Pixel-embedded signal processing for the next generation of radiation detectors at the LHC and FELs

Lodovico Ratti
(lodovico.ratti@unipv.it)
Università di Pavia and INFN Pavia, Italy
With the evolution of microelectronic technologies, pixel detectors have become essential devices in crucial applications in high energy physics (HEP) and photon science

- **HEP**: upgrades of the experiments at the LHC (CERN, Geneva, Switzerland) now in full swing
- **photon science**: 4th generation sources (X-ray free electron lasers, FELs) under construction around the world or already operative (e.g., EU-XFEL at DESY in Hamburg, Germany)

A pixel detector consists of

- an array of semiconductor sensors collecting charge released in the substrate
- a multichannel front-end chip with low-noise analog processing and some more or less intense digital processing on-board (channels in a 1 to 1 correspondence with the sensors)

**Different applications with different requirements**
**HEP: position sensing, momentum measurement**

- **Fundamental task:** measure the momentum of charged particles to reconstruct the decay vertex after beam-beam or beam target collision.

- **Detector interferes with measurement (multiple scattering) →** material budget issue: support and cooling has to be minimized → power dissipation to be reduced as much as possible.

- **To improve accuracy,** detector very close to the beam-pipe → reduce pixel size to improve position resolution and reduce occupancy.

- Precise amplitude information not so important → can be used to improve position measurement by center of gravity calculation of hit clusters.
Photon science: imaging $\rightarrow$ position + energy

- Fundamental task: measure the diffraction image of an X-ray pulse to reproduce the structure of a virus or a bio-molecule, or take a shot of a fast, nano-scopic phenomenon

- Accurate amplitude measurement is EXTREMELY important, from single photon to Poisson limit

- Need to store a large amount of images in the pixel (up to 3000 frames)

- The position of the detector with respect to the sample under measurement can be adjusted $\rightarrow$ pixel size is less demanding, but still small pixels are needed

- Single layer, no multiple scattering issues $\rightarrow$ room available for the cooling system $\rightarrow$ power dissipation is less of a problem, but still power is not unbounded
... but also some similarities

- **Multichannel, large area readout chips** → inter-channel and intra-channel (analog-digital) cross-talk, chip integration issues, power distribution

- **Complex circuits**: more than $10^8$ transistors, different IP blocks (voltage and current references, regulators, PLLs, ADCs and DACs, I/O pads), many different functions (data conversion and storage, sparsification, threshold adjustment, gain calibration)

- **Fast analog processing** to comply with the machine and the experiment requirements

- **Large amount of data** to transmit off-chip, with rates well in excess of $10$ Gbit/s
  - sparsification (suppression of uninteresting data) can reduce data rate in HEP applications to manageable levels
  - no sparsification possible in imaging applications

- **Unprecedented levels of radiation**: doses of ionizing radiation (TID, total ionizing dose) from several MGy to well above $10$ MGy
65 nm CMOS technology

After the 250 nm (LHC) and the 130 nm node (LHC upgrades, EU-XFEL,…), the HEP and X-ray FEL communities have adopted the 65 nm CMOS generation: several prototypes have been already fabricated and tested.

Among the wide choice of options of this technology, the Low Power flavor is less aggressive than other variants (thicker gate oxide, smaller gate current, higher voltage), and is more attractive for mixed-signal chips, where analog performance is an essential feature in a readout channel for semiconductor detectors.

65nm LP (Low Power) transistors (VDD = 1.2 V) are optimized for a reduced leakage (different level of gate oxide nitridation with respect to other flavours, different silicon stress,…)

Front-end electronics may benefit from scaling in terms of functional density (small pitch pixels) and digital performance - analog design remains a challenge (reduced supply voltage and dynamic range, statistical doping effects).

Design advances needed for full analog-digital integration: digital signal processing may be used to overcome analog limitations, analog circuits may be used to monitor the performance of digital circuits and their power consumption.
A digression on CMOS scaling

Why?
- faster and less power-hungry devices (shorter transit time, smaller parasitic capacitances)
- higher integration density
- increased radiation hardness

How?
- shrink device dimensions, increase substrate doping, reduce power supply, reduce gate oxide thickness

Classical scaling ended because of the reduced oxide thickness (direct tunneling at the gate)
L. Ratti, Pixel-embedded signal processing for the next generation detectors at the LHC and FELs
5th INFIERI International Summer School, Wuhan, China, May 12-26 2019

Nanoscale CMOS

to increase mobility


Fig. 2. NMOS cross-section. In addition to stress from cap layers and Ge raised source-drain (S-D) implants, device dimensions such as distance from source to drain, channel boundary to nearby STI (SA and SB), proximity and regularity of overlying metal patterns, and short distances to other device patterns within the local (≤ 2 μm) stress field induce transverse (F_y) and lateral (F_x and F_z) stress components, which affect threshold and mobility. Increasing the distance to P+ ties increases local tub (bulk) resistance components R1 and R2, which isolate the device MOS model substrate node from the device subcircuit symbol V_th and degrade RF performance. Hot carrier reliability stress is dependent on the sum of transverse and lateral fields E_y and E_x. These fields are increased near the drain by increasing source to bulk (V_sb) and drain (V_d) to gate (V_g) or source (V_s) voltages in various combinations. As hot carrier stress increases, damage to channel from interface trap density (N_it) affects threshold and mobility, while gate oxide thickness (ON) or high-dielectric-constant (Hi-K) insulator trap density (N_o) affects threshold and gate leakage.
Evolution of scaling

Alternative gate dielectrics

- With a high dielectric constant (high-k) material, a much thicker gate dielectric can be used, with the equivalent capacitance of much thinner SiO$_2$-based structures (in smaller than 45 nm CMOS).

