Scaling Impact on Analog Circuit Performance

Giovanni Anelli
CERN - European Organization for Nuclear Research
Experimental Physics Division
Microelectronics Group
CH-1211 Geneva 23 – Switzerland
Giovanni.Anelli@cern.ch
Outline

- Scaling down of CMOS technologies
- How scaling works for devices and interconnections
- Scaling impact on noise
- Scaling impact on matching
- Scaling impact on radiation tolerance
- Analog performance of submicron processes
- Substrate noise in mixed-mode integrated circuits
Outline

• Scaling down of CMOS technologies
  ➢ Moore’s law
  ➢ An example: Intel Microprocessors
  ➢ The International Technology Roadmap for Semiconductors: what will happen in the next 15 years?

• How scaling works for devices and interconnections
• Scaling impact on noise
• Scaling impact on matching
• Scaling impact on radiation tolerance
• Analog performance of submicron processes
• Substrate noise in mixed-mode integrated circuits
Moore’s law

1965: Number of Integrated Circuit components will double every year

1975: Number of Integrated Circuit components will double every 18 months

1996: The definition of “Moore’s Law” has come to refer to almost anything related to the semiconductor industry that
when plotted on semi-log paper approximates a straight line. I don’t want to do anything to restrict this
definition. - G. E. Moore, 8/7/1996

An example:
Intel’s Microprocessors

http://www.intel.com/
<table>
<thead>
<tr>
<th>Model</th>
<th>Release Date</th>
<th>Max. Clock Frequency</th>
<th>L1 Cache (Size, Type)</th>
<th>L2 Cache (Size, Type)</th>
<th>Process Size ( Micron)</th>
<th>Main Memory Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>4004</td>
<td>11 / 1971</td>
<td>108 KHz</td>
<td>2300</td>
<td>10 µm</td>
<td>108 MHz</td>
<td></td>
</tr>
<tr>
<td>8008</td>
<td>04 / 1972</td>
<td>200 KHz</td>
<td>3500</td>
<td>10 µm</td>
<td>6 MHz</td>
<td></td>
</tr>
<tr>
<td>8080</td>
<td>04 / 1974</td>
<td></td>
<td>4500</td>
<td>6 µm</td>
<td>8 MHz</td>
<td></td>
</tr>
<tr>
<td>8088</td>
<td>06 / 1979</td>
<td></td>
<td>29000</td>
<td>3 µm</td>
<td>12 MHz</td>
<td></td>
</tr>
<tr>
<td>8086</td>
<td>02 / 1982</td>
<td></td>
<td>134000</td>
<td>1.5 µm</td>
<td>2.5 GHz</td>
<td></td>
</tr>
<tr>
<td>Intel8086™</td>
<td>02 / 1985</td>
<td></td>
<td>275000</td>
<td>1 µm</td>
<td>16 MHz</td>
<td></td>
</tr>
<tr>
<td>Intel486™ DX</td>
<td>04 / 1989</td>
<td></td>
<td>1.2 M</td>
<td>1 µm</td>
<td>25 MHz</td>
<td></td>
</tr>
<tr>
<td>Pentium®</td>
<td>03 / 1993</td>
<td></td>
<td>3.1 M</td>
<td>0.8 µm</td>
<td>66 MHz</td>
<td></td>
</tr>
<tr>
<td>Pentium® Pro</td>
<td>11 / 1995</td>
<td></td>
<td>5.5 M</td>
<td>0.6 µm</td>
<td>150 MHz</td>
<td></td>
</tr>
<tr>
<td>Pentium® II</td>
<td>05 / 1997</td>
<td></td>
<td>7.5 M</td>
<td>0.35 µm</td>
<td>233 MHz</td>
<td></td>
</tr>
<tr>
<td>Pentium® III</td>
<td>02 / 1999</td>
<td></td>
<td>9.5 M</td>
<td>0.25 µm</td>
<td>500 MHz</td>
<td></td>
</tr>
<tr>
<td>Pentium® 4</td>
<td>11 / 2000</td>
<td></td>
<td>42 M</td>
<td>0.18 µm</td>
<td>1.5 GHz</td>
<td></td>
</tr>
</tbody>
</table>

http://www.intel.com/

Manila, December 2002
Giovanni Anelli, CERN
The correct size relationship.

