2019 IEEE SISC Conference, December 13, San Diego

Future Directions in > 100 GHz Devices

Mark Rodwell, Yihao Fang, Brian Markman, Hsin-Ying Tseng
University of California, Santa Barbara
rodwell@ece.ucsb.edu

This work was supported in part by the Semiconductor Research Corporation (SRC), DARPA, and the NSF

2019 IEEE SISC Conference, December 13, San Diego

~~Future Directions in > 100 GHz Devices~~

Materials Requirements for Future Transistors

Mark Rodwell, Yihao Fang, Brian Markman, Hsin-Ying Tseng
University of California, Santa Barbara
rodwell@ece.ucsb.edu

This work was supported in part by the Semiconductor Research Corporation (SRC), DARPA, and the NSF

Transistor design and materials requirements

Transistors for VLSI

Large I_{on}, small I_{off}, low V_{DD}, small footprint… hard !

MOSFETs, TFETs.

Transistors for wireless

High (f_{τ},f_{max}) , low noise, high power, high efficiency.

InP HBTs, InP MOS-HEMTs, (GaN HEMTs, SiGe HBTs)

Transistors for power switching:

high speed, high voltage, high current,

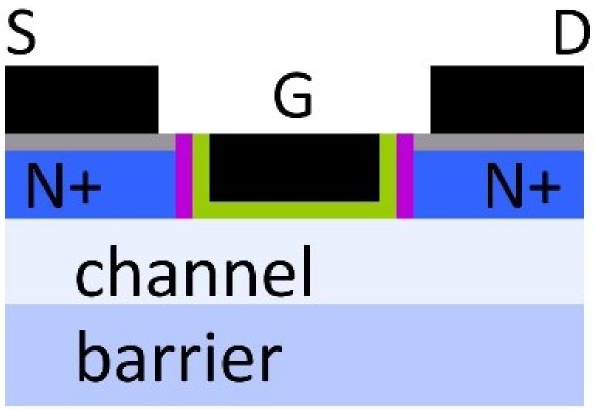
GaN, SiC, Si LDMOS

…not my field.

nm FETs:
MOS and TFETs

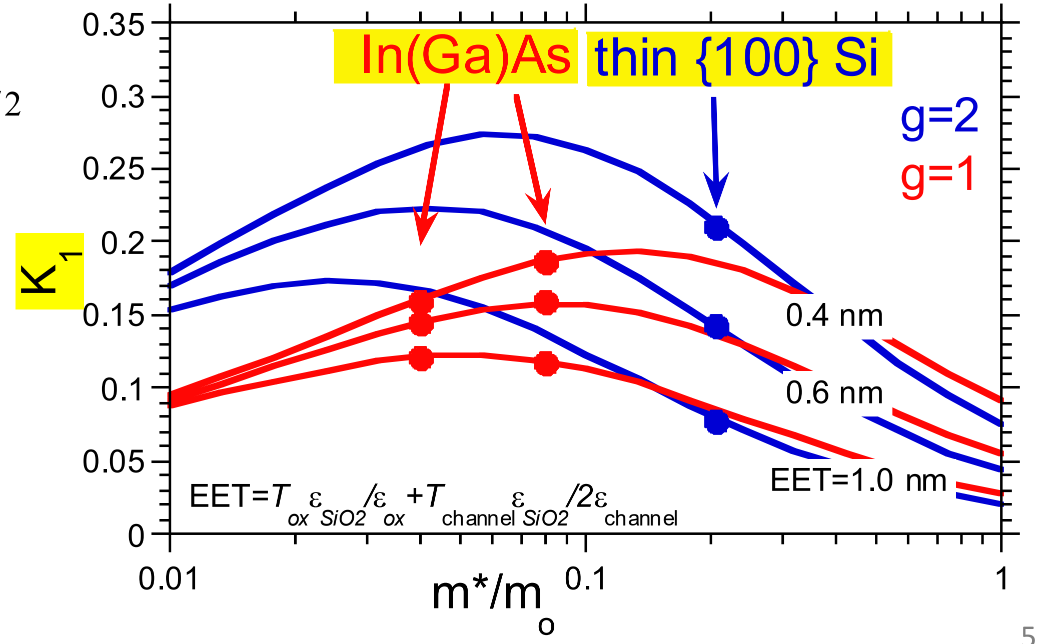
MOSFETs for VLSI

Goals: large I_{on}, small I_{off}, low V_{DD}, small footprint
Minimum L_g set by minimum equivalent oxide thickness. $Zr_xHf_{1-x}O_2$?
Minimum contact size set by contact resistivity ($<0.3\ \Omega\text{–}\mu m^2$)
Ballistic I_{on} set by band structure: EOT, m*, # valleys
I_{off} is degraded by low bandgap
Ideal: $m^*=0.1m_e$, 2-3 valleys, $E_g>1$eV, grades to small-E_g contact layers

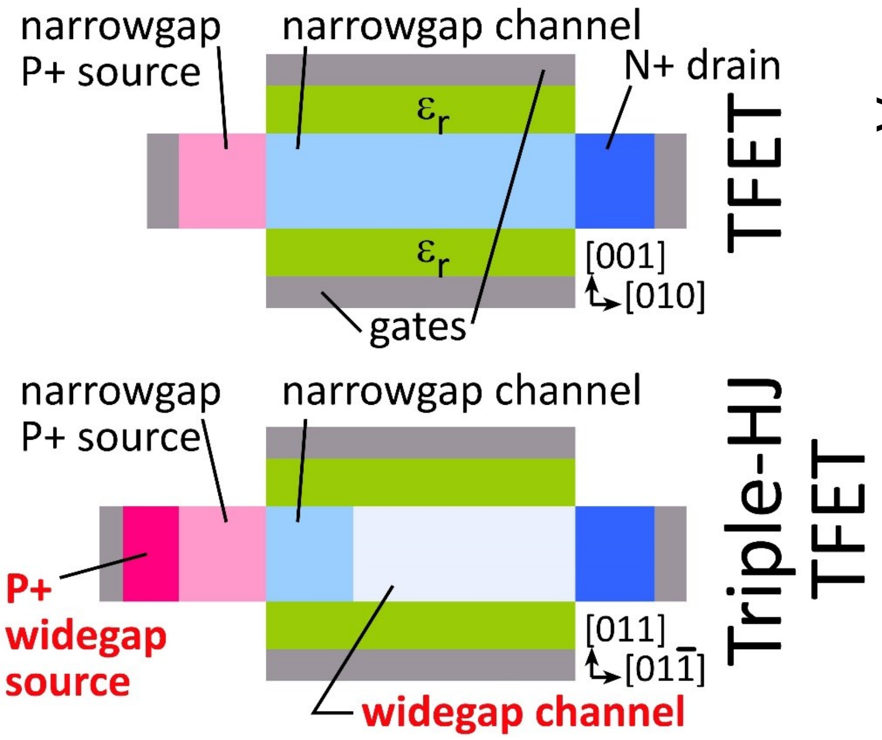


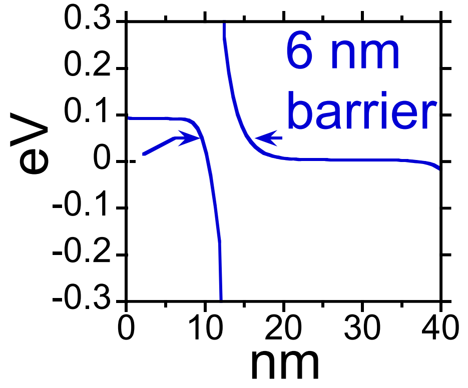
low-K dielectric spacer
high-K gate dielectric

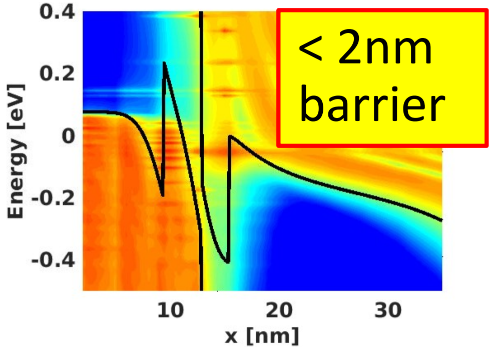
$$J = K_1 \cdot \left(\frac{84\text{mA}}{\mu\text{m}}\right) \cdot \left(\frac{V_{gs}-V_{th}}{1\text{V}}\right)^{3/2}$$

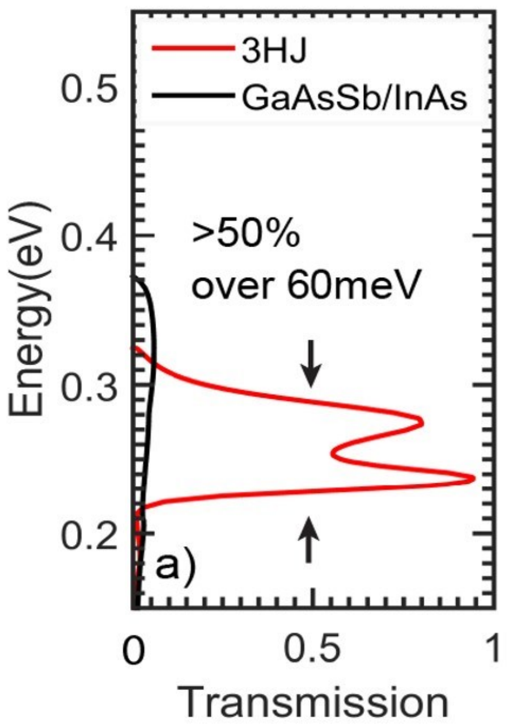


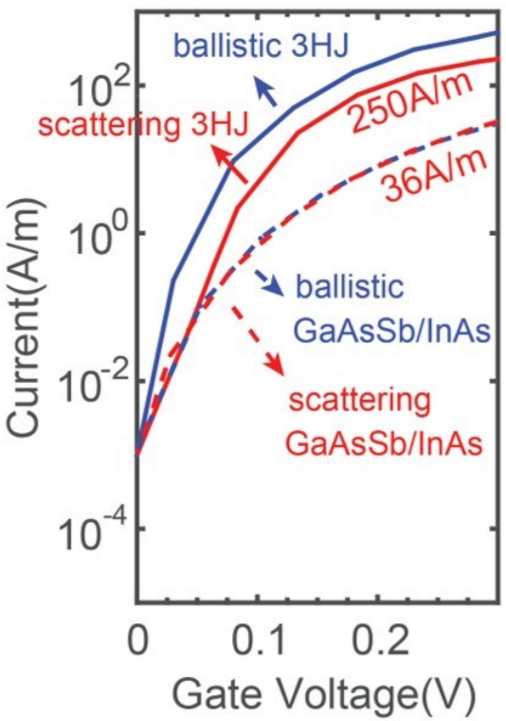
High-current triple-heterojunction TFETs for VLSI











Simulations: M. Povoloski, Purdue

TFETs: steep SS ✓, low I_{on}; ✗

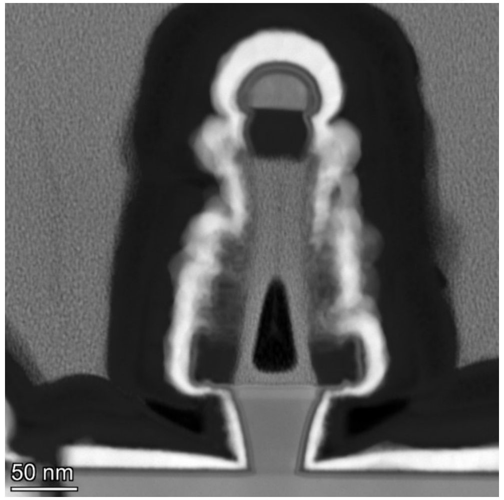
Triple-HJ TFET: steep SS ✓, high I_{on}; ✓

Materials needs:

Direct bandgap, large band offsets, small tunnel barrier, low m*

Low D_{it} for N & P materials. (N-InAs ✓, N-InP ✓, P-GaAsSb ?, P-InGaAs ?)

small dielectric EOT, low ρ contacts

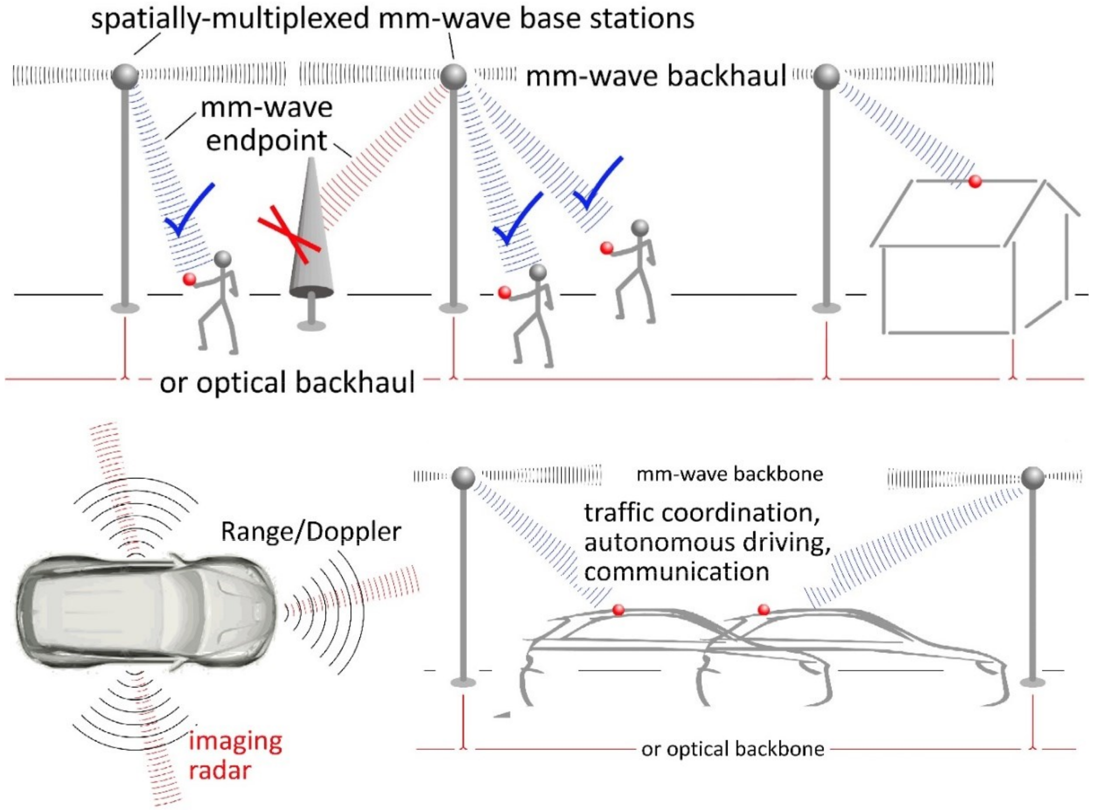


High-Frequency Transistors

Beyond-5G Wireless: 100-300GHz

10Gb mobile communications:
Unlimited information, anywhere.
Capacity well beyond 5G.

