
Paul O’Connor

BNL

- Technology
- Scaling
- Pitfalls
- Monolithic active pixel sensor
Custom Monolithics – technology options

- **Bipolar**
  - Workhorse of “old” analog
  - Available from a handful of vendors
  - Speed/power advantage over CMOS (diminishing)
  - Low integration density
- **Standard CMOS**
  - Suitable for most analog designs
  - Best for combining analog and digital
  - Highest integration density
  - Widely available
  - Short life cycle (3 years/generation)
- **BiCMOS**
  - Complex process, viability uncertain
- **Silicon on insulator (SOI)**
  - Modest speed advantage for digital
  - Drawbacks for analog
- **SiGe**
  - Exotic
  - Interesting for high frequency work
- **GaAs**
  - Unsuitable for wideband analog
Access to custom CMOS is easy

• Design tools available at low cost to universities

• Multiproject services (MOSIS, Europractice, ...) provide low cost access to foundries for prototyping
CMOS Economics

Each ASIC may need

- 1.5 engineer-years $300K
- 2 prototype runs 30K
- 1 minimum production run 150K

TOTAL $480K

Incremental cost per chip ~ $10 – 20 / cm²
CMOS Scaling

• Driven by digital VLSI circuit needs
• Goals: in each generation:
  2X increase in density
  1.5X increase in speed
  Control short-channel effects, threshold fluctuation
  < 1 failure in 10⁷ hours
<table>
<thead>
<tr>
<th>Year</th>
<th>85</th>
<th>88</th>
<th>91</th>
<th>94</th>
<th>97</th>
<th>00</th>
<th>02</th>
<th>04</th>
<th>07</th>
<th>10</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min. feature size [µm]</td>
<td>2</td>
<td>1.5</td>
<td>1.0</td>
<td>0.7</td>
<td>0.5</td>
<td>0.35</td>
<td>0.25</td>
<td>0.18</td>
<td>0.13</td>
<td>0.10</td>
<td>0.07</td>
</tr>
<tr>
<td>Gate oxide [nm]</td>
<td>44</td>
<td>33</td>
<td>22</td>
<td>16</td>
<td>11</td>
<td>7.7</td>
<td>5.5</td>
<td>4.0</td>
<td>2.9</td>
<td>2.2</td>
<td>1.6</td>
</tr>
<tr>
<td>Power supply [V]</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5/3.3</td>
<td>3.3</td>
<td>2.5</td>
<td>1.8</td>
<td>1.2</td>
<td>1</td>
<td>.7</td>
</tr>
<tr>
<td>Threshold voltage [V]</td>
<td>1.0</td>
<td>0.9</td>
<td>0.8</td>
<td>0.7</td>
<td>0.6</td>
<td>0.5</td>
<td>0.45</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
</tbody>
</table>
IBM Cu-11 Process (Blue Logic)

- $L_{eff} = 0.08\ \mu m$, $L_{cham} = 0.11\ \mu m$
- Up to 40 million wireable gates
- Trench capacitor embedded DRAM with up to 16 Mb per macro
- Dense high-performance, comparable SRAMs
- Power supply: 1.2 V with 1.5 V option
- I/O power supply: 3.3 V (dual oxide option)/ 2.5 V (dual oxide option)/ 1.8 V/1.5 V
- Power dissipation of 0.009 $\mu W$/MHz/gate
- Gate delays of 27 picoseconds (2-input NAND gate)
- Seven levels of copper for global routing
- Low-k dielectric for high performance and reduced power and noise
- HyperBGA (flip chip): 2577 total leads
CMOS scaling and charge amplifier performance

- Fundamental noise mechanisms
  - so far, no dramatic changes with scaling
- Noise
  - slight improvement with scaling
  - higher device $f_T$ reduces series thermal noise
- Weak- and moderate inversion operation more common
  - need different matching to detector capacitance.
- Reduced supply voltage
  - difficult to get high dynamic range
- Many difficulties with “end of the roadmap” devices

P. O’Connor, G. DeGeronimo, “Charge amplifiers in scaled CMOS”, NIM-A accepted for publication
Charge preamplifier noise vs. scaling

