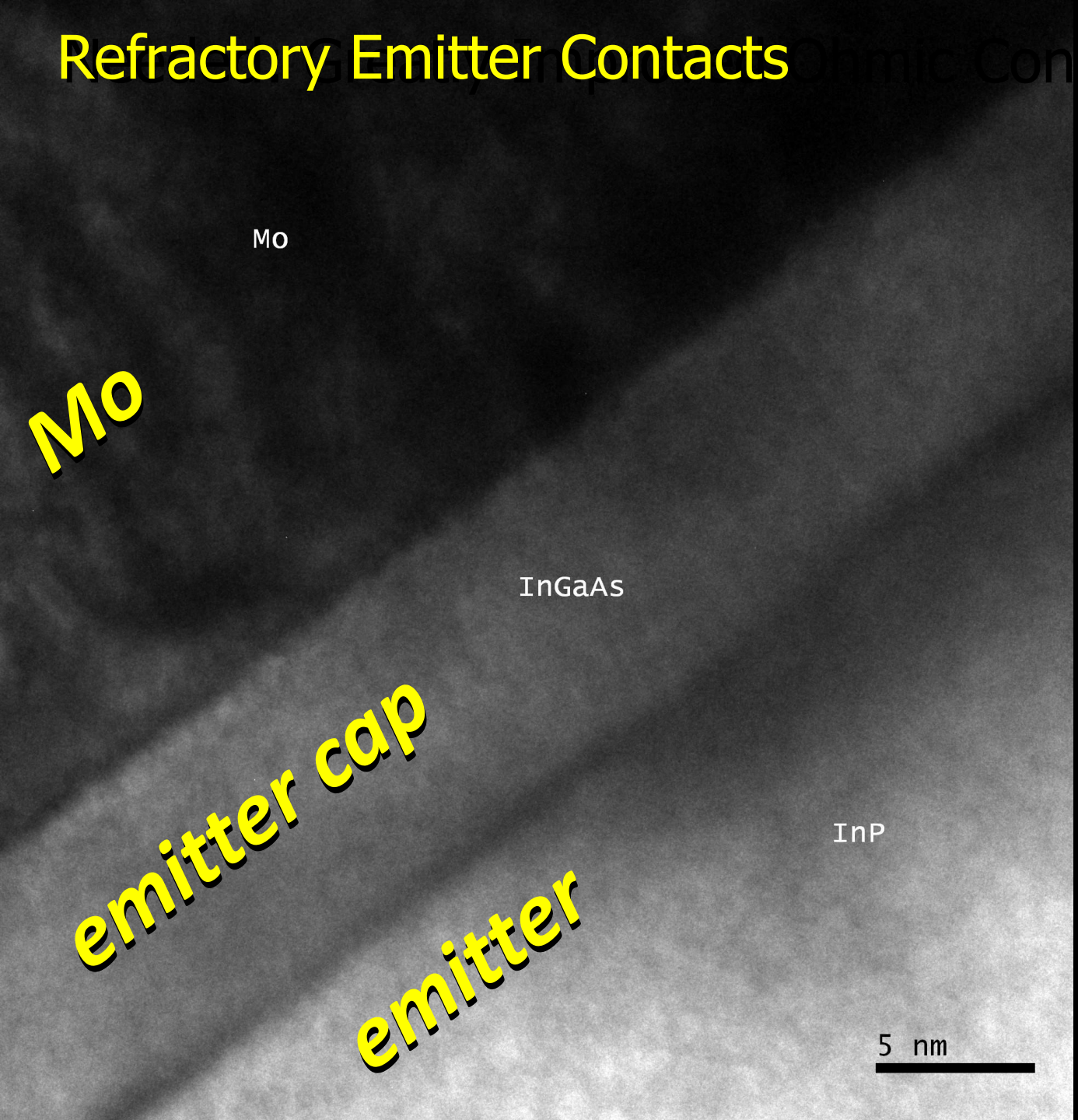
band-band tunneling: base bandgap
impact ionization: collector bandgap