TV-resolution wireless imaging:
See, fly, drive perfectly in any conditions.



COMMUNICATIONS SENSING TERAHERTZ

Debdeep Jena, Alyosha Molnar, Christoph Studer, Huili Xing*: Cornell University*

Dina Katabi*: MIT*

Sundeep Rangan*: New York University*

Amin Arbabian, Srabanti Chowdhury*: Stanford*

Elad Alon, Ali Niknejad, Borivoje Nikolic, Vladimir Stojanovic*: University of California, Berkeley*

Gabriel Rebeiz*: University of California, San Diego*

Jim Buckwalter, Upamanyu Madhow, Umesh Mishra, Mark Rodwell*: University of California, Santa Barbara*

Andreas Molisch, Hossein Hashemi*: University of Southern California*

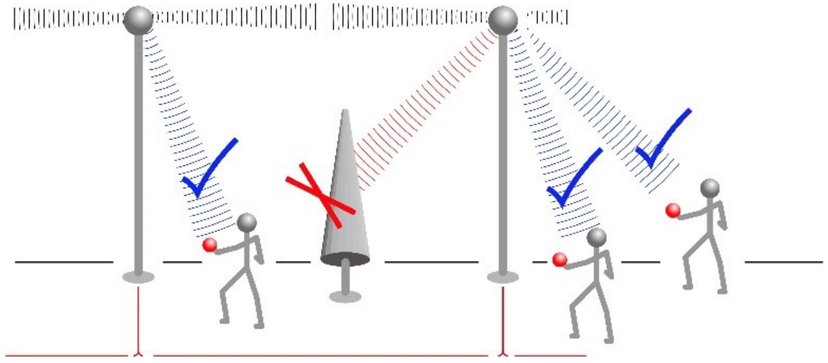
Harish Krishnaswamy*: Columbia University*

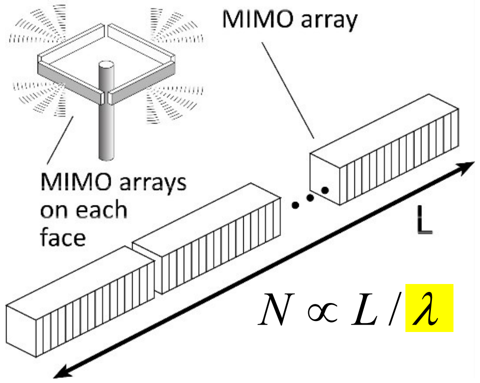
Kenneth O*: University of Texas, Dallas*

Target applications

MIMO hub:

140GHz, 128 beams/face, 10Gb/s/user, 100m

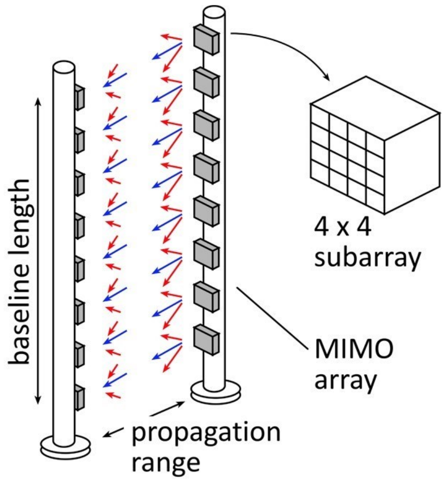




$N \propto L/\lambda$

Point-point MIMO:

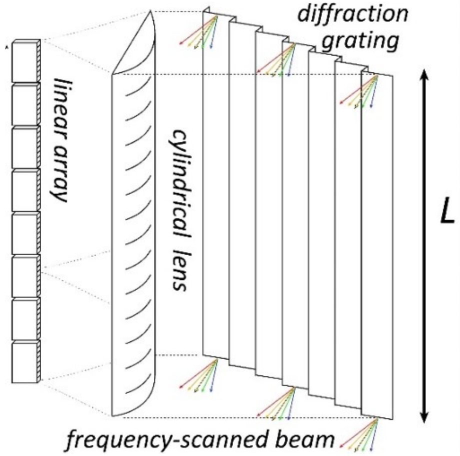
240GHz, 340GHz: 640Gb/s, 250-500m



$N \propto L^2/\lambda R$

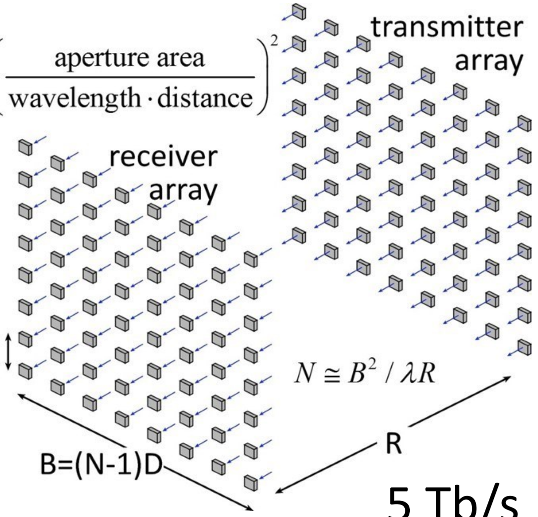
Hardware-efficient imaging:

240GHz, 340 GHz, 300m, 512×64 image, 60Hz, 15dB SNR



$\Delta\theta \propto \lambda / L$

$$\#\text{ channels} \propto \left(\frac{\text{aperture area}}{\text{wavelength}\cdot\text{distance}}\right)^2$$



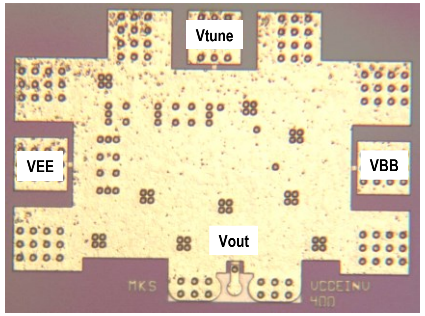
$N \cong B^2/\lambda R$

5 Tb/s !

A few high-frequency ICs

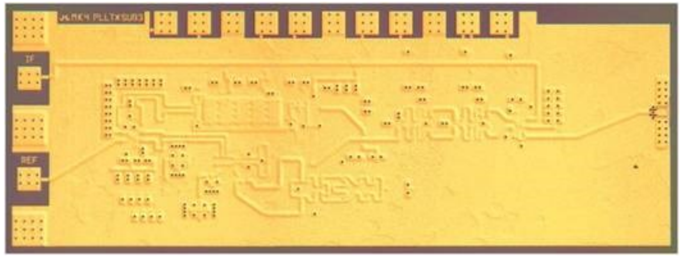
570 GHz fundamental oscillator; InP HBT

M. Seo, TSC / UCSB
JSSC, 2011



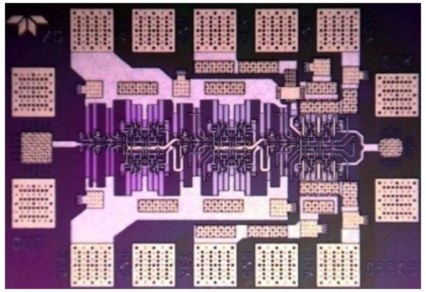
600 GHz Integrated Transmitter; InP HBT

PLL + Mixer+Amplifier
M. Seo TSC
2012 IMS



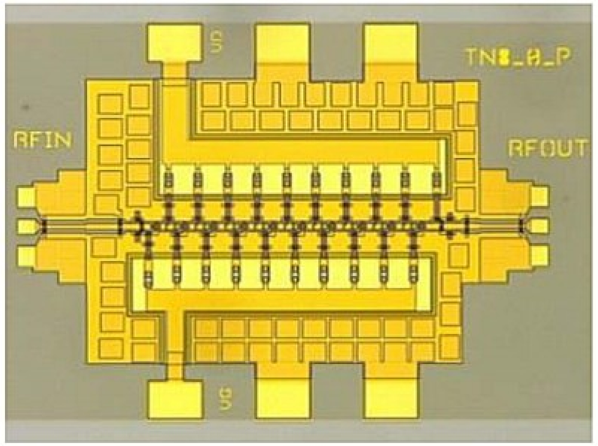
204 GHz static frequency divider (ECL master-slave latch); InP HBT

Z. Griffith, TSC / UCSB
CSIC 2010



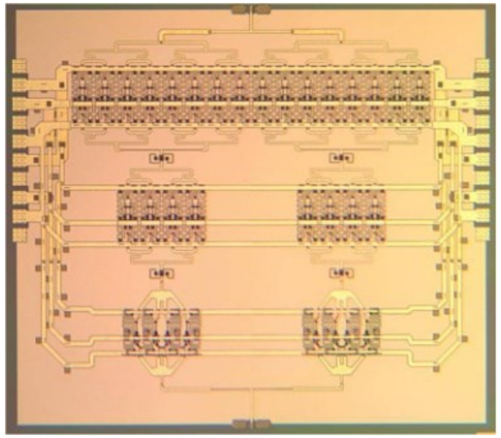
1.0 THz Amplifier; InP HEMT

Mei et al, Northrop-Grumman
EDL, 2015



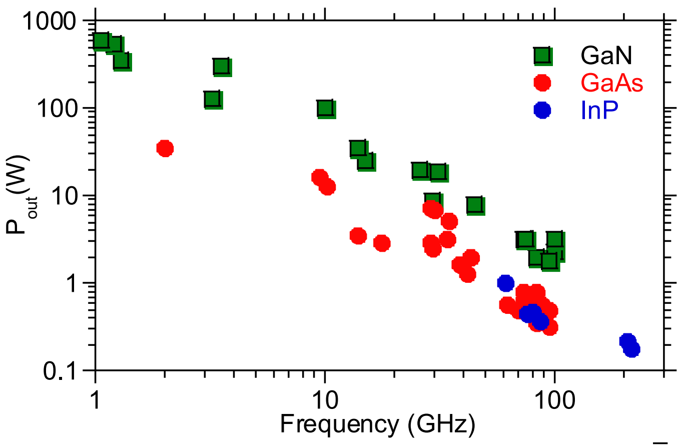
220 GHz 180 mW power amplifier; InP HBT

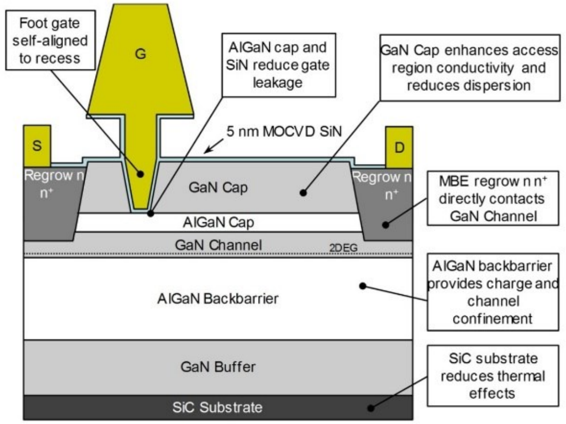
T. Reed, UCSB
Z. Griffith, TSC
CSICS 2013



Transistor development for 100-300GHz systems

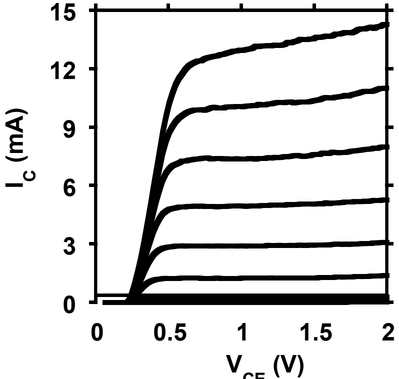
InGaN and GaN HEMTs:
High power from 100-340GHz
GaN: superior power density at all frequencies
UCSB/Mishra: InGaN for increased mobility
Cornell/Xing: AlN/GaN/AlN

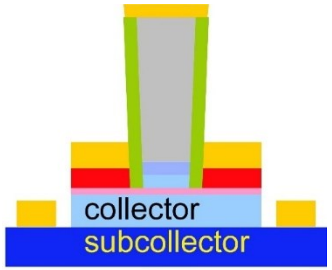




N-polar GaN: Mishra, UCSB

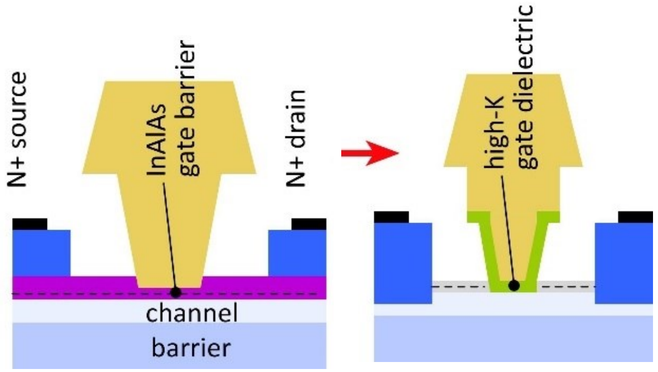
THz InP HBTs:
Efficient 100-650GHz power
more f_{max}: more efficient, higher frequencies
base regrowth: better contacts→ higher f_{max}.
status: working DC devices; moving to THz

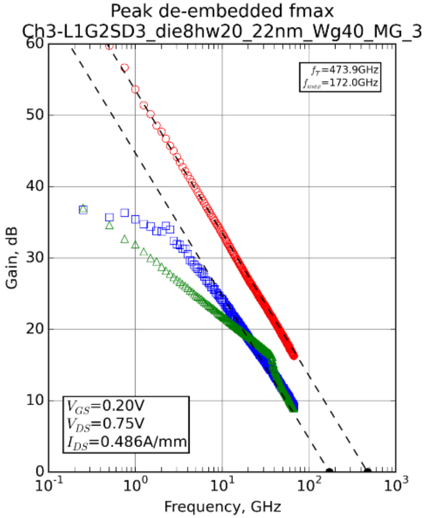




collector mesa, contacts, etc.