<table>
<thead>
<tr>
<th>System</th>
<th>$C_{ext}$</th>
<th>$t_s$</th>
<th>P</th>
<th>$I_{fmax}$</th>
<th>Detector</th>
<th>Typical Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>30</td>
<td>75</td>
<td>10</td>
<td>.001</td>
<td>Wire Chamber</td>
<td>Tracking, Imaging</td>
</tr>
<tr>
<td>b</td>
<td>15</td>
<td>25</td>
<td>0.2</td>
<td>10</td>
<td>Si Strip</td>
<td>Tracking</td>
</tr>
<tr>
<td>c</td>
<td>0.3</td>
<td>25</td>
<td>0.02</td>
<td>1</td>
<td>Si Pixel</td>
<td>Tracking</td>
</tr>
<tr>
<td>d</td>
<td>3</td>
<td>2500 – 500$^a$</td>
<td>10</td>
<td>0.01</td>
<td>Semiconductor</td>
<td>Spectroscopy</td>
</tr>
</tbody>
</table>

UNITS: pF ns mW nA

$^a$ For this system, the shaping time was varied at each $L_{min}$ to optimize the overall noise (i.e., to make the white series noise and parallel noise equal).
Charge amplifier power vs. scaling

\[ P_{\text{diss}} (\mu W) \]

\[ L_{\text{min}} (\text{um}) \]

\[ C_d = 0.3 \text{ pF} \]

\[ t_s = 100 \text{ ns} \]

\[ \text{ENC} = 100 \text{ e}^- \]
Commercial microelectronic components, what’s changed since 1991?

- Renewed development of analog catalog parts
  - Data converters
  - Computer components -- disk drive readout, phone/network interface, displays
  - Wireless communication
  - Handheld and consumer devices

- CMOS supplanting bipolar as the technology of choice for analog applications

- Advances in packaging, PCB, assembly technology
  - Thin- and fine-pitch leadeed SMT components; BGAs; chip-scale packages; packages with low thermal resistance
  - Flip-chip and chip-on-board assembly
  - Microvias, thin-core laminates, flex for high density integration (HDI)
  - Passive component miniaturization, arrays
Cellular telephone handset trends

- **1991 cell phone**
  - ¾ pound
  - 12V battery
  - 700 components
  - 8 hrs assembly time
  - $600

- **2001 cell phone**
  - 2 oz.
  - 3V battery
  - 4–5 modular components + passives integrated in substrate
  - 15 minute assembly time
  - < $150 or free

---

*Figure 4 - Capacitors Ranging from 1206, 0805, 0603, 0402 and 0201*

This is the world's first WRIST CAMERA. It features 1 MB of memory to hold up to 100 images.
Standard packages of 2001

National microSMD
1.41 x 1.67 x 0.85mm body size (8L)
“Silicon Dust”

Amkor thin BGA
Stacked Chip Scale Package

Double-decker
(in production now)

Triple
Monolithic front ends – what can go wrong

• Frequently overlooked problems in design
  – Good electrical model of detector
  – Statistical nature of signals
  – Unusual signal conditions:
    – turn-on
    – calibration
    – response to background events
  – Detector-preamp interface
  – Board-level issues:
    – power conditioning,
    – bias decoupling,
    – calibration,
    – input protection,
    – interface components,
    – cooling
Monolithic amplifier design: practical considerations

- Preamplifier reset
- High order filters
- Programmable pulse parameters
- Self-biasing
- Low-swing, differential I/O
- Circuits tolerant to variations in
  - temperature
  - process
  - power supply
  - DC leakage current
  - input & output loading

<table>
<thead>
<tr>
<th>Base</th>
<th>&lt; 0.1 mV</th>
<th>&lt; 30 mV</th>
<th>&lt; 0.04°C</th>
<th>&lt; 8 mA</th>
<th>-</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>&lt; 0.1 mV</td>
<td>&lt; 300 mV</td>
<td>&lt; 0.04°C</td>
<td>&lt; 8 mA</td>
<td>&lt; 0.04°C</td>
<td>-</td>
</tr>
<tr>
<td>C2</td>
<td>&lt; 0.1 mV</td>
<td>&lt; 300 mV</td>
<td>&lt; 0.04°C</td>
<td>&lt; 8 mA</td>
<td>&lt; 0.04°C</td>
<td>-</td>
</tr>
</tbody>
</table>

Pulse vs. Temperature

Gain variation

Pulse vs. I_leak

Loading

Programmable Dual Stage N=24x6
- $C_n = 1.5\text{pF}$
- $Q_n = 12\text{fC}$
- Gain = 200mV/fC
- $T = -30\degree\text{C}$ to $+50\degree\text{C}$

Programmable Dual Stage N=24x6
- $C_n = 1.5\text{pF}$, $Q_n = 12\text{fC}$
- Gain = 200mV/fC
- $I_{leak} = 250\text{pA} \div 70\text{nA}$
Monolithic front ends – what can go wrong

• Management issues
  – Isolation of chip designer, board designer, detector specialist
  – Managers not knowledgeable of chip design process:
    • CAD tools
    • Foundry capabilities
    • Documentation and review procedures
    • Timelines, iterations
  – Progress episodic rather than incremental
    • Harder to track progress
Commonly heard phrases

- “We prototyped all the functional blocks, now all we have to do is put them together on the same chip and wire them up.”