THz InP HBTs: Performance @ 130 nm Node

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

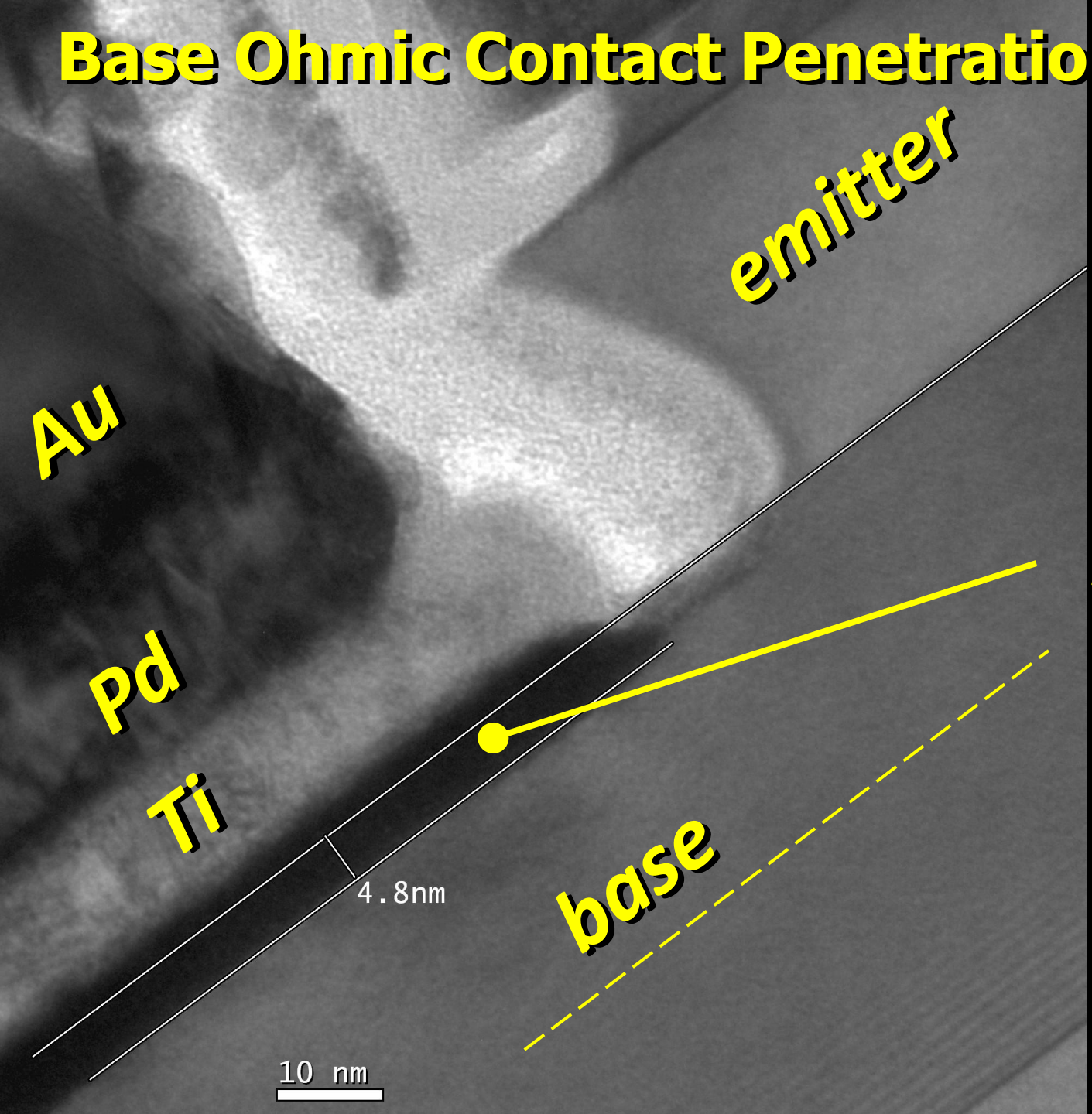


Refractory Emitter Contacts



negligible penetration

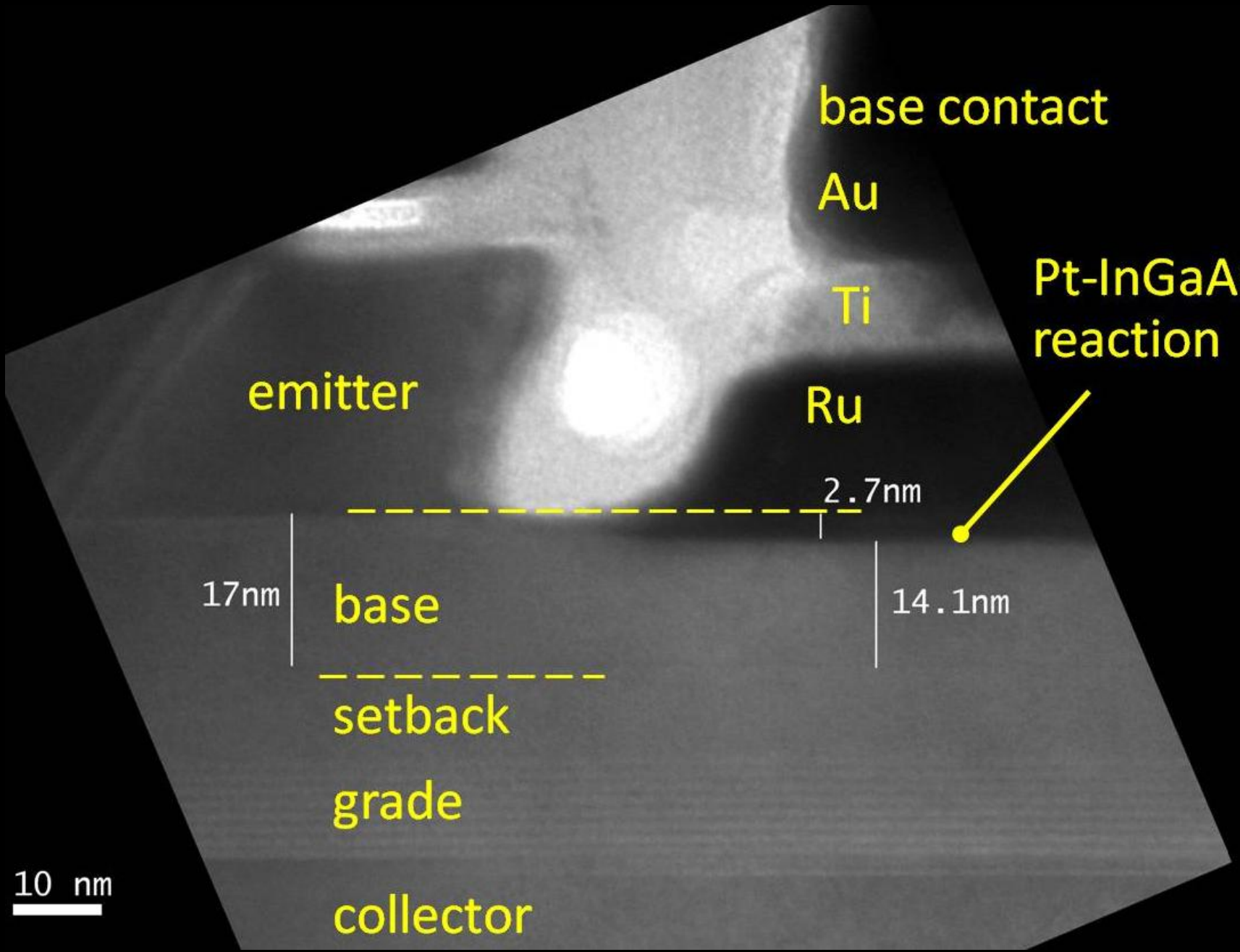
Base Ohmic Contact Penetration



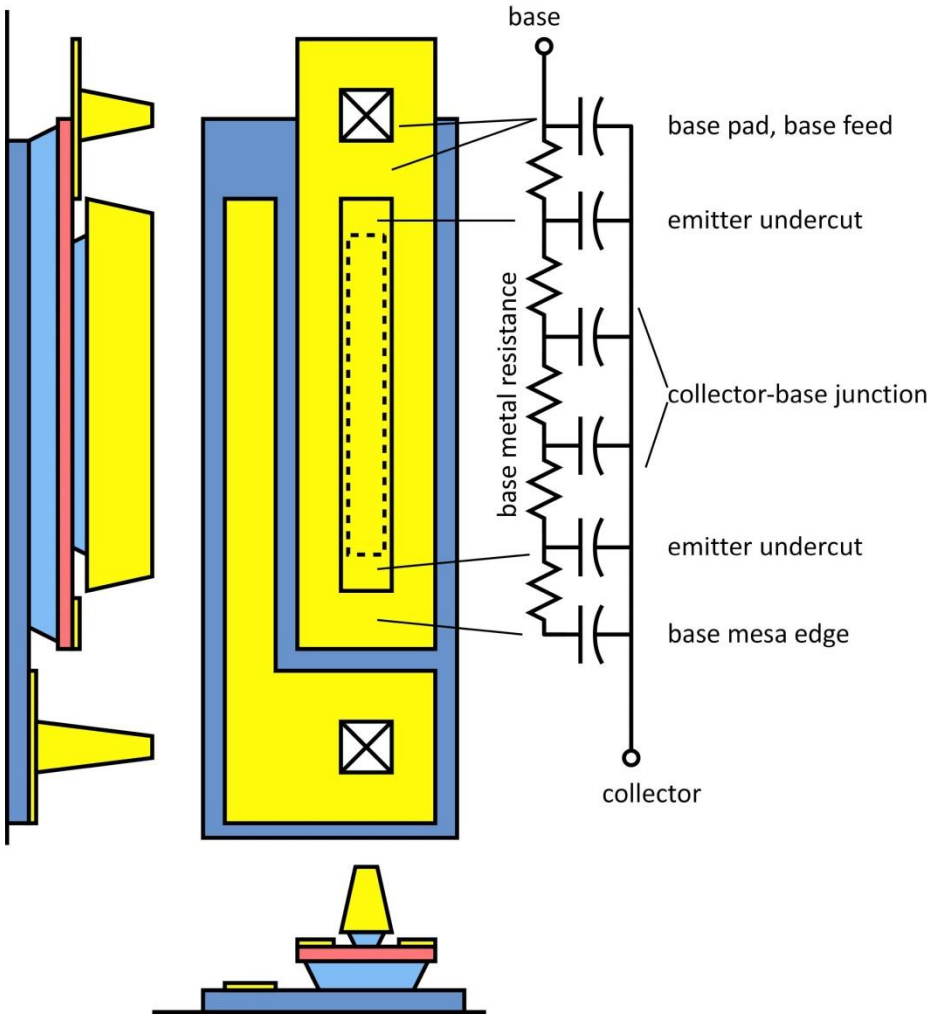
~5 nm Pt contact penetration

(into 25 nm base)

Blanket Base Metal Process



Parasitics along length of HBT emitter



Base pad & feed

increases C_{cb}

Emitter undercut

actual junction shorter than drawn.

→ excess C_{cb} , excess base metal resistance

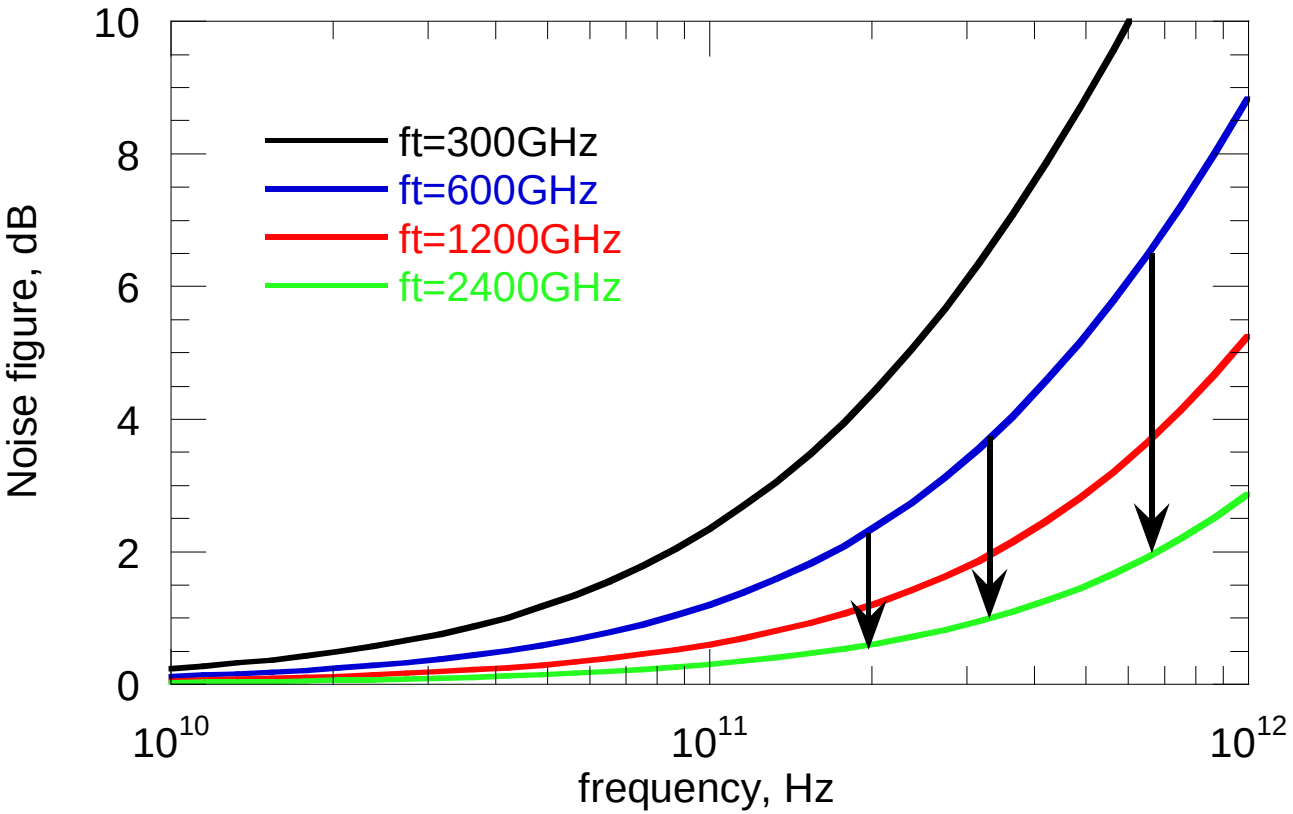
Base metal resistance

adds to R_{bb}

all these factors decrease f_{max}

Field-Effect Transistors ....for RF

HEMTs: Key Device for Low Noise Figure



$$F_{\min} \approx 1 + 2\sqrt{g_m(R_s + R_g + R_i)\Gamma} \cdot \left(\frac{f}{f_\tau}\right) + 2g_m(R_s + R_g + R_i)\Gamma \cdot \left(\frac{f}{f_\tau}\right)^2$$

$$\Gamma \approx 1$$

Hand-derived modified Fukui Expression, fits CAD simulation extremely well.

2:1 to 4:1 increase in f_τ → greatly improved noise @ 200-670 GHz.

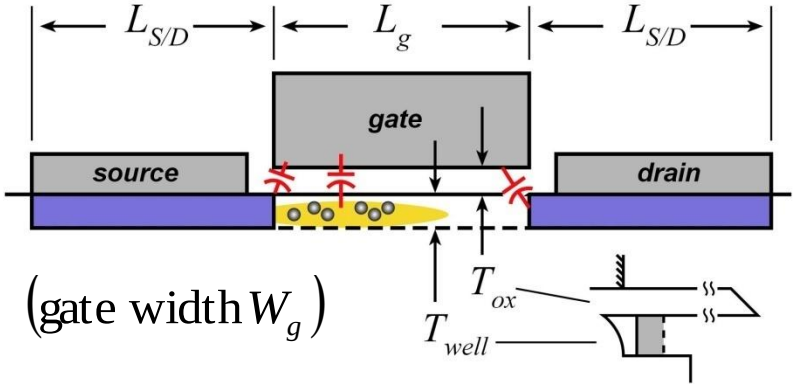
Better range in sub-mm-wave systems; or use smaller power amps.

Critical: Also enables THz systems beyond 820 GHz

FET Design

$$C_{gd} \cong C_{gs,f} \cong \varepsilon W_g$$

$$g_m = C_{g-ch} \cdot (v / L_g)$$



$$C_{g-ch} = \frac{L_g W_g}{T_{ox}/\varepsilon_{ox} + T_{\text{well}}/2\varepsilon_{\text{well}} + (q^2/\text{well state density})}$$

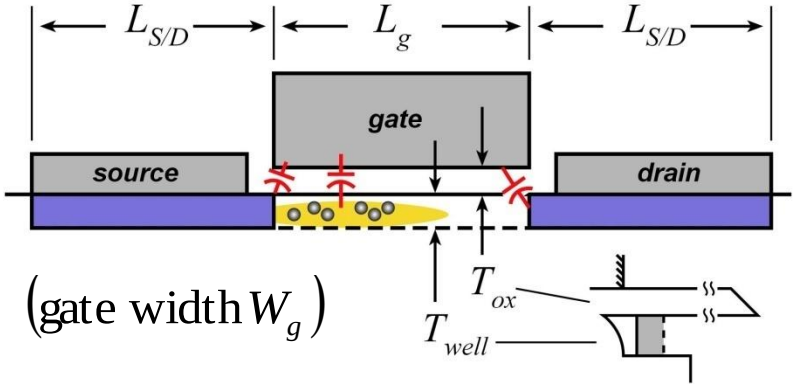
$$v \propto \left(\begin{array}{c} \text{voltage division ratio between} \\ \text{the above three capacitors} \end{array} \right)^{-1/2} \cdot \frac{1}{\sqrt{\text{transport mass}}}$$

$$R_{DS} \approx L_g /(W_g v \varepsilon) \qquad R_S = R_D = \frac{\rho_{\text{contact}}}{L_{\text{S/D}} W_g}$$

