THz & nm Transistor Electronics: It's All About The Interfaces.

M. Rodwell, Art Gossard
University of California, Santa Barbara

Collaborators (III-V MOS)
A. Gossard, S. Stemmer, C. Van de Walle
University of California Santa Barbara
P. Asbeck, A. Kummel, Y. Taur,
University of California San Diego
J. Harris, P. McIntyre,
Stanford University
C. Palmstrøm,
University of Minnesota
M. Fischetti
University of Massachusetts Amherst

rodwell@ece.ucsb.edu  805-893-3244, 805-893-5705 fax
TeraHertz and nanoMeter Electron Devices

How do we make very fast electron devices?
...by scaling

What are the limits to scaling?
- attainable contact resistivities,
- attainable thermal resistivities
- attainable contact stabilities
- and for FETs, attainable capacitance densities

How can the materials growth community help?
- work on interfaces (contacts and gate dielectrics)!

Guidance of utility of other device structures / features
- nanowire pillar devices
- access resistances & capacitances
- relevance and irrelevance of mobility
THz & nm
Semiconductor Device Design...

... is scaling
Frequency Limits and Scaling Laws of (most) Electron Devices

\[ \tau \propto \) thickness \]
\[ C \propto \text{area} / \text{thickness} \]
\[ R_{\text{top}} \propto \rho_{\text{contact}} / \text{area} \]
\[ R_{\text{bottom}} \propto 1 / \text{stripe length} \]
\[ I_{\text{max, space-charge-limit}} \propto \text{area} / (\text{thickness})^2 \]
\[ \Delta T \propto \frac{\text{power}}{\text{length}} \times \log \left( \frac{\text{length}}{\text{width}} \right) \]

To double bandwidth,
reduce thicknesses 2:1
reduce width 4:1, keep constant length
current density has increased 4:1
resistance  capacitance  transit time

device bandwidth

applies to almost all semiconductor devices:
transistors: BJT.s & HBT.s, MOSFET.s & HEMT.s, Schottky diodes, photodiodes, photo mixers, RTDs, ...

high current density, low resistivity contacts, epitaxial & lithographic scaling
FET.s only: high $\varepsilon_r\varepsilon_0/D$ dielectrics

THz semiconductor devices
Why aren't semiconductor lasers R/C/τ limited?

Dielectric waveguide mode confines AC field away from resistive bulk and contact regions.

AC signal is not coupled through electrical contacts.

Dielectric mode confinement is harder at lower frequencies.
Bipolar Transistor Design
Bipolar Transistor Design is Simple

\[ \tau_b \approx \frac{T_b^2}{2D_n} \]

\[ \tau_c = \frac{T_c}{2\nu_{sat}} \]

\[ R_{ex} = \frac{\rho_{contact}}{A_e} \]

\[ R_{bb} = \frac{\rho_{sheet}}{12L_e} + \frac{W_{bc}}{6L_e} + \frac{\rho_{contact}}{A_{contacts}} \]

\[ C_{cb} = \frac{\varepsilon A_e}{T_c} \]

\[ I_{c,Kirk} \propto \nu_{sat} A_e (V_{ce, operating} + V_{ce, punch-through}) / T_c^2 \]

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

Goal: double transistor bandwidth when used in any circuit
→ keep constant all resistances, voltages, currents
→ reduce 2:1 all capacitances and all transport delays

(emitter length $L_E$)
HBT scaling laws

Goal: double transistor bandwidth when used in any circuit
→ keep constant all resistances, voltages, currents
→ reduce 2:1 all capacitances and all transport delays

\[ \tau_b = \frac{T_b^2}{2D_n + T_b / \nu} \quad \rightarrow \text{thin base } \sim 1.414:1 \]
\[ \tau_c = \frac{T_c}{2\nu} \quad \rightarrow \text{thin collector } 2:1 \]

\[ C_{cb} \propto \frac{A_c}{T_c} \quad \rightarrow \text{reduce junction areas } 4:1 \]
\[ R_{ex} = \rho_c / A_e \quad \rightarrow \text{reduce emitter contact resistivity } 4:1 \]
\[ I_{c,Kirk} \propto \frac{A_e}{T_c^2} \quad \text{(current remains constant, as desired)} \]