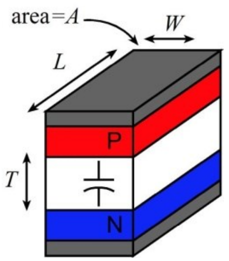
THz InP HBTs:
Sensitive 100-650GHz low-noise amplifiers
more f_τ: lower noise, higher frequencies
high-K gate dielectric → higher f_τ.
status: process modules



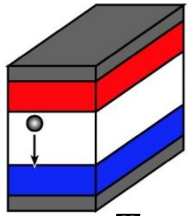


Transistor scaling laws: (V,I,R,C,τ) vs. geometry

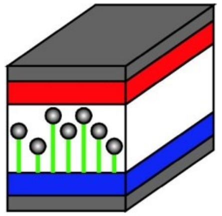
Depletion Layers



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

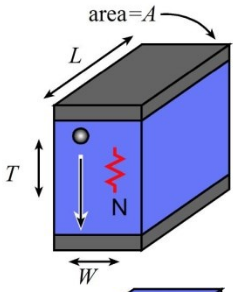


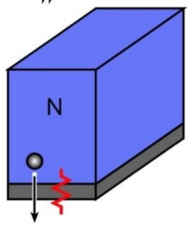
$$\tau = \frac{T}{2v}$$



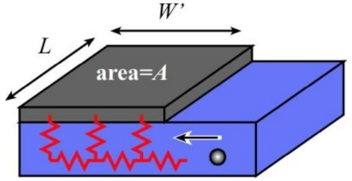
$$\frac{I_{\max}}{A} = \frac{4\varepsilon v_{sat}(V_{appl} + \phi)}{T^2}$$

Bulk and Contact Resistances





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

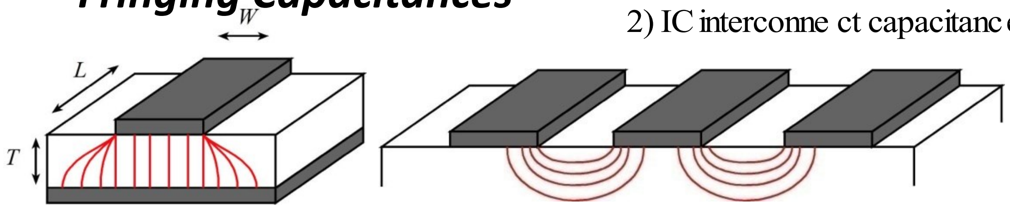


contact terms dominate

Fringing Capacitances

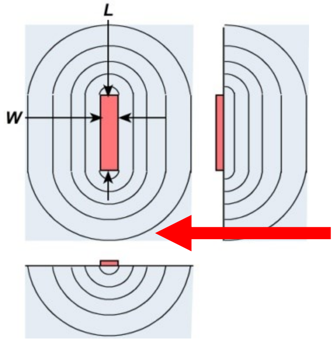
1) FET fringing capacitances

2) IC interconnect capacitances

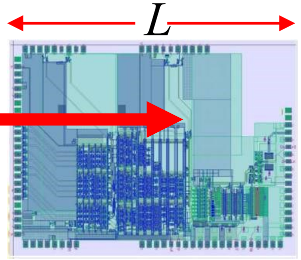


$$C_{finging} / L \sim \varepsilon \qquad C_{finging} / L \sim \varepsilon$$

Thermal Resistance

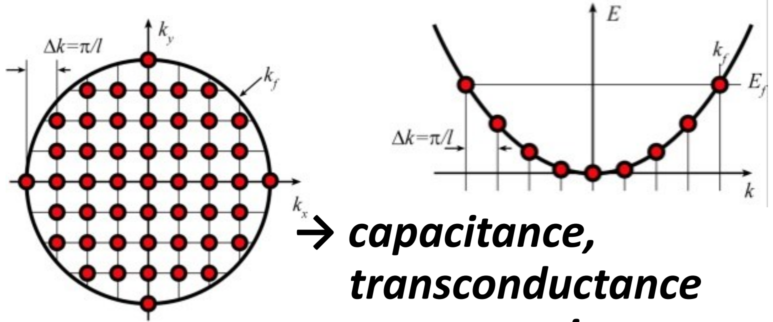


$$\Delta T_{IC} \propto \frac{P_{IC}}{K_{th} L}$$



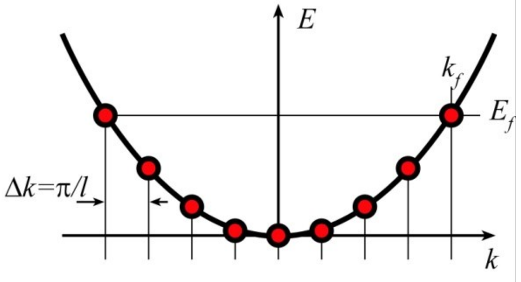
$$\Delta T_{\text{transistor}} \sim \frac{P}{\pi K_{th} L} \ln\left(\frac{L}{W}\right)$$

Available quantum states to carry current

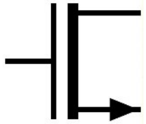


→ ***capacitance, transconductance contact resistance***

Degenerate State Density (Ballistic) Limits



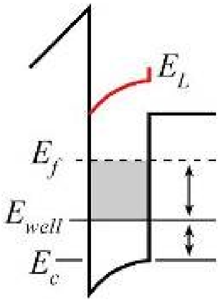
$$\text{Charge} = \int_{\text{Band edge}}^{\text{Fermi Energy}} q \cdot n(E)dE \qquad \text{Current} = \int_{\text{Band edge}}^{\text{Fermi Energy}} q \cdot v(E) \cdot n(E)dE$$



$$J_{sheet} \propto m^{1/2}(E_f - E_{well})^{3/2}$$

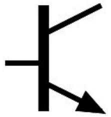
$$\text{not } (\mu c_{ox} / L_g)(V_{gs} - V_{th})^2$$

"ballistic limit"



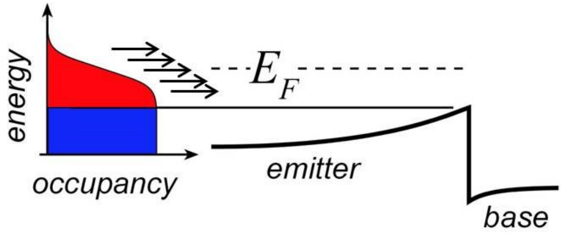
$$\rho_{sheet} = c_{dos}(V_{gs} - V_{th}) \propto m^{*}(E_f - E_{well})$$

"state density capacitance"

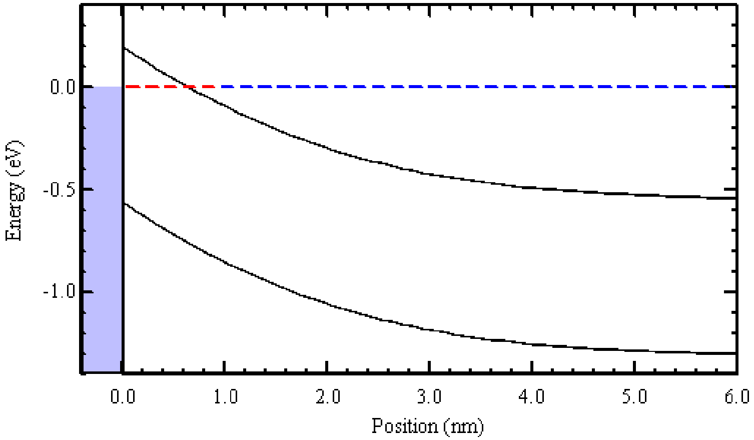


$$J \propto m^*(E_f - E_c)^2 \propto m^*(V_{be} - \varphi)^2$$

$$\text{not} \sim \exp(qV_{be} / kT)$$



Contacts



$$\rho_c \geq \frac{1}{n^{2/3}} \cdot \left(\frac{\hbar}{q^2}\right)\left(3\right)^{2/3}$$

Frequency Limits and Scaling Laws of (most) Electron Devices

$$\tau \propto \text{thickness}$$

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

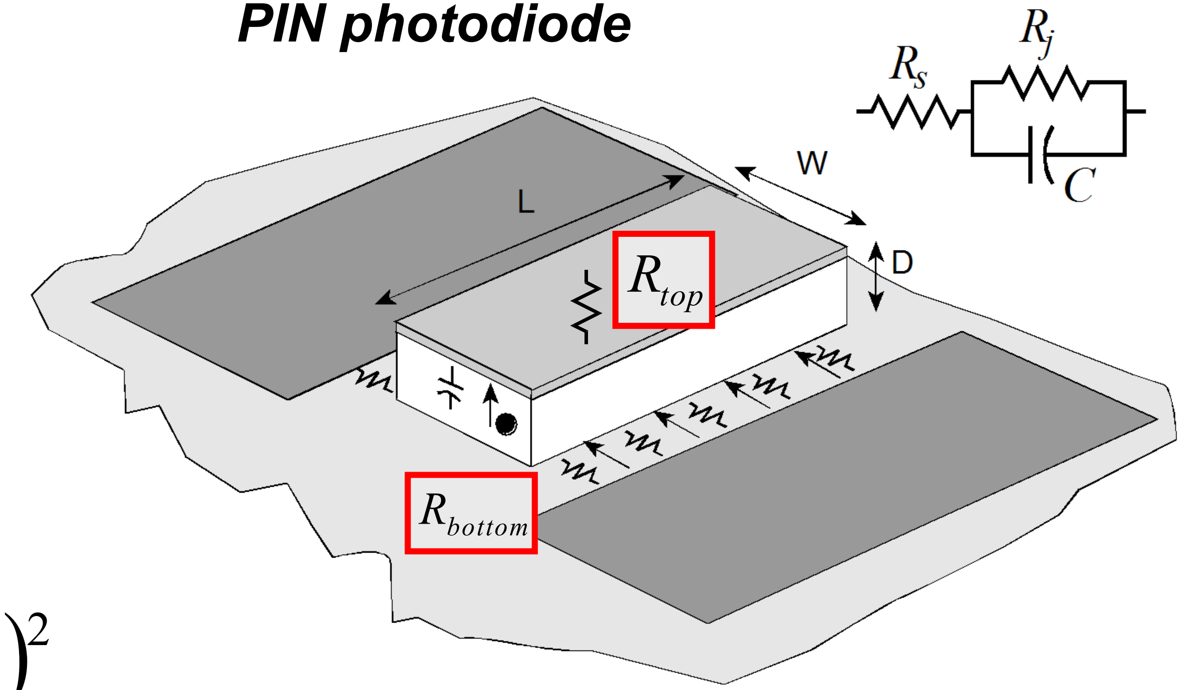
$$R_{top} \propto \rho_{contact} / \text{area}$$

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

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

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

PIN photodiode



To double bandwidth:

Reduce thicknesses 2:1
Improve contacts 4:1
Reduce width 4:1,
Keep constant length
Increase current density 4:1

