Plenary, ESSCIRC ESSDERC 2021 European Solid-state Circuits and Devices Conference, Grenoble and online.
September 15, 2021.

Transistors for 100-300GHz Wireless

Mark Rodwell, Brian Markman, Yihao Fang,
Logan Whitaker, Hsin-Ying Tseng, A. S. H. Ahmed
University of California, Santa Barbara
rodwell@ece.ucsb.edu

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

Acknowledgements

Systems


Sundeep Rangan — Networks, Applications, MIMO, Power


Upamanyu Madhow, UC Santa Barbara — MIMO algorithms, Imaging algorithms, Compressive imaging


Christoph Studer, Cornell — MIMO algorithms, VLSI MIMO, digital beamforming


Andreas Molisch, USC — 100-300GHz propagation measurements


Danijela Cabric, UCLA — MIMO algorithms (funding via CONIX)

ICs


Ali Niknejad, UC Berkeley — mm-wave CMOS: hub, mm-wave arrays, mm-wave MIMO


James Buckwalter, UC Santa Barbara — efficient PAs, III-V arrays


Kenneth O, UT Dallas — 140-300GHz SiGe ICs


Muhannad Bakir, Georgia Tech — high-frequency packaging


Gabriel Rebeiz, UC San Diego — mm-wave CMOS: handset, mm-wave arrays


Alyosha Molnar, Cornell — N-path mixers, MIMO ADCs


Elad Alon, UC Berkeley — design automation, equalizers


Tim Fisher, UCLA — advanced packaging materials


Andrew Kummel, UCSD — advanced packaging materials

Transistors


Umesh Mishra, UC Santa Barbara — N-polar GaN HEMTs for 140, 210GHz


Huili (Grace) Xing, Cornell — AlN/GaN HEMTs for 140, 210GHz


Susanne Stemmer, UC Santa Barbara — transistors in novel materials


Debdeep Jena, Cornell — GaN HEMTs on Si


Srabanti Chowdhury, UC Davis — Diamond cooling for GaN


Borivoje Nikolic, UC Berkeley — Massive MIMO demo.; VLSI design automation, VLSI MIMO processors


Amin Arbabian, Stanford — Compressive imaging; 140GHz radar chipsets and arrays


Mark Rodwell, UC Santa Barbara — 140/210/280GHz arrays for demos.; THz HBTs for PAs, THz HEMTs for LNAs

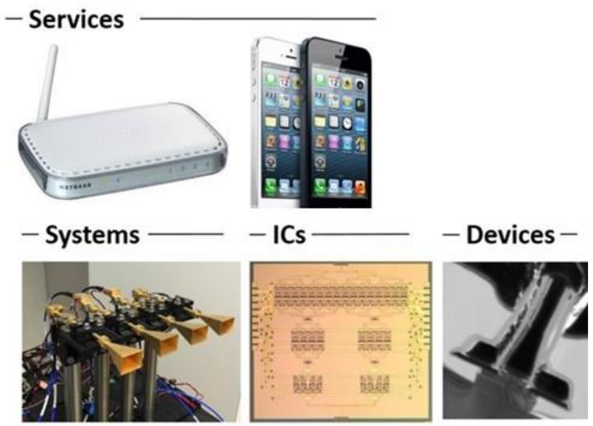
Also:
Kyocera: D. Kim, H. Horikawa, M. Imayoshi.
Samsung: G. Xu, N. Sharma, S. Abu-Surra, W. Choi
Pi-Radio: A. Dhananjay,

Transistors for 100-300GHz wireless

Wireless networks: exploding demand.

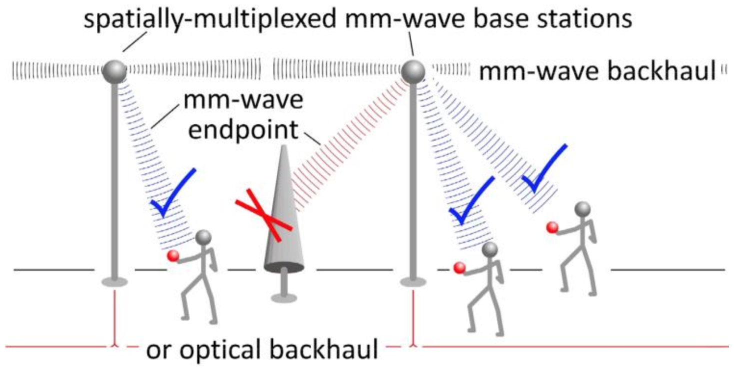
Immediate industry response: 5G.
~6~100GHz
increased spectrum, extensive beamforming

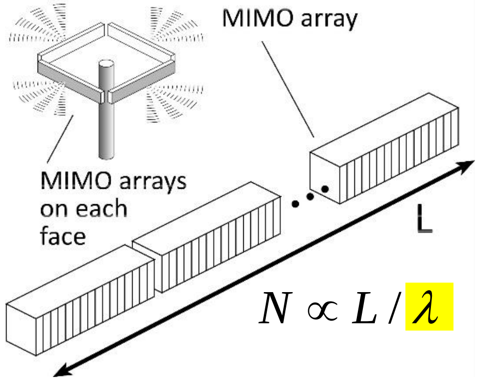
Next generation: 100-300GHz (???)
greatly increased spectrum, massive spatial multiplexing



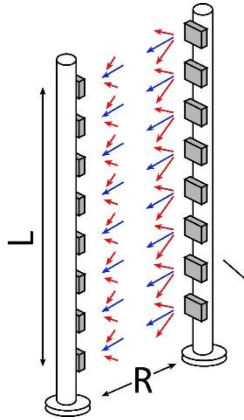
100-300GHz carriers, massive spatial multiplexing
→ Terabit hubs and backhaul links, high-resolution imaging radar

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

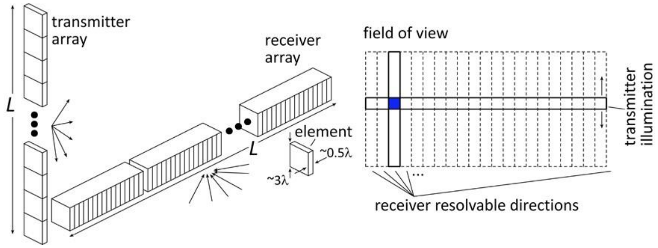




$N \propto L / \lambda$

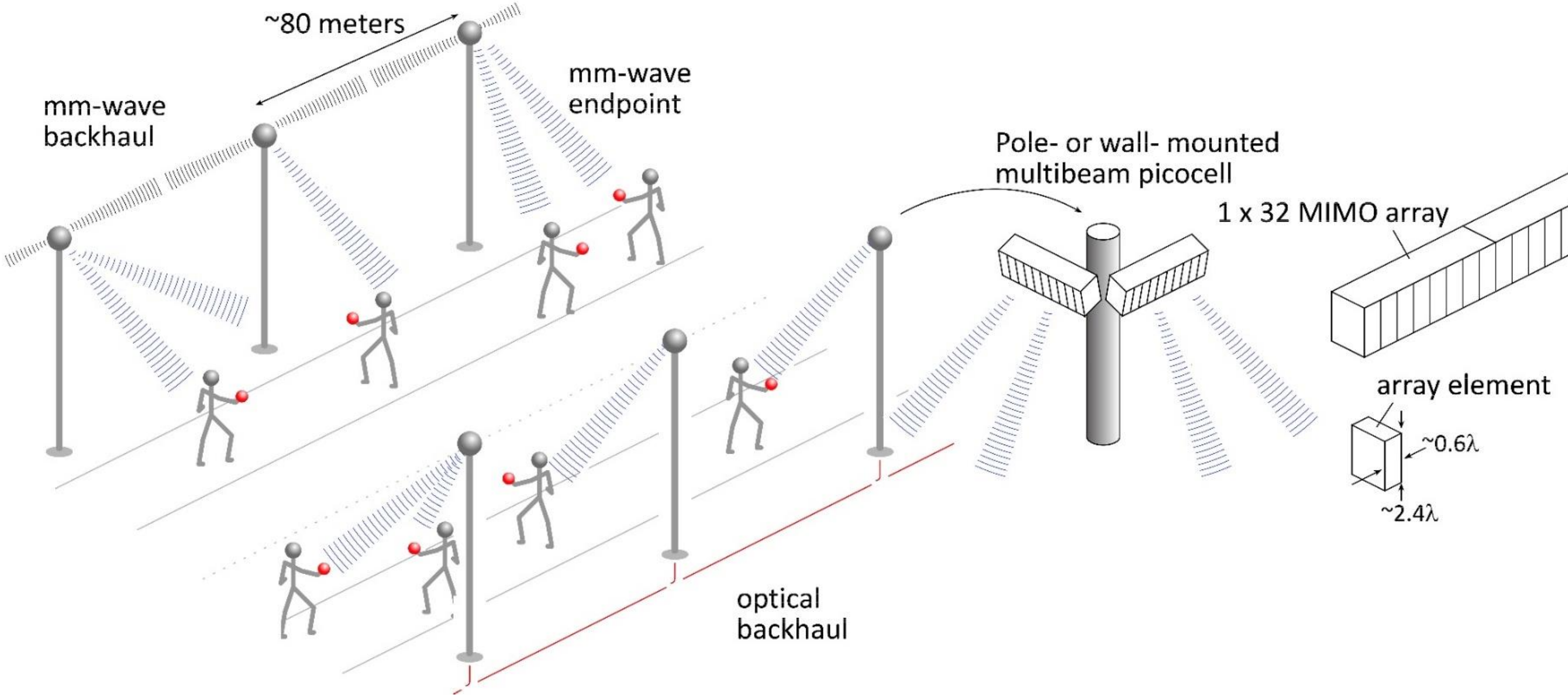


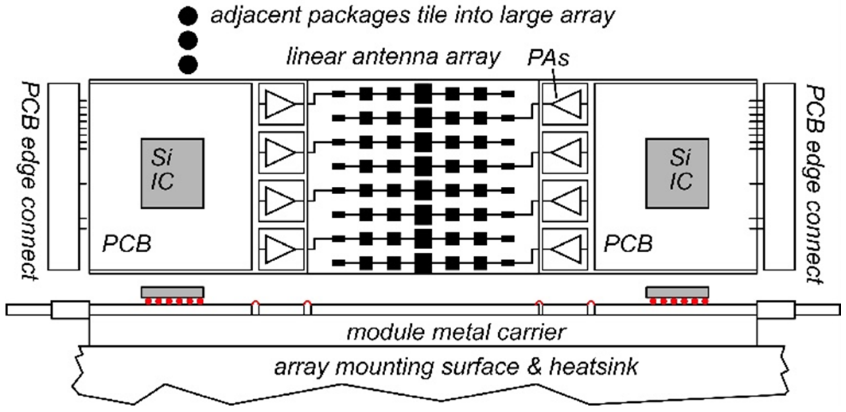
$N \propto L^2 / \lambda R$



What transistors do we need ?

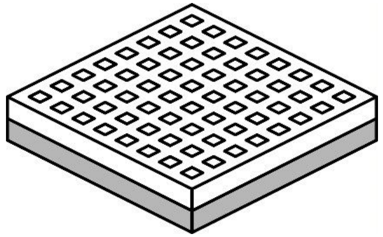
140GHz, 160 Gb/s MIMO network hub





Hub with 32-element array (four 1×8 modules):

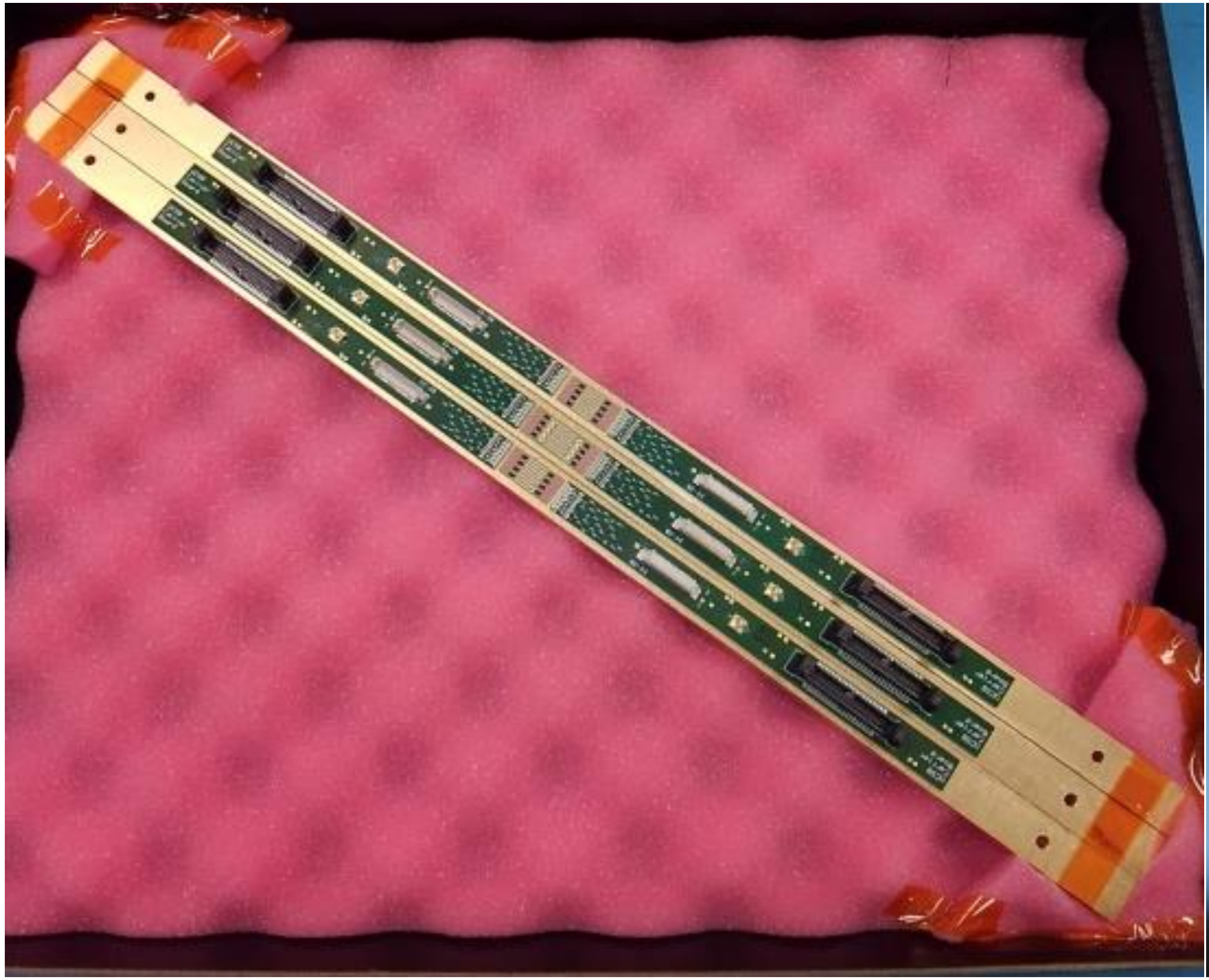
16 users/array. F=8dB LNAs, P_{1dB}=21 dBm PAs
10 Gb/s/beam→ 160 Gb/s total capacity
40 m range in 50mm/hr rain with **17**dB total margins

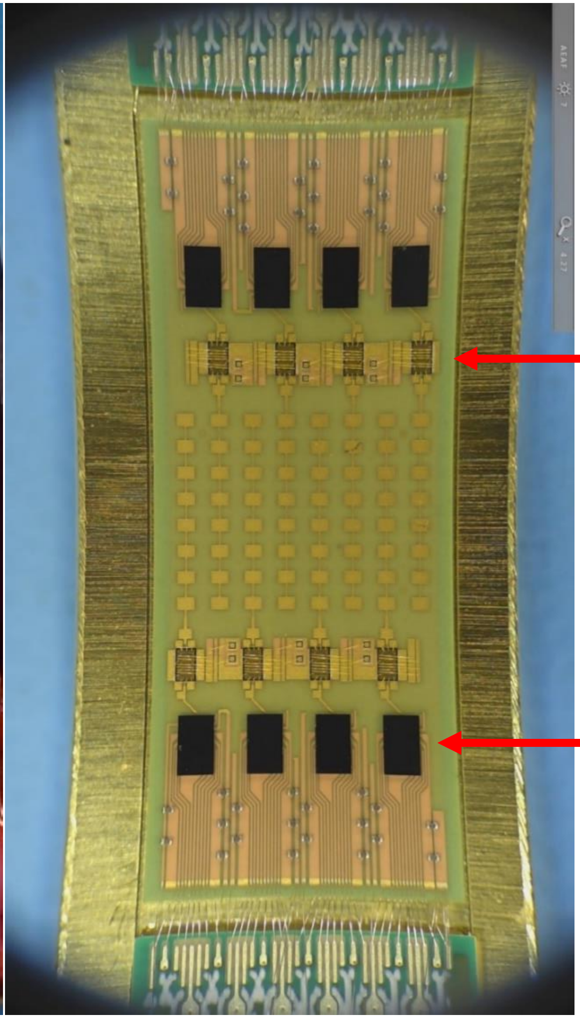


Handset:
8 × 8 array
(9×9mm)