\[ \Delta T \approx \frac{P}{\pi K_{lnP} L_e} \ln \left( \frac{L_e}{W_e} \right) + \frac{P}{\pi K_{lnP} L_e} \quad \text{need to reduce junction areas } 4:1 \]
\[ \text{reduce widths } 2:1 \& \text{reduce length } 2:1 \rightarrow \text{doubles } \Delta T \]
\[ \text{reducing widths } 4:1, \text{ keep constant length } \rightarrow \text{small } \Delta T \text{ increase } \checkmark \]

\[ R_{bb} \approx \frac{\rho_s W_e}{12L_e} + \frac{\rho_s W_{bc}}{6L_e} + \frac{\rho_c}{A_{contacts}} \quad \rightarrow \text{reduce base contact resistivity } 4:1 \]
\[ \text{reduce widths } 2:1 \& \text{reduce length } 2:1 \rightarrow \text{constant } R_{bb} \checkmark \]
\[ \text{reducing widths } 4:1, \text{ keep constant length } \rightarrow \text{reduced } R_{bb} \checkmark \checkmark \]

Linewidths scale as the inverse square of bandwidth because thermal constraints dominate.
Bipolar Transistor Scaling Laws

Changes required to double transistor bandwidth:

<table>
<thead>
<tr>
<th>parameter</th>
<th>change</th>
</tr>
</thead>
<tbody>
<tr>
<td>collector depletion layer thickness</td>
<td>decrease 2:1</td>
</tr>
<tr>
<td>base thickness</td>
<td>decrease 1.414:1</td>
</tr>
<tr>
<td>emitter junction width</td>
<td>decrease 4:1</td>
</tr>
<tr>
<td>collector junction width</td>
<td>decrease 4:1</td>
</tr>
<tr>
<td>emitter contact resistance</td>
<td>decrease 4:1</td>
</tr>
<tr>
<td>current density</td>
<td>increase 4:1</td>
</tr>
<tr>
<td>base contact resistivity</td>
<td>decrease 4:1</td>
</tr>
</tbody>
</table>

Linewidths scale as the inverse square of bandwidth because thermal constraints dominate.
Status of Bipolar Transistors: September 2007

Popular metrics:
- $f_t$ or $f_{max}$ alone
- $\frac{(f_t + f_{max})}{2}$
- $\frac{1}{f_t} + \frac{1}{f_{max}}$^{-1}

Much better metrics:
- Power amplifiers:
  - PAE, associated gain, mW/μm
- Low noise amplifiers:
  - $F_{min}$, associated gain, digital:
  - $f_{clock}$, hence
  - $\frac{C_{cb} \Delta V}{I_c}$
  - $\frac{R_{ex} I_c}{\Delta V}$
  - $\frac{R_{bb} I_c}{\Delta V}$
  - $(\tau_b + \tau_c)$
256 nm Generation InP DHBT

150 nm thick collector

- $H_{21}$
- $f_{\text{max}} = 780$ GHz
- $f = 424$ GHz

70 nm thick collector

- $H_{21}$
- $f_{\text{max}} = 560$ GHz
- $f = 560$ GHz

10 nm thick collector

- $H_{21}$
- $f_{\text{max}} = 218$ GHz
- $f = 660$ GHz

4.7 dB Gain at 306 GHz.

200 GHz master-slave latch design

Z. Griffith, E. Lind, J. Hacker, M. Jones
### InP Bipolar Transistor Scaling Roadmap

<table>
<thead>
<tr>
<th></th>
<th>industry</th>
<th>university</th>
<th>university</th>
<th>appears</th>
<th>maybe</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2007-8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>emitter</td>
<td>512</td>
<td>256</td>
<td>128</td>
<td>64</td>
<td>32 nm width</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>8</td>
<td>4</td>
<td>2</td>
<td>1 Ω⋅μm² access ρ</td>
</tr>
<tr>
<td>base</td>
<td>300</td>
<td>175</td>
<td>120</td>
<td>60</td>
<td>30 nm contact width, 1.25 Ω⋅μm² contact ρ</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>10</td>
<td>5</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>collector</td>
<td>150</td>
<td>106</td>
<td>75</td>
<td>53</td>
<td>37.5 nm thick, 72 mAμm² current density</td>
</tr>
<tr>
<td></td>
<td>4.5</td>
<td>9</td>
<td>18</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.9</td>
<td>4</td>
<td>3.3</td>
<td>2.75</td>
<td>2-2.5 V, breakdown</td>
</tr>
<tr>
<td>f_τ</td>
<td>370</td>
<td>520</td>
<td>730</td>
<td>1000</td>
<td>1400 GHz</td>
</tr>
<tr>
<td>f_max</td>
<td>490</td>
<td>850</td>
<td>1300</td>
<td>2000</td>
<td>2800 GHz</td>
</tr>
<tr>
<td>power amplifiers</td>
<td>245</td>
<td>430</td>
<td>660</td>
<td>1000</td>
<td>1400 GHz</td>
</tr>
<tr>
<td>digital 2:1 divider</td>
<td>150</td>
<td>240</td>
<td>330</td>
<td>480</td>
<td>660 GHz</td>
</tr>
</tbody>
</table>
How Can Material Scientists Help?