FET Design: Scaling

$$C_{gd} \cong C_{gs,f} \cong \varepsilon W_g$$

$$g_m = C_{g-ch} \cdot (v / L_g)$$

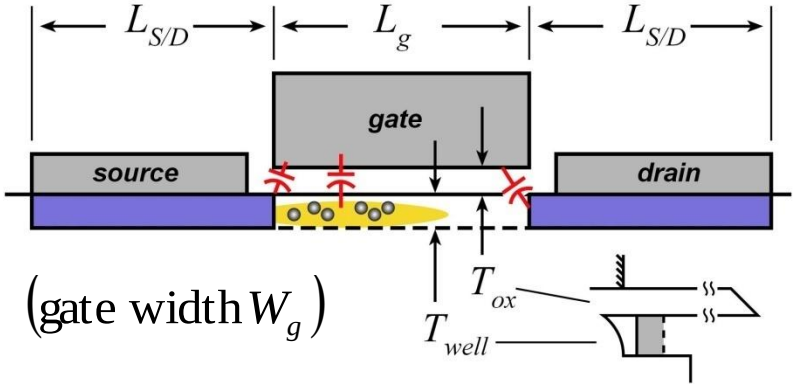


$$C_{g-ch} = \frac{L_g W_g}{T_{ox} / \varepsilon_{ox} + T_{\text{well}} / 2\varepsilon_{\text{well}} + (q^2 / \text{well state density})}$$

$$v \propto \left(\begin{array}{c} \text{voltage division ratio between} \\ \text{the above three capacitors} \end{array} \right)^{-1/2} \cdot \frac{1}{\sqrt{\text{transport mass}}}$$

$$R_{DS} \approx L_g / (W_g v \varepsilon) \qquad R_S = R_D = \frac{\rho_{\text{contact}}}{L_{\text{S/D}} W_g}$$

FET Design: Scaling



2:1 $C_{gd} \cong C_{gs,f} \cong \varepsilon W_g$ 2:1

constant 2:1

$$g_m = C_{g-ch} \cdot (v / L_g)$$ 2:1

2:1

$$C_{g-ch} = \frac{L_g W_g}{T_{ox}/\varepsilon_{ox} + T_{well}/2\varepsilon_{well} + (q^2/\text{well state density})}$$

2:1 (for $L_g W_g$) 2:1

2:1 (for T_{ox}) 2:1 (for T_{well}) 2:1 (for well state density)

constant

$$v \propto \left(\text{voltage division ratio between the above three capacitors}\right)^{-1/2} \cdot \frac{1}{\sqrt{\text{transport mass}}}$$

constant (voltage division ratio) constant (transport mass)

constant

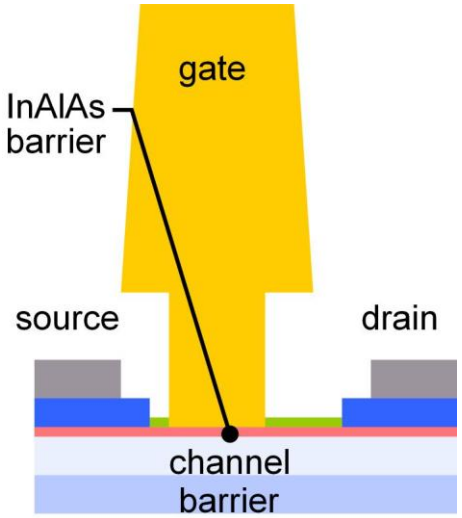
$$R_{DS} \approx L_g /(W_g v \varepsilon)$$

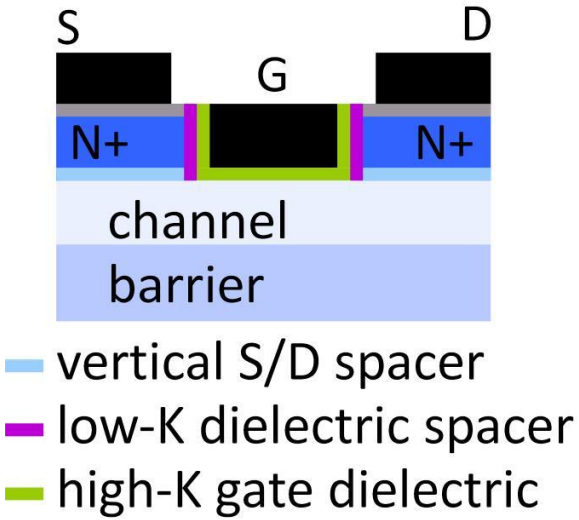
2:1 2:1

$$R_S = R_D = \frac{\rho_{\text{contact}}}{L_{S/D} W_g}$$

constant 4:1 2:1 2:1

Field-Effect Transistor Scaling Laws

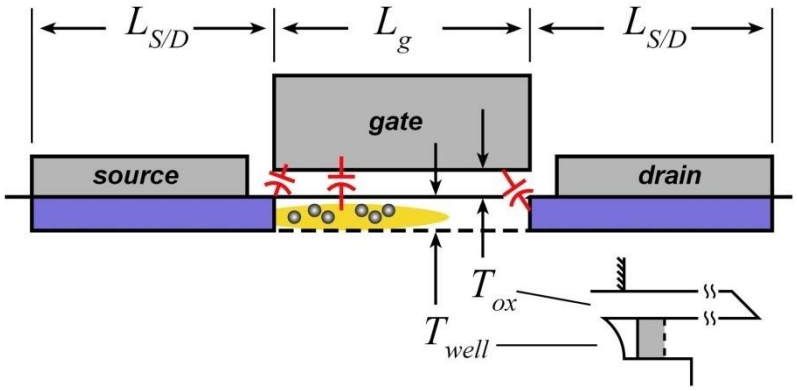




FET parameter	change
gate length	decrease 2:1
current density (mA/μm), g_m (mS/μm)	increase 2:1
transport effective mass	constant
channel 2DEG electron density	increase 2:1
gate-channel capacitance density	increase 2:1
dielectric equivalent thickness	decrease 2:1
channel thickness	decrease 2:1
channel density of states	increase 2:1
source & drain contact resistivities	decrease 4:1

fringing capacitance does not scale → linewidths scale as (1 / bandwidth)

Field-Effect Transistors No Longer Scale Properly



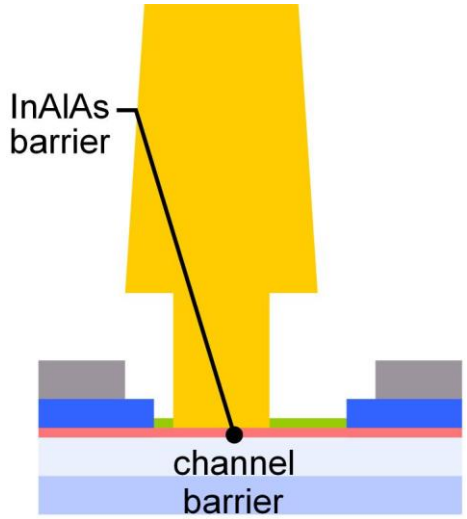
FET parameter	change
gate length	decrease 2:1
current density (mA/μm), g_m (mS/μm)	increase 2:1
transport effective mass	constant
channel 2DEG electron density	increase 2:1
gate-channel capacitance density	increase 2:1
dielectric equivalent thickness	decrease 2:1
channel thickness	decrease 2:1
channel density of states	increase 2:1
source & drain contact resistivities	decrease 4:1

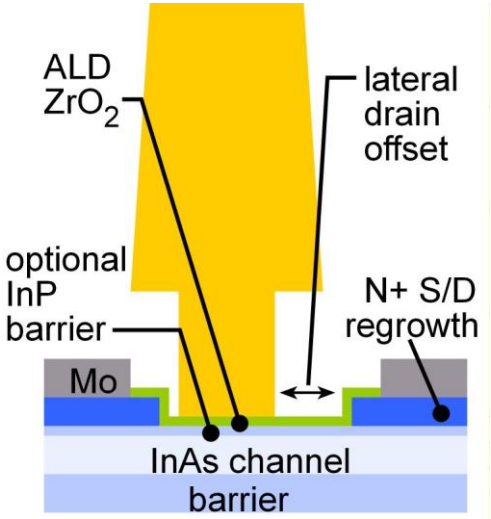
Gate dielectric can't be much further scaled.
Not in CMOS VLSI, not in mm-wave HEMTs

g_m/W_g (mS/μm) hard to increase→ C_{fringe}/g_m prevents f_τ scaling.

Shorter gate lengths degrade electrostatics→ reduced g_m/G_{ds}

Scaling roadmap for InP HEMTs





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

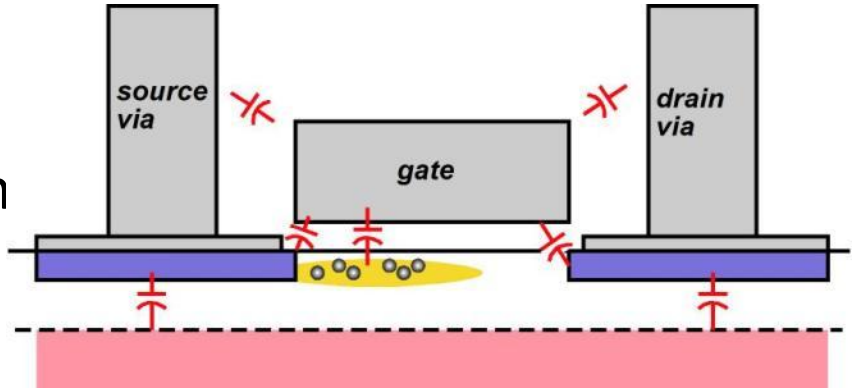
Field-Effect Transistors ....for logic

What goals for logic FETs ?

Low off-state current (nA to pA/μm) for low static dissipation
→ minimum subthreshold slope→ minimum L_g / T_{ox}
low gate tunneling, low band-band tunneling

Low delay C_{FET} $\Delta V/I_d$ in gates where transistor capacitances dominate.

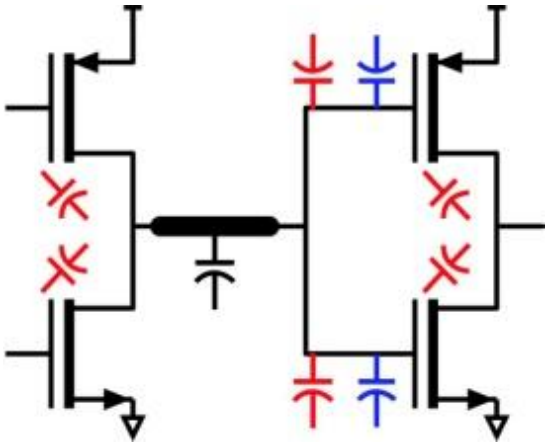
Parasitic capacitances are 0.5-1.0 fF/μm
→ low C , high I_d



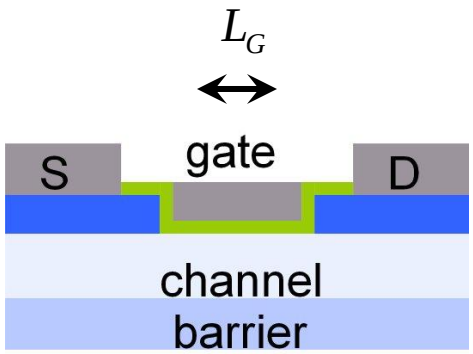
Low delay C_{wire} $\Delta V/I_d$ in gates where wiring capacitances dominate.