8-channel 140GHz MIMO hub modules

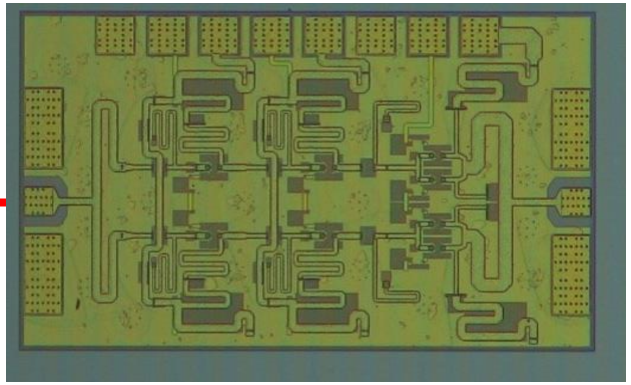
A. Farid et. al, in review



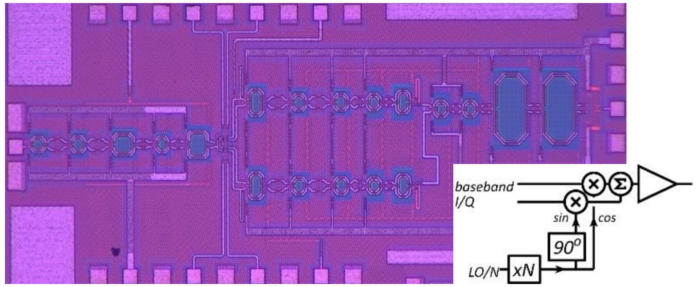


Kyocera LTCC carrier

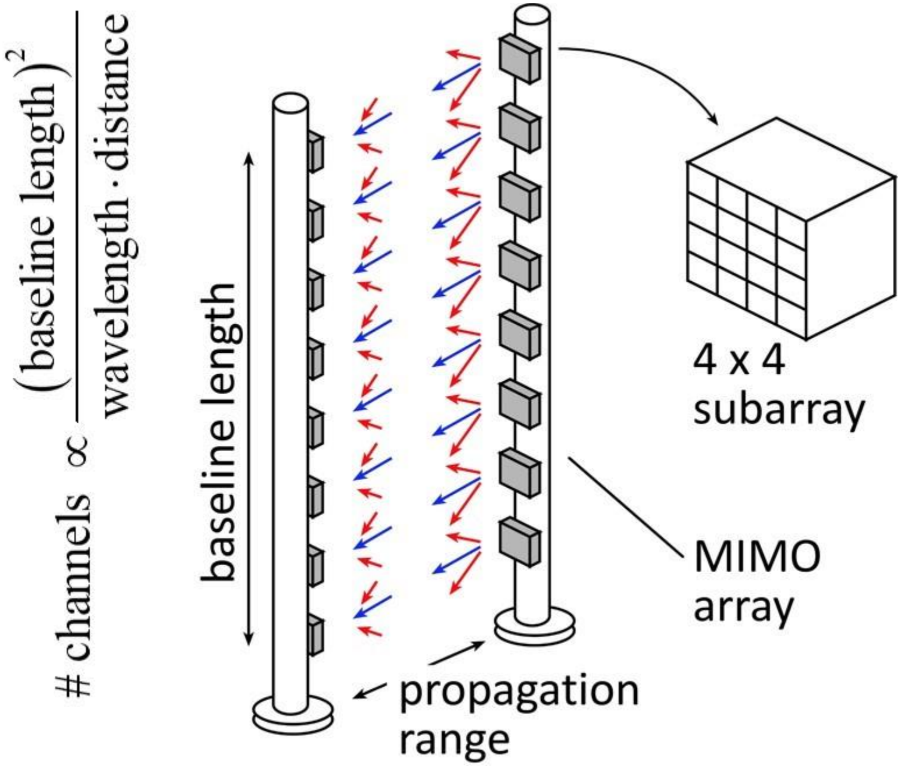
110mW power amplifier
Teledyne 250nm InP HBT
20.8% PAE

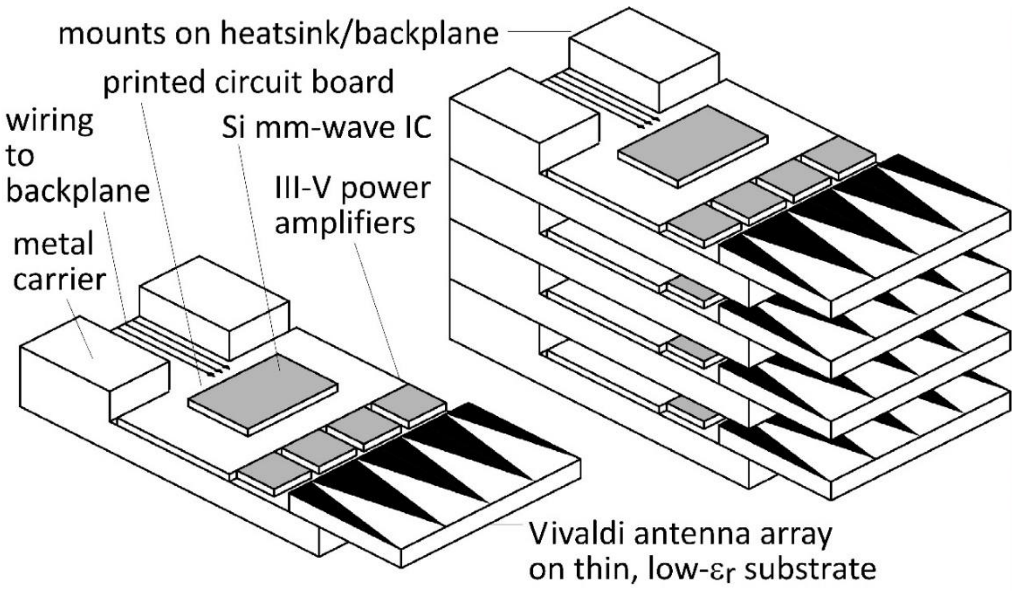


Transmitter IC
GlobalFoundries 22nm SOI CMOS



210 GHz, 640 Gb/s MIMO backhaul





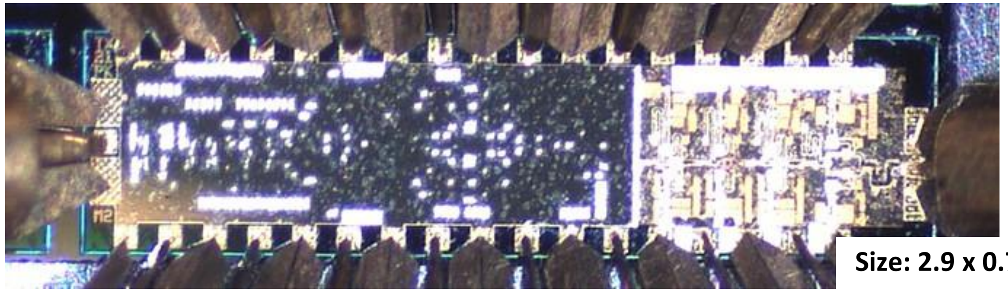
8-element MIMO array
3.1 m baseline for 500 meters range.
80Gb/s/subarray→ 640Gb/s total
4 × 4 sub-arrays → 8 degree beamsteering

Key link parameters
500 meters range in 50 mm/hr rain; 23 dB/km
20 dB total margins:
packaging loss, obstruction, operating, design, aging
LNAs: 6dB noise figure
PAs: 18dBm $=P_{1dB}$ (per element)

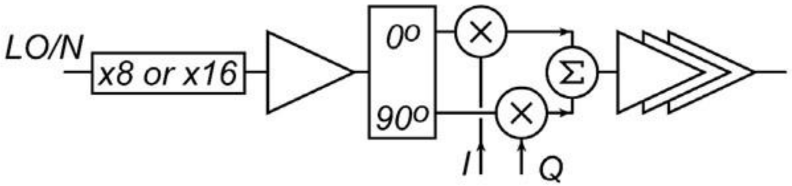
210 GHz transmitter and receiver ICs

210GHz transmitter: 20GHz bandwidth, 15.5-16.5dBm power

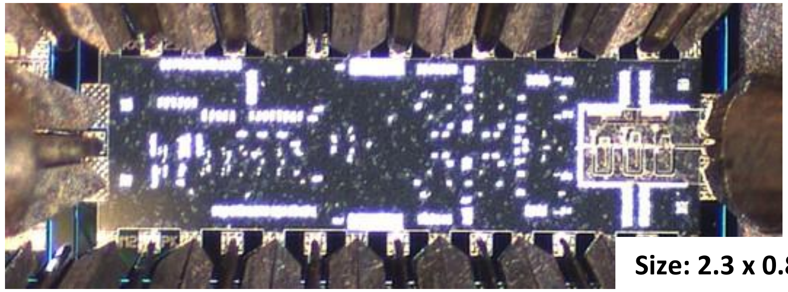
M. Seo et al, 2021 IMS; Teledyne 250nm InP HBT



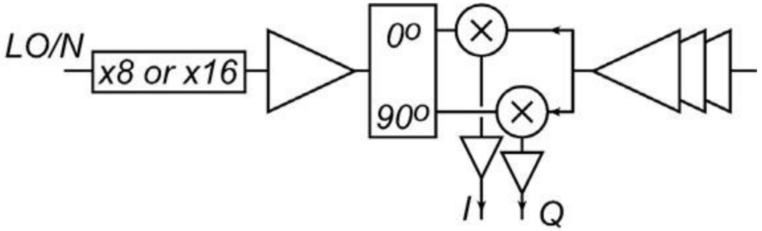
Size: 2.9 x 0.75 mm^2



210GHz receiver: 20GHz bandwidth, 7.7-9.5dB noise figure

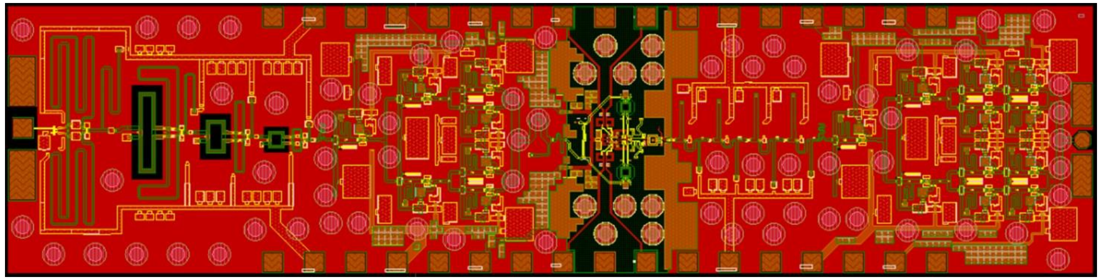


Size: 2.3 x 0.85 mm^2

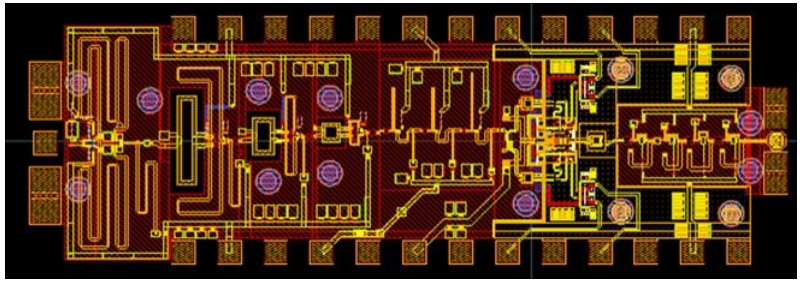


280GHz transmitter: 17dBm power (simulated)

Solyu, Alz, Ahmed, Seo; UCSB/Sungkyunkwan; Teledyne 250nm InP HBT



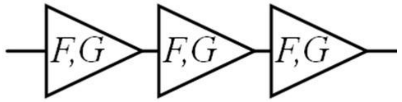
280GHz receiver: 11dB noise figure, 40GHz bandwidth (sim.)



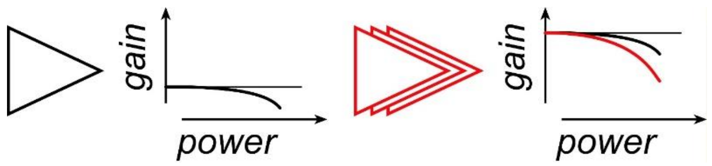
100-300GHz wireless: transistor requirements

Transmitters need:
high power-added efficiency $PAE = (P_{out} - P_{in})/P_{DC}$
high added power density (P_{out}-P_{in})/(gate width, emitter length)

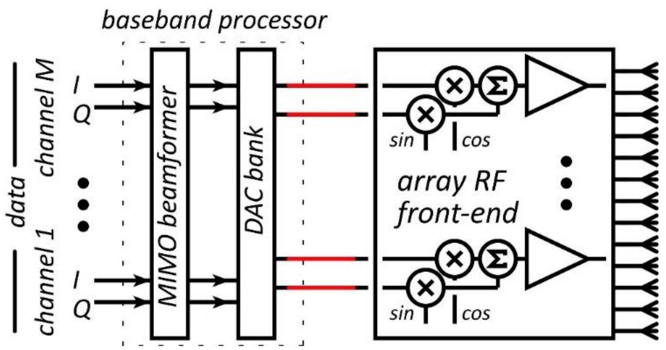
Receivers need:
low cascaded noise $F_{casc} = F + (F-1)/G + (F-1)/G^2 + \cdots$

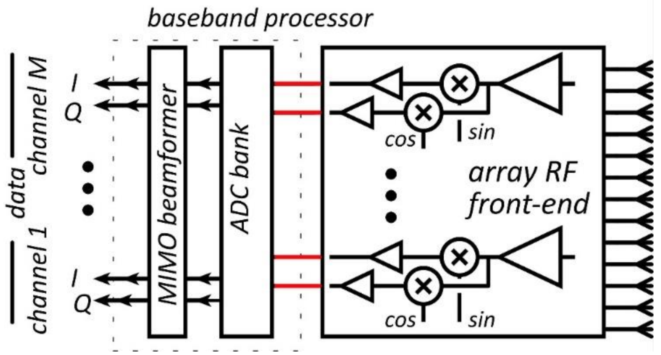


Need reasonable gain/stage.
die area, power,
accumulated gain compression



(gain in PAs, LNAs is less than MAG/MSG, U, ...)





Transistors for 100-300GHz

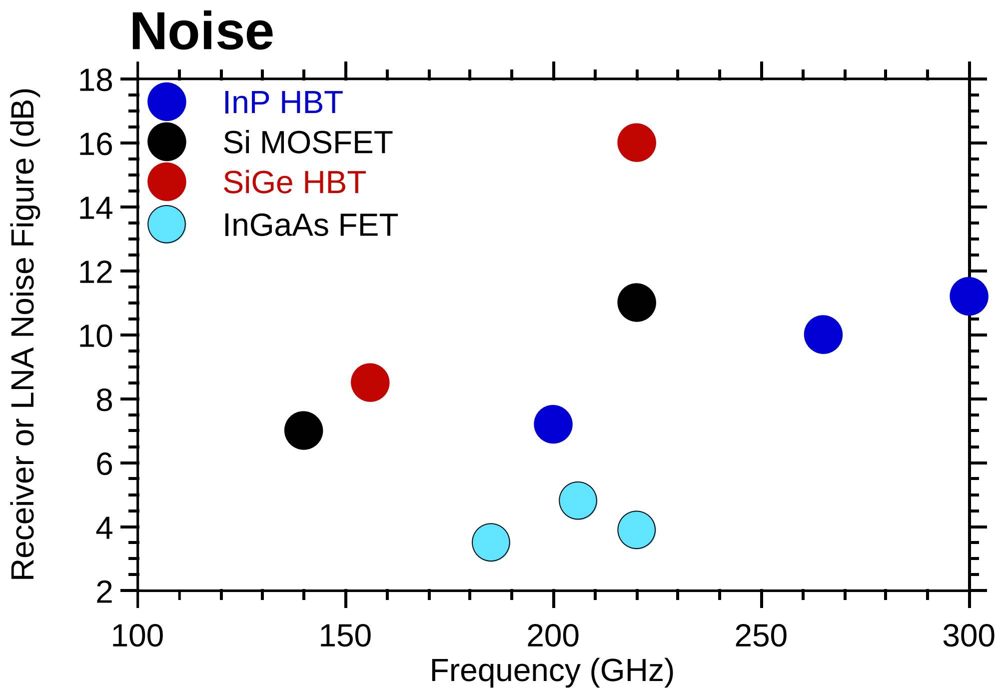
CMOS: good power & noise up to ~150GHz. Not much beyond. 65-32nm nodes are best.