http://www.intel.com/
The **Intel** 4004 processor

- Introduced: 15/11/1971
- Clock: 108 KHz
- 2300 Transistors
- 10 µm technology (NMOS)

http://www.intel.com/
The Pentium® 4

The first Pentium® 4
- Introduced: 20/11/2000
- Clock: 1.5 GHz
- 42 Million Transistors
- 0.18 μm technology

A more recent one
- Introduced: 26/08/2002
- Clock: 2.8 GHz
- 55 Million Transistors
- 0.13 μm technology
The National Technology Roadmap for Semiconductors (NTRS):
Sponsored by the Semiconductor Industry Association (SIA)

The International Technology Roadmap for Semiconductors (ITRS):
Sponsored by:
- Semiconductor Industry Association (SIA)
- European Electronics Component Manufacturers Association (EECA)
- Korea Semiconductor Industry Association (KSIA)
- Japan Electronics and Information Technology Industries Association (JEITA)
- Taiwan Semiconductor Industry Association (TSIA)

These documents always contained a 15-year outlook of the major trends of the semiconductor industry
Future perspectives

Outline

• Scaling down of CMOS technologies
• How scaling works for devices and interconnections
  ➢ Constant field scaling
  ➢ Generalized scaling
  ➢ Scaling of interconnections
• Scaling impact on noise
• Scaling impact on matching
• Scaling impact on radiation tolerance
• Analog performance of submicron processes
• Substrate noise in mixed-mode integrated circuits
Why scaling??

Scaling improves density, speed and power consumption of digital circuits

Example: CMOS inverter

\[ P_{\text{static}} = I_{\text{leakage}} \cdot V_{DD} \]

\[ P_{\text{dynamic}} = C_L \cdot V_{DD}^2 \cdot f \]

\[ P_{\text{DPP}} = C_L \cdot V_{DD}^2 \]

Power-delay product

- \( t_{\text{ox}} \)
- \( V_{DD} \)
- \( C_L \)
The aim of scaling is to reduce the device dimensions (to improve the circuit performance) without introducing effects which could disturb the good operation of the device.

\[ x_d = \sqrt{\frac{2\epsilon_{\text{Si}}}{qN_A}} \sqrt{\phi_{\text{bi}}} + V \]

## Constant field scaling

### Summary of the scaling factors for several quantities

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Scaling Factor</th>
<th>Quantity</th>
<th>Scaling factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device dimensions ($L, W, t_{ox}, x_D$)</td>
<td>$1/\alpha$</td>
<td>Capacitances</td>
<td>$1/\alpha$</td>
</tr>
<tr>
<td>Area</td>
<td>$1/\alpha^2$</td>
<td>Capacitances per unit area</td>
<td>$\alpha$</td>
</tr>
<tr>
<td>Devices per unit of chip area (density)</td>
<td>$\alpha^2$</td>
<td>Charges</td>
<td>$1/\alpha^2$</td>
</tr>
<tr>
<td>Doping concentration ($N_A$)</td>
<td>$\alpha$</td>
<td>Charges per unit area</td>
<td>$1$</td>
</tr>
<tr>
<td>Bias voltages and $V_T$</td>
<td>$1/\alpha$</td>
<td>Electric field intensity</td>
<td>$1$</td>
</tr>
<tr>
<td>Bias currents</td>
<td>$1/\alpha$</td>
<td>Body effect coefficient ($\gamma$)</td>
<td>$1/\sqrt{\alpha}$</td>
</tr>
<tr>
<td>Power dissipation for a given circuit</td>
<td>$1/\alpha^2$</td>
<td>Transistor transit time ($\tau$)</td>
<td>$1/\alpha$</td>
</tr>
<tr>
<td>Power dissipation per unit of chip area</td>
<td>$1$</td>
<td>Transistor power-delay product</td>
<td>$1/\alpha^3$</td>
</tr>
</tbody>
</table>
Subthreshold slope and width of the moderate inversion region do not scale. This can have a devastating impact on the static power consumption of a digital circuit.