→ need high I_d / W_g

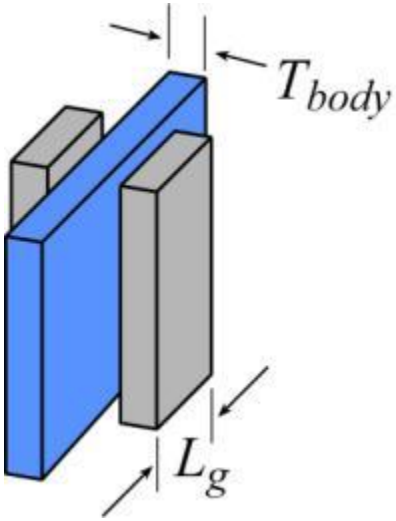
and small !



nm/VLSI MOSFET Scaling: Ideal and Feasible

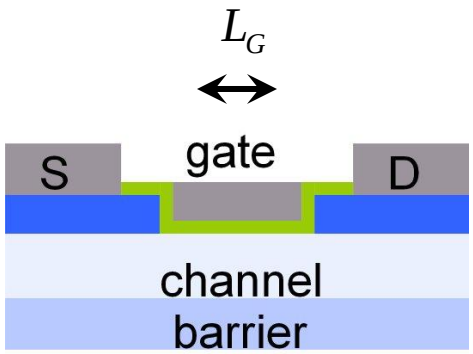


$(\text{gate width } W_G)$

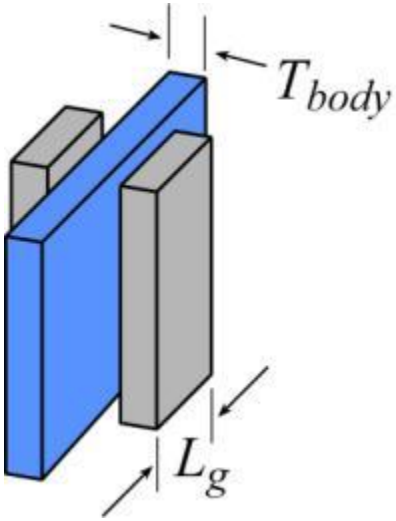


FET parameter	change
gate length	decrease 2:1
current density (mA/mm)	increase 2:1
transport mass	constant
2DEG electron density	increase 2:1
gate-channel capacitance density	increase 2:1
dielectric equivalent thickness	decrease 2:1
channel thickness	decrease 2:1
channel state density	increase 2:1
contact resistivities	decrease 4:1

nm/VLSI MOSFET Scaling: Ideal and Feasible

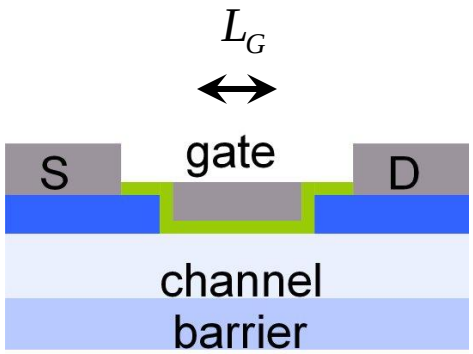


(gate width W_G)



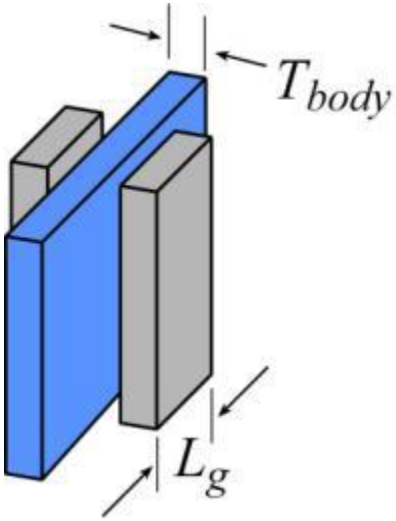
FET parameter	change
gate length	decrease 2:1
current density (mA/mm)	increase 2:1
transport mass	constant
2DEG electron density	increase 2:1
gate-channel capacitance density	increase 2:1
dielectric equivalent thickness	decrease 2:1
channel thickness	decrease 2:1
channel state density	increase 2:1
contact resistivities	decrease 4:1

nm/VLSI MOSFET Scaling: Goals



$(\text{gate width } W_G)$

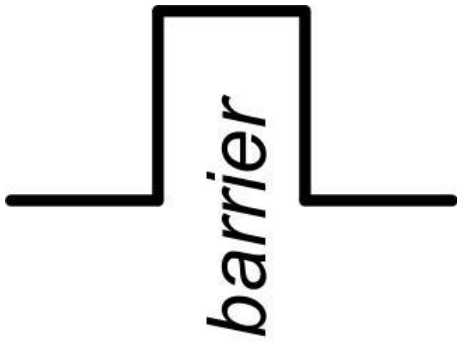
FET parameter	
gate length	8 nm
current density (mA/mm)	1 mA/μm @0.5V
transport mass	
2DEG electron density	$3*10^{12}/cm^2$
gate-channel capacitance density	
dielectric equivalent thickness	~~0.4 nm~~ (0.8 nm fin)
channel thickness	2 nm (4 nm fin)
channel state density	
contact resistivities	0.3 Ω-μm^2



Minimum Dielectric Thickness & Gate Leakage

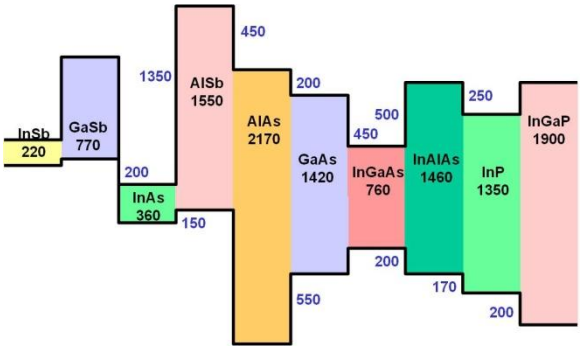
Attenuation coefficient:

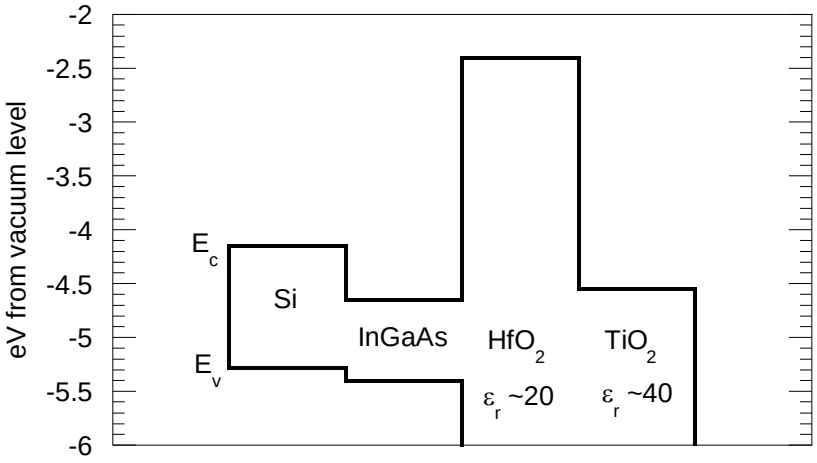
$$\alpha \cong \frac{\sqrt{2m^* E_{\text{barrier}}}}{\hbar}$$



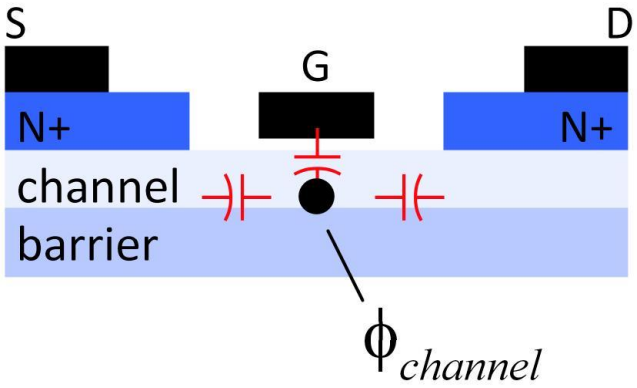
Transmission Probability (WKB approximation)

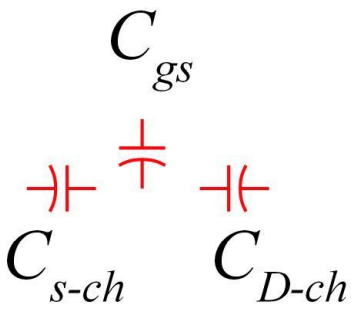
$$P \cong \exp(-2\alpha T_{\text{barrier}})$$





Aspect ratio and subthreshold swing





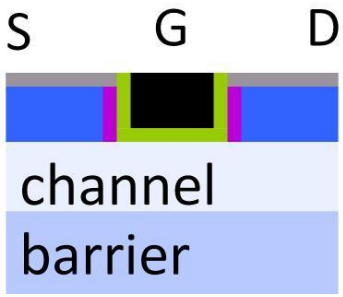
$$C_{gs} \sim \varepsilon A / D$$

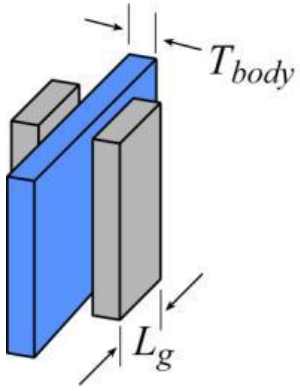
$$C_{d-ch} \sim \varepsilon_r \varepsilon_0 W_g$$

$$\frac{\partial \phi}{\partial V_g} \cong \frac{C_{gs}}{C_{gs} + C_{s-ch} + C_{D-ch}} \cong \frac{C_{gs}}{C_{gs} + 2\varepsilon_r \varepsilon_0 W_g} = \frac{1}{1 + 2\varepsilon_r \varepsilon_0 W_g / C_{gs}}$$

$$I_D \propto \exp\left(-\frac{q\phi}{kT}\right) = \exp\left(\frac{-V_{gate}}{(kT/q)(1 + 2\varepsilon_r \varepsilon_0 W_g / C_{gs})}\right)$$

Contact Resistance Scaling



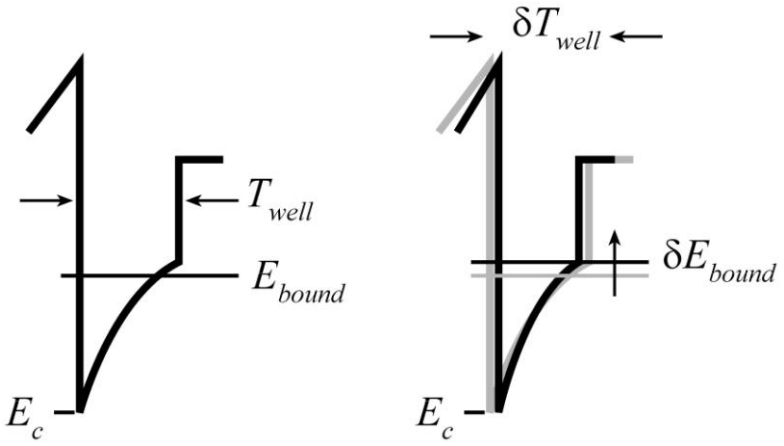


FET parameter	?? nm Node
gate length	~10 nm
current density (mA/mm)	1 mA/μm @0.5V
transport mass	
2DEG electron density	$3*10^{12}/cm^2$
gate-channel capacitance density	
dielectric equivalent thickness	0.5 nm (fin: 1.0 nm)
channel thickness	2.5 nm (fin: 5 nm)
channel state density	
contact resistivities	0.4 Ω-μm^2