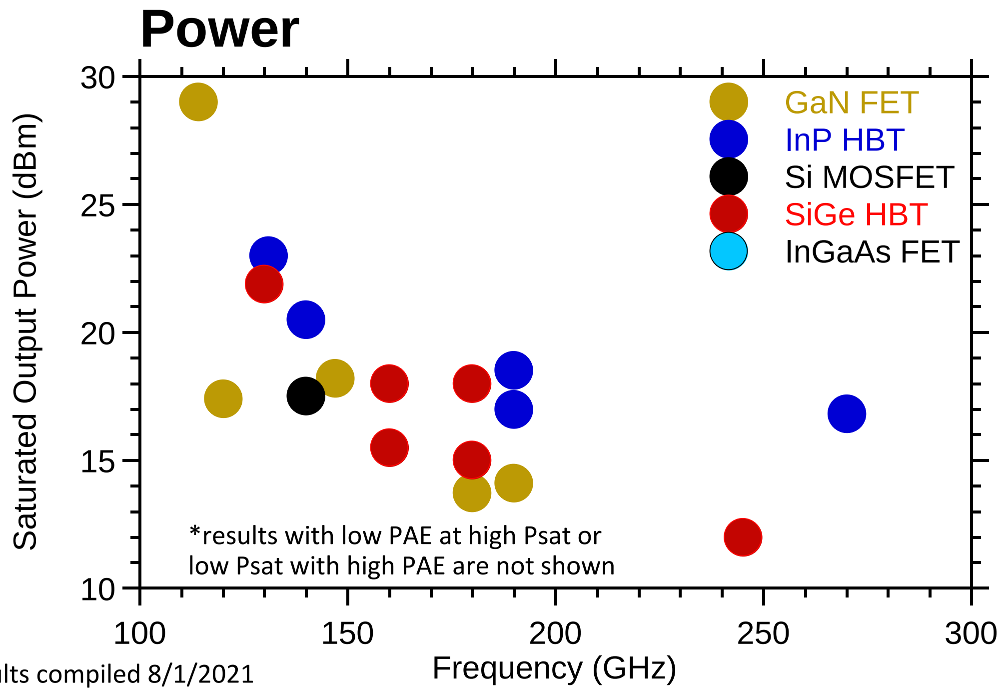
InP HBT: record 100-300GHz PAs

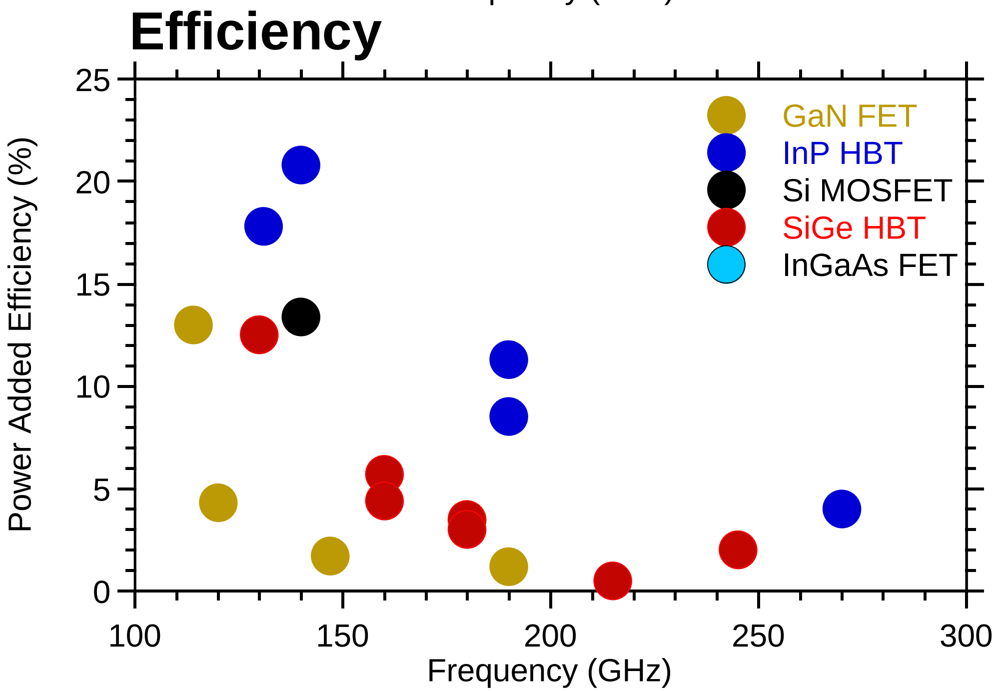
SiGe HBT: outperforms CMOS above 200GHz

GaN HEMT: record power below 100GHz. Bandwidth improving

InGaAs-channel HEMT: world's best low-noise amplifiers







Results compiled 8/1/2021

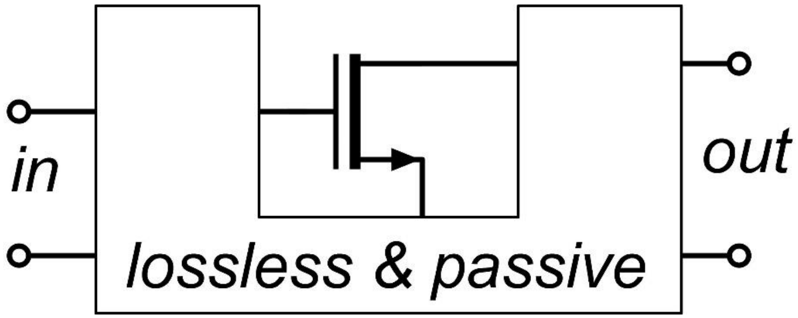
Where the IC designer can't help us.

mm-wave transistor gain is low: gain-boosting is common
Common-source vs. common-gate.
Capacitive neutralization. Controlled positive feedback (Singhakowinta, Int. J. Electronics, 1966)

Such circuits don't improve the parameters that matter the most.

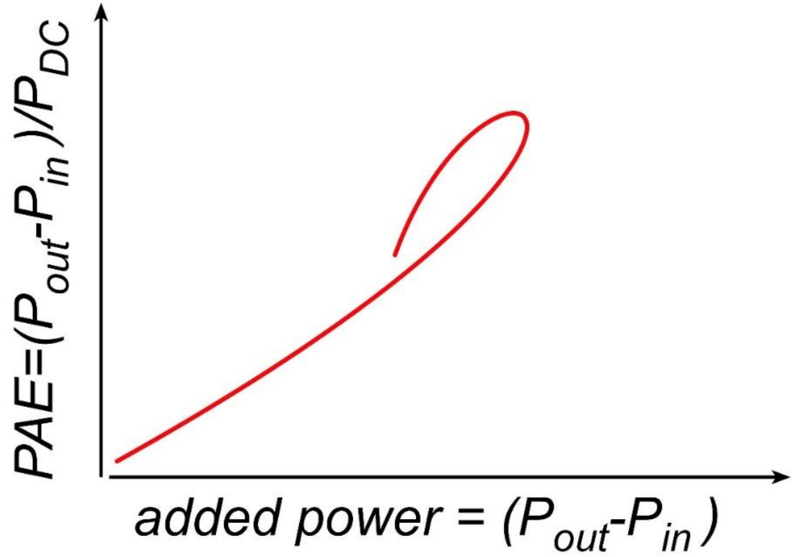
The circuit* doesn't change the transistor minimum cascaded noise figure. (Haus, Adler, Proc. IRE, 1958)

The circuit* doesn't change the transistor maximum efficiency vs. added power curve.



$$F_{casc} = F + (F-1)/G + (F-1)/G^2 + \cdots$$

*If lossless, and given the correct source and load impedances.

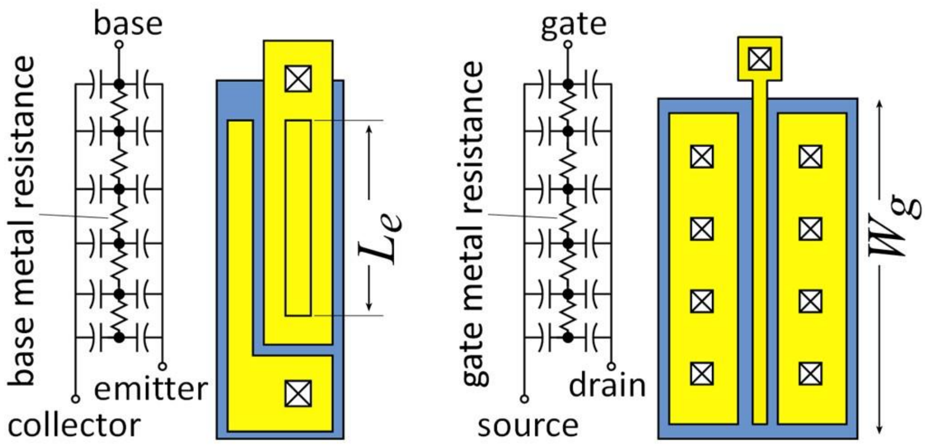


Current density, finger pitch limit cell output power

$$\text{Electrode } RC \text{ charging time} \propto (\text{finger length})^2$$

$$\text{Maximum finger length} \propto 1/\sqrt{\text{frequency}}$$

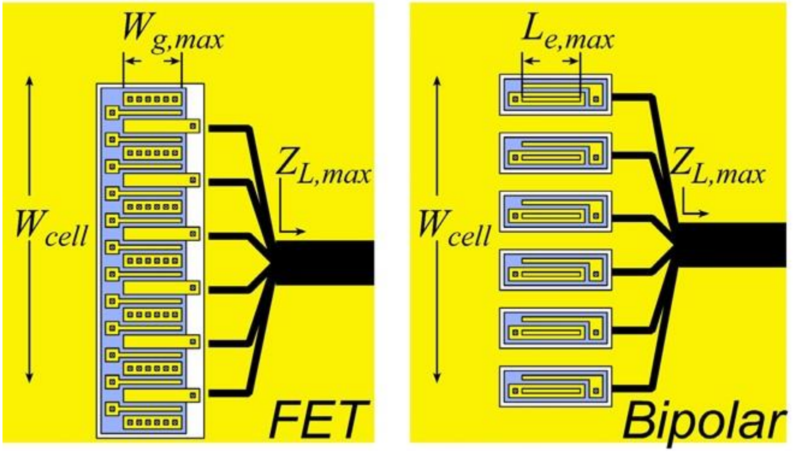
$$\text{Current per finger} \propto 1/\sqrt{\text{frequency}}$$



$$\text{Maximum cell width} \propto 1/\text{frequency}$$

$$\text{Maximum number fingers} \propto 1/\text{frequency}$$

$$\text{Maximum current per cell} \propto 1/\text{frequency}^{3/2}$$



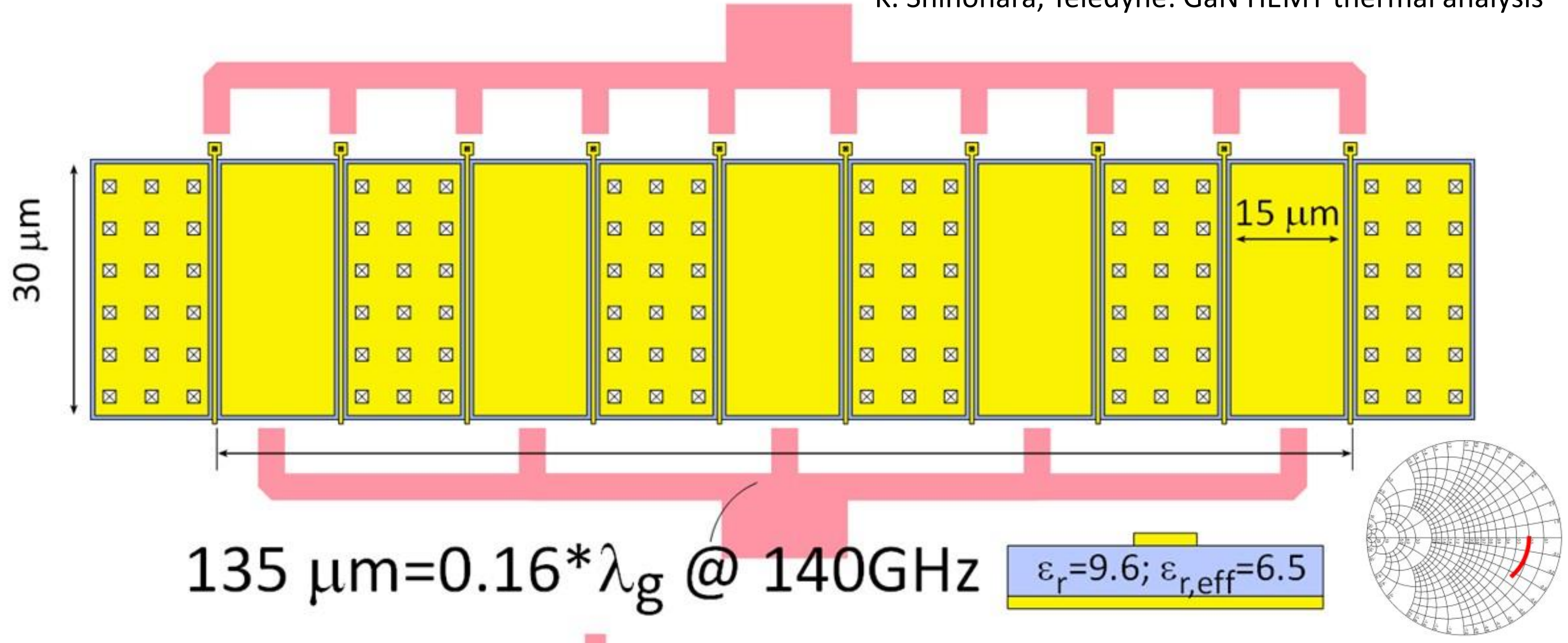
$$\text{Maximum RF power per cell} \propto (\text{maximum load resistance})\cdot(\text{maximum current})^2 \propto 1/(\text{frequency})^3$$

Compare to Johnson F.O.M.: $\text{maximum power per cell} \propto (\text{maximum voltage})^2/(\text{minimum load resistance}) \propto 1/(\text{frequency})^2$

Current density, finger pitch limit cell output power

K. Shinohara, Teledyne: GaN HEMT thermal analysis

50Ω GaN PA cell @ 140GHz (1.6W)
25V swing, 1.67mA/μm,
gates: 30 μm width, 15 μm pitch



50Ω InP HBT PA cell @ 280GHz (40mW)
4V swing, 3.3mA/μm,
emitters: 6 μm length, 6 μm pitch

High V_{br}, low I_{max} ? Device sized to drive 50Ω might approach λ_g/4 width.
Small finger pitch is critical; limited by thermal design

Current density, finger pitch limit power combining

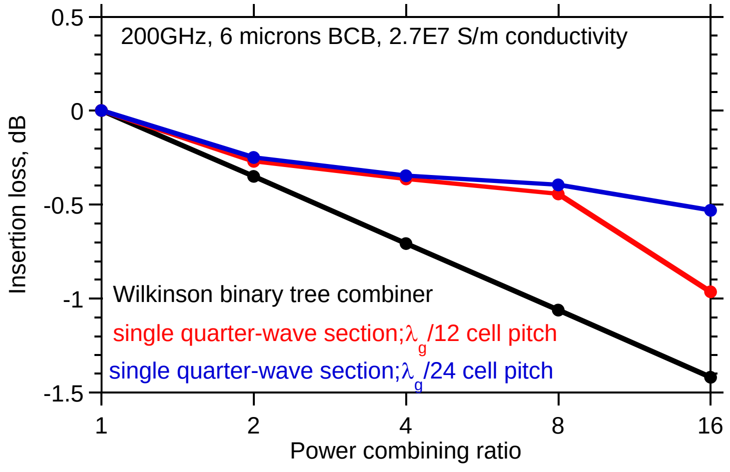
More cells: more output power

Number of cells limited by combining losses.

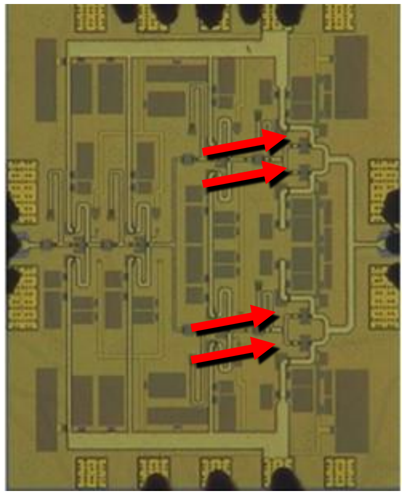
Losses mostly limited by size.

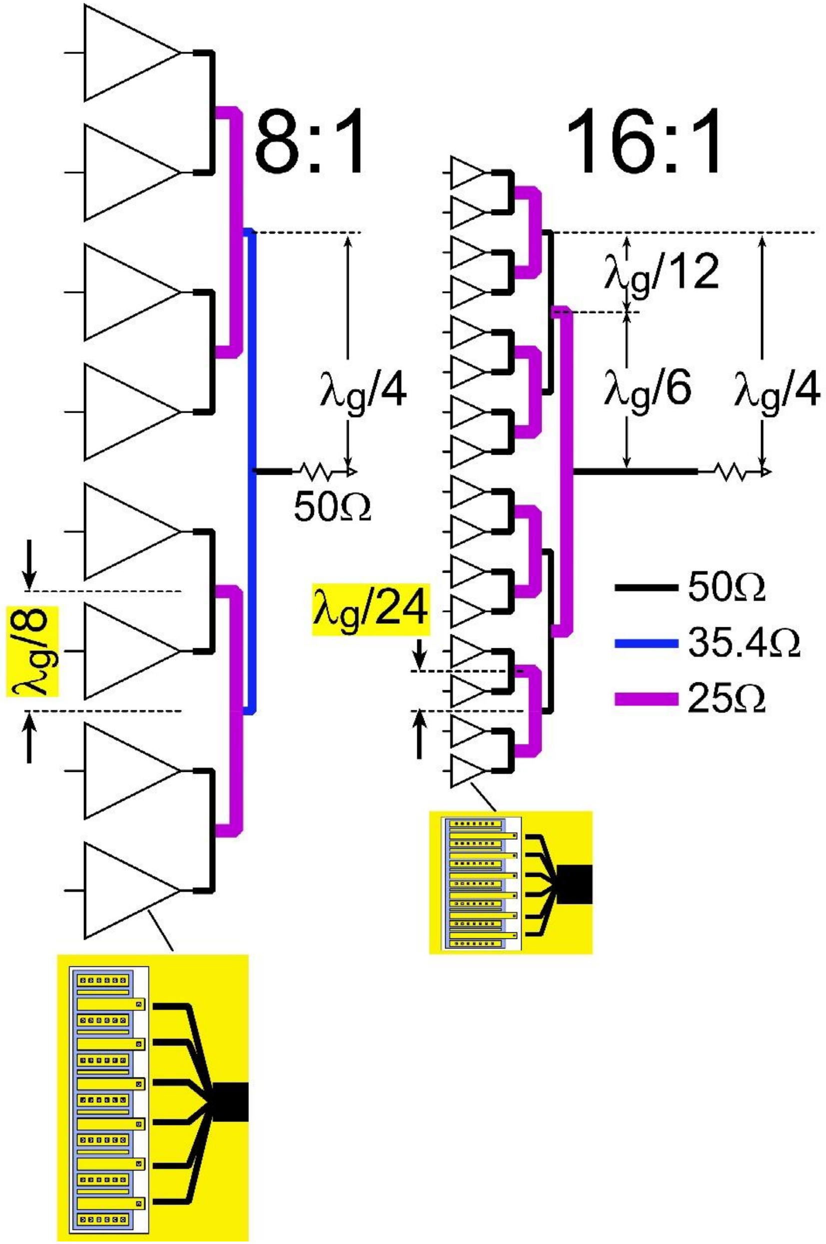
Can 50Ω cell fit in $\lambda_g/8$ (120 μm) pitch ?
→8:1 combining with 0.4dB loss @200GHz

Can 50Ω cell fit in $\lambda_g/24$ (40 μm) pitch ?
→16:1 combining with 0.5dB loss @200GHz



200GHz InP HBT PA: 80 μm cell pitch

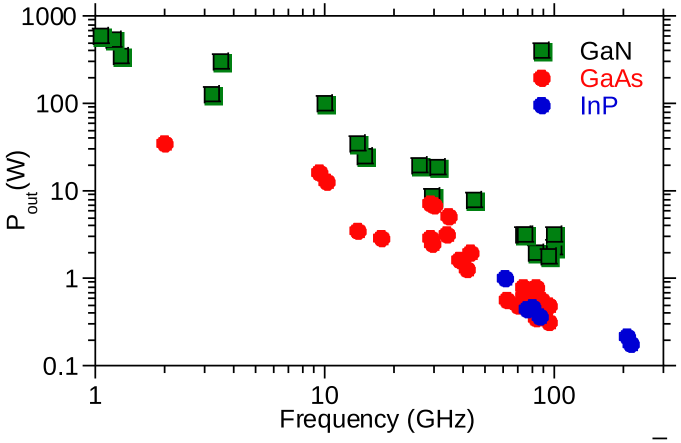


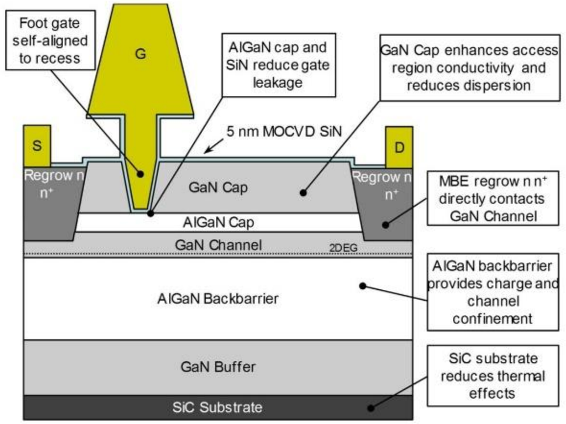


mm-Wave Transistor Development

InGaN and GaN HEMTs:

Leading power technology to ~110GHz
Efforts to extend this to 140, 220GHz.