\[ \log I_D \]

\[ V_{GS} \]

\[ 0 \text{ V} \]

nA

pA
Generalized scaling

- The dimensions in the device scale as in the constant field scaling
- $V_{dd}$ scales to have reasonable electric fields in the device, but slower than $t_{ox}$, to have a useful voltage swing for the signals
- The doping levels are adjusted to have the correct depletion region widths
- To limit the subthreshold currents, $V_T$ scales more slowly than $V_{dd}$

An accurate scaling of the interconnections is needed as well, so that we can profit at the circuit level of the improvements made at the device level. Interconnections are becoming more and more important in modern technologies because the delay they introduce is becoming comparable with the switching time of the digital circuits.

"Reverse" scaling

The scaling method is different from the one applied to devices

If $W$, $L$, $t_m$ and $t_{ox}$ are scaled by $1/\alpha$

- Current density scales by $\alpha$
- $R$ scales by $\alpha$, $C$ scales by $1/\alpha$
- RC (delay) does not scale!!!

In practice, wires dimensions are reduced only for local interconnections (but not $t_m$). At the chip scale, $t_m$ and $t_{ox}$ are increased (reverse scaling).


Manila, December 2002

Giovanni Anelli, CERN
Hierarchical scaling

Figure 36 Cross-section of Hierarchical Scaling

### Generalized Selective Scaling

<table>
<thead>
<tr>
<th>Physical parameter</th>
<th>Constant-Scaling Factor</th>
<th>Generalized Scaling Factor</th>
<th>Generalized Selective Scaling Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channel length, Insulator thickness</td>
<td>$1/\alpha$</td>
<td>$1/\alpha$</td>
<td>$1/\alpha_d$</td>
</tr>
<tr>
<td>Wiring width, channel width</td>
<td>$1/\alpha$</td>
<td>$1/\alpha$</td>
<td>$1/\alpha_w$</td>
</tr>
<tr>
<td>Electric field in device</td>
<td>$1$</td>
<td>$\varepsilon$</td>
<td>$\varepsilon$</td>
</tr>
<tr>
<td>Voltage</td>
<td>$1/\alpha$</td>
<td>$\varepsilon/\alpha$</td>
<td>$\varepsilon/\alpha_d$</td>
</tr>
<tr>
<td>On-current per device</td>
<td>$1/\alpha$</td>
<td>$\varepsilon/\alpha$</td>
<td>$\varepsilon/\alpha_w$</td>
</tr>
<tr>
<td>Doping</td>
<td>$\alpha$</td>
<td>$\varepsilon\alpha$</td>
<td>$\varepsilon\alpha_d$</td>
</tr>
<tr>
<td>Area</td>
<td>$1/\alpha^2$</td>
<td>$1/\alpha^2$</td>
<td>$1/\alpha_d^2$</td>
</tr>
<tr>
<td>Capacitance</td>
<td>$1/\alpha$</td>
<td>$1/\alpha$</td>
<td>$1/\alpha_w$</td>
</tr>
<tr>
<td>Gate delay</td>
<td>$1/\alpha$</td>
<td>$1/\alpha$</td>
<td>$1/\alpha_d$</td>
</tr>
<tr>
<td>Power dissipation</td>
<td>$1/\alpha^2$</td>
<td>$\varepsilon^2/\alpha^2$</td>
<td>$\varepsilon^2/\alpha_w\alpha_d$</td>
</tr>
<tr>
<td>Power density</td>
<td>$1$</td>
<td>$\varepsilon^2$</td>
<td>$\varepsilon^2\alpha_w/\alpha_d$</td>
</tr>
</tbody>
</table>

$\alpha$ is the dimensional scaling parameter, $\varepsilon$ is the electric field scaling parameter, and $\alpha_d$ and $\alpha_w$ are separate dimensional scaling parameters for the selective scaling case. $\alpha_d$ is applied to the device vertical dimensions and gate length, while $\alpha_w$ applies to the device width and the wiring.

