Scaling of RF CMOS

S. Simon Wong
Electrical Engineering Department, Stanford University
802.11 Evolution

Prism Chip Set
1998

Single Chip
180nm CMOS


CMOS Scaling

- For Digital Applications
  Motivations – Faster; Lower cost/Hz-bit-function

- For RF Applications
  Spectra (1-6GHz) do not change; Are faster devices needed?
  Motivations – Smaller; Lower cost/function; Lower Power
Outline

- Transistor Scaling
- Passive Components
- Future Integration Trends
- Conclusions
Sample RF Circuits

Low Noise Amplifier

Power Amplifier

Extensive usage of inductor to tune out capacitance
MOSFET Frequency Response

\[ f_t = \frac{1}{2\pi} \frac{g_m}{C_{gg} + C_{gso} + C_{gdo} + C_{par}} \rightarrow \frac{1}{2\pi} \frac{v_{sat}}{L_g} \]

\[ f_{\text{max}} \approx \frac{f_t}{2\sqrt{(R_g + R_i)(g_{ds} + 2\pi f_t C_{gdo})}} \]

- \( g_m \) = transconductance
- \( C_{gg} \) = gate-channel capacitance
- \( C_{gdo} \) = gate-drain overlap capacitance
- \( C_{gso} \) = gate-source overlap capacitance
- \( C_{par} \) = gate parasitic capacitance
- \( V_{sat} \) = carrier saturation velocity
- \( L_g \) = effective gate length
- \( R_g \) = gate resistance
- \( R_i \) = equivalent input resistance for non-quasi static effect
- \( g_{ds} \) = output conductance
Theoretical Limit
$L_g = 30 \text{ nm, } f_t = 500 \text{ GHz}$
MOSFET $f_T$

$f_T \gg$ typical RF spectra

Effect of Layout on $f_T$ and $f_{\text{max}}$

Parallel $R_g$ improves $f_{\text{max}}$
Gate connected at both ends
Source drain metals do not overlap
Bulk contacts surround device
Optimum Finger Width

\( f_T \) scales with \( 1/L \), independent of \( T_{ox} \).

\( f_T \) depends weakly on finger width.


Optimum Finger Width

\[ f_{\text{max}} \text{ depends on finger width.} \]


Dependence of $f_T$ on Bias Current

$f_T$ depends on $I_{DS}$.
$I_{DS}$ can be reduced for similar $f_T$ with scaled technology.

MOSFET Noise Model

Drain Thermal Noise

Drain current noise is dominated by thermal noise at high frequency. \( \gamma \) models the excessive noise observed in short channel devices. The excessive noise is believed to be due to substrate noise.

\[
S_{I_D} = 4k_B T g_{do} \gamma
\]

\[
e_n^2 = \frac{i_n^2}{g_{do}^2}
\]
Induced Gate Noise

Gate noise induced by drain thermal noise is only a portion of the total noise. \( \beta \) models the excessive noise observed in short L devices.
Noise Figure

• Noises due to parasitic resistances associated with drain, source, bulk and especially gate need to be accounted for.

• Other noise sources, such as short noise due to gate leakage and avalanche noise at high $V_{DS}$, are usually insignificant.

• Noise figure describes the noise performance of a circuit.

• Noise figure depends on the source impedance matching. When $Z_s$ is low, $v_n$ dominates. When $Z_s$ is high, $i_n$ dominates.
For typical common source type LNA, \( \text{NF} \sim 1 + \frac{K f}{f_T} \), NF improves with scaled technology.

In practice, it is very difficult to achieve the minimum NF because of other constraints.
NF depends on $f_T$, and hence $I_{DS}$. $I_{DS}$ can be reduced for similar NF with scaled technology.

Drain 1/f Noise
Drain 1/f Noise

Drain current noise is dominated by 1/f noise at low f.