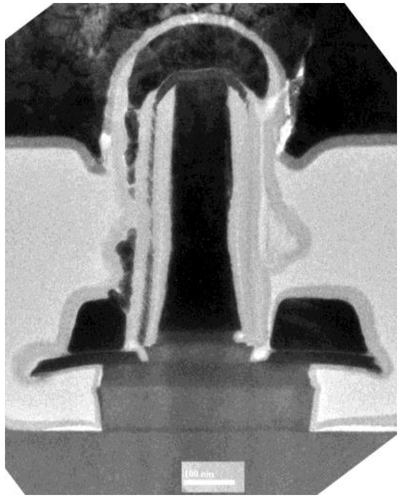


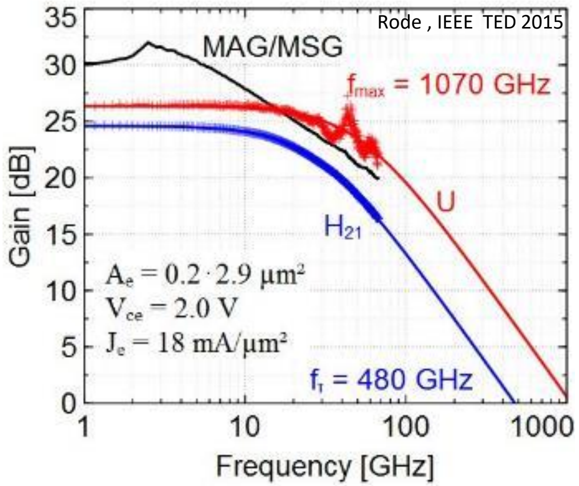


N-polar GaN: Mishra, UCSB

THz InP HBTs:

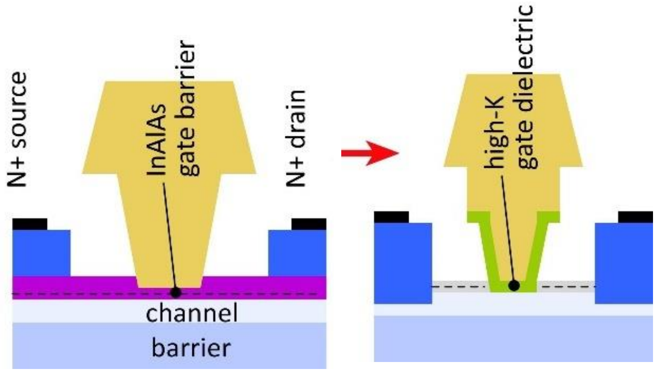
State-of-art: 1.1THz f_{max} @ 130nm node (Teledyne: Urteaga, DRC 2011)
Efficient 100-650GHz power

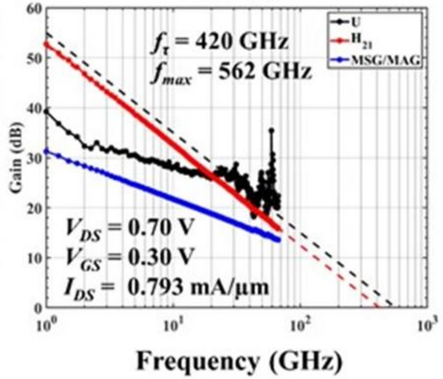




THz InP HEMTs:

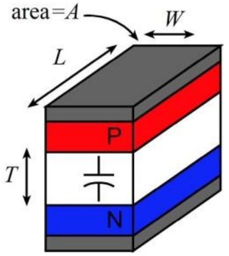
State-of-art: 1.5THz f_{max} @ 32nm node (NGST: X. Mei, EDL 2015)
Sensitive 100-650GHz low-noise amplifiers
high-K gate dielectric *might* permit further scaling.



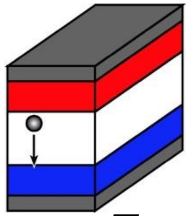


Transistor scaling laws: (V,I,R,C,τ) vs. geometry

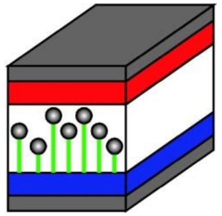
Depletion Layers



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

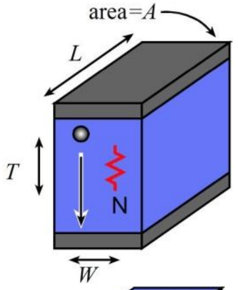


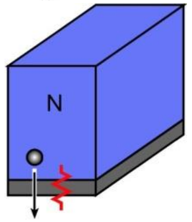
$$\tau = \frac{T}{2v}$$



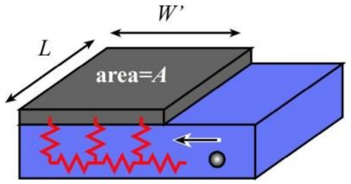
$$\frac{I_{\max}}{A} = \frac{4\varepsilon v_{sat}(V_{appl} + \phi)}{T^2}$$

Bulk and Contact Resistances





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

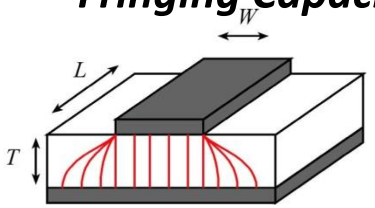


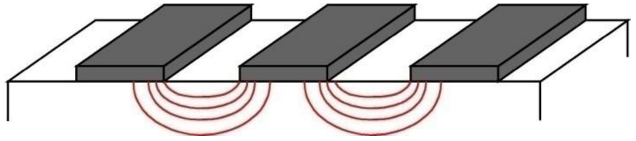
contact terms dominate

Fringing Capacitances

1) FET fringing capacitances

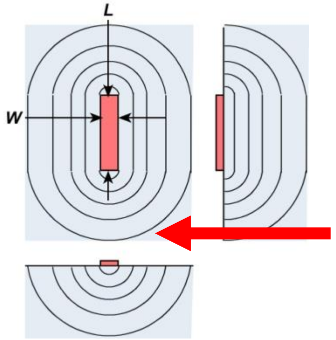
2) IC interconnect capacitances



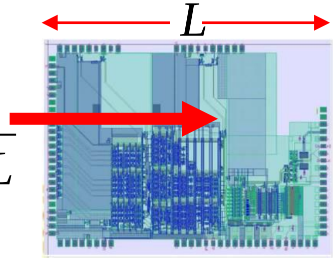


$$C_{finging} / L \sim \varepsilon \qquad C_{finging} / L \sim \varepsilon$$

Thermal Resistance

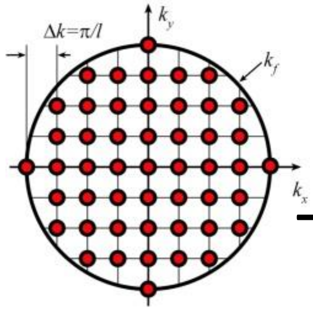


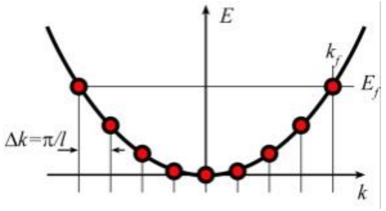
$$\Delta T_{IC} \propto \frac{P_{IC}}{K_{th} L}$$



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

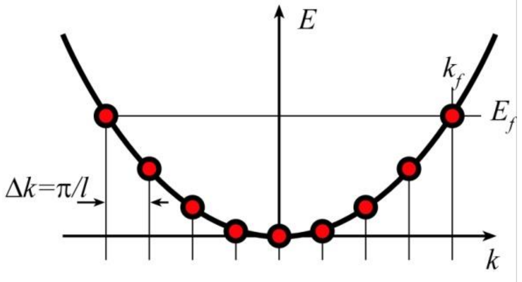
Available quantum states to carry current



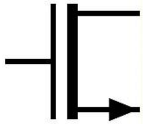


→ ***capacitance, transconductance contact resistance***

Degenerate State Density (Ballistic) Limits



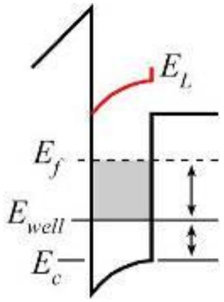
$$\text{Charge} = \int_{\text{Band edge}}^{\text{Fermi Energy}} q \cdot n(E)dE \qquad \text{Current} = \int_{\text{Band edge}}^{\text{Fermi Energy}} q \cdot v(E) \cdot n(E)dE$$



$$J \propto m^{1/2}(E_f - E_{well})^{3/2} \propto (V_{gs} - V_{th})^{3/2}$$

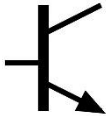
$$\text{not } (\mu c_{ox} / L_g)(V_{gs} - V_{th})^2$$

"ballistic limit"



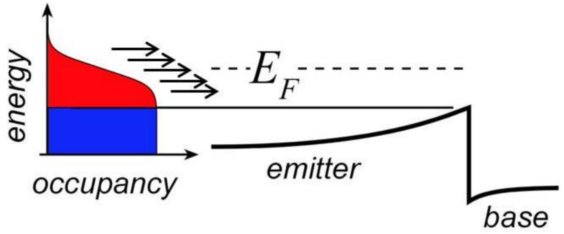
$$\rho_{sheet} = c_{dos}(V_{gs} - V_{th}) \propto m^*(E_f - E_{well})$$

"state density capacitance"

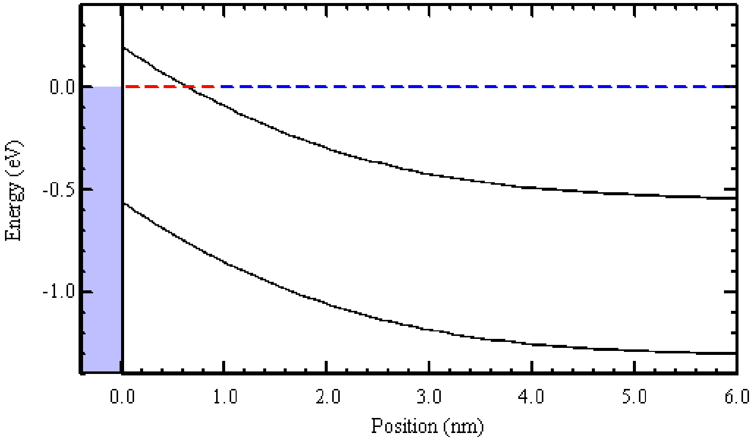


$$J \propto m^*(E_f - E_c)^2 \propto m^*(V_{be} - \varphi)^2$$

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



Contacts



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

Bipolar Transistor Design: Scaling

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

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

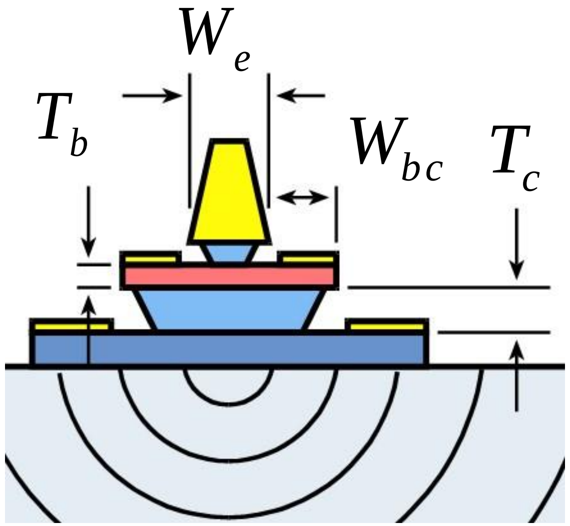
$$C_{cb} = \varepsilon A_c / T_c$$

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

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

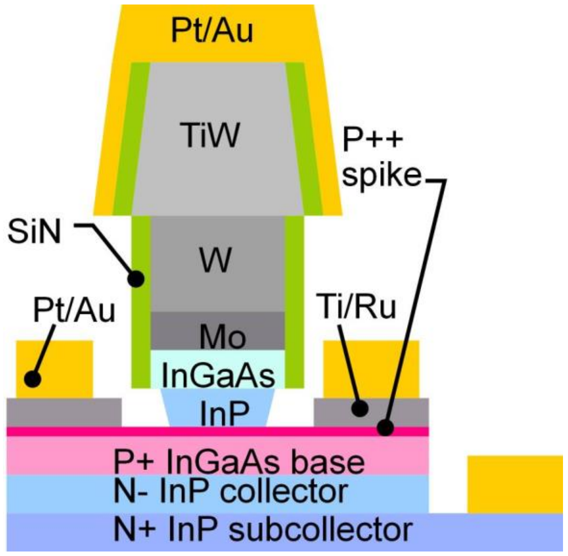
$$R_{ex} = \rho_{\text{contact}} / A_e$$

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



(emitter length L_E)

Bipolar Transistor Scaling Laws



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

Narrow junctions.