THz Bipolar Transistors

Bipolar Transistor Design: Scaling

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

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

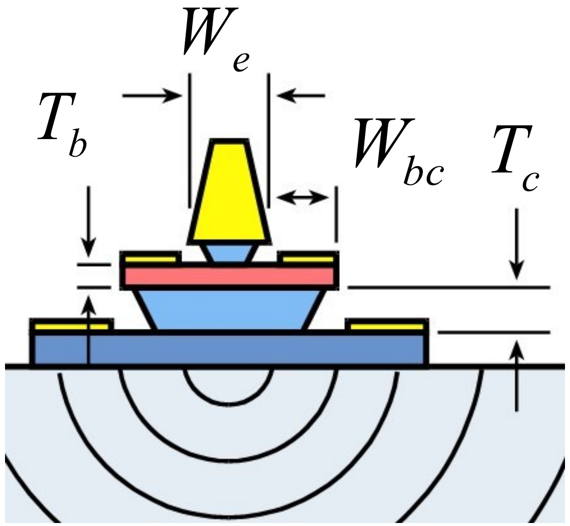
$$C_{cb} = \varepsilon A_c / T_c$$

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

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

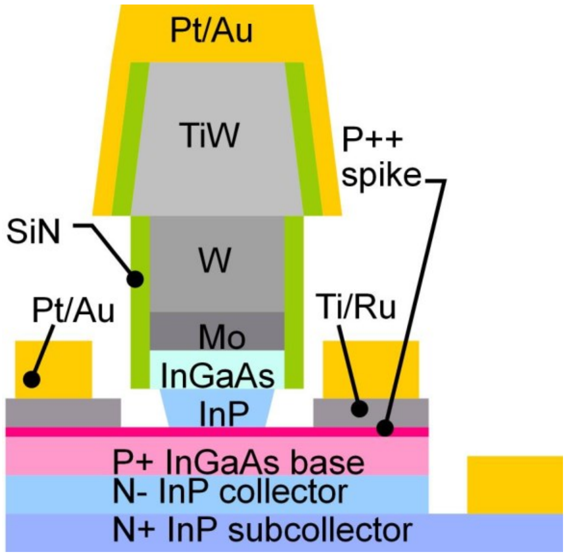
$$R_{ex} = \rho_{\text{contact}} / A_e$$

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



(emitter length L_E)

Bipolar Transistor Scaling Laws



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

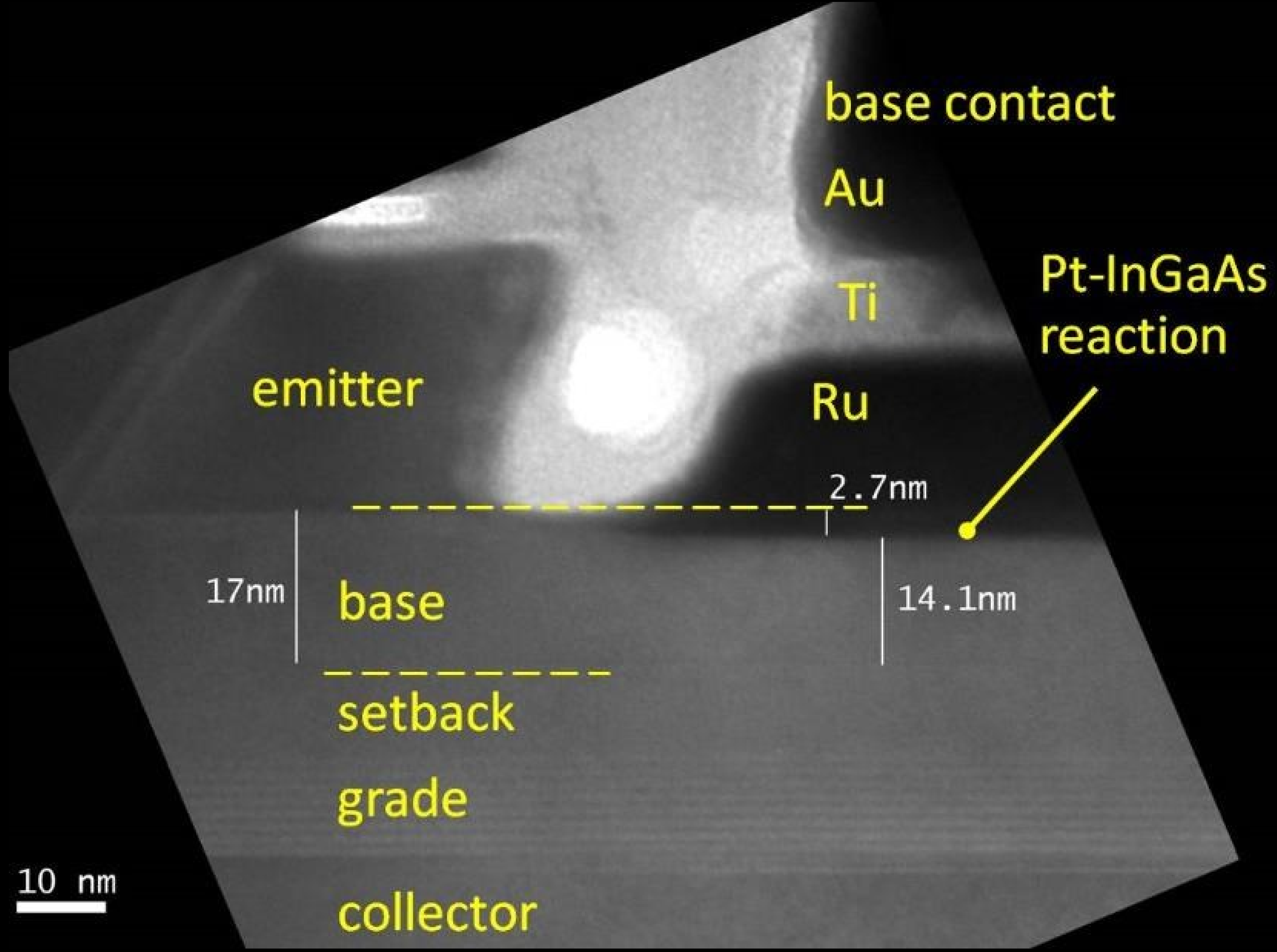
Narrow junctions.

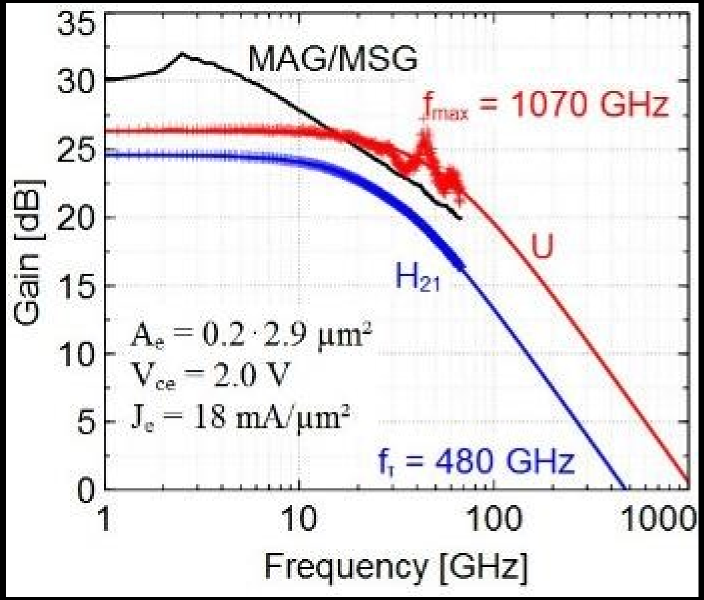
Thin layers

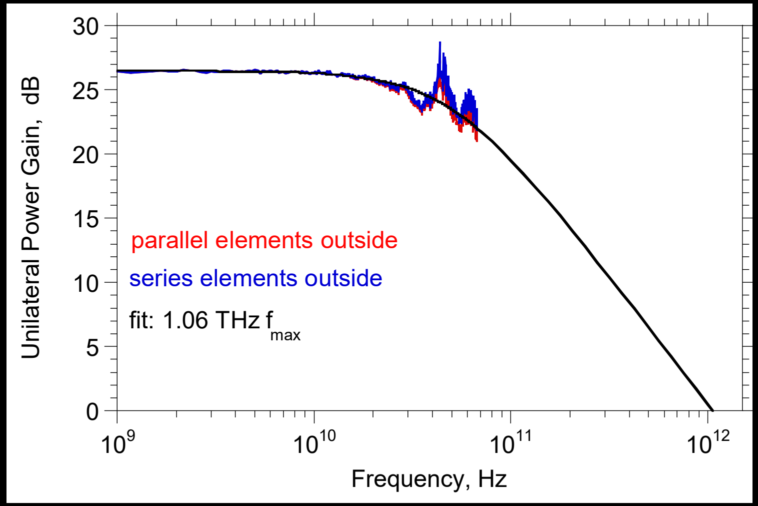
High current density

Ultra low resistivity contacts

InP HBTs: 1.07 THz @200nm







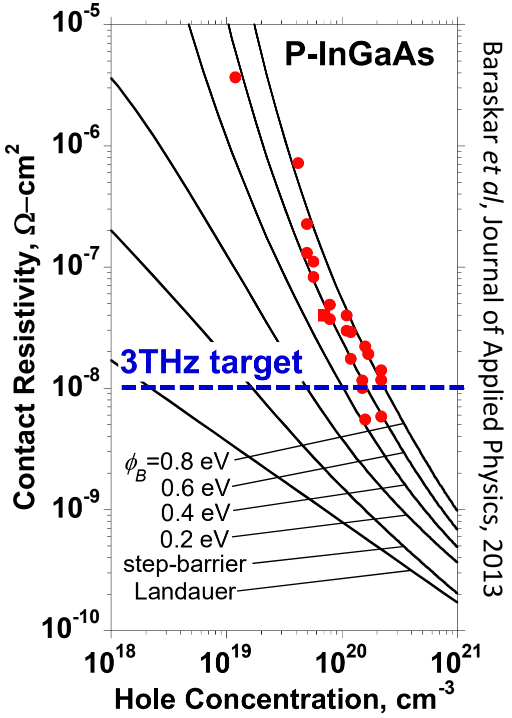
Rode et al., IEEE TED, Aug. 2015

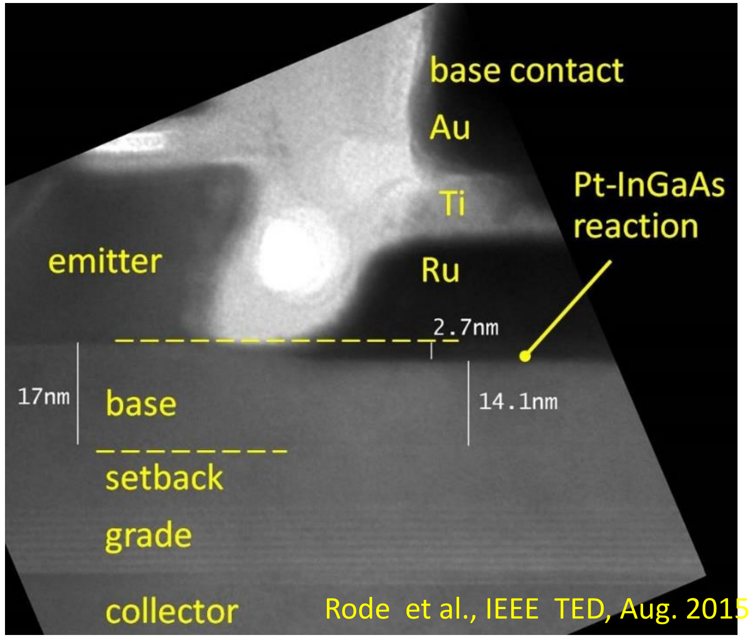
Challenges at the 64nm/2THz & 32nm/3THz Nodes

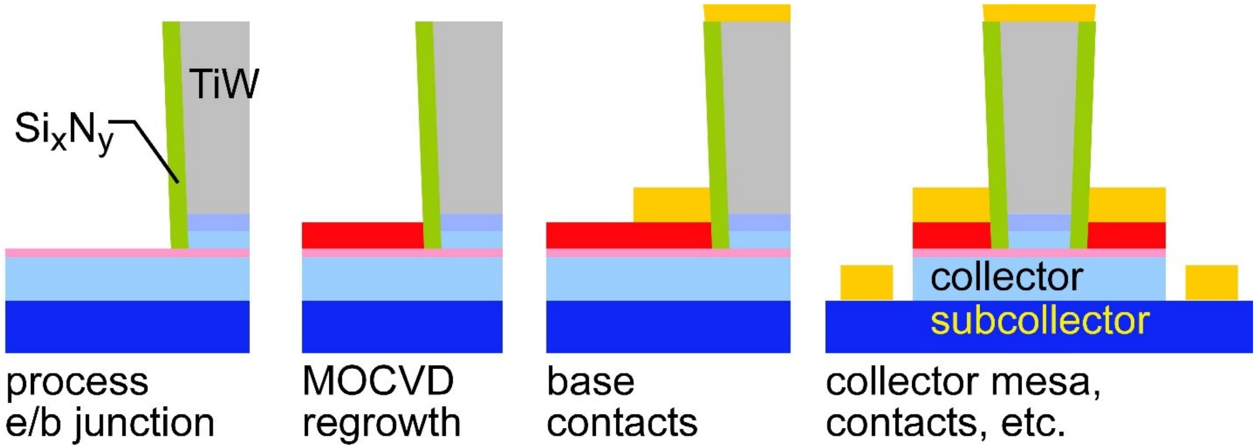
Need high base contact doping
$>10^{20}/\text{cm}^3$ for good contacts
high Auger recombination
very low β.