Although not critical for LNA, very important for VCO, A/D

\[ S_{id} = K \frac{g_m^2}{W L C_{ox}^2 f} \]
Degradation of 1/f Noise

Characteristics of Scaled MOSFET

<table>
<thead>
<tr>
<th>Process</th>
<th>0.25 μm ('98)</th>
<th>0.18 μm ('00)</th>
<th>0.13 μm ('02)</th>
<th>90 nm ('04)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_{dd} ) (V)</td>
<td>2.5 (1x)</td>
<td>1.8 (0.7x)</td>
<td>1.2 (0.5x)</td>
<td>1.0 (0.4x)</td>
</tr>
<tr>
<td>( I_{dsat} ) (μA/μm)</td>
<td>600 (1x)</td>
<td>600 (1x)</td>
<td>550 (1x)</td>
<td>850 (1.4x)</td>
</tr>
<tr>
<td>( I_{off} ) (nA/μm)</td>
<td>0.01 (1x)</td>
<td>0.02 (2x)</td>
<td>0.32 (32x)</td>
<td>7 (700x)</td>
</tr>
<tr>
<td>( I_{gate} ) (nA/μm)</td>
<td>2.5e-5 (1x)</td>
<td>1.8e-3 (100x)</td>
<td>0.65 (5e4)</td>
<td>6.3 (70000x)</td>
</tr>
<tr>
<td>( I_{on}/I_{off} ) (10^6)</td>
<td>60 (1x)</td>
<td>30 (0.5x)</td>
<td>1.7 (0.03x)</td>
<td>0.12 (0.002x)</td>
</tr>
<tr>
<td>( g_m ) (mS/μm)</td>
<td>0.3 (1x)</td>
<td>0.4 (1.3x)</td>
<td>0.6 (2x)</td>
<td>1.0 (3.3x)</td>
</tr>
<tr>
<td>( g_{ds} ) (μS/μm)</td>
<td>7.7 (1x)</td>
<td>15 (2x)</td>
<td>42 (5.4x)</td>
<td>100 (13x)</td>
</tr>
<tr>
<td>( g_m / g_{ds} )</td>
<td>39 (1x)</td>
<td>27 (0.7x)</td>
<td>14 (0.36x)</td>
<td>10 (0.26x)</td>
</tr>
<tr>
<td>( f_T ) (GHz)</td>
<td>30 (1x)</td>
<td>60 (2x)</td>
<td>80 (2.7x)</td>
<td>140 (4.7x)</td>
</tr>
<tr>
<td>Delay (ps/gate)</td>
<td>45 (1x)</td>
<td>30 (0.7x)</td>
<td>15 (0.3x)</td>
<td>11 (0.24x)</td>
</tr>
<tr>
<td>( C_g ) (fF/gate)</td>
<td>0.47 (1x)</td>
<td>0.35 (0.7x)</td>
<td>0.25 (0.5x)</td>
<td>0.16 (0.34x)</td>
</tr>
<tr>
<td>( C_j ) (fF/gate)</td>
<td>0.83 (1x)</td>
<td>0.80 (1x)</td>
<td>0.88 (1.1x)</td>
<td>0.66 (0.8x)</td>
</tr>
</tbody>
</table>
Linearity of MOSFET

Effect of Non-Linearity
High voltage required for the cascode transistor.
Power Combining

Typical RF Transceiver
MOSFET-Based T/R Switch

Solution:
Bias the gate through a resistor – allows the gate to be bootstrapped with incoming signal – improve linearity
Bias the substrate through a L-C tank – allows the substrate to be bootstrapped with incoming signal,
Linearity is greatly improved.

Switch Linearity

![Graph showing the relationship between \( P_{\text{out}} \) (dBm) and \( P_{\text{in}} \) (dBm) for Tx and Rx modes, with markers for \( P_{1\text{dB}} \) in each mode.]
Outline

• Transistor Scaling
• Passive Components
• Future Integration Trends
• Conclusions
Capacitor Options

Inter-level Metal-Oxide-Metal (MOM)

Intra-level Metal-Oxide-Metal (MOM)

High K Metal-Insulator-Metal (MIM)