To build a 5-THz bipolar Transistor...
...we need $0.25 \, \Omega \mu m^2$ Ohmic contacts,
& these must be stable at $300 \, mA/\mu m^2$.

...Can you help?
Ohmic Contacts
**Ex-Situ** Ohmic Contacts are a Mess

**textbook contact**

<table>
<thead>
<tr>
<th>metal</th>
<th>semiconductor</th>
</tr>
</thead>
</table>

**with surface oxide**

<table>
<thead>
<tr>
<th>metal</th>
<th>oxides, etc</th>
<th>semiconductor</th>
</tr>
</thead>
</table>

**with metal diffusion**

<table>
<thead>
<tr>
<th>metal</th>
<th>semiconductor</th>
</tr>
</thead>
</table>

Surface contaminated by semiconductor oxides
On InGaAs surface: Indium and Gallium Oxides, elemental As

Metals Interdiffuse with Semiconductor
- TiPtAu contacts: Ti diffusion
- Pt contacts: reaction
- Pd contacts: reaction

Interface is degraded → poor conductivity
Interface is badly-controlled → hard to understand → hard to improve
Our HBT Base Contacts Today Use Pd or Pt to Penetrate Oxides

TEM: Lysczek, Robinson, & Mohney, Penn State
Sample: Urteaga, RSC

Wafer first cleaned in reducing
Pd & Pt react with III-V semiconductor
Penetrate surface oxide
Provide ~5 $\Omega \cdot \mu m^2$ resistivity
(InGaAs base, 8*10^{19}/cm^3)
reaction depth is a problem for HBT base

Improvements in HBT Emitter Access Resistance

125 nm generation requires 5 Ω - μm² emitter resistivities

65 nm generation requires 1-2 Ω - μm²

Recent Results

<table>
<thead>
<tr>
<th>Material</th>
<th>Technique</th>
<th>Resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>ErAs/Mb</td>
<td>MBE in-situ</td>
<td>1.5 Ω - μm²</td>
</tr>
<tr>
<td>Mo</td>
<td>MBE in-situ</td>
<td>0.6 Ω - μm²</td>
</tr>
<tr>
<td>TiPdAu</td>
<td>ex-situ</td>
<td>0.5 Ω - μm²</td>
</tr>
<tr>
<td>TiW</td>
<td>ex-situ</td>
<td>0.7 Ω - μm²</td>
</tr>
</tbody>
</table>

Degeneracy contributes 1 Ω - μm²

20 nm emitter-base depletion layer contributes 1 Ω - μm² resistance

\[
\frac{\partial E_{fn}(x)}{\partial x} = - \frac{J}{qn(x)}
\]

Fermi-Dirac
Boltzmann

Equivalent series resistance approximation
**In-situ ErAs-InGaAs Contacts**

Epitaxially formed, no surface defects, no Fermi level pinning (?)
In-situ, no surface oxides, coherent interface, continuous As sublattice
Thermodynamically stable
ErAs/InAs Fermi level should be above conduction band

Results nevertheless disappointing:
1.5 Ω - μm²
Low-Resistance Refractory Contacts to N-InGaAs

Results initially by luck: control samples for ErAs experiments

Mo contacts: deposition by MBE immediately after InGaAs growth

TiW contacts: sputter deposition after UV-Ozone & 14.8-normality ammonia soak

Both give ~ 1 $\Omega$-$\mu$m$^2$ resistivity

\[ \begin{align*}
\text{in-situ Mo contact} & \quad \text{ex-situ TiW contact}
\end{align*} \]
Chris Palmstrom suggests materials such as Fe$_3$Ga, CoGa, NiAl

It might be possible to grow these with low interfacial densities on InGaAs or InAs.