Need moderate contact penetration
Pd or Pt contacts
react with 3++ nm of base
penetrate surface contaminants
too deep for thin base

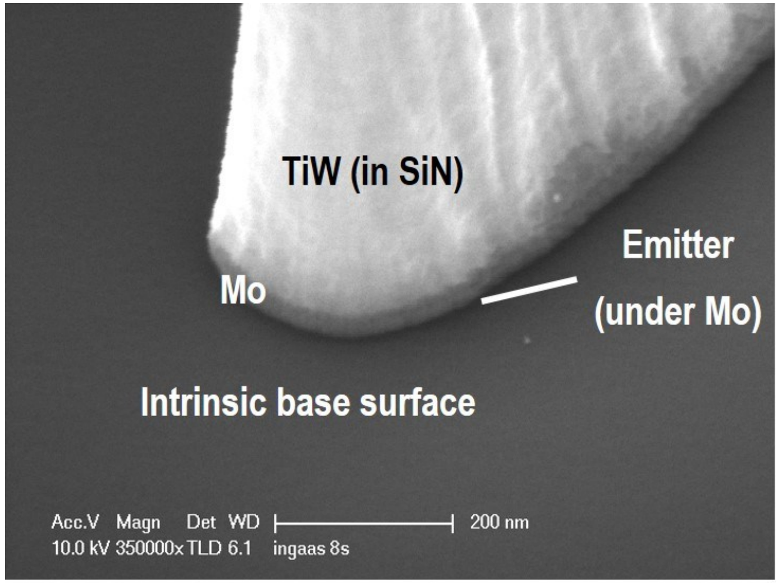
Solution: base regrowth:
thin, moderately-doped intrinsic base
InGaAs or GaAsSb @ 10^{19}-$10^{20}/\text{cm}^3$
thick, heavily-doped extrinsic base
P-GaAs, ~$10^{21}/\text{cm}^3$



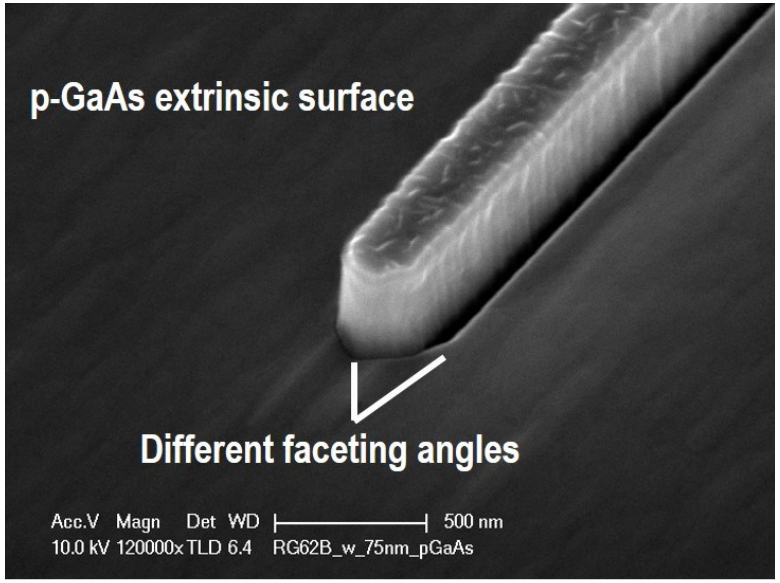




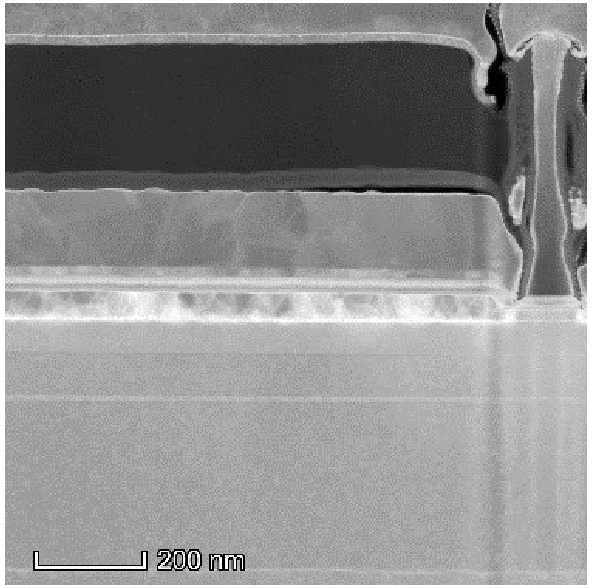
Regrown-Base InP HBTs: Images



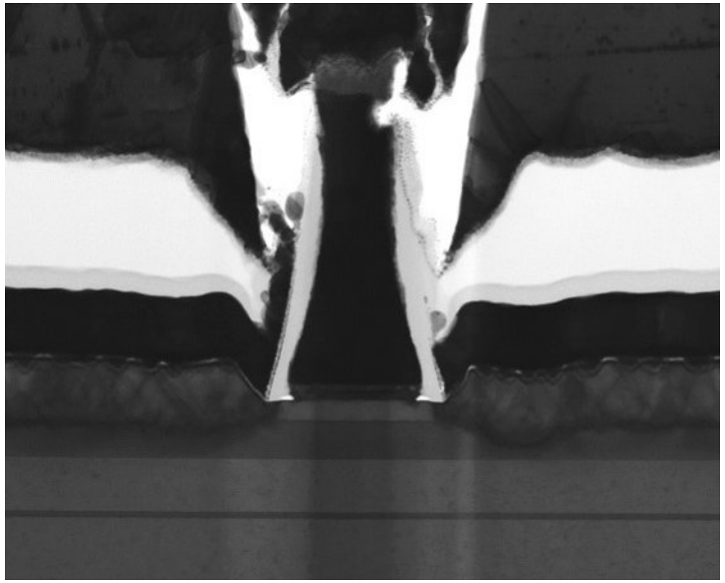
Before regrowth

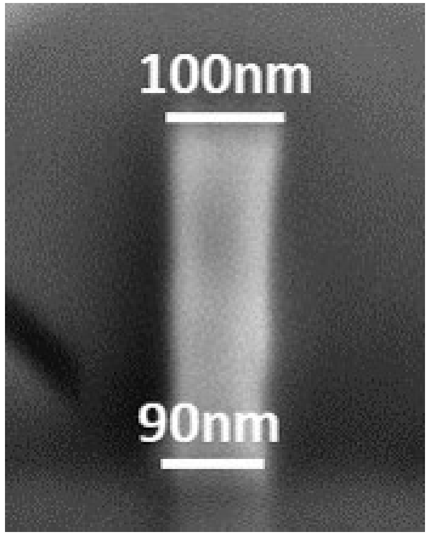


After 100nm p-GaAs regrowth



Cross-sections



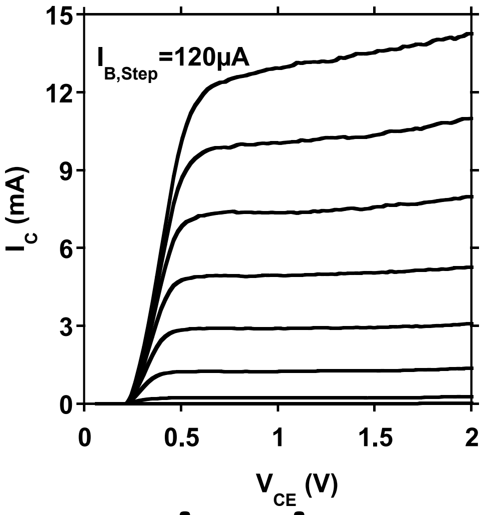


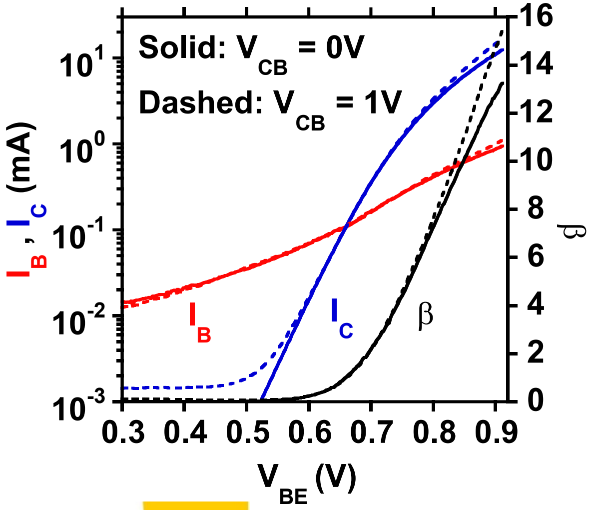
Dry-etched
TiW emitter contact

Regrown-Base InP HBTs: Status

Good DC data: even given regrowth

refractory Mo/W/TiW emitter contact maintains low ρ_c.





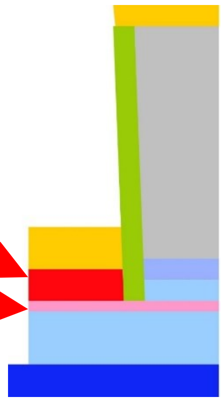
Excellent base contacts; but hydrogen base passivation

0.4 Ω–µm² resistivity for GaAs/metal contact ✓

290 Ω sheet resistivity for regrown base ✓

0.60Ω–µm² resistivity for InGaAs/GaAs contact ✗

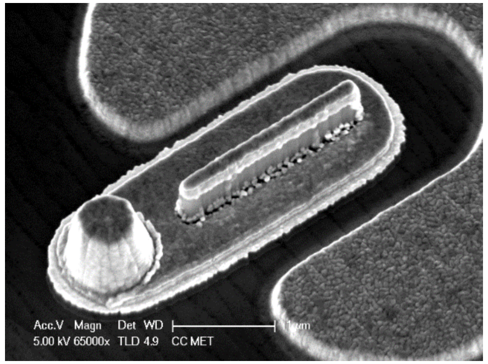
1940Ω/ sheet resistivity for intrinsic base ✗



Current process runs: GaAsSb intrinsic base

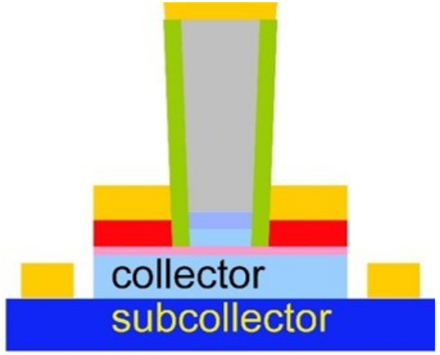
resistant to hydrogen passivation of carbon base dopant

(current reference sample: GaAsSb base, no regrowth)→→



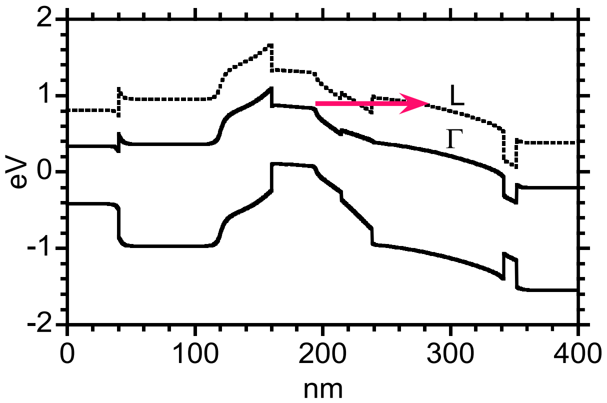
Materials for Improved THz HBTs

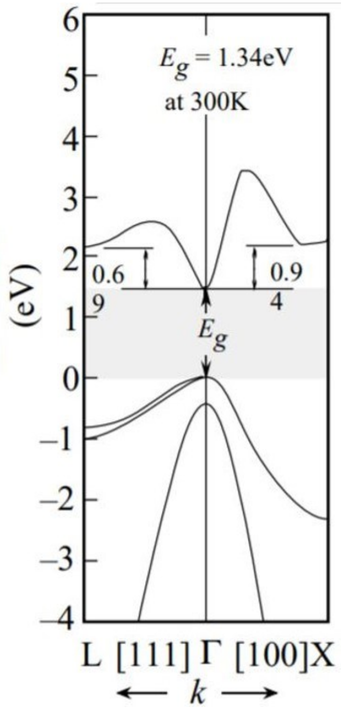
Contacts: existing materials are sufficient.
N-InAs for emitter, regrown ~10^{21} cm^{-3} GaAs for base



Base transit time is no problem: regrowth permits very thin base

Desire: improved voltage at a given bandwidth
1.4eV bandgap→ 30V/μm breakdown
But: increased transit time at high V_{ce}:
Γ-(L,X) scattering, *E*-*k* dispersion





Prof. Jasprit Singh; eecs320 notes, Univ Michigan;

Is there a better collector material ?:
Larger bandgap
Larger Γ-(L,X) energy separation
Low optical phonon scattering rates (GaN suffers here).
High thermal conductivity.