With the above #s, contacts degrade on-current by ~15%

A 2.4 Ω–μm^2 contact would reduce the current 2:1

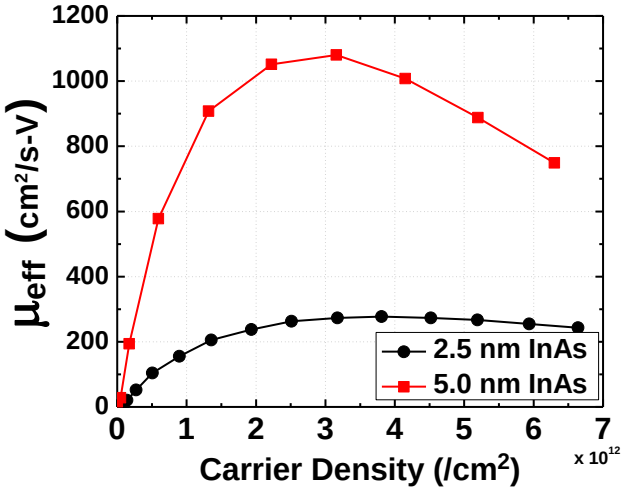
Mobility in Thin Channels: Surface Roughness Scattering

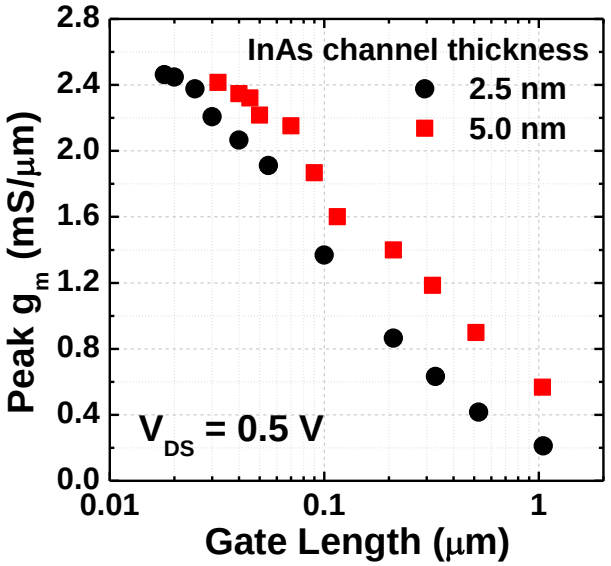


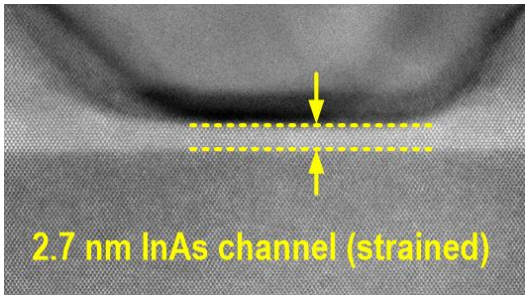
$$\text{Mobility} \propto \frac{m_q^2 T_{well}^6}{\delta T_{well}^2}$$

Sakaki

mobility in FET channels







Mobility is high if surfaces are smooth

JOURNAL OF APPLIED PHYSICS 115, 123711 (2014)

Two dimensional electron transport in modulation-doped $In_{0.53}Ga_{0.47}As/AlAs_{0.56}Sb_{0.44}$ ultrathin quantum wells

Cheng-Ying Huang,[1,a)] Jeremy J. M. Law,[1] Hong Lu,[1,2] Debdeep Jena,[3] Mark J. W. Rodwell,[1] and Arthur C. Gossard[1,2]

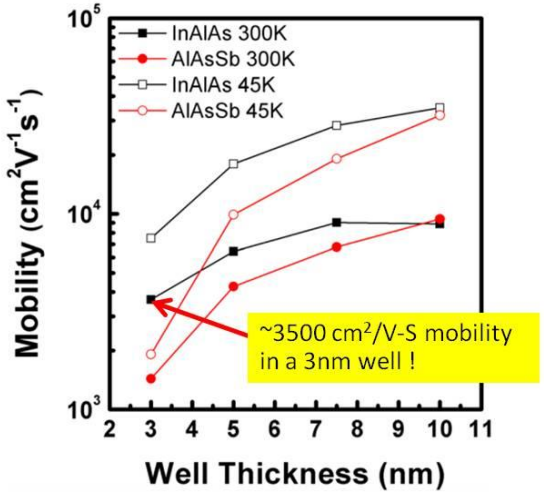
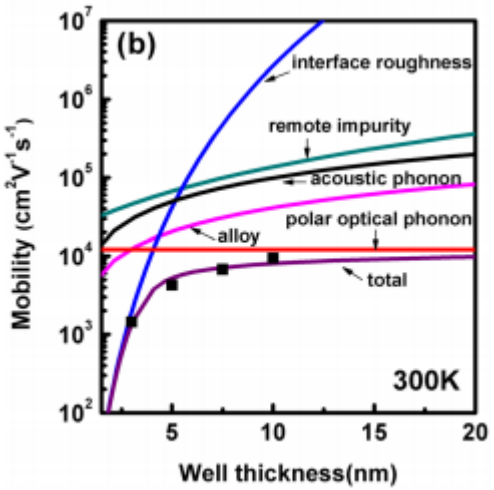


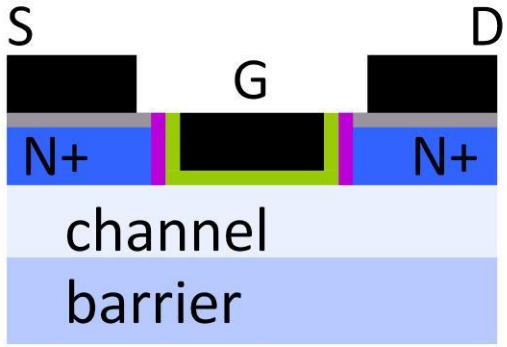
FIG. 4. Measured low temperature (45 K) and room temperature (300 K) mobilities of InGaAs/InAlAs and InGaAs/AlAsSb 2DEGs as a function of the InGaAs well thickness.



$$\text{Mobility} \propto \frac{m_q^2 T_{well}^6}{\delta T_{well}^2}$$

Quantum well: smooth surfaces

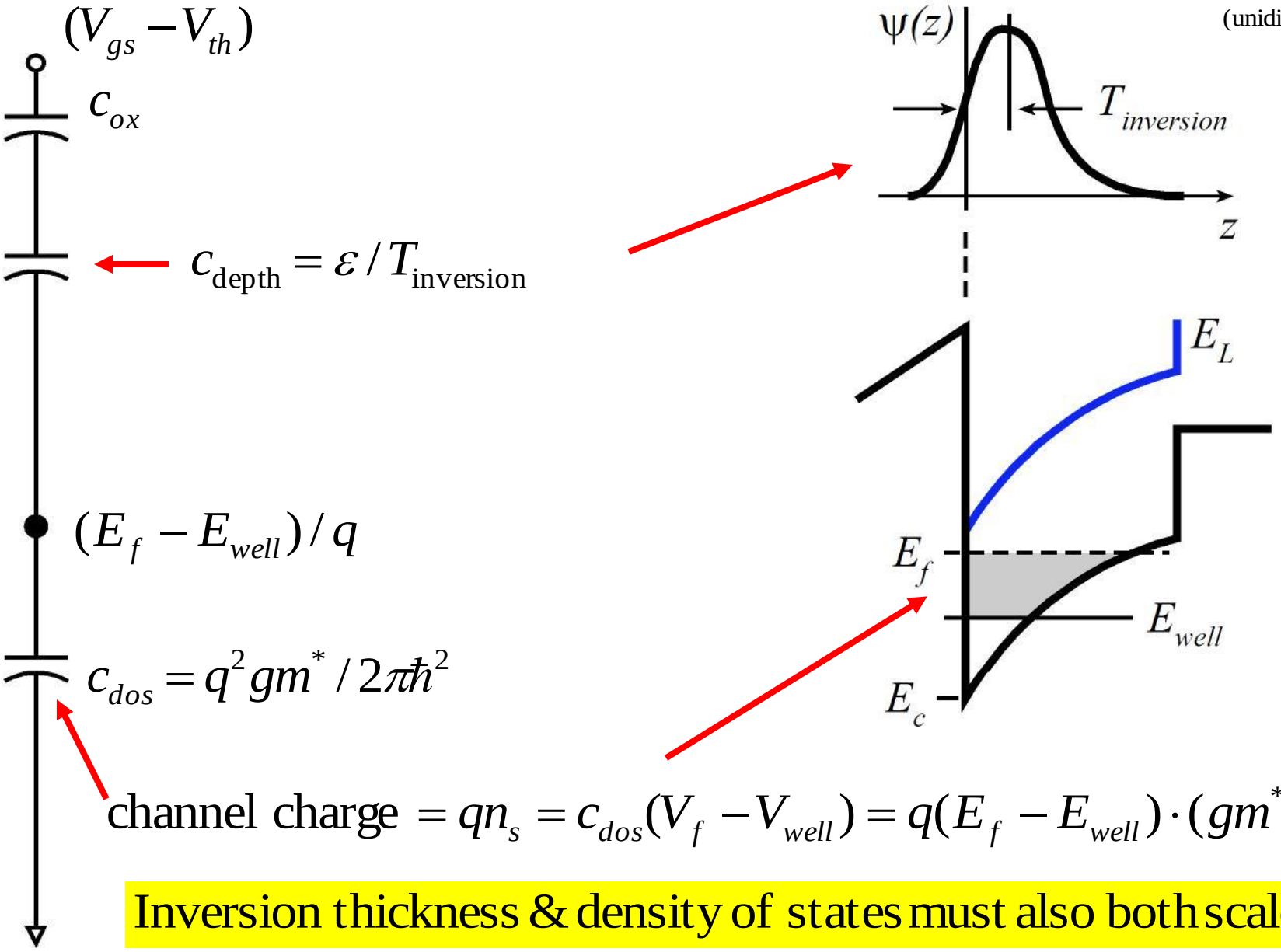
FET: rough surfaces



low-K dielectric spacer

high-K gate dielectric

Terms in gate-channel capacitance

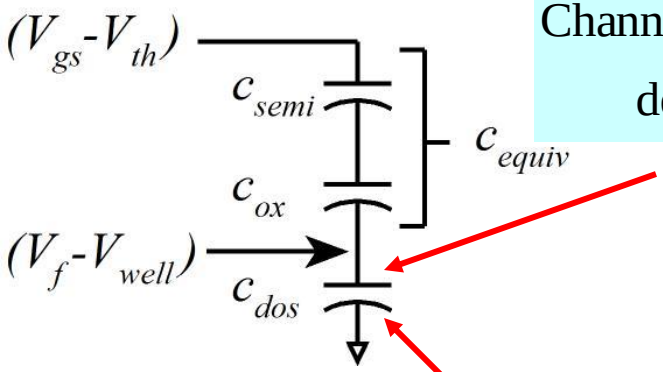


$$\text{channel charge} = qn_s = c_{dos}(V_f - V_{well}) = q(E_f - E_{well}) \cdot (gm^* / 2\pi\hbar^2)$$

Inversion thickness & density of states must also both scale.