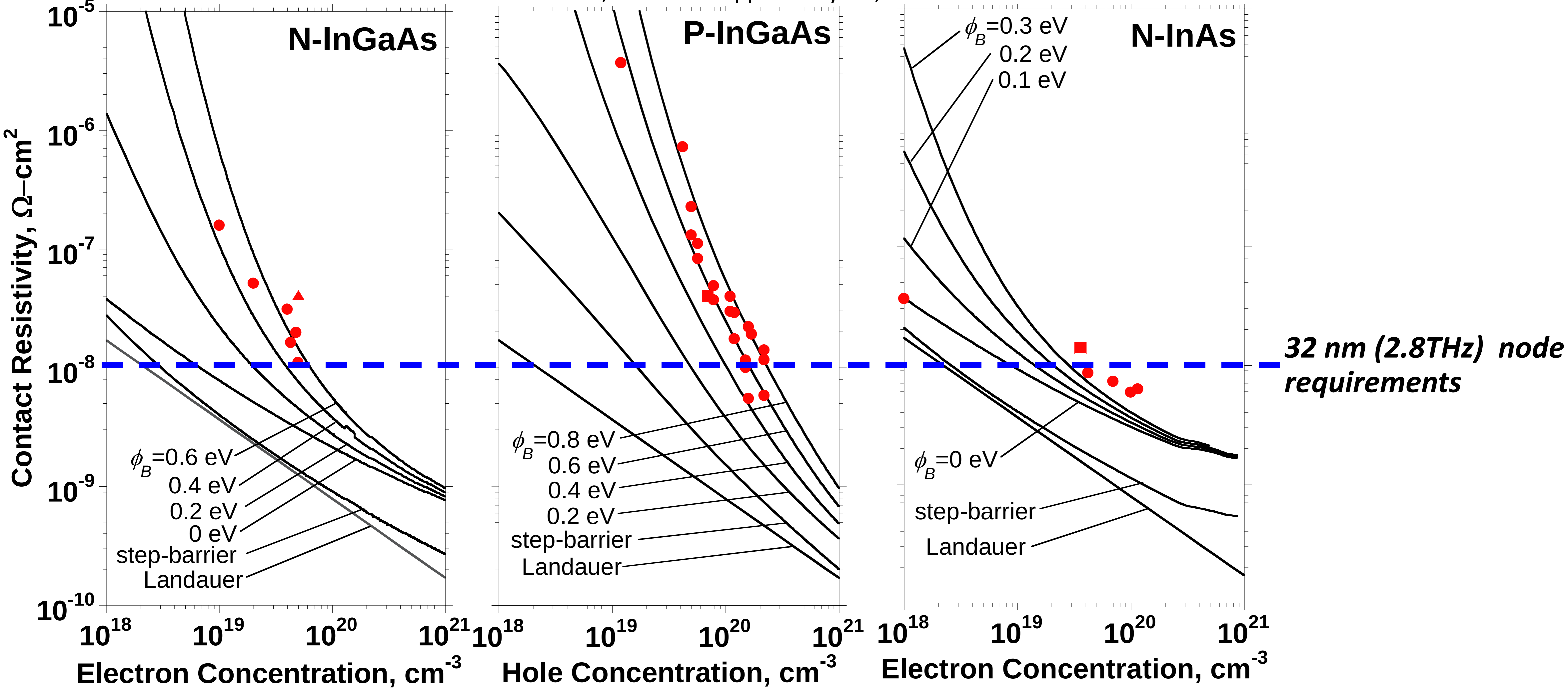
Thin layers

High current density

Ultra low resistivity contacts

Refractory Ohmic Contacts to In(Ga)As

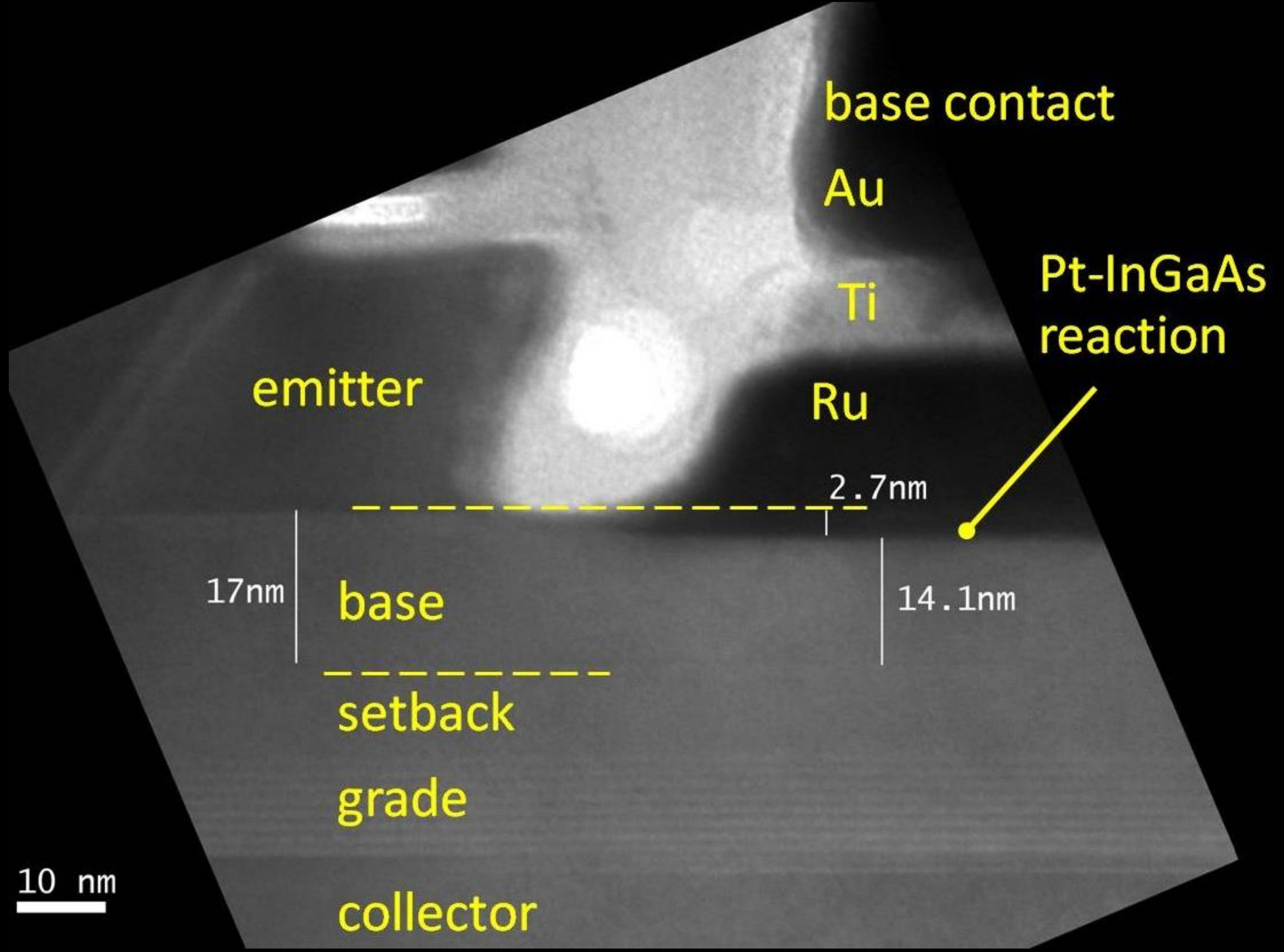
Baraskar *et al*, Journal of Applied Physics, 2013



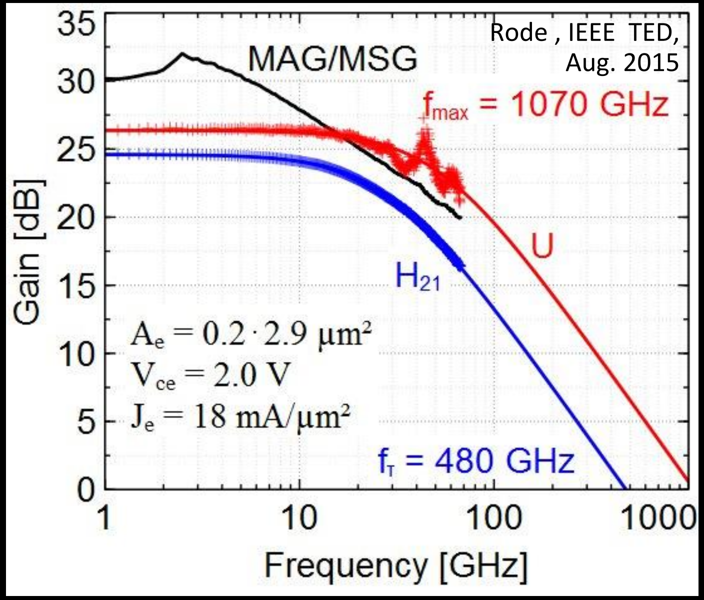
Refractory: robust under high-current operation / Low penetration depth: ~ 1 nm

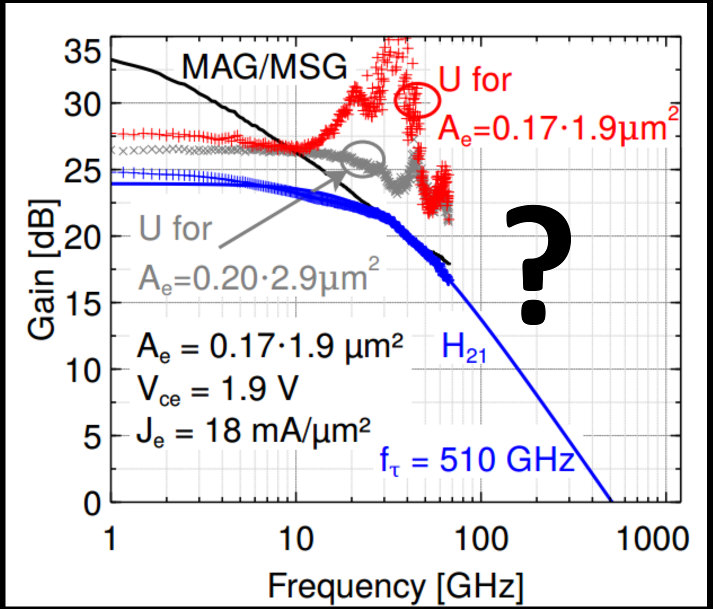
Why no ~2THz HBTs today? Problem: reproducing these base contacts in full HBT process flow

InP HBTs: 1.07 THz @200nm, ?? @ 130nm

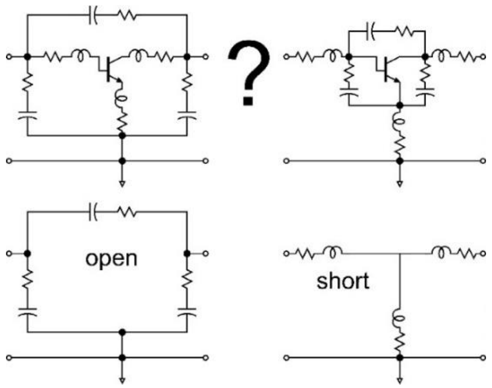


Rode et al., IEEE TED, Aug. 2015

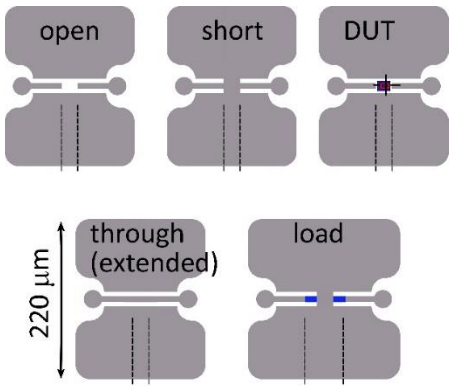


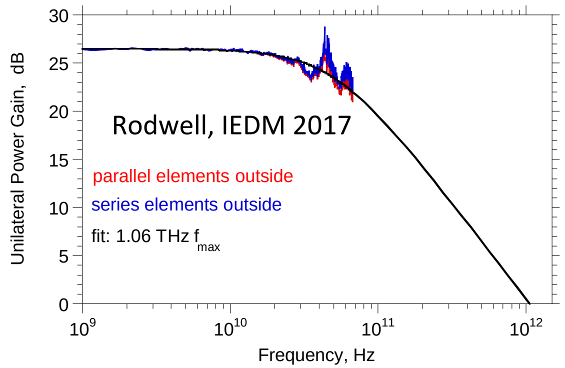


THz Transistor Measurements

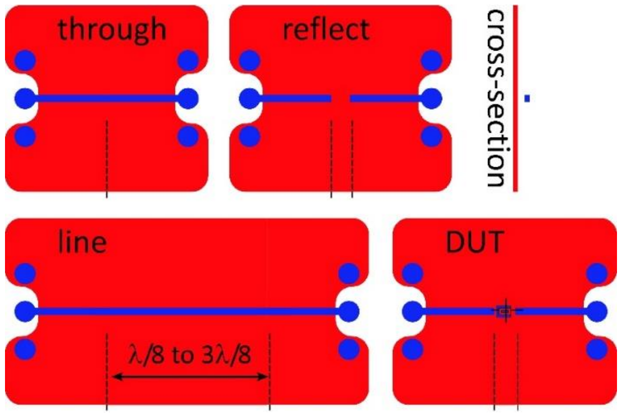


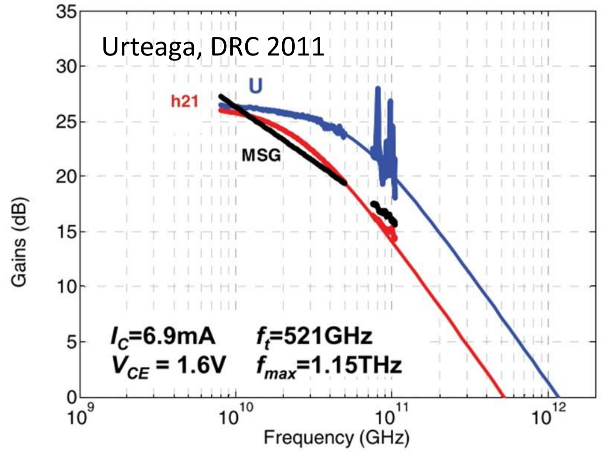
Simple pads:
Substrate coupling: need small pads, narrow CPW
Ambiguity in pad stripping order.
UCSB 130nm HBTs: order not important.
Add through & load to remove ambiguity





On-wafer through-reflect-line:
No ambiguity from pad stripping.
Calibration to line Zo
Still must avoid substrate mode coupling
CPW particularly vulnerable.
better: thin-film microstrip
or ~25 µm substrate with TSV's





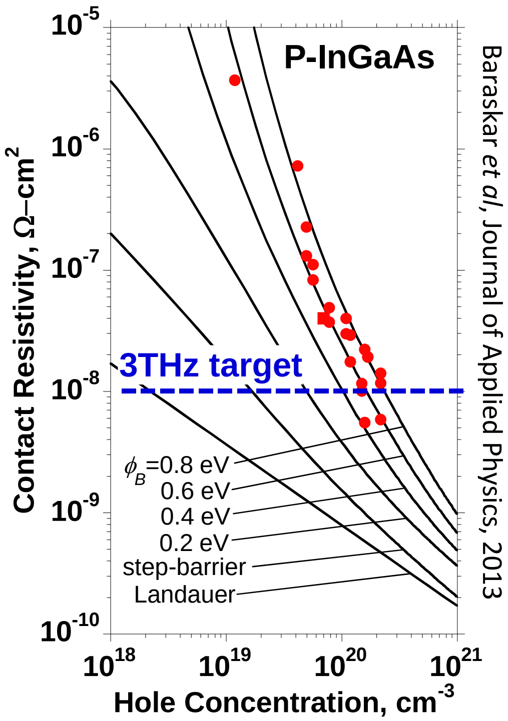
D. F. Williams, A. Young, M. Urteaga, "A Prescription for Sub-Millimeter-Wave Transistor Characterization," IEEE Transactions on Terahertz Science and Technology, July 2013

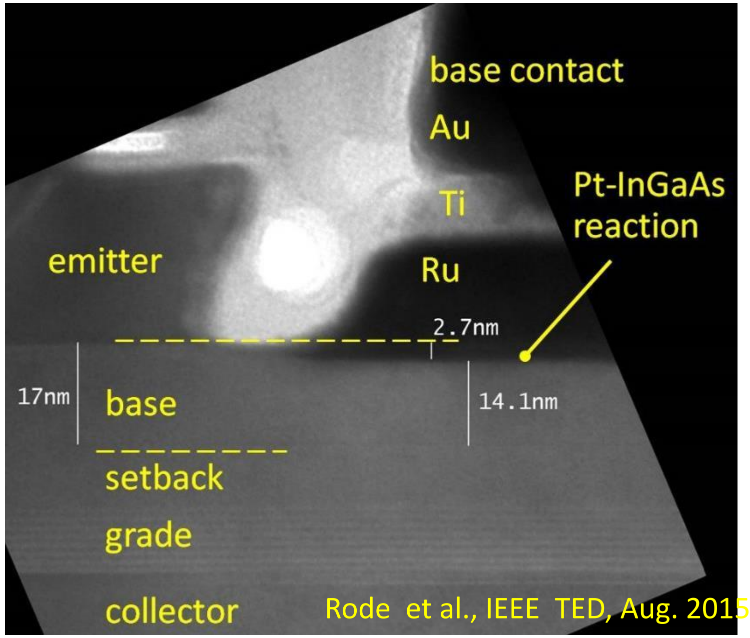
Challenges @ 64nm/2THz, 32nm/3THz Nodes

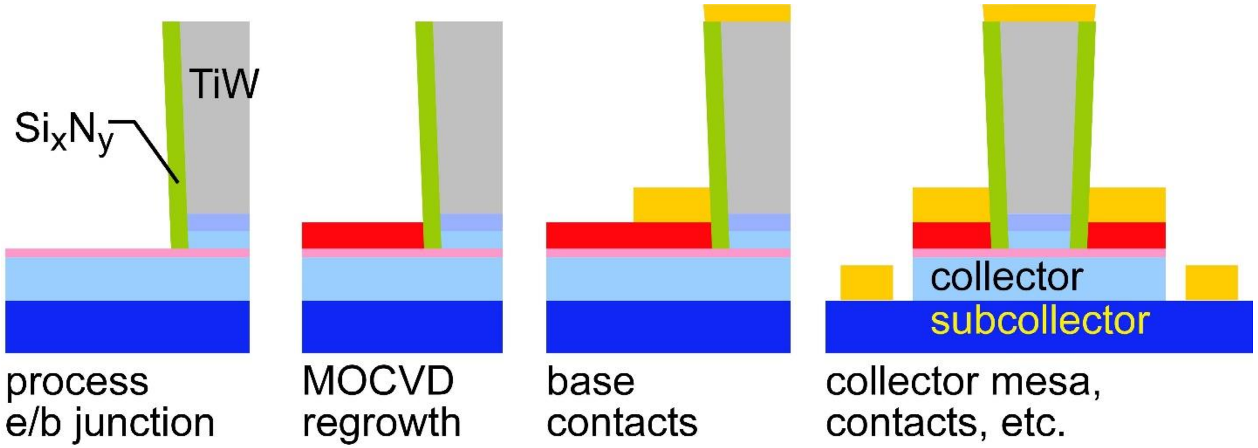
Need high base contact doping
$>10^{20}/cm^3$ for good contacts
high Auger recombination
very low β.