THz Field-Effect Transistors

FETs (HEMTs): key for low noise

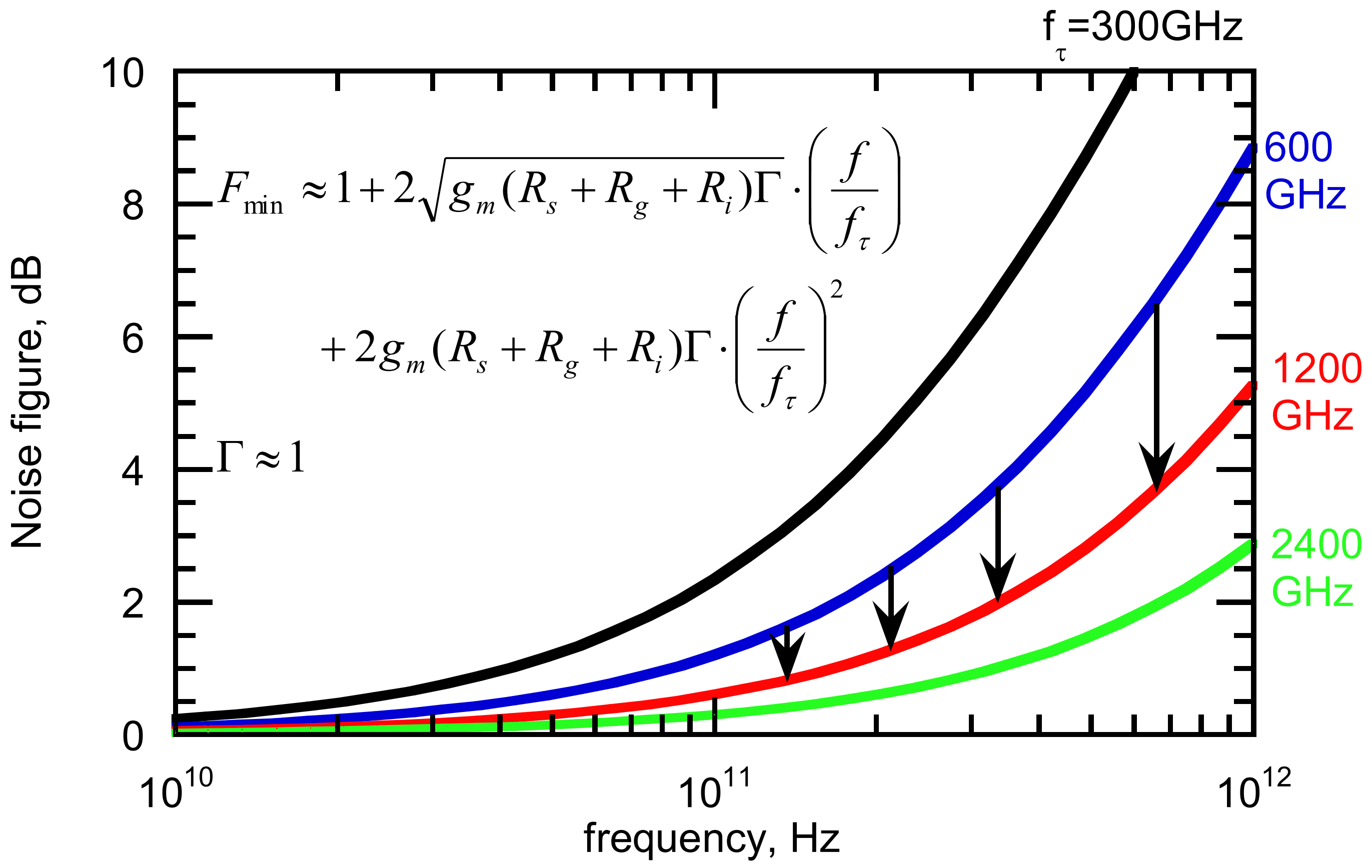
2:1 to 4:1 increase in f_τ :

improved noise

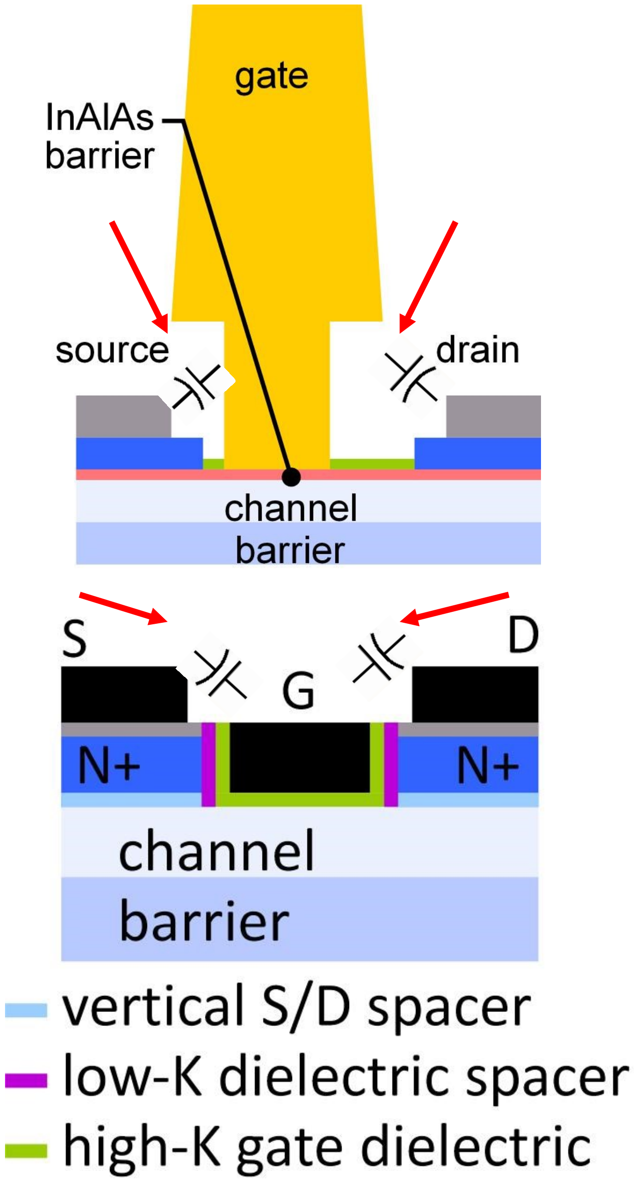
less required transmit power

smaller PAs, less DC power

or higher-frequency systems



High-Frequency FET Scaling

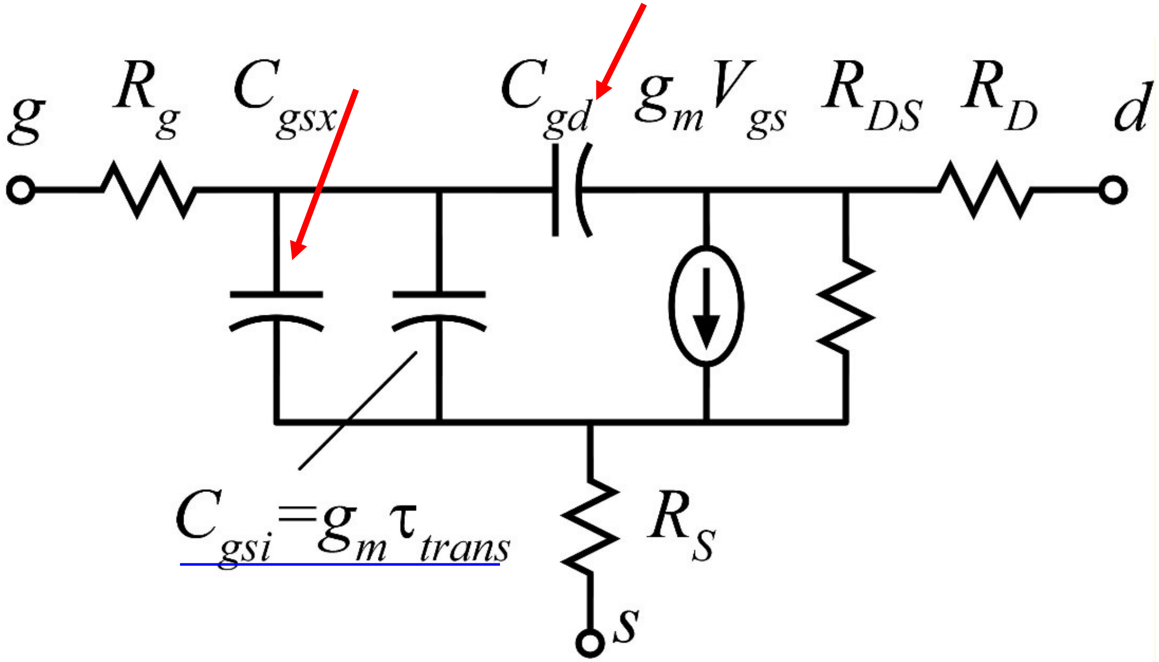


To double f_τ, reduce L_g 2:1, but this is not enough

Must also reduce C_{gsx}/g_m , C_{gd}/g_m time constants 2:1

→ g_m/W_g must be doubled

Must also thin dielectric and channel by 2:1 ($g_m R_{ds}$)



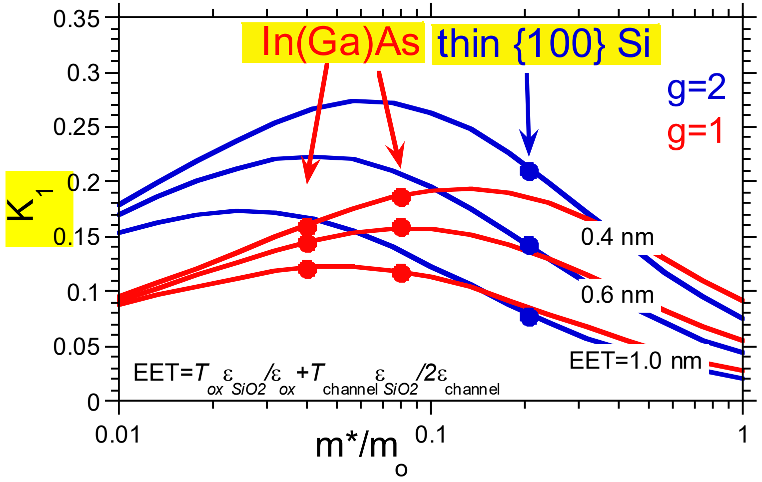
FET Current and Transconductance

Fermi velocity from $(E_f - E_{well}) = m^* v_f^2 / 2$

current $\propto$ Fermi velocity $\cdot$ charge

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

$$g_m \propto \boxed{K_1} \cdot (V_{gs} - V_{th})^{1/2}$$

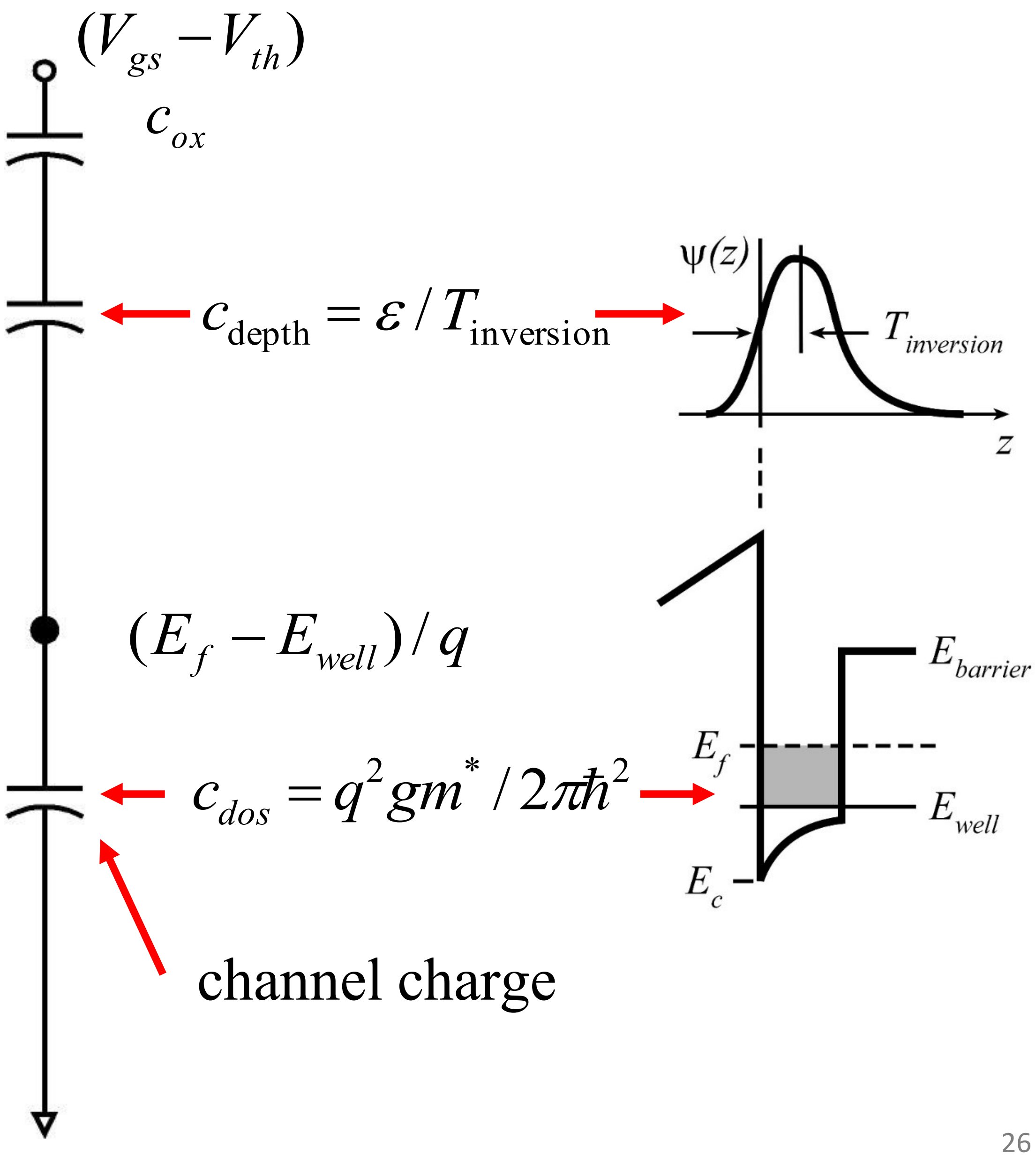


To increase g_m:

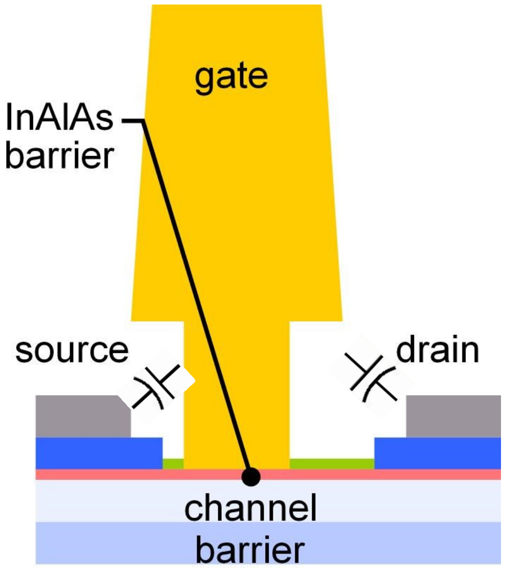
thin the oxide & channel

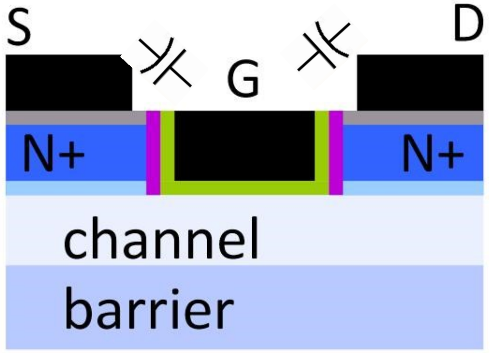
and increase K_1 (mass, # valleys)...**hard**

or increase $(V_{gs}-V_{th})$...**also hard**



FET Scaling Laws (these now broken)





- vertical S/D spacer
- low-K dielectric spacer
- high-K gate dielectric

FET parameter	change
gate length	decrease 2:1
current density (mA/mm)	increase 2:1
specific transconductance (mS/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
either (channel state density)	increase 2:1
or (V_{gs}-V_{th})	increase 4:1
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_{end}/g_m prevents f_τ scaling.