<table>
<thead>
<tr>
<th></th>
<th>Metal Thickness (μm)</th>
<th>Metal Space (μm)</th>
<th>Dielectric Thickness (μm)</th>
<th>Approx K</th>
</tr>
</thead>
<tbody>
<tr>
<td>250 nm</td>
<td>0.6</td>
<td>0.4</td>
<td>1</td>
<td>4.1</td>
</tr>
<tr>
<td>180 nm</td>
<td>0.5</td>
<td>0.3</td>
<td>0.8</td>
<td>3.8</td>
</tr>
<tr>
<td>130 nm</td>
<td>0.4</td>
<td>0.2</td>
<td>0.5</td>
<td>3.8</td>
</tr>
<tr>
<td>90 nm</td>
<td>0.3</td>
<td>0.15</td>
<td>0.3</td>
<td>3.2</td>
</tr>
</tbody>
</table>
Capacitor Performance

3D Capacitor

Varactor Options

MOS accumulation mode varactor offers higher C/Area, more tuning range, comparable Q, and less temp sensitivity.

On-Chip Spiral Inductor

Quality Factor

\[ Q = 2\pi \frac{|\text{Peak Magnetic Energy} - \text{Peak Electric Energy}|}{\text{Energy Loss in One Oscillation Cycle}} \]

\[
\frac{\omega L_s}{R_s} \times \frac{R_p}{R_p + [(\omega L_s/R_s)^2 + 1] \cdot R_s} \times \left(1 - \frac{R_s^2(C_p + C_s)}{L_s} - \omega^2 L_s(C_p + C_s)\right)
\]

Substrate Loss Factor \hspace{2cm} Self-Resonance Factor
Substrate Loss due to E-field Penetration

[Diagram showing the penetration of E-field into a substrate with labeled components such as $H$, $E$, $C_{si}$, $C_{ox}$, $R_{si}$, $L_s$, and $R_s$.]
Ground Shield to Stop E-field Penetration

- Lenz’s Law:
  Induced eddy currents with opposing magnetic field
Patterned Ground Shield

- Patterned Ground Shield
- Inserted between the inductor and substrate
- PGS fingers orthogonal to spiral for form a “star” shape
- No interference with the inductor magnetic field
- Use silicided polysilicon with Metal1 strap for low resistance path to ground.

# Comparison of Back End Technology

<table>
<thead>
<tr>
<th>Metal Layers</th>
<th>Top Metal Thickness (μm)</th>
<th>Top Metal to Sub (μm)</th>
<th>M5 to Sub (μm)</th>
<th>Approx K</th>
</tr>
</thead>
<tbody>
<tr>
<td>250 nm</td>
<td>5</td>
<td>1 Al</td>
<td>7.5</td>
<td>7.5</td>
</tr>
<tr>
<td>180 nm</td>
<td>6</td>
<td>1 Al</td>
<td>8</td>
<td>6.5</td>
</tr>
<tr>
<td>130 nm</td>
<td>8</td>
<td>0.9 Cu</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>90 nm</td>
<td>9</td>
<td>0.85 Cu</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>
Optimized Inductor Q

- With all metal layers
- With five metal layers

3nH, 2GHz, side < 500 µm

Technology
65nm offers more options on metal and dielectric, including multiple thick metal layers
Outline

- Transistor Scaling
- Passive Components
- Future Integration Trends
- Conclusions
Integration of Off-Chip Components

Significant reduction in system cost
Chip Comparison

180nm 802.11abg

130nm 802.11n 2X2
Chip Comparison

130nm GSM

90nm GSM
Chip Comparison

130nm WLAN 2X2 N

130nm GSM

130nm Bluetooth
Radio Scaling

[Graph showing the relationship between technology node and radio area for different types of WLAN systems.]

- **WLAN - single band**
- **WLAN - dual band**
- **WLAN - N per channel**

The graph indicates a preliminary not integrated scaling trend.
Radio Scaling

![Graph showing radio area vs technology node for Bluetooth, GPS, and GSM technologies, with a line indicating scaling.](image-url)
Conclusions

– Scaled CMOS should improve NF and reduce power consumption; reduction in linearity and voltage capability complicates the design
– Advanced technology offers more options for passive components; quality of passive components may be compromised if not all metal layers are used
– Further Integration of off-chip components to reduce system cost
– Technology selection will depend on analog/digital mix
– Chip with mostly RF functions will scale moderately with technology
– Chip with largely digital functions will scale with technology
– Advanced technology may benefit the integration of multiple radios