---

Outline

• Scaling down of CMOS technologies
• How scaling works for devices and interconnections
  • Scaling impact on noise
  • Scaling impact on matching
  • Scaling impact on radiation tolerance
• Analog performance of submicron processes
• Substrate noise in mixed-mode integrated circuits
Scaling impact on noise

\[
\frac{V_{in}^2}{\Delta f} = 4kTn_\gamma \frac{1}{g_m} + \frac{K_a}{C_{ox}^2 WL f^\alpha} \quad g_m = \sqrt{2^{-n+1}} \frac{C_{ox}}{W/L} I_{DS} \quad C_{ox} = \frac{\varepsilon_{SiO_2}}{t_{ox}}
\]

White noise: keeping the same W/L ratio and the same current, we have an improvement in the noise since \( C_{ox} \) (and therefore \( g_m \)) increases with scaling.

1/f noise: we suppose that the constant \( K_a \) does not change with scaling. In this case, we have an improvement in the noise if we keep the same device area (WL), or we have the same noise if we scale both W and L. Data taken from the Roadmap foresee that \( K_a \) will remain more or less constant even for the most advanced CMOS processes. This must, of course, be verified…
Scaling impact on noise

For the same device dimensions and current, both the channel thermal noise and the flicker (1/f) noise should decrease

BUT

there can be other effects in submicron MOSFETs that tend to increase the noise, such as, for example, carriers heating and parasitic resistances.

The constant $K_a$ is HIGHLY technology dependent. It might be difficult to keep it under control in new advanced processes. Moreover, the effect on 1/f noise on new dielectric materials is not yet known.

Another (possibly serious) source of problems in the future will be the leakage current through the gate oxide. Thinner gate oxides will have a much higher leakage current, which will have a higher shot noise.
Gate leakage current shot noise

\[ \sqrt{\frac{i_n^2}{\Delta f}} = 2ql \]

In the hypothesis:
- current density = 1 A/cm²
- W = 1000 μm
- L = 0.1 μm

We have
- I = 1 μA
- 2ql = 0.56 pA/√Hz

which are NOT negligible values!


Manila, December 2002

Giovanni Anelli, CERN
1/f Noise parameter $K_a$

0.25 $\mu$m CMOS technology

$K_a \times 10^{-27}$ C$^2$/m$^2$

Gate Length [ $\mu$m ]

- NMOS moderate inversion
- PMOS moderate inversion
- NMOS strong inversion
- PMOS strong inversion
Excess noise factor $\Gamma$

**0.25 $\mu$m CMOS technology**

- **NMOS weak inversion**
- **PMOS weak inversion**
- **NMOS moderate inversion**
- **PMOS moderate inversion**
- **NMOS strong inversion**
- **PMOS strong inversion**

**Gate Length [ $\mu$m ]**

Manila, December 2002

Giovanni Anelli, CERN
Outline

- Scaling down of CMOS technologies
- How scaling works for devices and interconnections
- Scaling impact on noise
- Scaling impact on matching
- Scaling impact on radiation tolerance
- Analog performance of submicron processes
- Substrate noise in mixed-mode integrated circuits
The ion implantation process follows Poisson statistics. Therefore, the uncertainty in the number of dopant implanted is given by the square root of the number.