Shorter gate lengths degrade electrostatics→ reduced g_m/G_{ds} → reduced f_{max}, f_τ

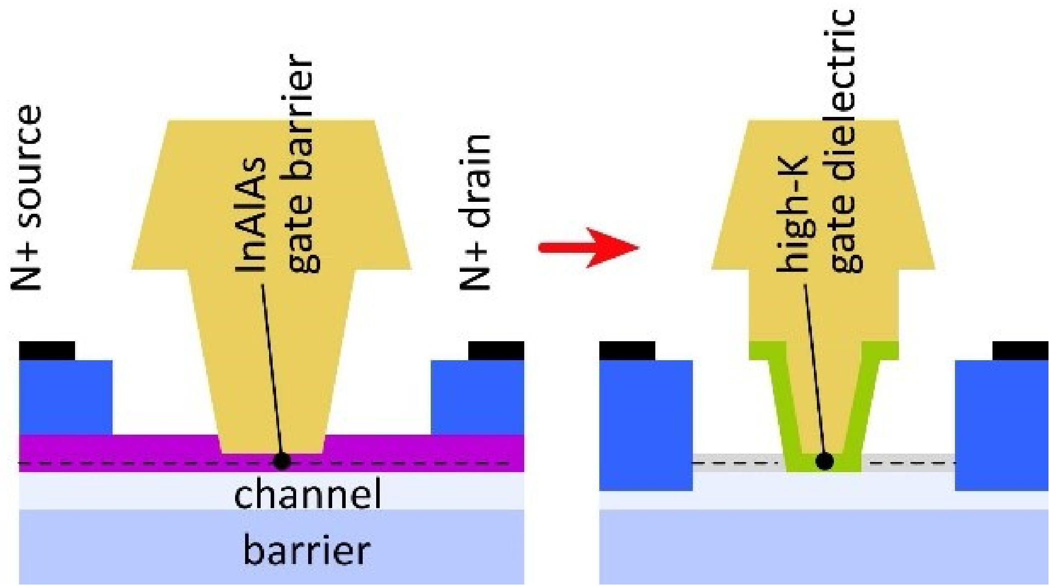
Towards faster HEMTs: InAs MOS-HEMTs

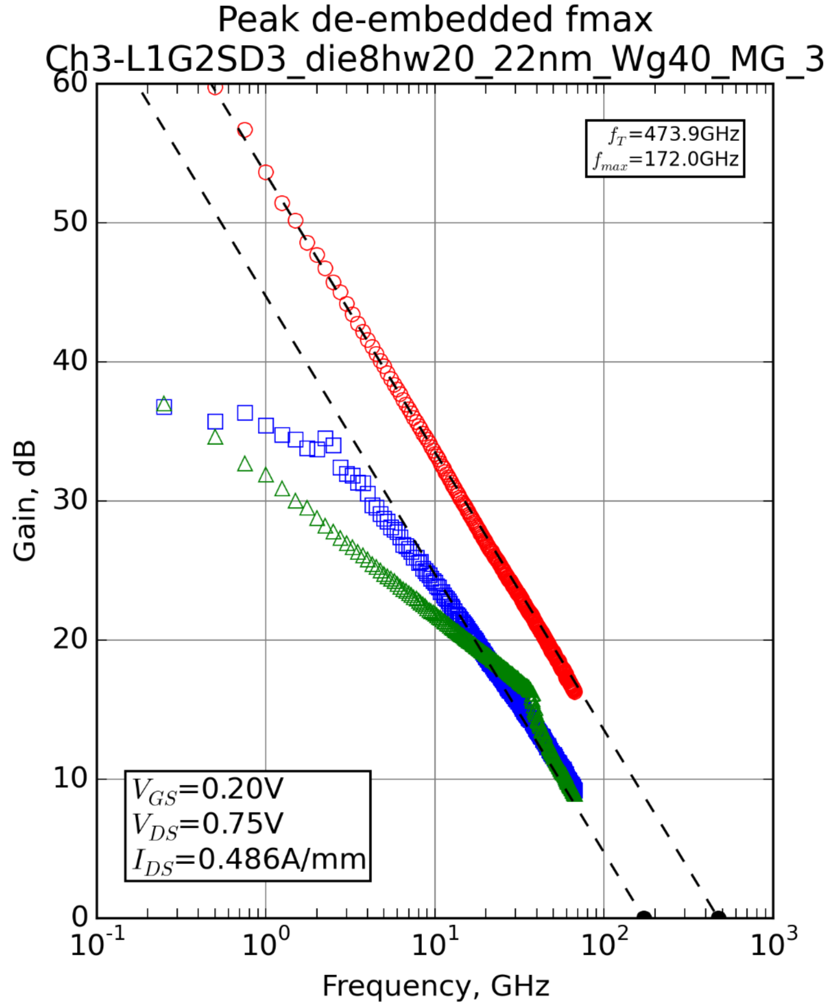
1st MOS-HEMT demonstration: Fraunhofer IAF

Scaling limit: gate insulator thickness
HEMT: InAlAs barrier: tunneling, thermionic leakage
solution: replace InAlAs with high-K dielectric
2nm ZrO_2 (ε_r=25): adequately low leakage

Scaling limit: source access resistance
HEMT: InAlAs barrier is under N+ source/drain
solution: regrowth, place N+ layer on InAs channel

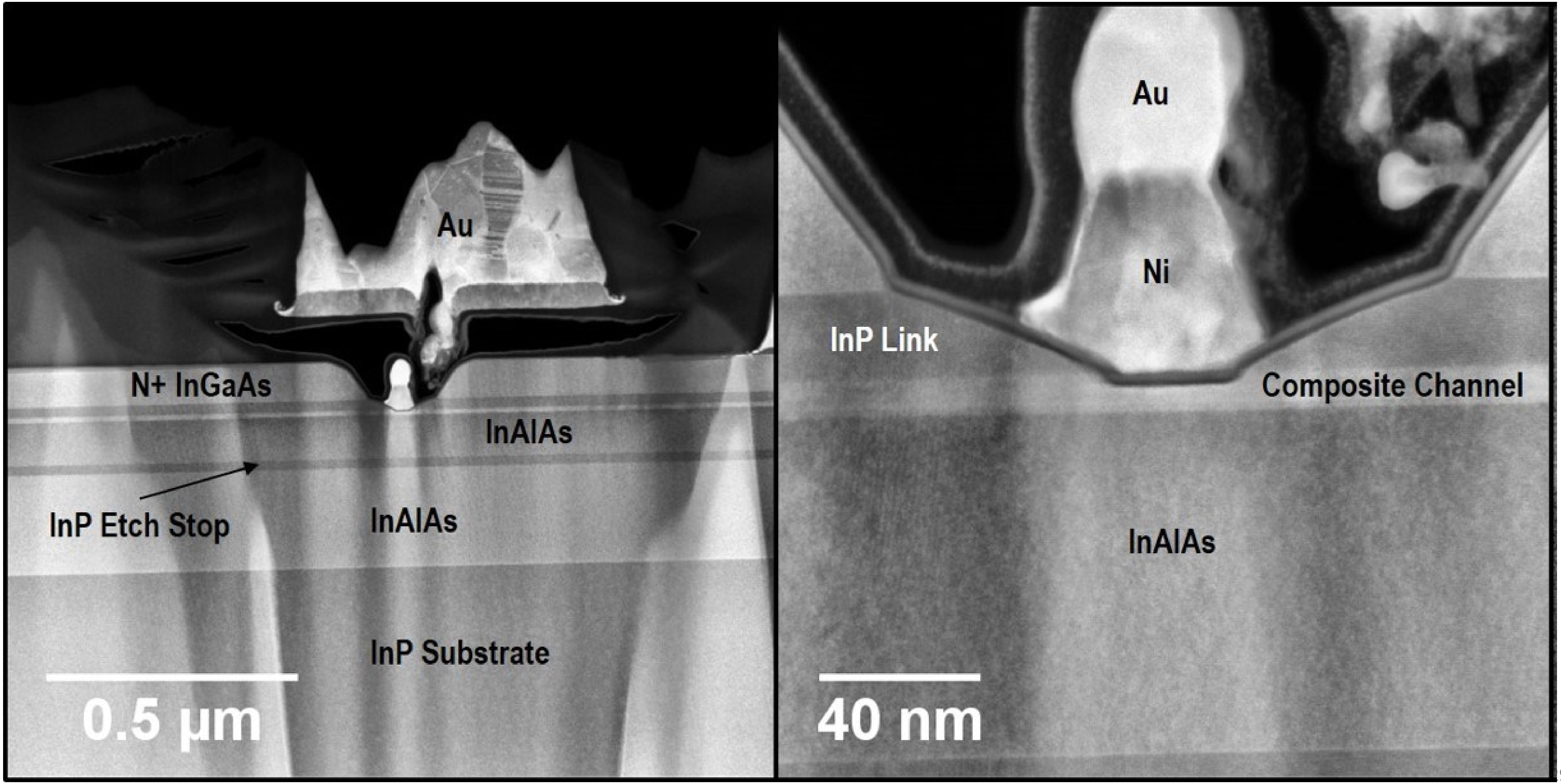
Target ~10nm node
~0.3nm EOT, 3nm thick channel
1.2 to 1.5 THz f_τ.



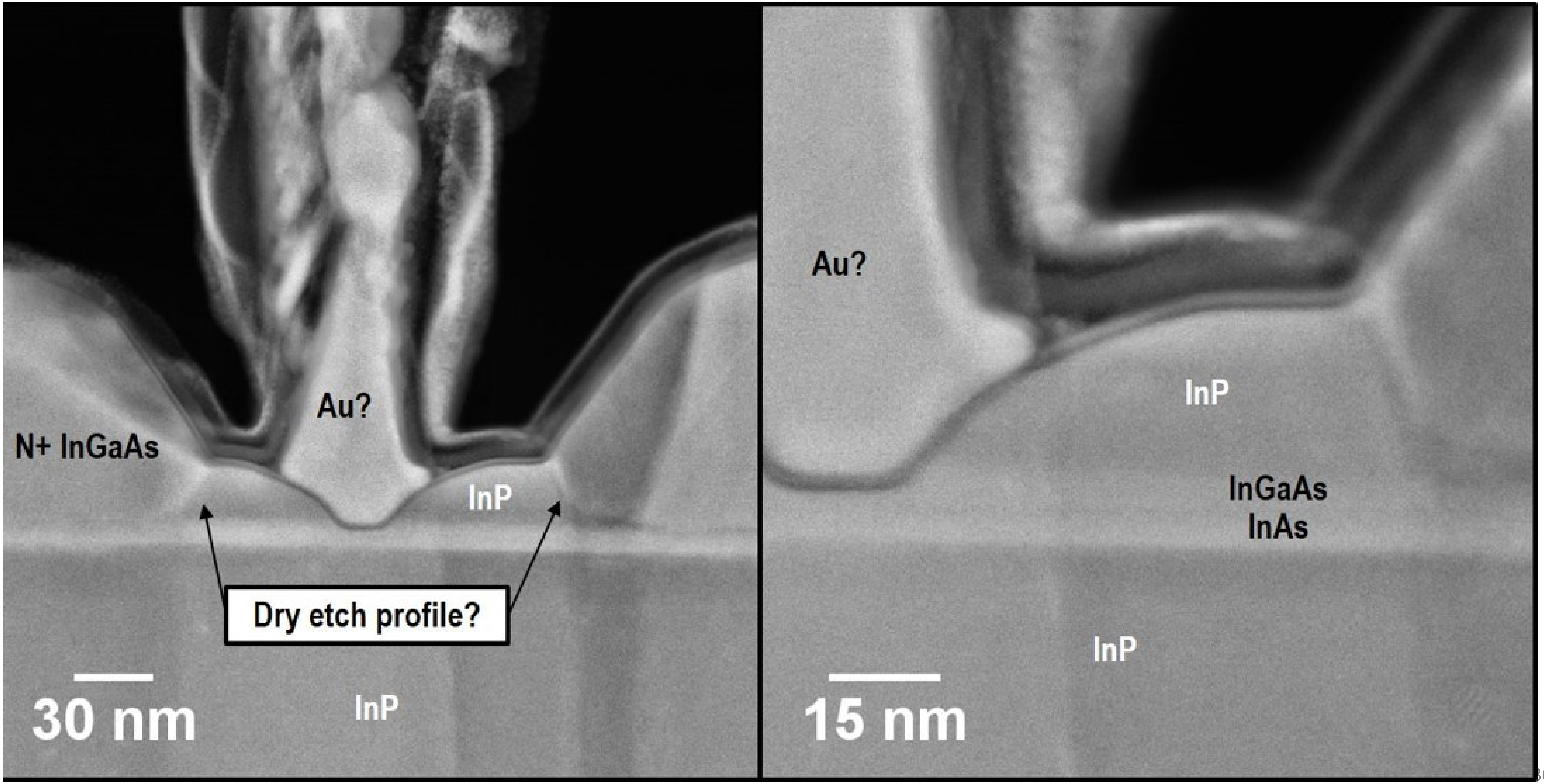


8/2019: process working: 470GHz f_τ. @ ~28nm L_g.
f_{max} very low, f_τ high but not high enough.
Now fixing obvious process problems.