Key question: what resistivity would we expect for a zero-defect, zero-barrier metal-semiconductor interface?

If we introduce a small difference in Fermi Level between metal and semiconductor, what current do we compute from integration of $N(E) v(E) F(E) T(E)$?
Shape as Substitute for Low-Resistance Contacts: SiGe HBTs

- Wide emitter contact: low resistance
- Narrow emitter junction: scaling (low $R_{bb}/A_e$)
- Thick extrinsic base: low resistance
- Thin intrinsic base: low transit time
- Wide base contacts: low resistance
- Narrow collector junction: low capacitance

These are planar approximations to radial contacts:

$\rightarrow$ reduced access resistance

$R_{bulk} = \frac{2\rho_{bulk}}{\pi L} \ln \left( \frac{\sqrt{2} \cdot r}{W} \right)$

$R_{contact} = \frac{2\rho_c}{\pi L r}$

should help less with small devices:

...widths scale faster than thicknesses $\rightarrow$ trench fringing capacitance
dielectric trench conducts heat badly
Field-Effect Transistors
Simple FET Scaling

Goal double transistor bandwidth when used in any circuit

→ reduce 2:1 all capacitances and all transport delays
→ keep constant all resistances, voltages, currents

All lengths, widths, thicknesses reduced 2:1

S/D contact resistivity reduced 4:1

\[
C_{gs} \sim \varepsilon W_g L_g / T_{ox} \\
C_{gs,f} \sim C_{gd} \sim \varepsilon W_g \\
C_{sb} \sim C_{db} \sim \varepsilon W_g L_c / T_{sub}
\]

If \( T_{ox} \) cannot scale with gate length,
\( C_{\text{parasitic}}/C_{gs} \) increases,
\( g_m/W_g \) does not increase

hence \( C_{\text{parasitic}}/g_m \) does not scale

\[
\tau \sim L_g / v \\
g_m \sim C_{gs} / \tau \sim (\varepsilon L_g W_g / T_{ox}) / \tau \\
G_{ds} \sim C_{d-ch} / \tau \sim \varepsilon W_g / \tau
\]

If \( T_{ox} \) cannot scale with gate length,
\( G_{ds}/g_m \) increases
## Well-Known: Si FETs no longer Scale Well

### EOT is not scaling as $1/L_g$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>2.2</th>
<th>2.1</th>
<th>2.0</th>
<th>1.9</th>
<th>1.6</th>
<th>1.5</th>
<th>1.4</th>
<th>1.4</th>
<th>1.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{ox}$ (nm) [2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gate Length (nm) [2]</td>
<td>75</td>
<td>65</td>
<td>53</td>
<td>45</td>
<td>37</td>
<td>32</td>
<td>25</td>
<td>28</td>
<td>22</td>
</tr>
<tr>
<td>$g_m/g_d$ at $5 \cdot L_{min\text{-}digital}$ [3]</td>
<td>47</td>
<td>40</td>
<td>32</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>1/f-noise ($\mu V^2 \cdot \mu m^2/Hz$) [4]</td>
<td>190</td>
<td>180</td>
<td>160</td>
<td>140</td>
<td>100</td>
<td>90</td>
<td>80</td>
<td>80</td>
<td>70</td>
</tr>
<tr>
<td>$\sigma V_{th}$ matching (mV-um) [5]</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>$I_{ds}$ ($\mu A/\mu m$) [6]</td>
<td>19</td>
<td>15</td>
<td>13</td>
<td>11</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Peak $F_i$ (GHz) [7]</td>
<td>120</td>
<td>140</td>
<td>170</td>
<td>200</td>
<td>240</td>
<td>280</td>
<td>320</td>
<td>360</td>
<td>400</td>
</tr>
<tr>
<td>Peak $F_{max}$ (GHz) [8]</td>
<td>200</td>
<td>220</td>
<td>270</td>
<td>310</td>
<td>370</td>
<td>420</td>
<td>480</td>
<td>530</td>
<td>590</td>
</tr>
</tbody>
</table>

(ITRS roadmap copied from Larry Larson's files)

### High-K gate dielectrics: often significant SiO2 interlayer, can limit EOT scaling

### S/D access resistance also a challenge: about $1 \Omega \cdot \mu m^2$ required for 20 nm

### Because gate equivalent thickness is not scaling, present devices scale badly

(output conductance is degrading with scaling)

### other capacitances are not scaling in proportion to $C_{gs}$

(hence are starting to dominate high frequency performance)
How Can Materials Scientists Help ?