The error becomes proportionally more important for smaller devices! ($=1/\sqrt{N}$)
Scaling & dopant fluctuations

\[ \sigma_{\Delta V_{th}} = C \cdot \frac{t_{ox} \cdot 4\sqrt{N}}{\sqrt{W L}} \]

- For the same device dimensions, matching improves
- For minimum size devices, matching might be worse

<table>
<thead>
<tr>
<th>( L_{\text{min}} [\mu\text{m}] )</th>
<th>( t_{\text{ox}} [\text{nm}] )</th>
<th>( N_a [\text{cm}^{-3}] )</th>
<th>( \frac{A_N}{t_{\text{ox}}} [\text{mV} \cdot \mu\text{m} / \text{nm}] )</th>
<th>( A_N [\text{mV} \cdot \mu\text{m}] )</th>
<th>( \sigma_{\Delta V_{th}} [\text{mV}] )</th>
<th>( 6 \cdot \sigma_{V_{th}} [\text{mV}] )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2</td>
<td>25</td>
<td>5.10^{16}</td>
<td>0.328</td>
<td>8.2</td>
<td>6.84</td>
<td>29</td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>6.10^{16}</td>
<td>0.344</td>
<td>6.9</td>
<td>6.89</td>
<td>29.2</td>
</tr>
<tr>
<td>0.8</td>
<td>15</td>
<td>7.5.10^{16}</td>
<td>0.365</td>
<td>5.5</td>
<td>6.84</td>
<td>29</td>
</tr>
<tr>
<td>0.5</td>
<td>10</td>
<td>1.2.10^{17}</td>
<td>0.414</td>
<td>4.1</td>
<td>8.28</td>
<td>35.1</td>
</tr>
<tr>
<td>0.25</td>
<td>5.5</td>
<td>2.4.10^{17}</td>
<td>0.498</td>
<td>2.7</td>
<td>11</td>
<td>46.5</td>
</tr>
<tr>
<td>0.18</td>
<td>4</td>
<td>3.3.10^{17}</td>
<td>0.542</td>
<td>2.2</td>
<td>12</td>
<td>51.1</td>
</tr>
</tbody>
</table>


Manila, December 2002

Giovanni Anelli, CERN

29
Matching will have a very important impact on the performance of deep submicron CMOS circuits.
Matching data from the Roadmap

Data taken from The International Technology Roadmap for Semiconductors (2001 Edition)
Outline

- Scaling down of CMOS technologies
- How scaling works for devices and interconnections
- Scaling impact on noise
- Scaling impact on matching
- Scaling impact on radiation tolerance
  - Radiation effects in MOS transistors
  - Scaling impact on the radiation tolerance
  - Enclosed Layout Transistor
- Analog performance of submicron processes
- Substrate noise in mixed-mode integrated circuits
Ionizing particles through a MOST

- Threshold voltage shift
- Mobility degradation
- Swing degradation
- Threshold voltage shift

Other degradations:
- Transconductance
- Noise
- Matching


Manila, December 2002
Giovanni Anelli, CERN
Transistor level leakage (NMOS)

Transistor level leakage

“CENTRAL” MOS TRANSISTOR

This is for LOCOS, what about STI? Things do not improve!

Transistor level leakage: example

NMOS - 0.7 μm technology - \( t_{ox} = 17 \) nm

Threshold voltage shift

Prerad

After 1 Mrad

Transistor level leakage: example

NMOS - 0.7 \( \mu \)m technology - \( t_{ox} = 17 \) nm

Threshold voltage shift

Prerad

After 1 Mrad
Total Dose damage and scaling

\[ \frac{-\Delta V_{FB}}{1 \text{Mrd} (\text{Si})} \propto t_{ox}^2 \]

\[ -\Delta D_{it} / \text{Mrad} \]

Decreasing $t_{ox}$ we decrease the degradation of:

- Transconductance
- Subthreshold slope
- Noise

And the threshold voltage shift for n-channel transistors might not be negative anymore...
$\Delta V_T$ and $t_{ox}$ scaling

![Graph showing $\Delta V_{th}/\text{Mrad}(\text{SiO}_2)$ vs. $t_{ox}$ with various markers for different values of $t_{ox}$.](image)
1/f Noise and irradiation

W/L = 2000/0.78, I_{DS} = 0.5 mA

Total Dose

Annealing

0 Mrad 30 Mrad 60 Mrad 100 Mrad

K_a [x10^{-27} C^2/m^2]
White noise and irradiation

N-channel, ELT, $W = 2$ mm, $I_{DS} = 0.5$ mA, $V_{DS} = 0.8$ V, $V_{BS} = 0$ V

![Graph showing white noise and irradiation effects on a N-channel transistor.](image)
White noise and irradiation