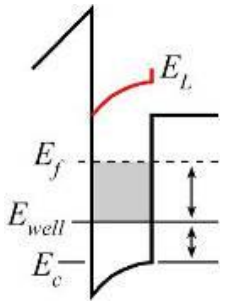
Calculating Current in Ballistic Limit

Natori



Channel Fermi voltage = voltage applied to c_{dos}

determines Fermi velocity v_f through $E_f = qV_f = m^* v_f^2/2$



mean electron velocity $= \bar{v} = (4/3\pi)v_f$

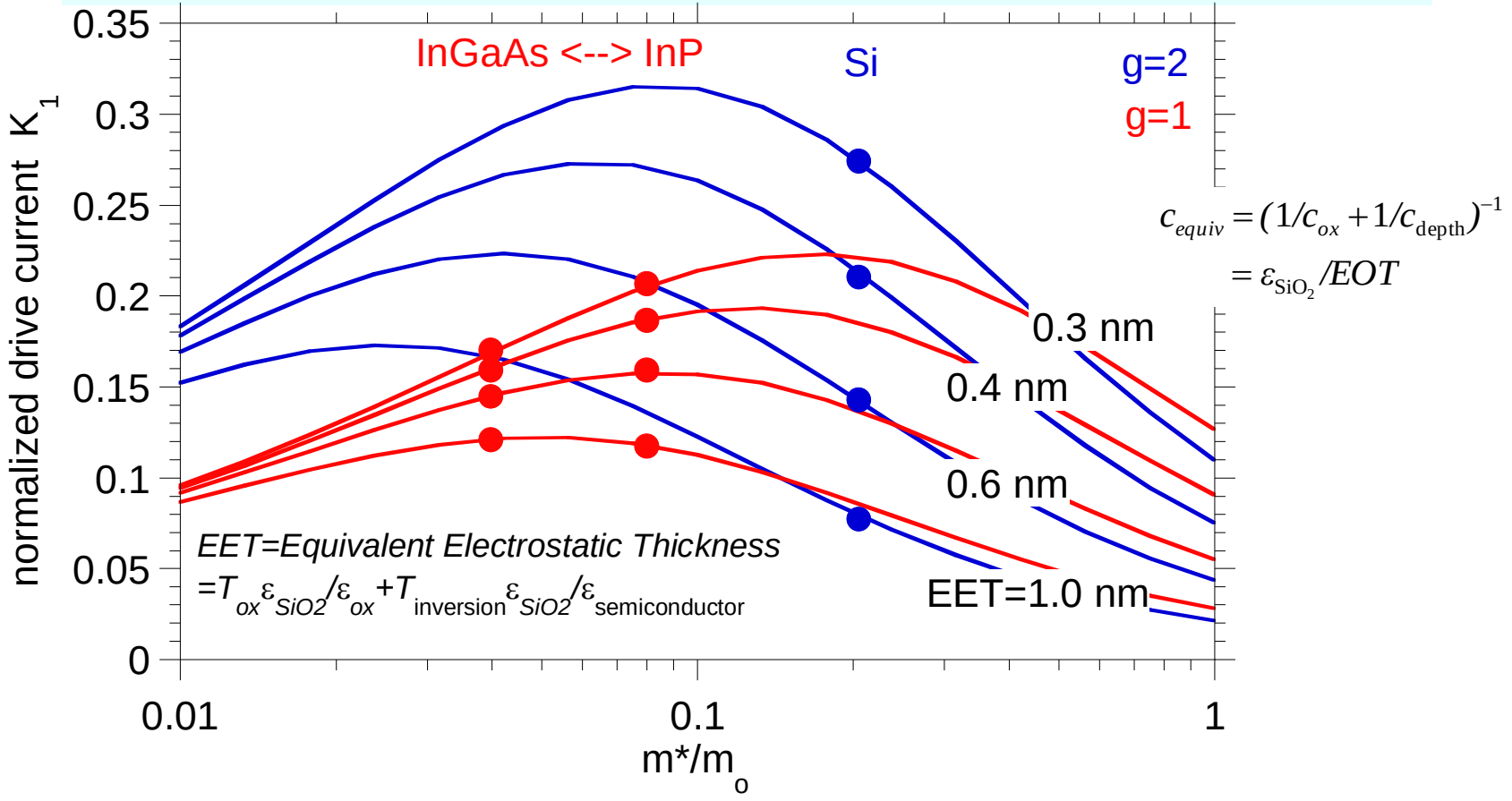
Channel charge : $\rho_s = c_{dos}(V_f - V_c) = \dfrac{c_{dos}c_{equiv}}{c_{equiv}+c_{dos}}(V_{gs}-V_{th})$

$c_{dos} = q^2 g m^*/2\pi\hbar^2 = c_{dos,o}\cdot g\cdot(m^*/m_o)$, where g is the # of band minima

$$\Rightarrow \quad J = \left(84\,\frac{\text{mA}}{\mu\text{m}}\right)\frac{g\cdot(m^*/m_o)^{1/2}}{\left(1+(c_{dos,o}/c_{ox})\cdot g\cdot(m^*/m_o)\right)^{3/2}}\left(\frac{V_{gs}-V_{th}}{1\,\text{V}}\right)^{3/2}$$

Drive current versus mass, # valleys, and EOT

$$J = K_1 \cdot \left(84\frac{\text{mA}}{\mu\text{m}}\right) \cdot \left(\frac{V_{gs} - V_{th}}{1\text{V}}\right)^{3/2}, \quad \text{where } K_1 = \frac{g \cdot \left(m^*/m_o\right)^{1/2}}{\left(1 + (c_{dos,o}/c_{equiv}) \cdot g \cdot (m^*/m_o)\right)^{3/2}}$$



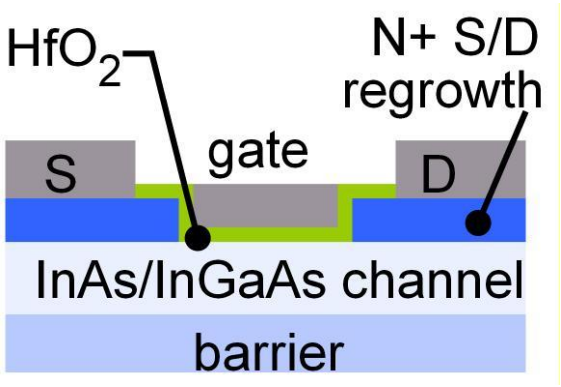
$$c_{equiv} = (1/c_{ox} + 1/c_{\text{depth}})^{-1} = \varepsilon_{SiO_2}/EOT$$

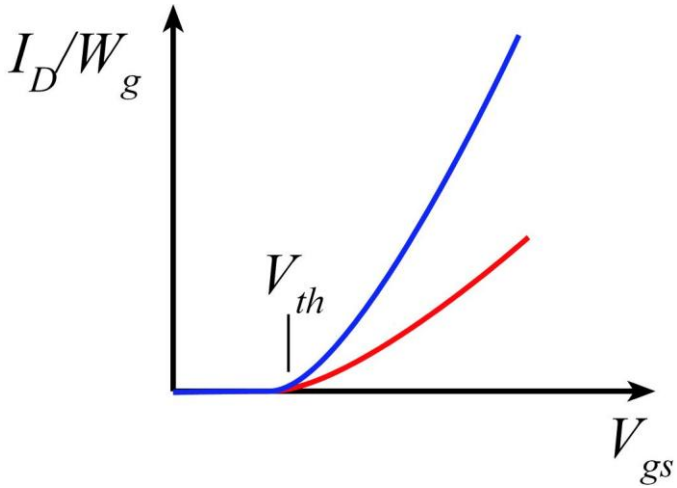
InGaAs MOSFETs: superior I_d to Si at large EOT.

InGaAs MOSFETs: inferior I_d to Si at small EOT.

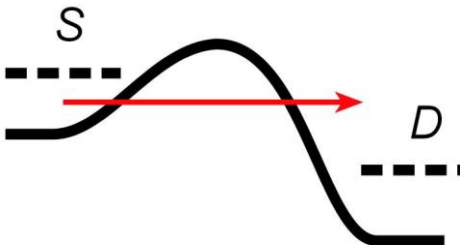
Why III-V MOS ?

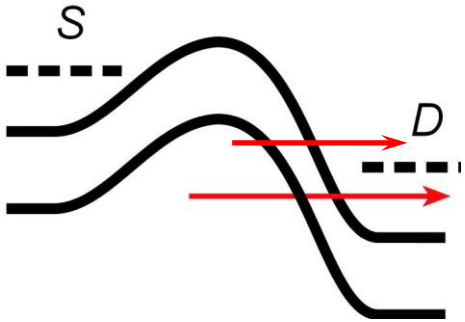
III-V vs. Si: Low m*→ higher velocity. Fewer states→ less scattering → higher current. Can then trade for lower voltage or smaller FETs.



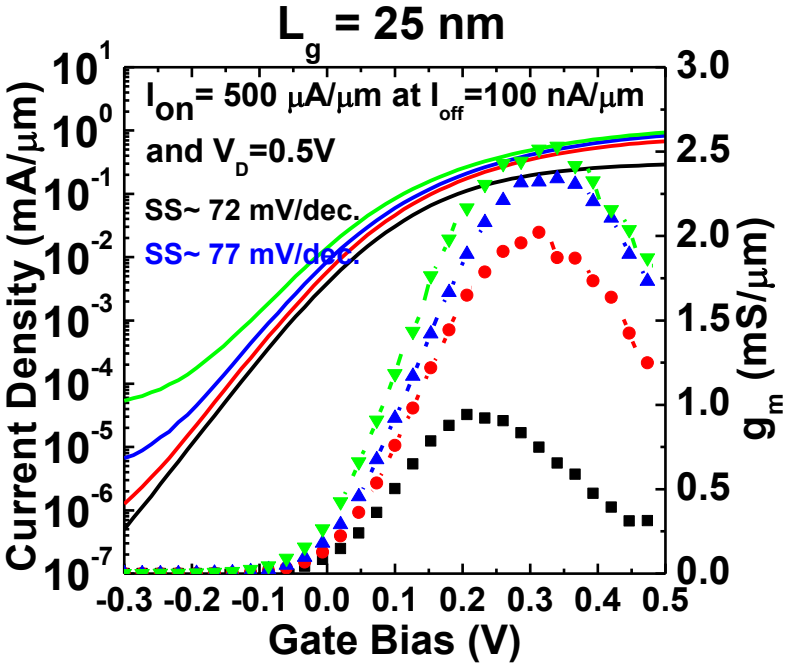


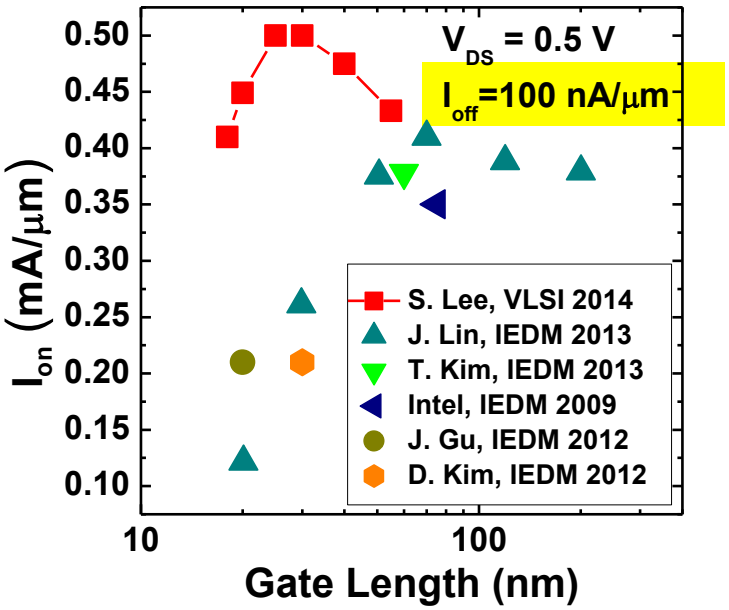
Problems: Low m*→ less charge. Low m* → more S/D tunneling. Narrow bandgap→ more band-band tunneling, impact ionization.