High K-dielectrics for Si CMOS are still extremely important

Self-aligned (Salicide-like) contacts of very low resistivity are needed

...for 2 mA/micron operation at 700 mV gate overdrive, we want ~300 Ohm-micron lateral access resistivity
→ about 0.7 Ohm-micron^2 resistivity in a 25 nm wide contact
Why consider III-V (InGaAs/InP) CMOS?

**Low access resistance:**  
1 Ω-μm², 10 Ω-μm

Light electron → high electron velocity (thermal or Fermi injection)  
→ increased $I_d/W_g$ at a given oxide thickness (?)  
→ decreased $C_{gs}/g_m$ at a given gate length

**Challenge:**

**Low density of states**

$$C_{dos} = \frac{q^2 m^*}{\pi \hbar^2}$$

<table>
<thead>
<tr>
<th>$C_{ox}$</th>
<th>$C_{dos}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.4 μF/cm²  
@ 1 nm EOT  
limits $n_s$ to ~ 6*10¹²/cm²  
limits $I_d/W_g$  
limits $g_m/W_g$

**Challenge:**

filling of low-mobility satellite valleys

limits $n_s$ to ~ 8*10¹²/cm²  
limits $I_d/W_g$  

**Challenge:**

light electron limits vertical scaling  
~1.5-2.5 nm minimum  
mean electron depth
III-V MOS: What might be accomplished

Drive current simulation - ideal (ballistic) assumptions

22 nm gate length, 5 nm thick InGaAs / InP channel

under similar assumptions, silicon channels show 3-4 mA/μm

intrinsic $C_{gs} \approx 350$ fF/mm — comparable to fringing and stray capacitances
S/D Contact Process Flow For III-V MOSFETs

**selective-area S/D regrowth**

1. (starting material)
2. CBE in-situ etch
3. CBE in-situ selective-area regrowth
4. Electroplate S/D metal

**non-selective-area S/D regrowth**

1. (starting material)
2. Recess etch nonselective regrowth
3. Planarize
4. Etch
5. Strip planarization material
III-V MOSFETs Can Provide Very Low S/D Access Resistance

Objective: \( I_d/W_g \approx 5 \text{mA/\text{\mu m}} @ (V_{gs} - V_{th}) = 0.7 \text{V} \)
\( \Rightarrow \) transconductance \( g_m/W_g > 7 \text{mS/\mu m} \)
\( \Rightarrow 14 \Omega - \mu \text{m source resistance will reduce } g_m \text{ and } I_d \text{ by 10\%}. \)

With 50 nm wide contacts, this requires \( \rho_c < 0.7 \Omega - \mu \text{m}^2 \)

Modern III-V HEMTs have ~ 10:1 larger (~ 100 \Omega - \mu \text{m}) source resistance... because of the poor extrinsic source access region.

Extremely High \( g_m > 2.2 \text{S/mm} \) and \( f_t > 550 \text{GHz in 30-nm Enhancement-Mode InP-HEMTs with Pt/Mo/Ti/Pt/Au Buried Gate} \)

Kotakoz Shihoda, 'Neha H', Mark J.W. Redman, and Berndt Erfr
Improving FETs by Developing Other Materials

Other materials may offer high mobilities but...

\[ I_D \approx c_{ox} W_g v_{injection} (V_{gs} - V_{th} - \Delta V) \text{ for } (V_{gs} - V_{th}) / \Delta V >> 1 \]

where \( \Delta V = \frac{v_{injection} L_g}{\mu} \)

→ mobilities above ~ 1000 cm²/V-s of little benefit at 22 nm Lg

increased injection velocities are of value...

...but not at sacrifice in density of states
Nanopillar and Nanowire Devices

Nanopillar devices might have improved 2-D electrostatics... but only if wire diameter is ~10 nm or less

Access resistances are serious issue

Capacitances to source-drain pad regions a serious concern

III-V Nanowires FETs still must address defect density
dielectric-semiconductor interface

III-V nanopillar devices experience same DOS, confinement challenges as planar III-V devices
Conclusion
THz & nm Transistor Electronics is all about the interfaces

Bipolar Transistors:
- P and N ohmic contacts with very low resistivity
- Stability at high current density

FETs
- Gate dielectrics
- Contact resistance
- Density of states