P-channel, $W = 2 \text{ mm}$, $I_{DS} = 0.5 \text{ mA}$, $V_{DS} = 0.8 \text{ V}$, $V_{BS} = 0 \text{ V}$

![Graph showing noise vs. frequency for different conditions: Prerad, After 100 Mrad, and After Annealing.](image-url)
Enclosed Layout Transistor (ELT)

ELTs solve the leakage problem in the NMOS transistors
At the circuit level, guard rings are necessary
Effectiveness of ELTs

NMOS - 0.7 μm technology - $t_{ox} = 17$ nm

![Graph showing the effect of radiation on MOSFET characteristics.](image-url)
ELT & deep submicron

NMOS - 0.25 $\mu$m technology - $t_{ox} = 5$ nm

Prerad and after 13 Mrad

No leakage
No $V_T$ shift

Manila, December 2002
Giovanni Anelli, CERN
Outline

- Scaling down of CMOS technologies
- How scaling works for devices and interconnections
- Scaling impact on noise
- Scaling impact on matching
- Scaling impact on radiation tolerance
- Analog performance of submicron processes
  - Scaling impact on analog performance
  - Analog design in “digital” processes
  - Integrated capacitors
- Substrate noise in mixed-mode integrated circuits
Scaling impact on $\mu C_{\text{ox}}$

Due to the scaling of the gate oxide thickness, the gate capacitance $C_{\text{ox}}$ increases with scaling. This increases the driving capability of the transistor. For a given $W/L$ ratio and a fixed bias current, the transconductance also increases with scaling.

$$g_m = \sqrt{2 \frac{\beta}{n} I_{DS}} = \frac{\beta}{n} (V_{GS} - V_T)$$

$$\beta = \mu C_{\text{ox}} \frac{W}{L}$$

The following data are taken from the design manuals of different CMOS technologies:

<table>
<thead>
<tr>
<th>$L_{\text{min}}$ [$\mu$m]</th>
<th>$t_{\text{ox _physical}}$ [nm]</th>
<th>$t_{\text{ox _effective}}$ [nm]</th>
<th>$C_{\text{ox}}$ [fF/$\mu$m$^2$]</th>
<th>$\mu C_{\text{ox}}$ [$\mu$A/V$^2$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2</td>
<td>24</td>
<td>---</td>
<td>1.44</td>
<td>68</td>
</tr>
<tr>
<td>0.8</td>
<td>17</td>
<td>---</td>
<td>2.03</td>
<td>90</td>
</tr>
<tr>
<td>0.5</td>
<td>10</td>
<td>---</td>
<td>3.45</td>
<td>134</td>
</tr>
<tr>
<td>0.25</td>
<td>5.5</td>
<td>6.2</td>
<td>5.5</td>
<td>280</td>
</tr>
<tr>
<td>0.13</td>
<td>2.2</td>
<td>3.15</td>
<td>10.9</td>
<td>555</td>
</tr>
</tbody>
</table>

The quantity $g_m r_0$ is called intrinsic gain of the transistor. It represents the maximum gain obtainable from a single transistor, and it is a very useful figure of merit in analog design.
Scaling impact on the intrinsic gain

\[ g_m = \frac{\beta}{n} (V_{GS} - V_T) \]

\[ \beta = \mu C_{ox} \frac{W}{L} \]

\[ I_{DS_{SAT}} = \frac{\beta}{2n} (V_{GS} - V_T)^2 \]

\[ \lambda = \frac{1}{V_{DS}} \cdot \frac{\Delta L}{L - \Delta L} \approx \frac{1}{V_{DS}} \cdot \frac{\Delta L}{L} \]

\[ \Delta L \approx \sqrt{\frac{2\varepsilon_{Si}}{qN_a}} \sqrt{V_{DS}} \]

\[ g_{out} = \lambda \cdot I_{DS_{SAT}} \]

\[ r_0 = \frac{1}{g_{out}} = \frac{1}{\lambda \cdot I_{DS_{SAT}}} \]