Seem to need 1-3nm contact penetration
Pd or Pt contacts
react with 3++ nm of base
penetrate surface contaminants
too deep for thin base

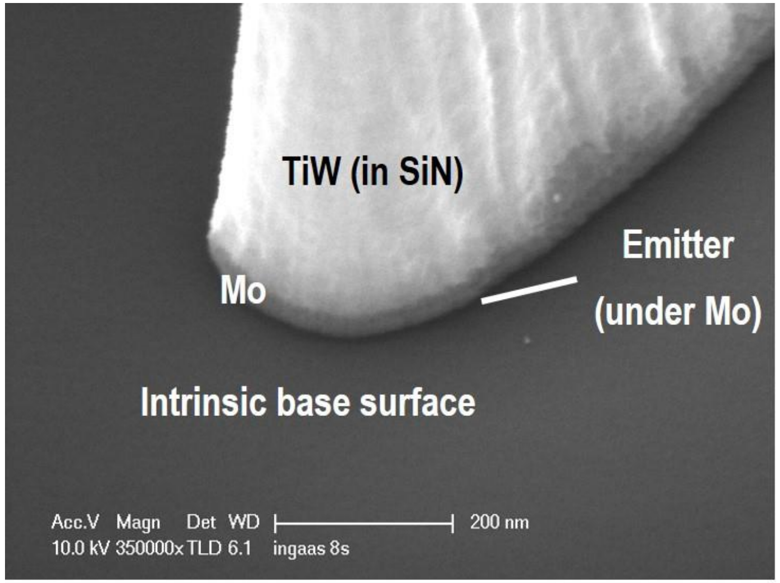
Base regrowth as possible solution
thin, moderately-doped intrinsic base
InGaAs or GaAsSb @ 10^{19}-$10^{20}/cm^3$
thick, heavily-doped extrinsic base
P-GaAs, ~$10^{21}/cm^3$



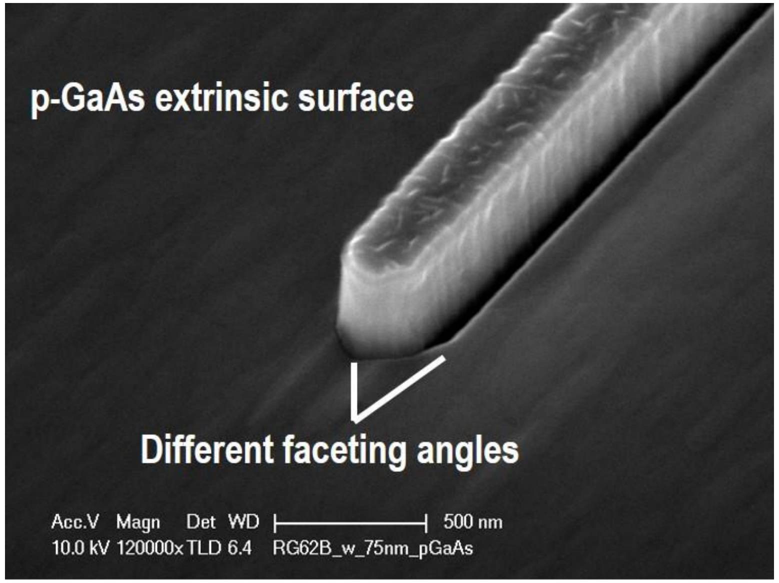




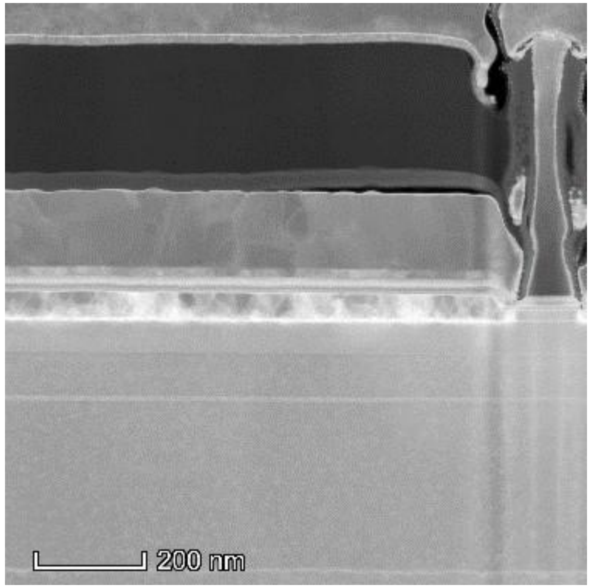
Regrown-Base InP HBTs: Images



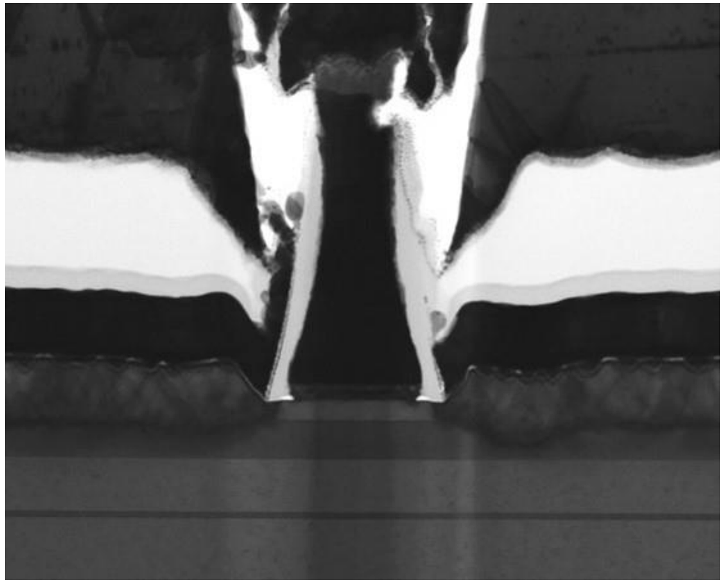
Before regrowth

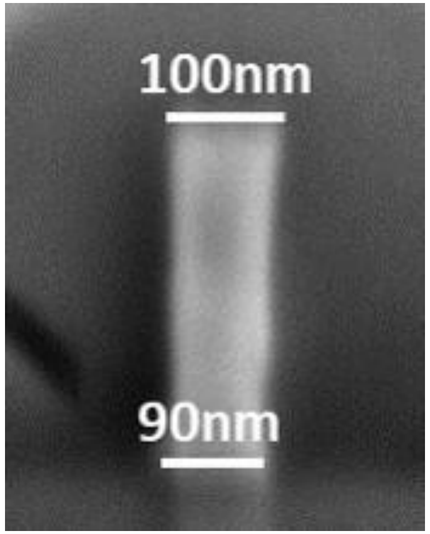


After 100nm p-GaAs regrowth



Cross-sections



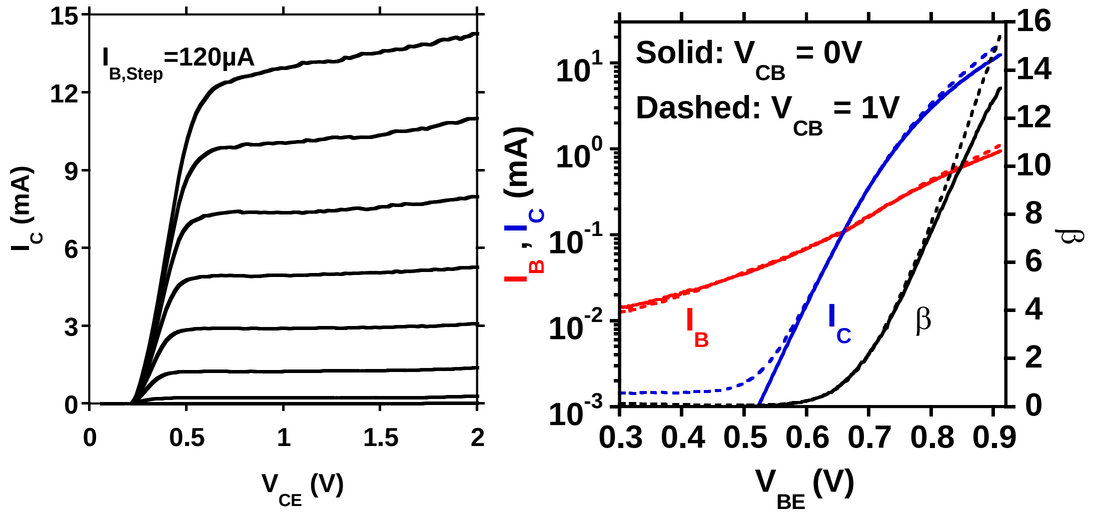


Dry-etched
TiW emitter contact

Regrown-Base InP HBTs: Status

Good DC data: even given regrowth

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



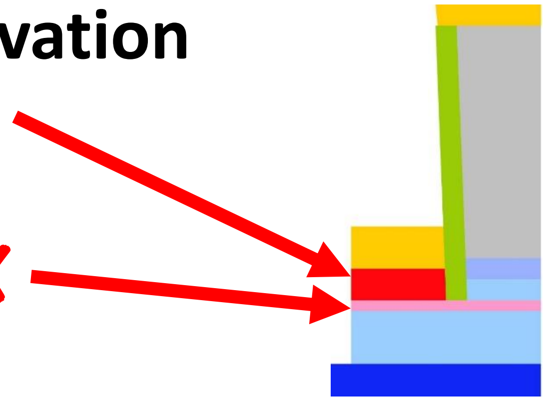
Excellent base contacts; but hydrogen base passivation

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 ✗



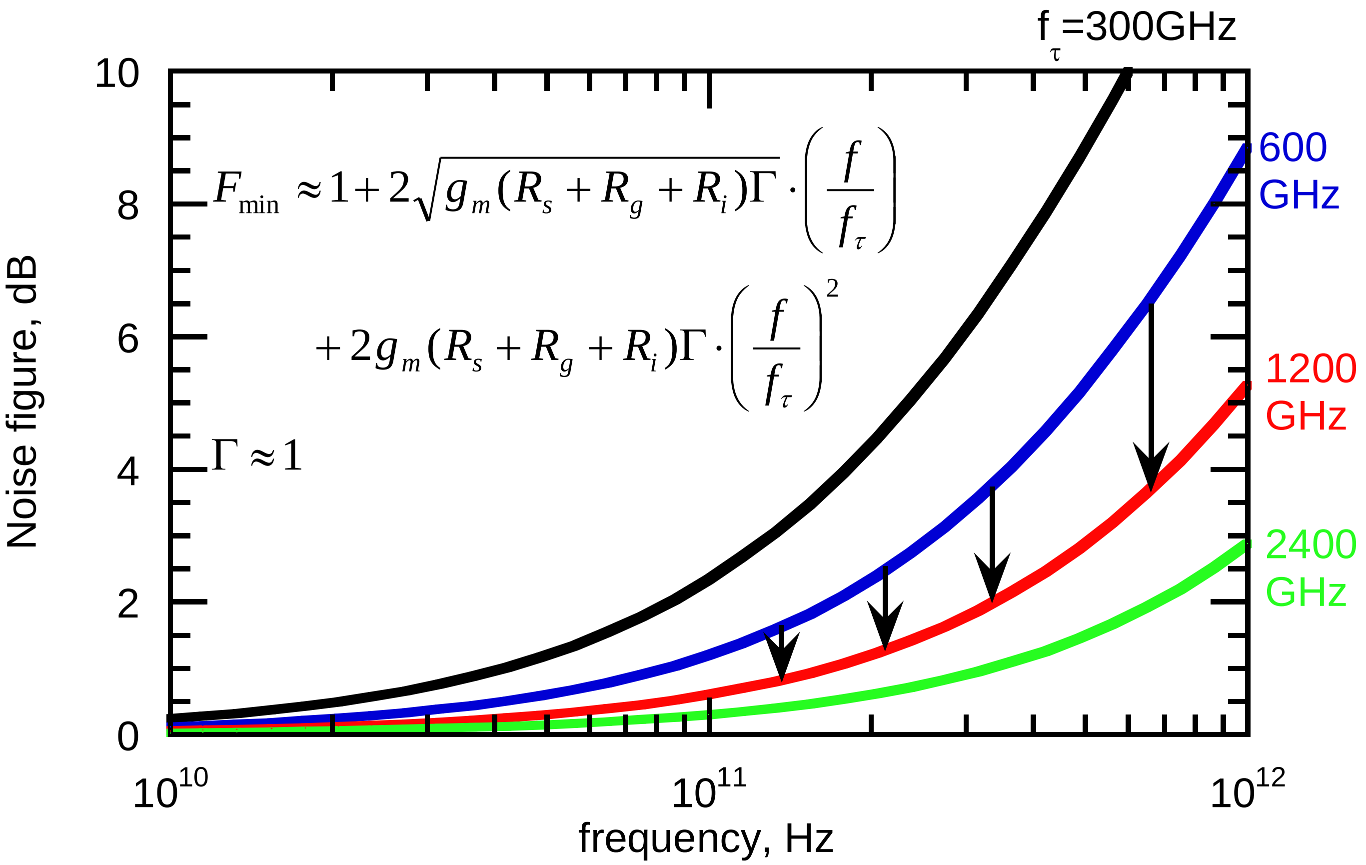
Recent efforts: in-situ MOCVD hydrogen anneal

Preliminary results: marginal ~300GHz f_{max} (still excessive hydrogen)

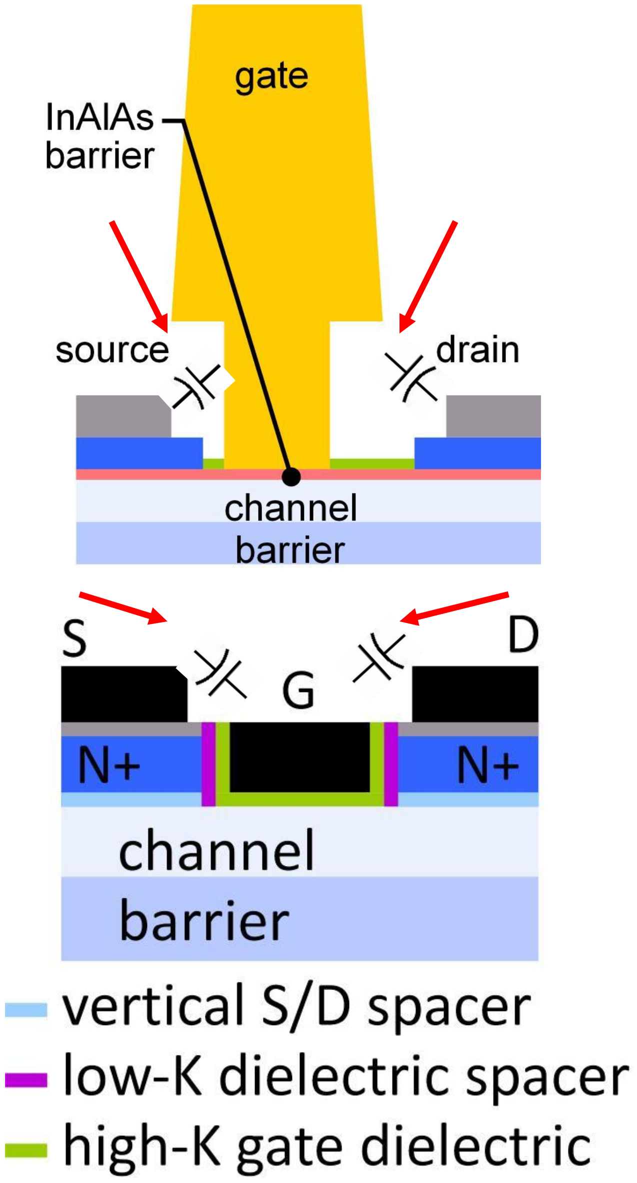
FETs (HEMTs): key for low noise

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

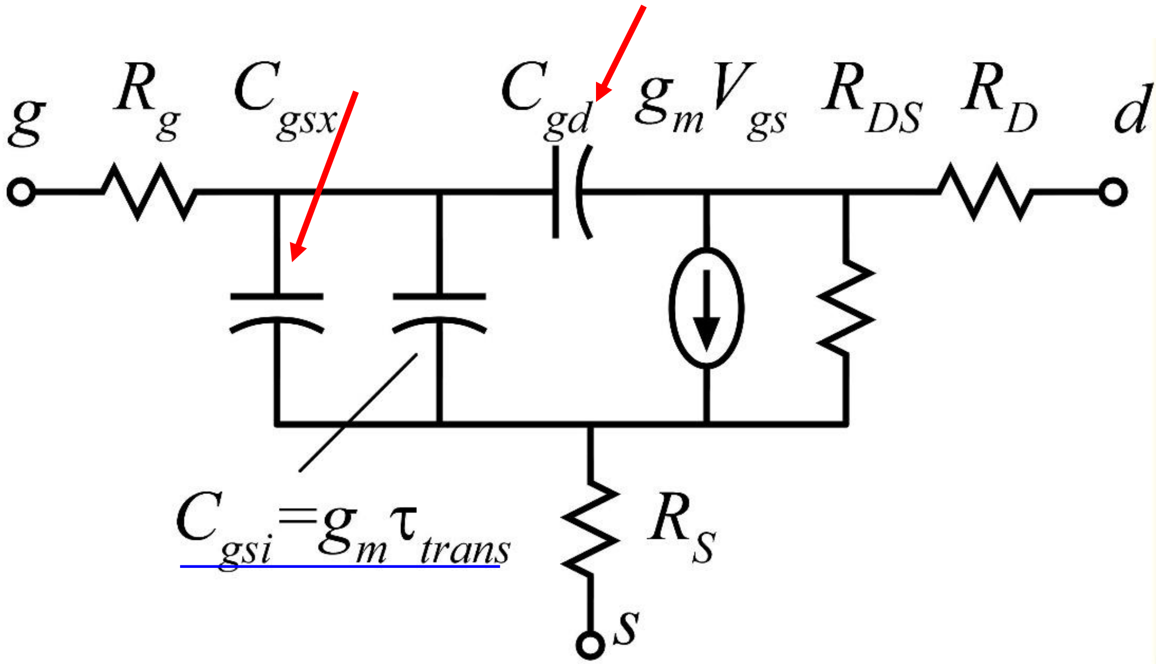


To double f_τ, reduce L_g 2:1, but this is not enough

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 ($g_m R_{ds}$)