InGaAs/InAs FETs are leaky!





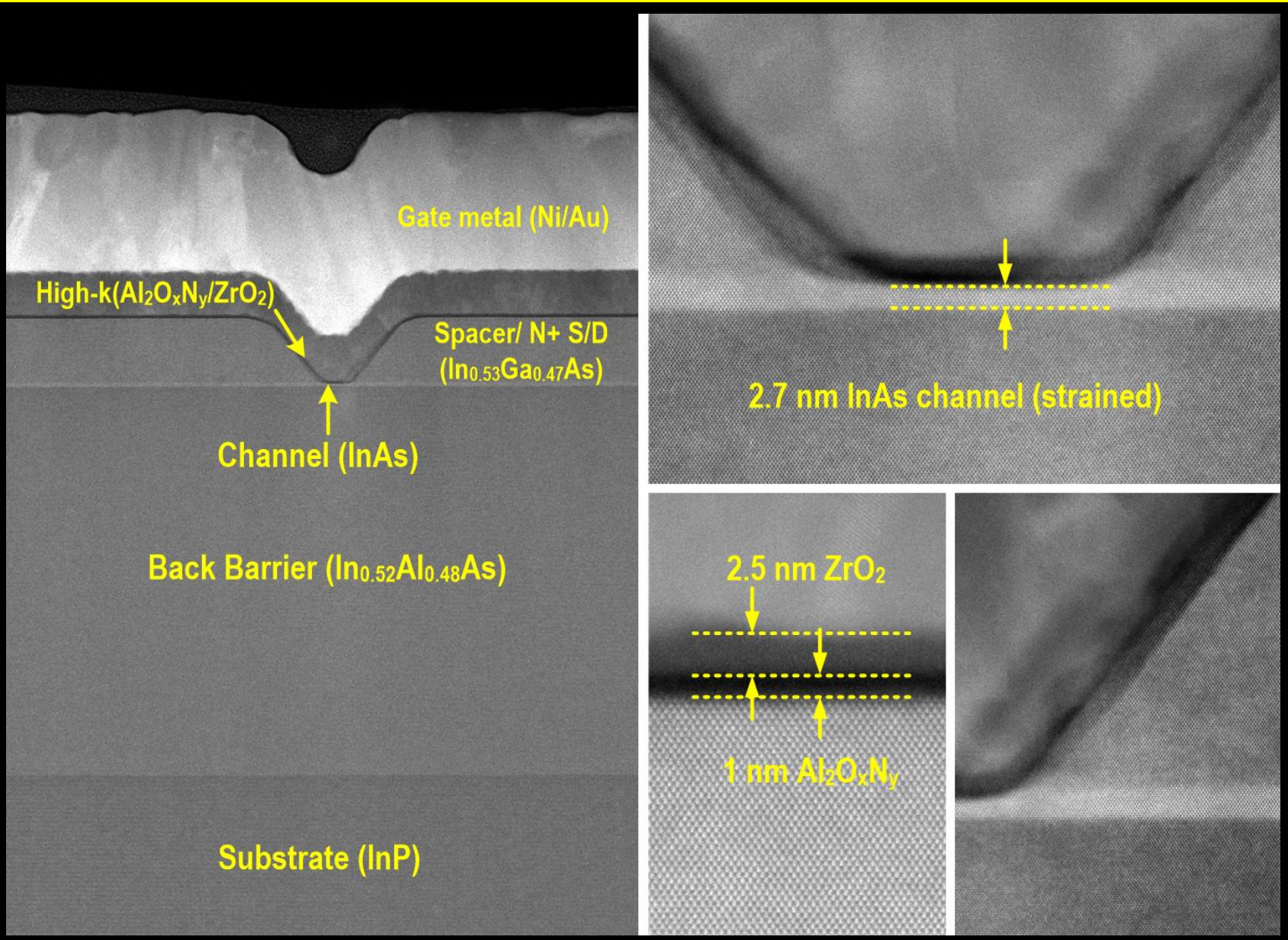
HP = High Performance: I_{off}=100 nA/μm

GP = General Purpose: I_{off}=1 nA/μm

LP = Low Power: I_{off}=30 pA/μm

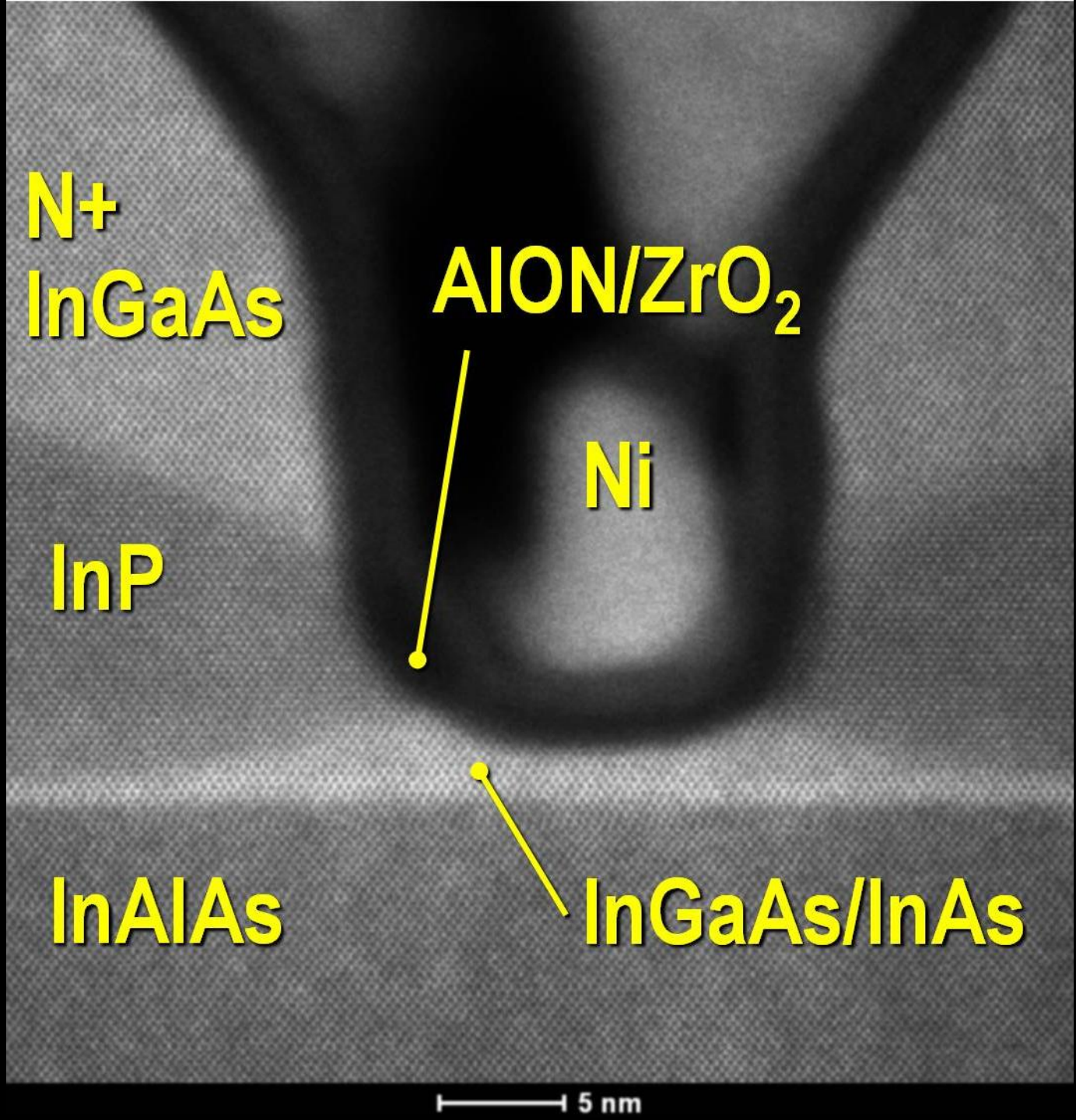
ULP = Ultra Low Power: I_{off}=10 pA/μm

III-V MOSFET



Courtesy of S. Kraemer (UCSB) *Heavy elements look brighter Lee *et al.*, 2014 VSLI Symposium

III-V MOSFET

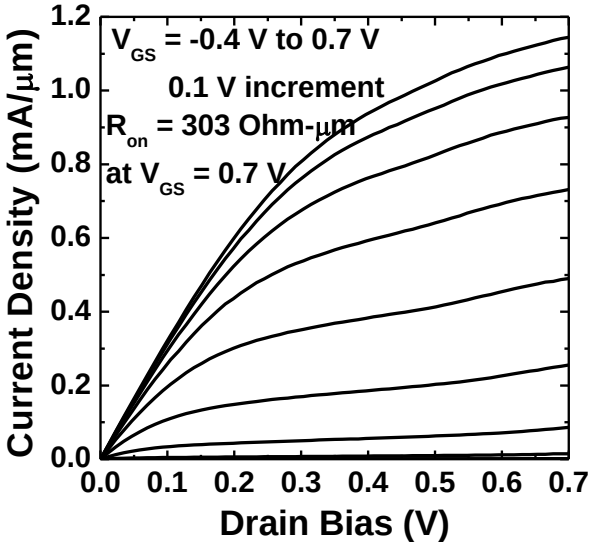


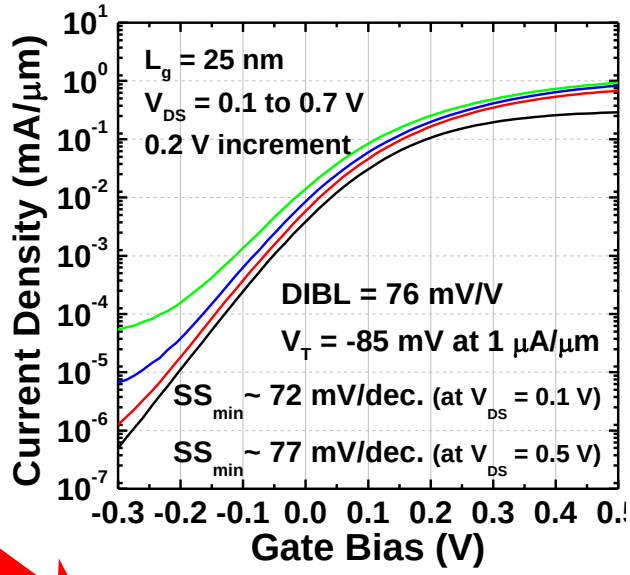
Huang *et al.*, 2015 DRC

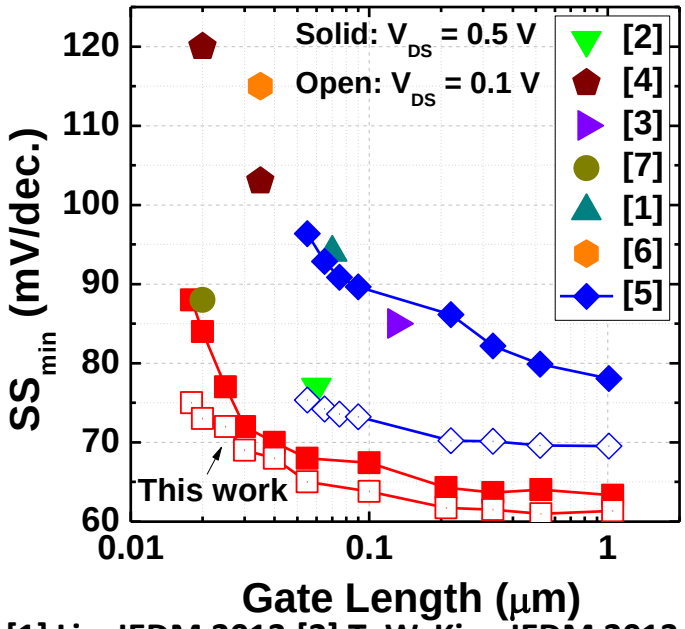
Courtesy of
S. Kraemer (UCSB)