- “All the chips work let’s go to production”

- “The chip works fine in simulation”

- “We already have a chip that does that, all we have to do is…”
CMS silicon strip readout

Prototype of layer 3 in the Inner Tracker

1 of the 128 channels

230 m$^2$ Si
12 million strips
92,000 APV-25 chips
APV-25: 0.25 um CMOS
128 chan X 192 bucket P/S, SCA, mux
246 + 36.3 e/pf, 2.3mW/chan, 2% nonlinearity to 5 MIP
7.2 X 6.5 mm, 85% yield
Si pixel readout

- Binary readout

Pixel needs to tell whether particle passed or not in a 25 ns time slot. It does NOT need to tell how much charge was collected over a certain time.
CMOS APS for particle detection/tracking

- Monolithic – special assembly technology not required
- Low cost
- Low multiple scattering
- Good spatial resolution (few μm)
- Random access
- Integration of control and DSP
- Radiation tolerance (?)

- Special process
- Collection time scales with pixel size
- Circuit architecture embryonic
Simple monolithic active pixel

- n+ pixel circuitry
- p-well
- p- well
- n- well
- p- epitaxial layer
- charged particle
- p++ substrate

- VDD
- RE_SEL
- ROW_SEL
- COLUMN LINE

“photo” diode

diffusion isochron
Comparison of bump-bonded and active pixel sensors for tracking

<table>
<thead>
<tr>
<th></th>
<th>Bump-bonded sensor</th>
<th>Active pixel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology</td>
<td>hybrid</td>
<td>monolithic</td>
</tr>
<tr>
<td>MIP signal charge</td>
<td>&lt; 24000</td>
<td>800</td>
</tr>
<tr>
<td>ENC noise charge</td>
<td>100 – 300</td>
<td>20 – 50</td>
</tr>
<tr>
<td>Pixel area</td>
<td>20,000 µm²</td>
<td>&lt; 400 µm²</td>
</tr>
<tr>
<td>Sensor capacitance</td>
<td>200</td>
<td>&lt; 10 fF</td>
</tr>
<tr>
<td>Detector bias</td>
<td>100</td>
<td>1 V</td>
</tr>
<tr>
<td>Charge collection time</td>
<td>&lt; 20</td>
<td>depends on pixel area</td>
</tr>
</tbody>
</table>
Summary

- PHENIX upgrade program can take advantage of a decade of progress in microelectronics.
- A study of the monolithic active pixel sensor as a vertex detector is warranted.
- By avoiding known pitfalls in the ASIC development process, cost and performance goals should be met.
BASIC APS READOUT WITH CORRELATED DOUBLE SAMPLING.

\[ \text{VCDS} = \frac{C_1}{C_2} \frac{V_{SIG} + (1 + C_1/C_2) (V_{OFF}/(1+A))}{1} \]
APS Readout with Zero Suppression

- Readout of each row followed by threshold discrimination and zero suppression in columns.
- No additional logic in pixels.
- Minimal periphery in one dimension allows close abutting.
- Achieves substantial reduction in readout time compared to non-sparse readout, but with much less overhead than full pixel-level zero suppression.

S. Kleinfelder
Video chip

Stuart Kleinfeld, SukHwan Lim, Xinqiao Liu, Abbas El Gamal

Figure 3: Pixel schematic.

10,000 fps, every 4th frame displayed propeller speed ~ 2000 rpm
Active Pixel Sensor (APS)

- 20 μm square pixels
- 5 chips per slat
- 90 million pixels
- 40 μm thick chips
Example: Hybrid Pixel Detectors
(CERN RD-19 E. Heijne et al.)

Pixel Ladders
(6 chips)

Half plane ~ 50 000 sensing elements

LHC1: 2000 CMOS readout channels

WA97 NA57 Experiment
1.2 M channels

153 Tracks!
- Sensor and FE-chips are directly connected with bumpbonds
- Signals are routed in a (flex-) capton layer on the sensor
- MCC and passive components are mounted on the flex
- Wirebonds connect FE-chips and MCC with the flex