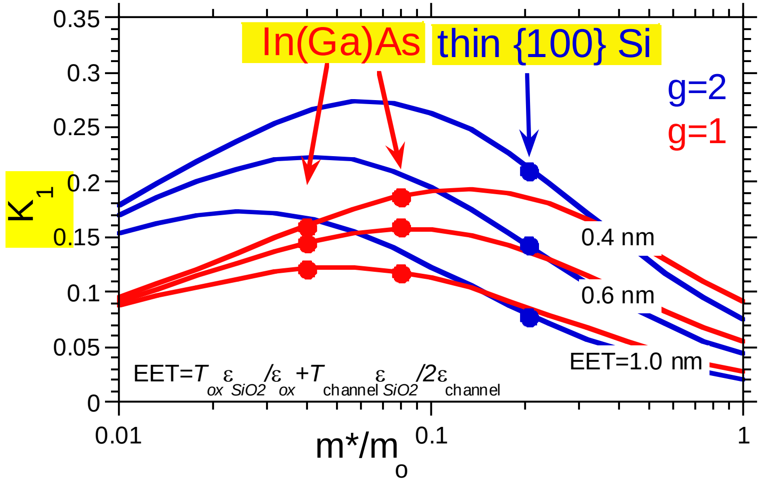
FET Current and Transconductance

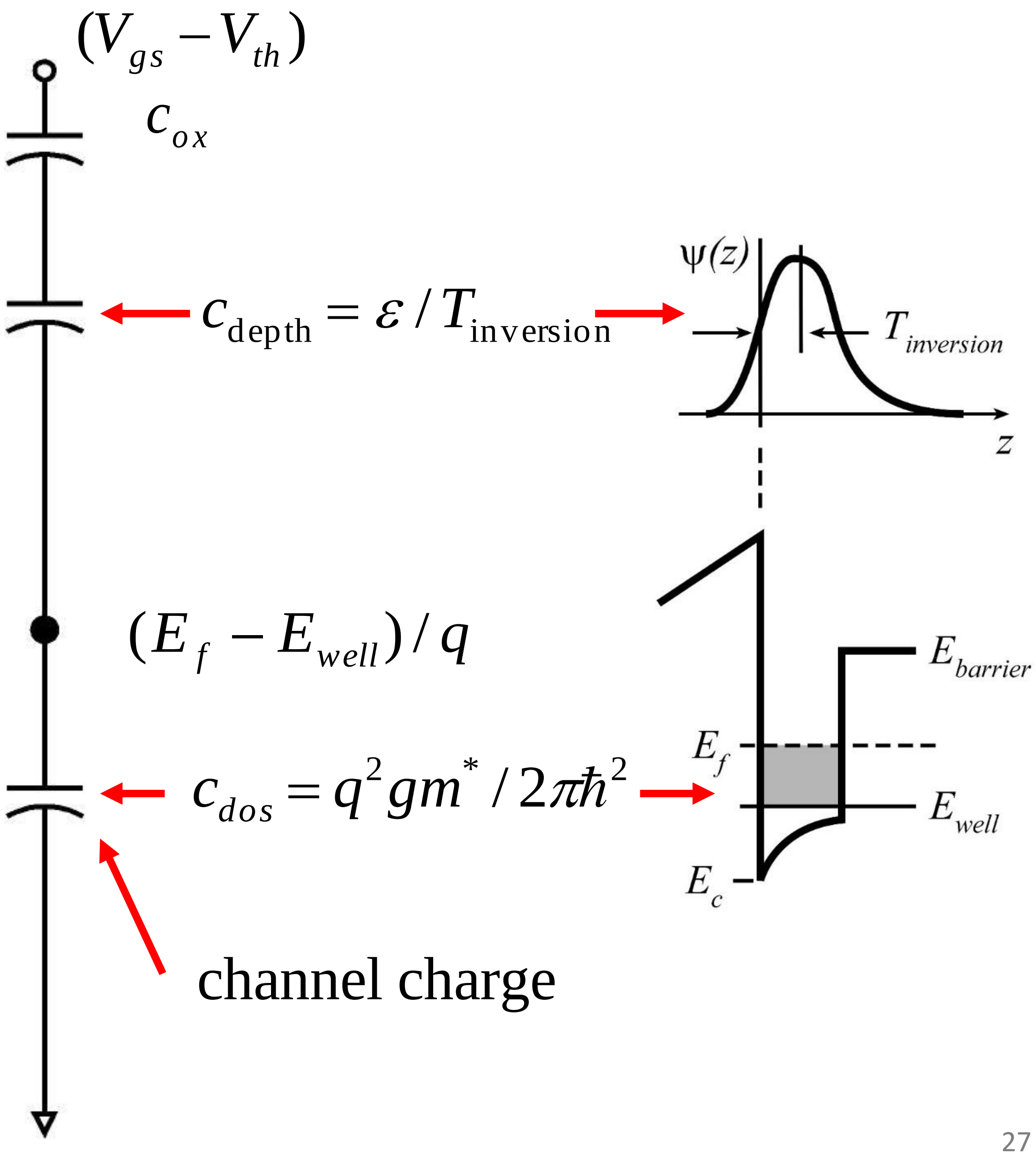
Fermi velocity from $(E_f - E_{well}) = m^* v_f^2 / 2$

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

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

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





To increase g_m:

thin the oxide & channel

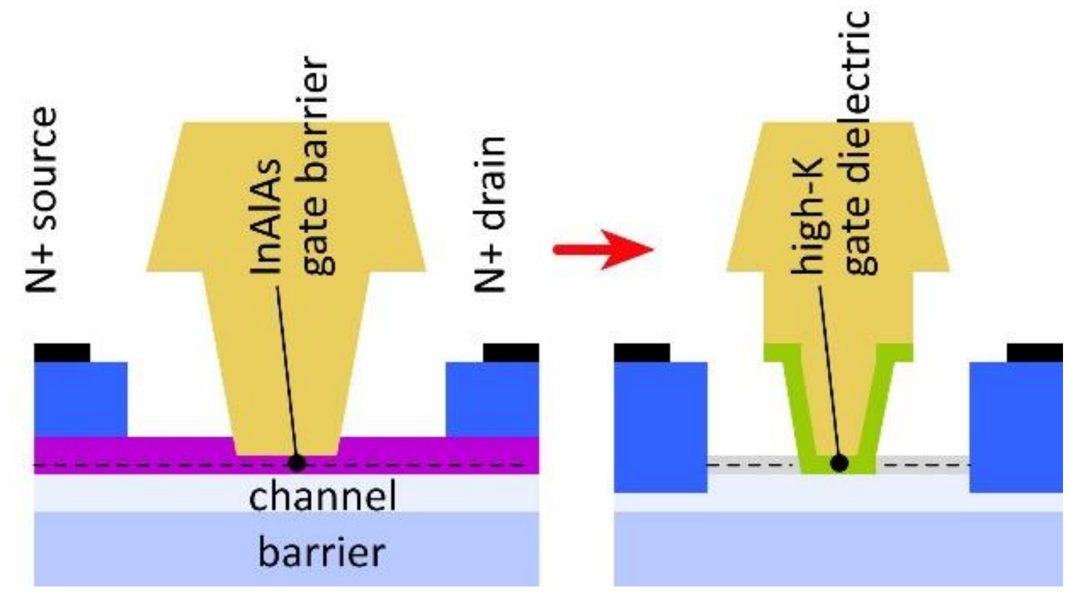
and increase K_1 (mass, # valleys)...**hard**

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

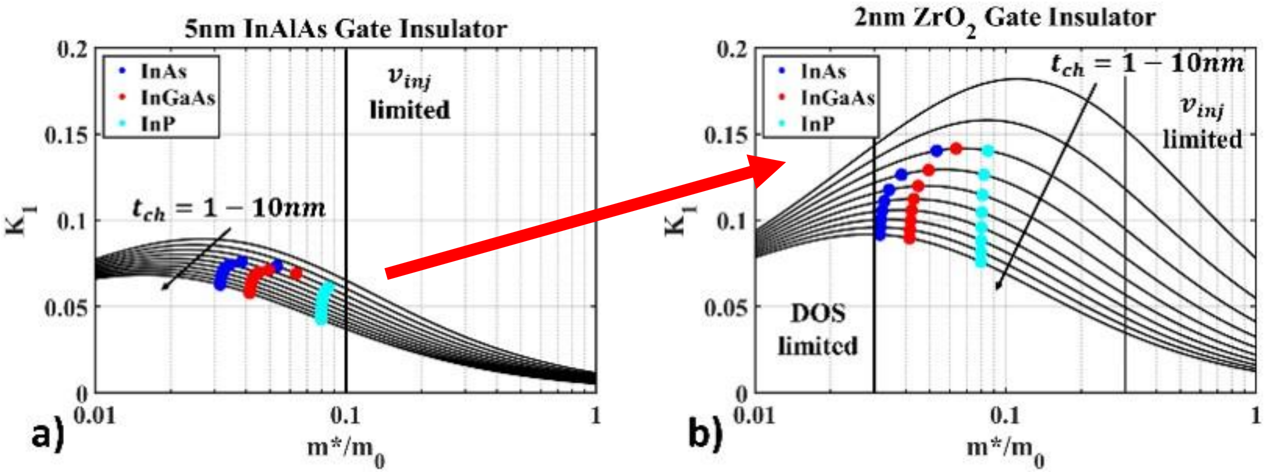
Towards faster HEMTs: InAs MOS-HEMTs

Thinner gate insulator
HEMT: ~6nm InAlAs (ε_r=12), limited by tunneling
MOS-HEMT**:** 2nm ZrO_2 (ε_r=25)

Less source resistance
HEMT: InAlAs barrier under N+ source/drain
MOS-HEMT: N+ layer on InAs channel



Simple ballistic theory: thin dielectric→ increased g_m.
HEMT: InAlAs barrier: tunneling, thermionic leakage



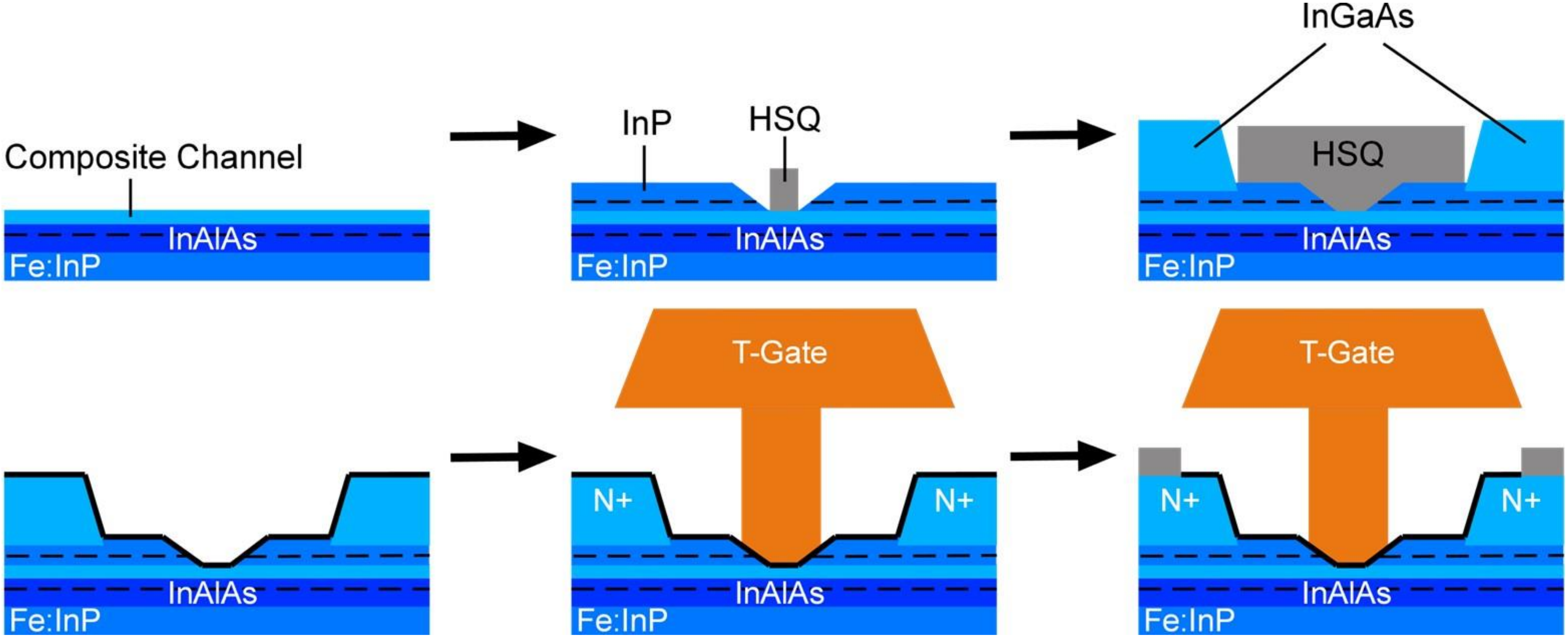
$$J = K_1 \cdot (84\text{mA}/\mu\text{m}) \cdot \left((V_{gs} - V_{th}) / 1\text{V} \right)^{3/2}$$

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

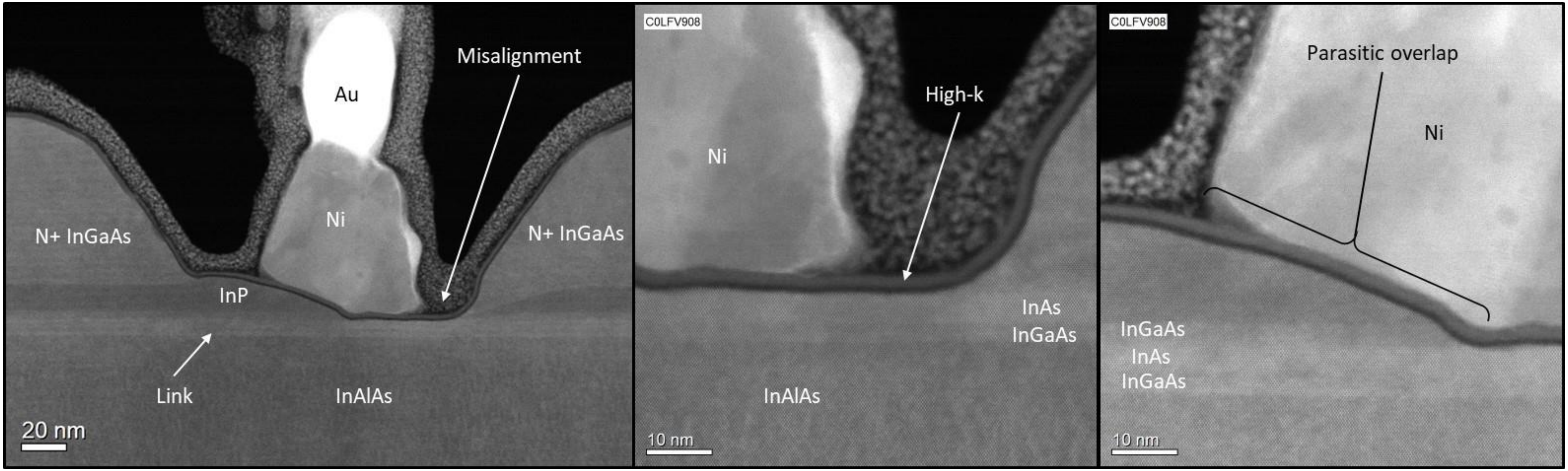
Limitations to theory:
Assumes parabolic *E-k* dispersion: unrealistic
Ignores effect of maximum gate overdrive (V_{gs}-V_{th})