Intrinsic Gain = \( g_m \cdot r_0 \)

Supposing to have constant field scaling for the technology, we obtain:

<table>
<thead>
<tr>
<th>L</th>
<th>W</th>
<th>( \beta )</th>
<th>( V_{GS} - V_T )</th>
<th>( g_m )</th>
<th>( V_{DS} )</th>
<th>( \Delta L )</th>
<th>( \lambda )</th>
<th>( I_{DS_{SAT}} )</th>
<th>( g_{out} )</th>
<th>( r_0 )</th>
<th>( g_m \cdot r_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/( \alpha )</td>
<td>1/( \alpha )</td>
<td>( \alpha )</td>
<td>1/( \alpha )</td>
<td>1</td>
<td>1/( \alpha )</td>
<td>1/( \alpha )</td>
<td>( \alpha )</td>
<td>1/( \alpha )</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1/( \alpha )</td>
<td>1</td>
<td>1/( \alpha )</td>
<td>1/( \alpha )</td>
<td>1/( \alpha )</td>
<td>1</td>
<td>1/( \alpha^2 )</td>
<td>1/( \alpha^2 )</td>
<td>( \alpha^2 )</td>
<td>( \alpha )</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>( \alpha )</td>
<td>1/( \alpha )</td>
<td>1</td>
<td>1/( \alpha )</td>
<td>1/( \alpha )</td>
<td>1/( \alpha )</td>
<td>1</td>
<td>( \alpha )</td>
<td>( \alpha )</td>
<td></td>
</tr>
<tr>
<td>1/( \alpha )</td>
<td>( \alpha )</td>
<td>( \alpha^3 )</td>
<td>1/( \alpha )</td>
<td>( \alpha^2 )</td>
<td>1/( \alpha )</td>
<td>( \alpha )</td>
<td>( \alpha )</td>
<td>( \alpha^2 )</td>
<td>1/( \alpha^2 )</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
Analog power consumption

\[ P_{\text{min}} = 8 \pi kT \cdot \text{SNR} \cdot f_{\text{sig}} \cdot \frac{V_{\text{DD}}}{V_{\text{DD}} - \Delta V} \]

\( \Delta V \) is the fraction of the \( V_{\text{DD}} \) not used for signal swing

Optimal analog power/performance trade-off for 0.35 - 0.25 \( \mu \text{m} \) technologies

---


Manila, December 2002

Giovanni Anelli, CERN
Weak inversion region

$t_{ox}$ scales for the same device dimensions the boundary between weak inversion and strong inversion moves towards higher currents

\[ g_m = \sqrt{\frac{2}{n} I_{DS}} \]

\[ g_m = \frac{I_{DS}}{n\phi_t} \]

\[ I_{DS_{\_W \_to \_S}} = 2\beta n\phi_t^2 \]
Scaling impact on analog circuits

$t_{ox}$ scales for the same device dimensions

- Threshold voltage matching improves
  \[ \sigma_{\Delta V_{th}} = \frac{\text{Const} \cdot t_{ox}}{\sqrt{W \cdot L}} \]

- $1/f$ noise decreases
  \[ \frac{v_{in^{-1/f}}^2}{\Delta f} = \frac{K_a}{C_{ox}^2 \cdot WL} \cdot \frac{1}{f^\alpha} \]

- Transconductance increases (same current)
  \[ g_m = \sqrt{2n \cdot \mu C_{ox} \cdot \frac{W}{L} \cdot I_{DS}} \]

- White noise decreases
Scaling impact on analog circuits

- New noise mechanisms
- Modeling difficulties
- Lack of devices for analog design
- Reduced signal swing (new architectures needed)
- Substrate noise in mixed-signal circuits
- Velocity saturation. Critical field: $3 \, \text{V/\mu m}$ for electrons,
  $10 \, \text{V/\mu m}$ for holes
  $$g_{m\_vel.sat.} = W C_{ox} \, v_{sat}$$
The integrated circuits market is driven by digital circuits, such as memories and microprocessors. This led to an increasing interest in integrating analog circuits together with digital functions in processes optimized for digital circuits, making what it is called a System on a Chip (SoC). This approach has several advantages and disadvantages.