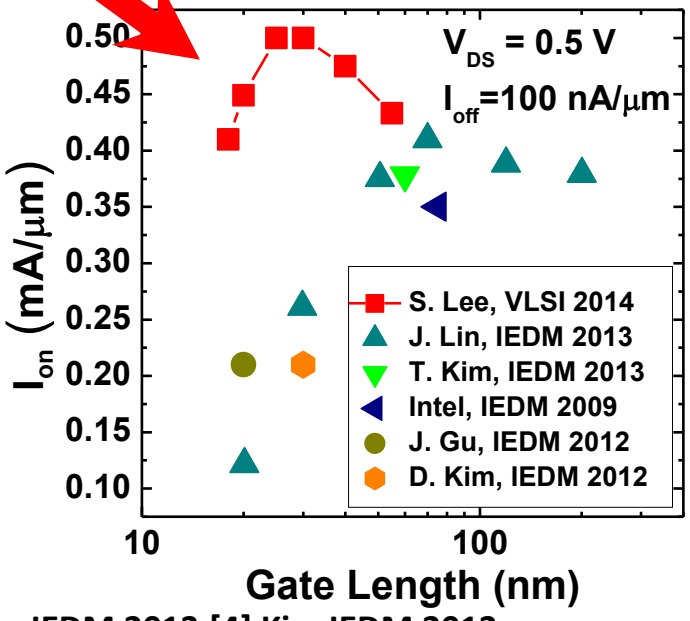
Reducing leakage: Ultra-thin channel

S. Lee et al., VLSI 2014



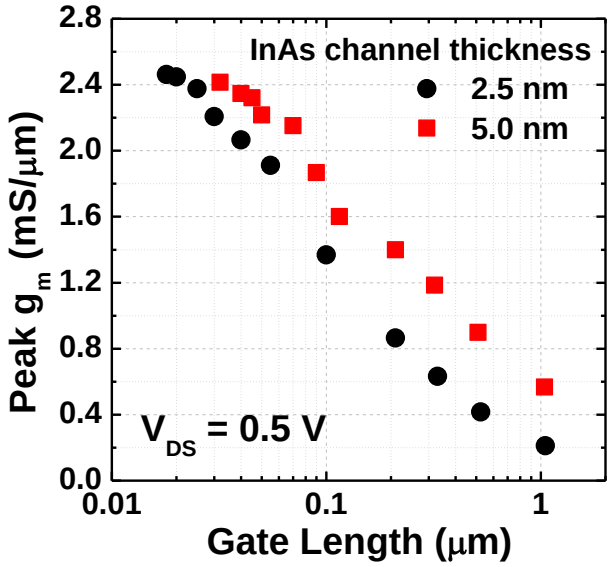


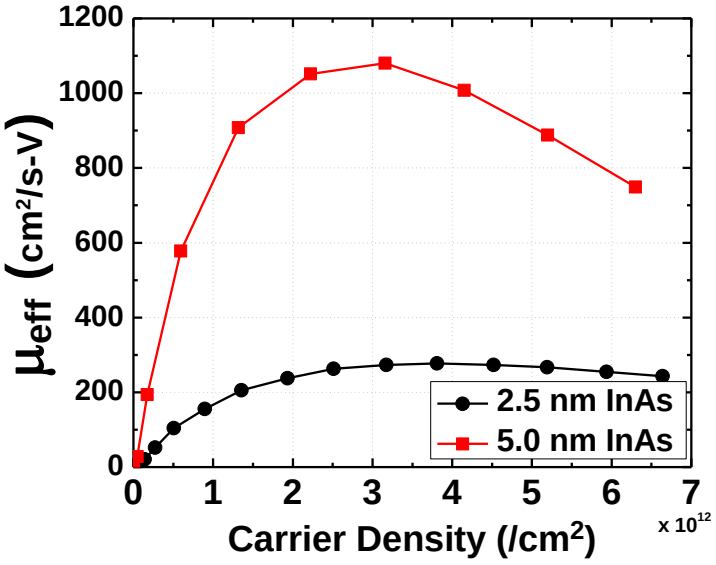


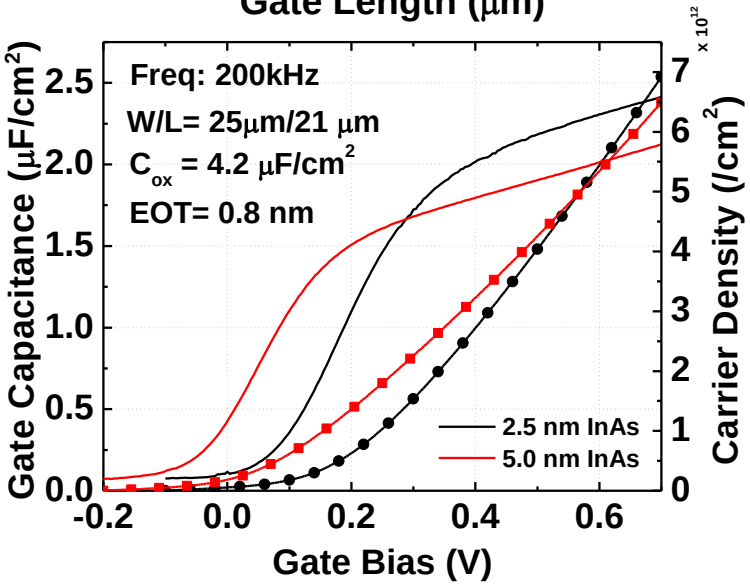


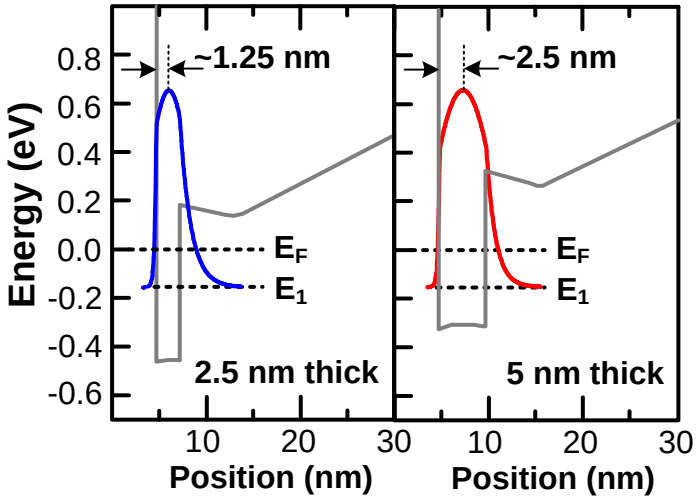
[1] Lin IEDM 2013,[2] T.-W. Kim IEDM 2013,[3] Chang IEDM 2013,[4] Kim IEDM 2013
[5] Lee APL 2013 (UCSB), [6] D. H. Kim IEDM 2012,[7] Gu IEDM 2012,[8] Radosavljevic IEDM 2009

On-state comparison: 2.5 nm vs. 5.0 nm-thick InAs channel









1D-Possion Schrodinger solver
(coded by W. Frensley, UT Dallas)

Wrap-Up

What are the challenges

Small dimensions

Thin semiconductor layers (2-3nm)

Extremely low resistance contacts

High current densities

Very thin dielectrics

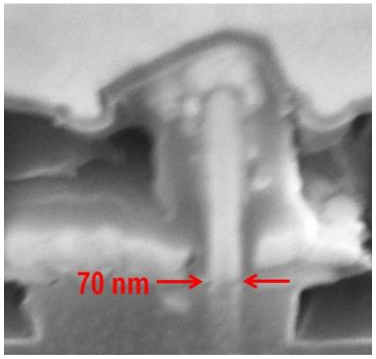
Available semiconductor states (III-Vs)

Resistances in interconnects and electrodes

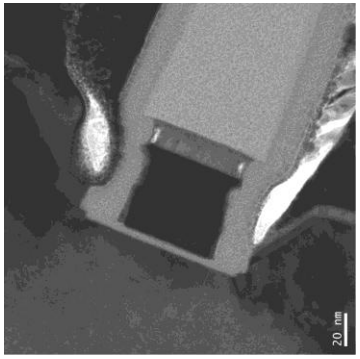
Where lies the future of electronics ?

"End of the scaling roadmap" ***"More than Moore"***

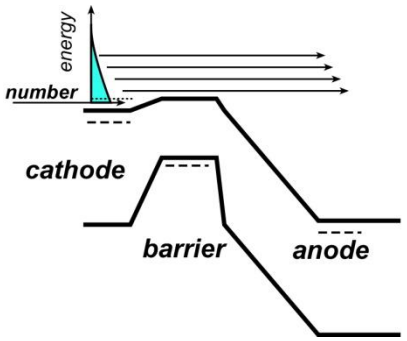
tubes



bipolar transistors

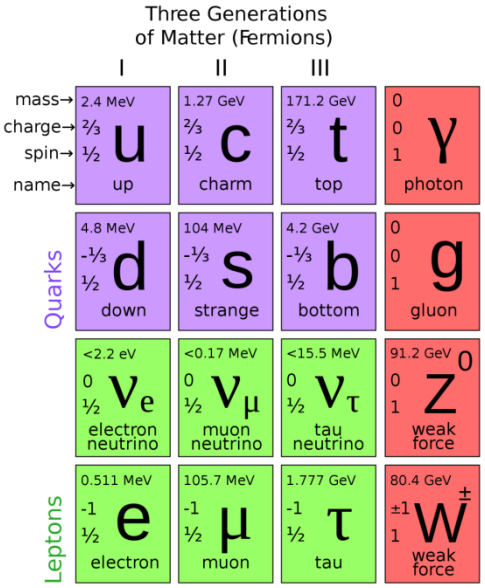


field-effect transistors



electrostatic barrier

Time for a new approach ?



Gravitonics ?

Quarkonics ?

Neutrinonics ?

Mechanics ?

1000's of heavy Nuclei ...

Magnetics ?

$F \sim v_1 v_2 / c^2 \rightarrow$ weak!

Charge-control: fundamentally best

electrons are light
electrostatic force: strong & long-range
electromagnetic waves: strong, long-range
→ electromagnetic interconnects
(wires !) are best: efficient, dense

The future of electronics:

classic charge-control devices
few-nm dimensions,
10^{12}-scale integration
multi-THz bandwidths

(backup slides follow)