1st MOS-HEMT demonstration:
Fraunhofer IAF / IBM Zurich

MOS-HEMT: fabrication flow



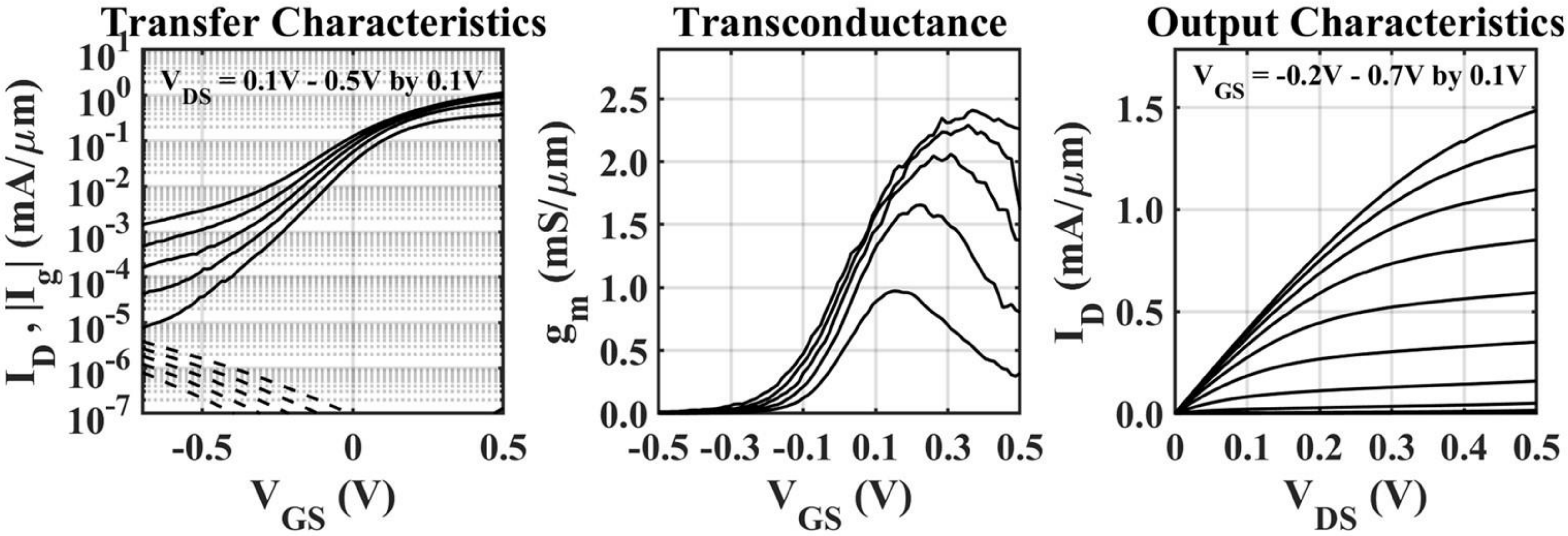
MOS-HEMT: device structure



t_{ch}	Channel Material	ZrO_2 Cycles
7.0 nm	InAs / InGaAs	30

B. Markman et. al, 2021 DRC

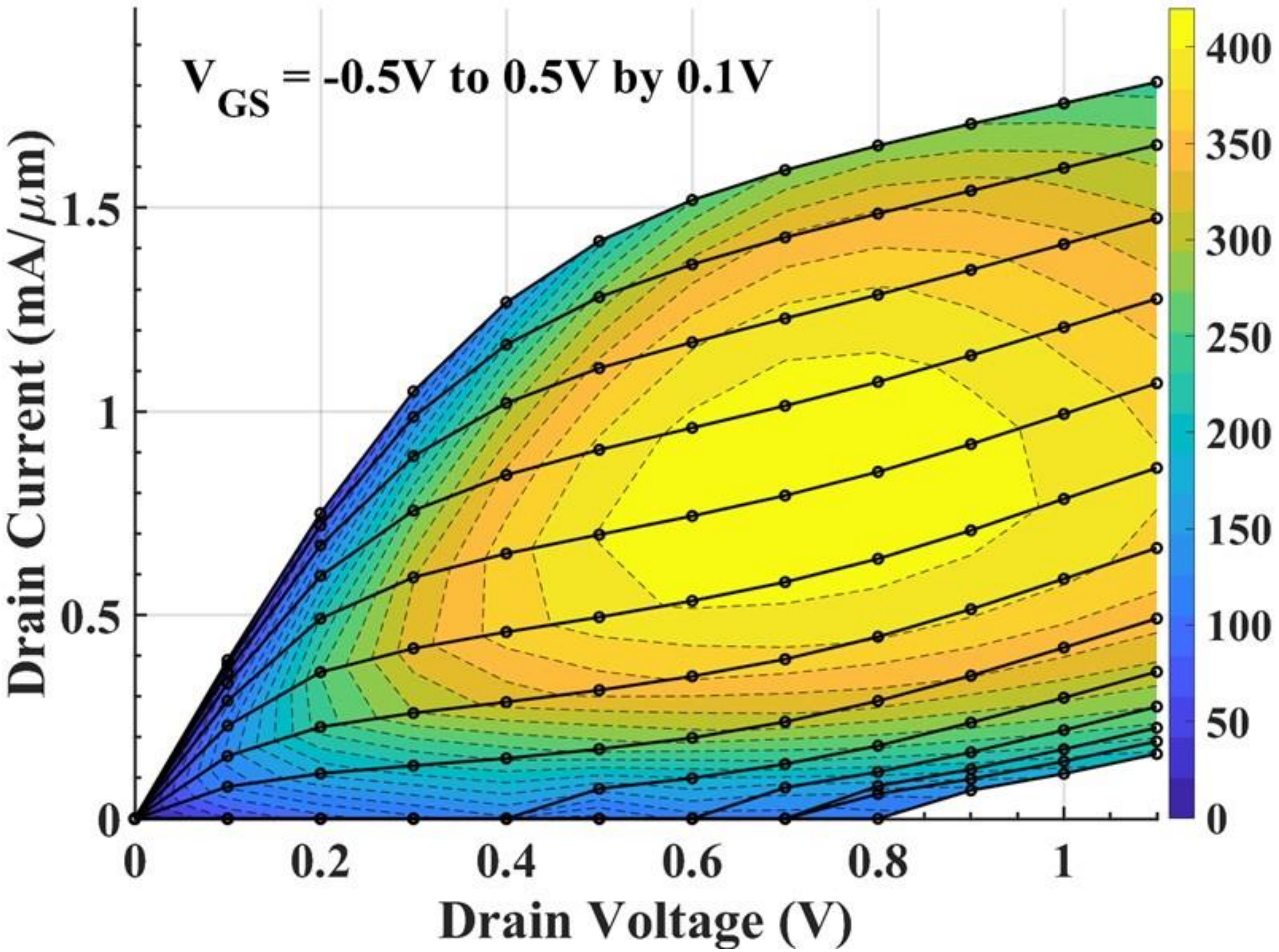
DC characteristics @ 40 nm L_g , 2 × 10μm W_g

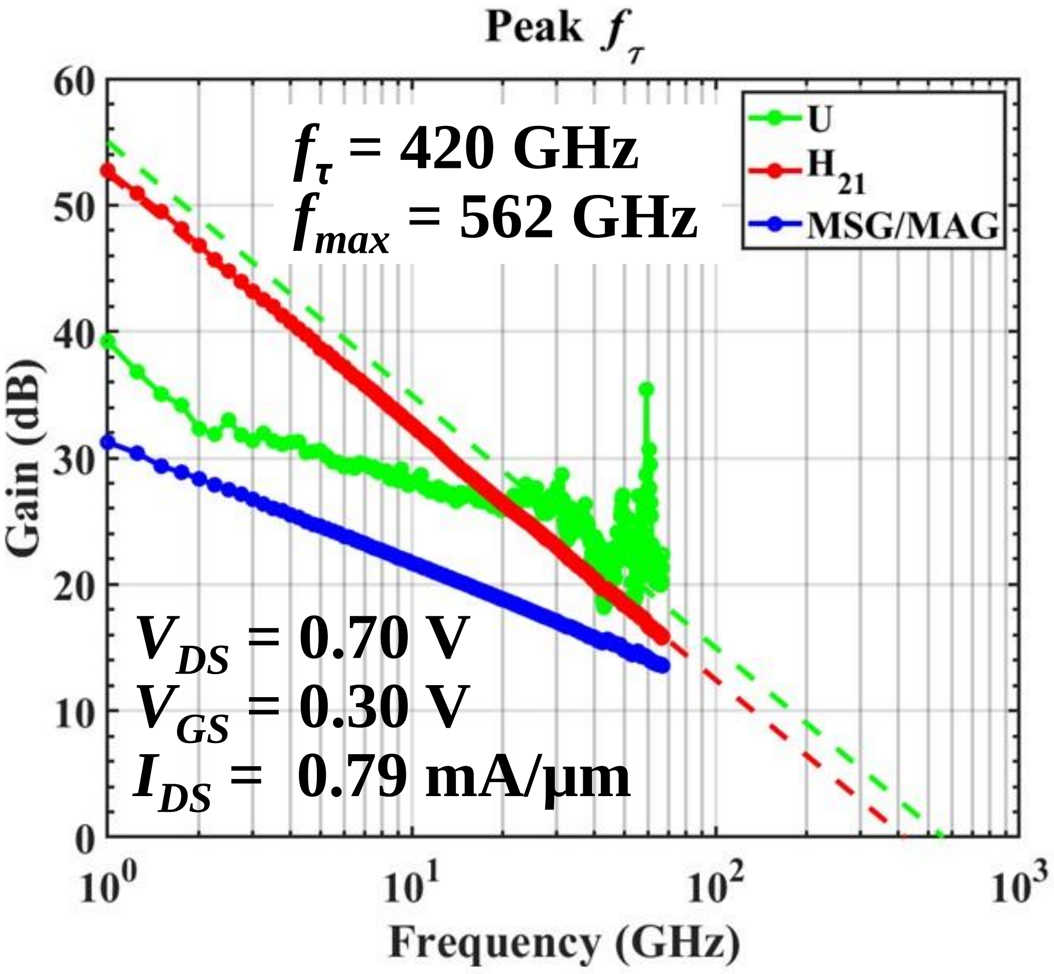


Peak g_m	R_A	I_{on}	I_{off}	I_g	Long L_g SS_{min}
2.4 mS/μm	49 – 55 Ω•μm	> 1.45 mA/μm	< 10 nA/μm	< 10 nA/μm	76 mV/dec

B. Markman et. al, 2021 DRC

RF characteristics @ 40 nm L_g , 2 × 10μm W_g





Peak f_τ = 420 GHz on L_g = 40 nm $(0\bar{1}1)$ conduction device, peak f_{max} at L_g = 50 nm

f_{max} extrapolation difficult because of peaks in U; calibration artifacts or negative resistance

Need for higher energy barriers

To increase transconductance:

thin the oxide, thin the well

→ increased eigenstate energy

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

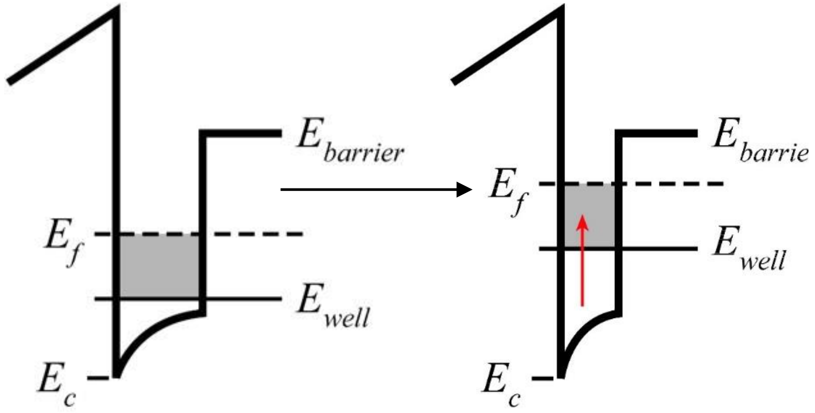
→ constrains maximum transconductance: $I_D \propto (V_{gs}-V_{th})^{3/2} \rightarrow g_m \propto (V_{gs}-V_{th})^{1/2}$

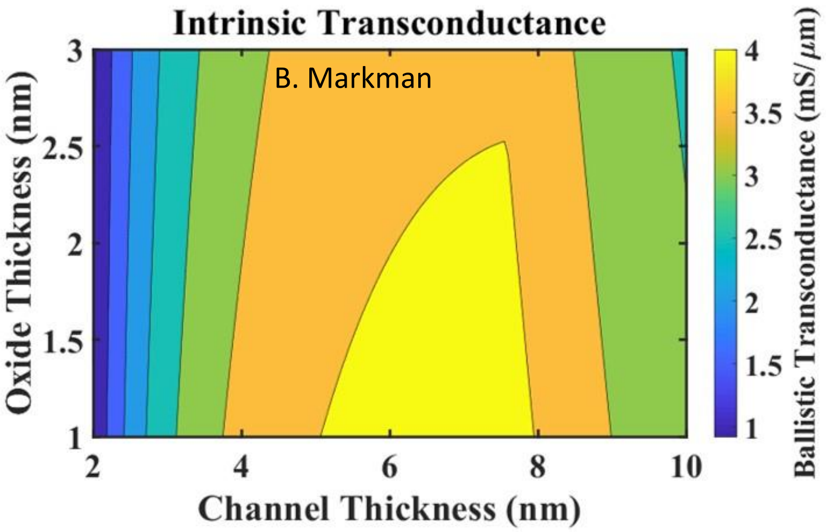
→ maximum achievable g_m.

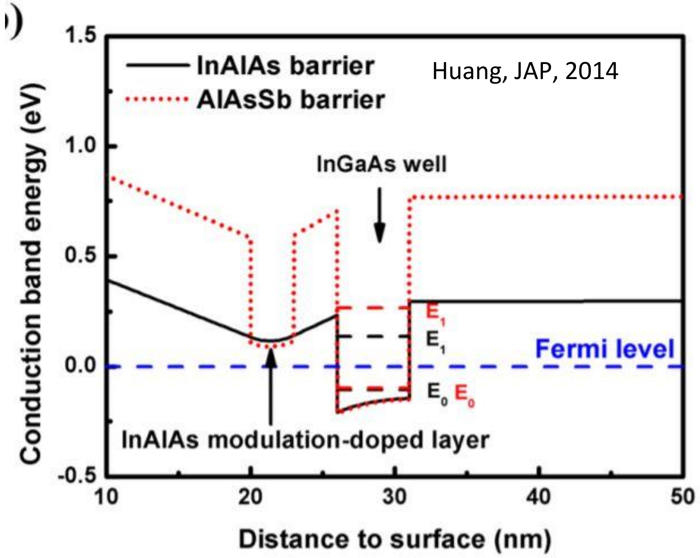
Need high barrier energies

InAs/InAlAs vs InAs/AlAsSb

InP/AlAsSb ????







Transistors for 100-300GHz wireless

Systems

Multi-beam (MIMO) endpoint and backhaul links.
Imaging radar

Transistor parameters

LNAs: cascaded noise figure
PAs: high PAE, high power density (W/mm)
high $f_\tau \times V_{br}$
PAs need high A/mm & closely-spaced fingers

Today's available IC technologies

CMOS: good to ~150GHz.   65-32nm nodes are best.
SiGe: surpasses CMOS above 200GHz.
InP HBT: record 100-300GHz PAs
GaN HEMT: record power below 100GHz. Improving
InGaAs FETs: record LNAs

Improved InP HBTs

goal: improved PAE in 100-300GHz PAs.
challenge: base contact resistivity scaling. Process complexity.

Improved InGaAs FETs

High-K gate dielectric may permit further scaling.
High-K / InP / AlAsSb ?

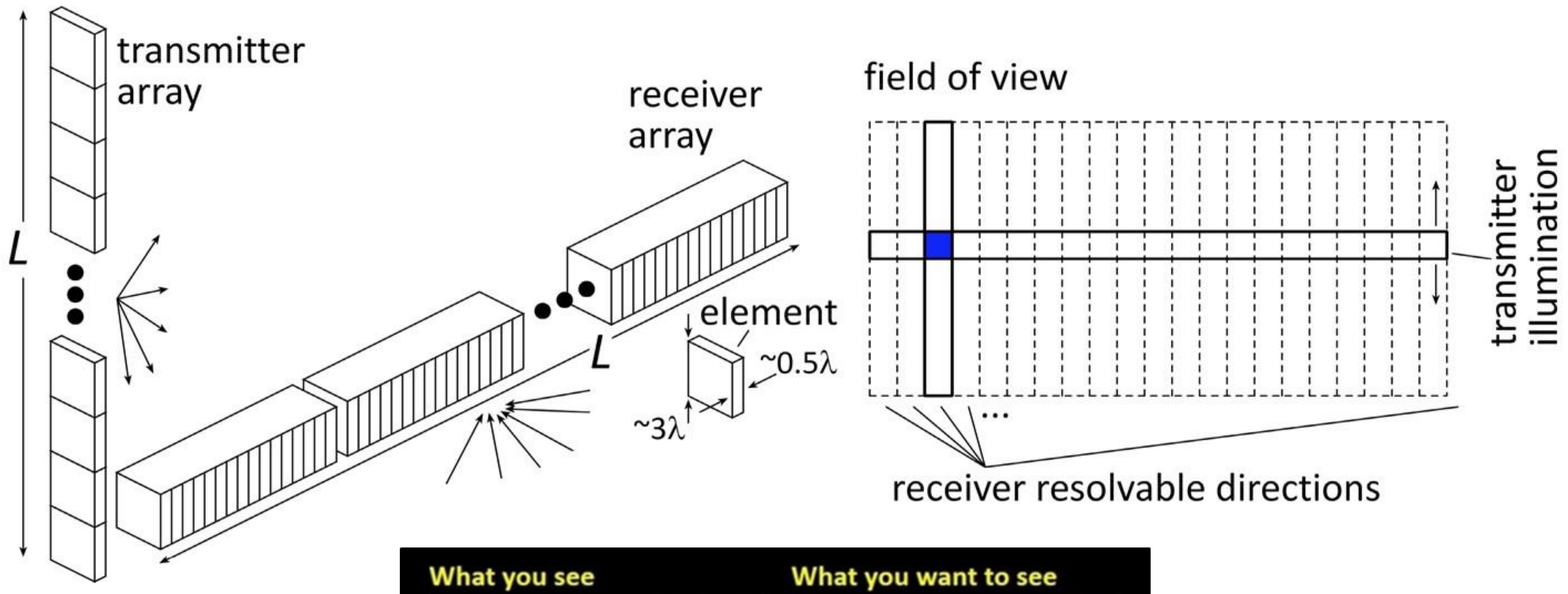
In case of questions

210 GHz FMCW crossed-array imaging car radar

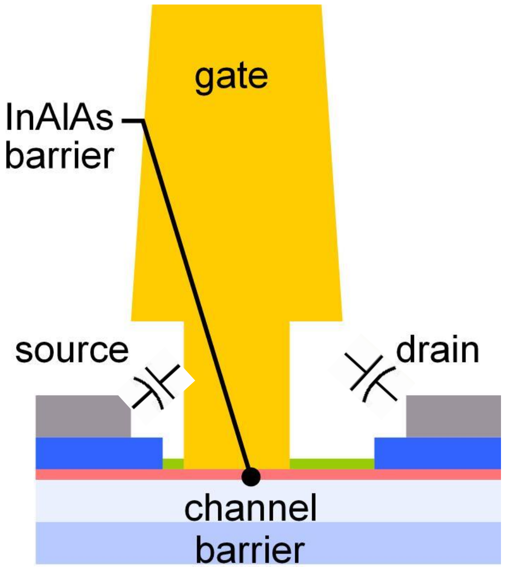
Array:
36×1 transmit,   1×216 receive
36 (v) × 216 (h) image
length: 15cm (6 inches),
beamwidth: 0.27º,
view: 10º (v) ×90º (h).
scan: 40Hz

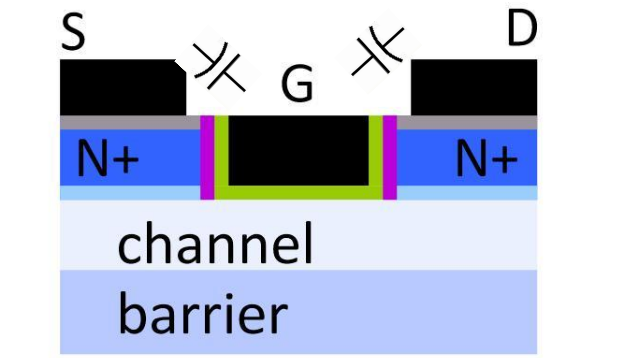
Electronics
transmit power/element: 50mW
receiver noise: 6dB
packaging losses: 2dB TX, 2dB RX

Sees:
22cm diameter target (a soccer ball)
@ -10dB reflectivity
200m range,
with 10dB SNR
in heavy fog/rain @ 22dB/km
with 4dB operating margins.



FET Scaling Laws (these now broken)





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