**ADVANTAGES:**

- Lower wafer cost
- Higher yield
- Higher speed
- Lower power consumption (not always)
- Complex digital functions on chip (DSP)

**DISADVANTAGES:**

- Low power supplies
- Lack of “analog” components
- Inadequate modeling
  - Output conductance
  - Different inversion regions
- “Digital” noise

---

Analog design needs high-quality passive components. These are not present in processes optimized for digital design, or at least not in the first stages of the process development. These analog “options” are:

- High-resistivity poly for resistors
- Diffusion resistors
- Trimming options
- Linear and dense capacitors
  - Metal to metal (at least one special metal layer required)
  - Metal to poly
  - Poly to poly
Integrated capacitors

- High-linearity capacitors can be obtained with metal-to-metal structures. This generally requires adding a special metal layer to the technology, in order to reduce the dielectric thickness between the metal plates. The density reached is generally below 1 fF/μm² (not very high...).

- Dense capacitors can be obtained exploiting the high capacitance density of the thin gate oxide. This allows having dense and precise capacitors, good matching but very poor linearity. This solution can be adopted in any process.

- A third possible solution is suggested by the availability of many interconnection layers. Exploiting the parasitic capacitance between metal wires in a clever way, one can obtain linear capacitors with good matching and linearity and densities up to 1.5 fF/μm². These capacitors can be integrated in any process!!

Multi-metal-layer capacitors

This solution is a possibility, but it does not exploit the fact that in deep submicron processes the highest parasitic capacitance can be obtained “horizontally” rather than vertically, i.e. $t_{ox} > s$.

Fig. 3. Ratio of metal thickness to horizontal metal spacing versus technology (channel length).

Multi-metal-layer capacitors

MOS capacitors

NMOS in an N well
Accumulation Region

MOS structure
Accumulation Region

![Diagram showing NMOS in an N well and MOS structure with accumulation regions.](https://example.com/diagram.png)
C-V characteristics

The NMOS in an N well capacitor is in accumulation for \( V > 0 \) V.

The NMOS capacitor is in inversion for \( V > 0 \).
Outline

• Scaling down of CMOS technologies
• How scaling works for devices and interconnections
• Scaling impact on noise
• Scaling impact on matching
• Scaling impact on radiation tolerance
• Analog performance of submicron processes
• Substrate noise in mixed-mode integrated circuits
Integrating analog blocks on the same chip with digital circuits can have some serious implications on the overall performance of the circuit, due to the influence of the “noisy” digital part on the “sensitive” analog part of the chip.

The switching noise originated from the digital circuits can be coupled in the analog part through:

• The power and ground lines
• The parasitic capacitances between interconnection lines
• The common substrate

The substrate noise problem is the most difficult to solve.

Different types of substrates

Substrate noise: how to reduce it

To minimize the impact of disturbances coming from the substrate on the sensitive analog blocks, we have mainly three ways:

• Separate the “noisy” blocks from the “quiet” blocks. This is effective especially in uniform lightly doped substrates. For heavily doped substrates, it is useless to use a separation greater that about 4 times the epitaxial layer thickness.

• In n-well processes, p+ guard rings can be used around the different blocks. Unfortunately, this is again effective mainly for lightly doped substrates. Guard rings (both analog and digital) should be biased with separate pins.

• The most effective way to reduce substrate noise is to ground the substrate itself in the most “solid” possible way (no inductance between the substrate and ground). This can be done using many ground pins to reduce the inductance, or, even better, having a good contact on the back of the chip (metallization) and gluing the chip with a conductive glue on a solid ground plane.

• Separate the ground contact from the substrate contact in the digital logic cells, to avoid to inject the digital switching current directly into the substrate.