Device Images



More Device Images



Increasing FET transconductance

For 1-2 THz f_τ, seek g_{mi}= 4-8 mS/μm.

thin the oxide, thin the well

→ increased eigenstate energy

→ loss of confinement at large (V_{gs}-V_{th})

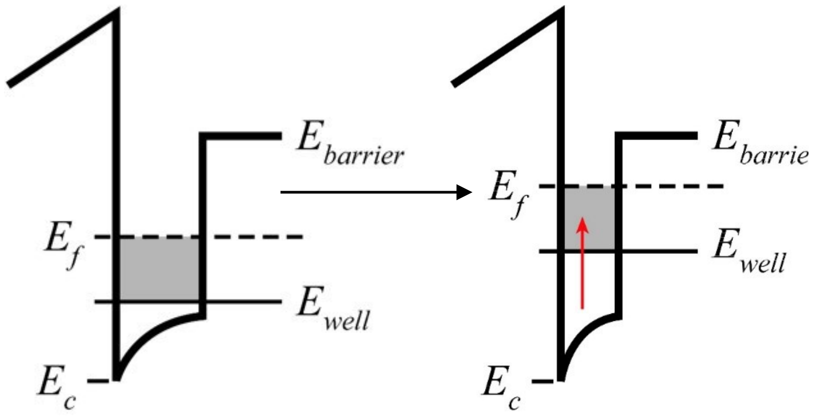
→ constrains maximum transconductance: $g_m \propto (V_{gs}-V_{th})^{1/2}$

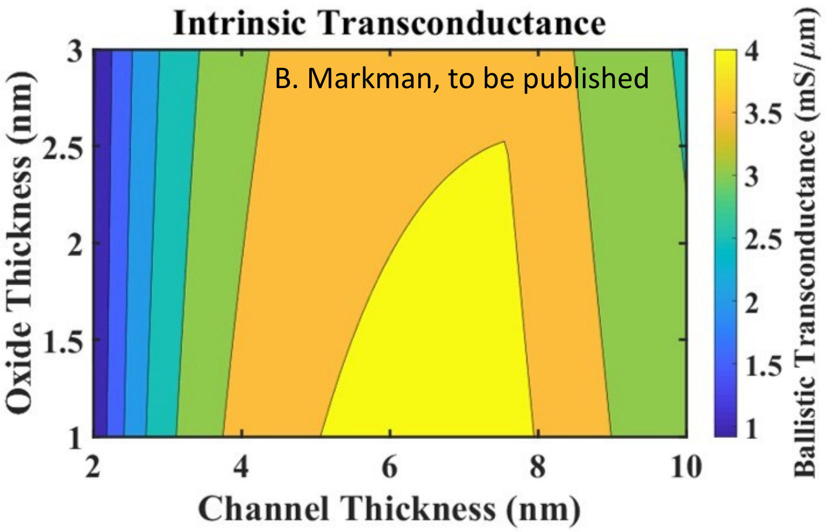
→ maximum achievable g_m.

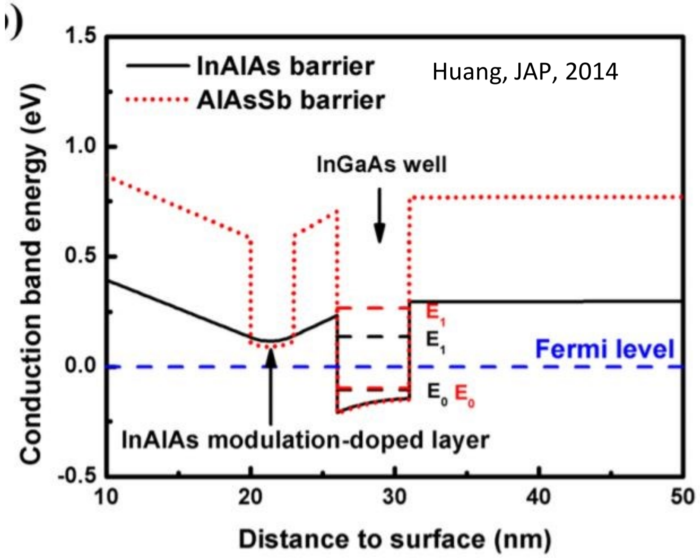
Need high barrier energies

InAs/InAlAs vs InAs/AlAsSb

InP/AlAsSb ????







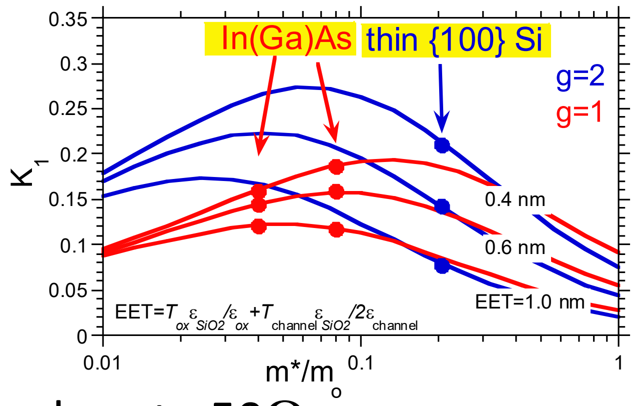
mm-Wave CMOS also won't scale much further

High-frequency Si MOSFET design follows the same principles as high-frequency III-V FET design.
Same physics, so same limits. But, the carrier velocities are lower

Difficult to further thin the gate dielectric (increased gate leakage)→ g_m can't increase

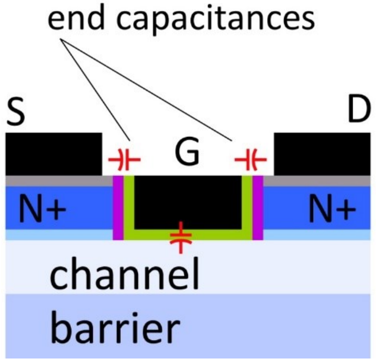
Shorter gates don't significantly reduce the capacitance: dominated by ends; ~1fF/μm total

Maximum g_m, minimum C→ upper limit on f_τ.
about 350-400 GHz**.** At the *bottom* of the wiring stack



Usable bandwidth is lower than this
high-frequency circuits need controlled-impedance Z_o=50Ω interconnects
microstrip lines with signal lines at top of wiring stack.
high-frequency transistors must interface to ~50 Ohm external impedances
~20-30 FET fingers must be tied in parallel to bring device port impedances close to 50Ω.
The necessary wiring further reduces ft, fmax: c.a. 300GHz in leading CMOS technologies.

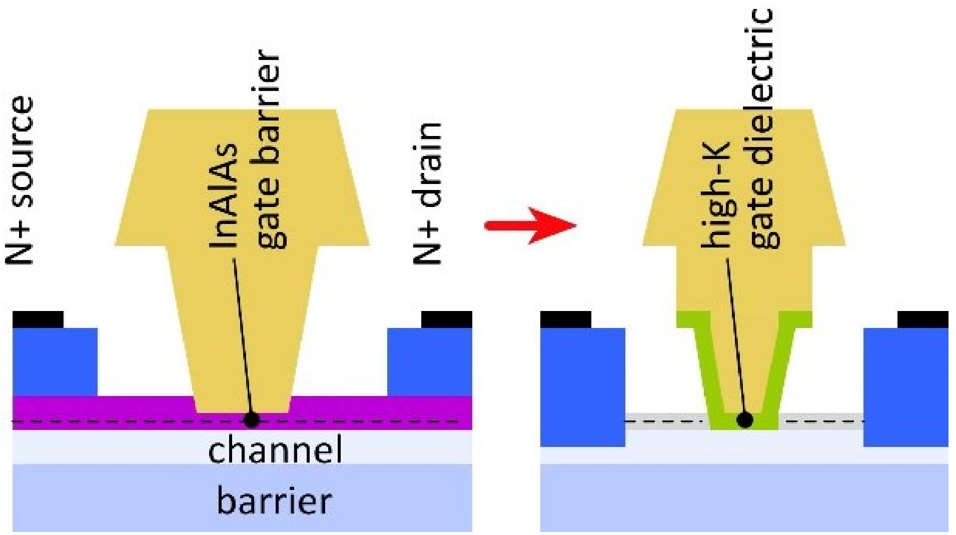
The best CMOS node for mm-wave is 65nm.
45nm SOI, 22nm SOI are close to this performance
Intel's RF-optimized 22nm finFET is also close to this performance



Materials for THz (MOS) HEMTs

Low resistance in source-gate access region:
moderate to high mobility

High transconductance
sufficient mobility to reach ballistic current.
$m^* \approx 0.08 \cdot m_o$ to $0.16 \cdot m_o$ high ballistic v_{inj}
desirable/hard: multiple band minima (large n_s)
large band offsets for large $(E_f - E_c) \rightarrow$ large $(V_{gs} - V_{th})$



Low S/D access resistances
S/D materials with high doping, low barriers. Grade heterointerfaces

High velocities and high breakdown fields in gate-drain drift region
m* need not be particularly low
low phonon scattering rates
large intervalley energy separation
wide bandgap

Closing

Materials Requirements for Future Transistors

VLSI (MOSFETs, TFETs):

It's mostly about the interfaces: gate dielectrics, S/D contacts

The channel also matters: high ballistic I_{on}, sufficient E_g for low I_{off}

MOS-HEMTs for mm-wave ICs, 50-500GHz low-noise amplifiers

Same as VLSI MOSFETs: gate dielectrics, S/D contacts

Large ballistic I_{on}.

How to increase g_m ? Large $(V_{gs}-V_{th}) \rightarrow$ large $(E_{barrier}-E_{well})$

HBTs for mm-wave ICs, 50-500 GHz power

Contacts are critically important…and achievable

Breakdown vs. transit time: ballistic overshoot matters.

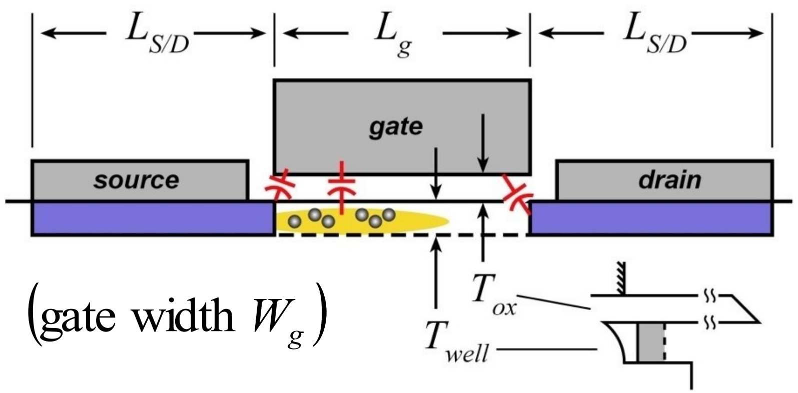
Large intervalley separation, low scattering rates

In case of questions

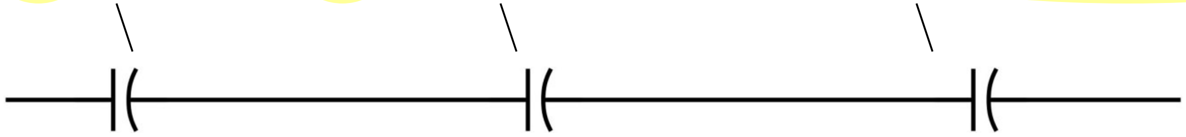
FET Design: Scaling

$$C_{gd} \cong C_{gs,f} \cong \varepsilon_r \varepsilon_0 W_g$$

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



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



$$v \propto \begin{pmatrix} \text{voltage division ratio between} \\ \text{the above three capacitors} \end{pmatrix}^{-1/2} \cdot \frac{1}{\sqrt{\text{transport mass}}}$$

$$R_{DS} \approx L_g / (W_g v \varepsilon_r \varepsilon_0) \quad R_S = R_D = \frac{\rho_{\text{contact}}}{L_{\text{S/D}} W_g}$$