- Thicker dielectrics are more sensitive to ionizing radiation; as always, actual behavior will depend on process details - hafnium-based dielectrics with good radiation tolerance have been reported.

Multiple gate (MUG) devices

- 3D gate structures have been devised as a way to avoid short-channel effects in aggressively scaled MOSFETs (≤ 22 nm). Control of lateral gates on silicon channel may be beneficial in terms of radiation tolerance (no lateral leakage). However, radiation effects in these advanced devices may be more complex than in bulk MOS.

Carbon-based electronics (beyond CMOS)

- Carbon nanotubes and graphene have generated much interest: not yet clear if they will be a replacement for Si CMOS. Because of their extremely low volumes (few atomic layers), their radiation response may mostly depend on interfaces and surrounding materials.
Coming next: FinFET and UTB FD SOI

- Ultra-thin body fully depleted SOI and non-planar FinFET device structures promise to be capable of extending the CMOS scaling trend

https://www.eetimes.com

https://blog.globalfoundries.com
Anticipate changes (if you can)

Transition from planar, single-gate to vertical, multiple-gate structure does not affect analog performance significantly.

Increase, as compared to old CMOS, of the threshold voltage over VDD ratio may force the analog designer to favour simpler architectures and/or look for innovative design solutions.

FinFET 14 nm, VDD=900 mV

NMOS
W/L=600 μm/80 nm
V_{ds}=0.4V

PMOS
W/L=600 μm/80 nm
V_{sd}=0.4V
Outline

Signal processing in HEP
- Experiment upgrades at the LHC: the RD53 collaboration
- Classical readout chain for capacitive detectors
- Linear front-end
- Synchronous front-end
- Differential front-end

Signal processing in photon science
- X-ray free electron lasers (FELs)
- Dynamic range and compression
- Dynamic compression with MOS capacitors
- Time-variant shaping
- In-pixel A to D conversion
Processing the signal from pixel detectors in high energy physics
Very challenging requirements for the innermost layers of the pixel detectors in ATLAS and CMS

- very high particle rate: ~500 MHz/cm² → hit rates of 2-3 GHz/cm²
- smaller pixels: 25 x 100 or 50 x 50 um² → increased resolution, improved track separation
- increased trigger rate: 1 MHz
- low mass, low power, <0.5W/cm²
- harsh radiation environment: 10 MGy(SiO₂) TID, 10¹⁶ 1MeV eq. n/cm² fluence
- low threshold: 600-1000 e⁻ → severe requirements on noise and dispersion
CHIPIX65 and RD53

Activity funded by INFN through the CHIPIX65 project, carried out also in the framework of the RD53 collaboration

**CHIPIX65**
- Italian collaboration funded by INFN (~700 kEuro/3 years, started 2014), led by Lino Demaria (INFN, Torino, Italy)
- INFN units from Bari, Milano, Padova, Pavia, Perugia, Pisa & Torino
- develop a chip for pixel detectors using a 65 nm CMOS process
- organized in 5 WP: radiation hardness, digital electronics, analog electronics, chip integration and management
- develop analog front-end circuits, IP blocks, explore new digital readout architectures, fabricate & test a 64x64 readout channel array

**RD53**
- joint CMS-ATLAS effort to develop pixel front-end for the phase 2 upgrades
- 22 institutions from Europe and US, >100 members
- 65 nm CMOS is the common technology platform
- about 50% of the people are microelectronic designers
- synergies with other collaborations (e.g., CLIC)
- study radiation effects, develop a simulation & verification environment, design & share rad-hard libraries, design small size prototypes & a full size pixel array (>1 cm²) from a common engineering run and test it on a beam
RD53 chip architecture

- 95% digital
- Charge digitization (ToT)
- 256k pixel channels per chip
- Pixel regions with buffering
- Data compression in end-of-column
- Chip size > 20 mm x 20 mm
Hybrid pixel detector (HPD) approach

Each layer can be fabricated using the best technology for each function

- detector layer: direct conversion of charged particle (or photon) energy to charge \(\Rightarrow\) good spatial resolution
- readout chip layer: CMOS technology offers large integration density, low power dissipation, radiation hardness, high processing speed

Generally long development times and large costs
Optimal readout channel for semiconductor detectors

- Large bandwidth charge to voltage conversion
- Band limitation introduced by the filtering stage (shaper)

\[ \frac{Q}{C_F} \]

reset network

\[ T(s) \]

\[ \mu \frac{Q}{C_F} \]

\[ V_S \]

\[ V_O \]

\[ I_L \]

\[ C_D \]

\[ C_F \]

detector

charge preamplifier

shaper

L. Ratti, Pixel-embedded signal processing for the next generation detectors at the LHC and FELs
5th INFIERI International Summer School, Wuhan, China, May 12-26 2019
Charge digitization

- **Reset network**
- **V_{th}**
- **V_{o}**
- **V_{s}**
- **T(s )**
- **C_{F}**
- **latch**
- **AND**
- **counter**
- **peak & hold**
- **ADC**

- **binary readout**
- **time over threshold (ToT)**
- **A/D conversion**

Only when amplitude measurement is required.
Time over Threshold (ToT) method

- Provides a direct amplitude-to-time conversion
- Signal at the shaper output compared to a fixed voltage at the input of threshold discriminator
- The duration of the discriminator signal is the time during which the signal at the shaper output exceeds the threshold
- Digitization achieved by computing the logic AND between the discriminator pulse and a reference clock and by counting the number of clock pulses
- Preampli output returns to the baseline with a constant slope → linear relationship between peak amplitude and the ToT duration (the rise time of the shaper output signal is assumed to be negligible)
Channel specifications for CMS/ATLAS pixels

- **Input dynamic range**: $3 \times 10^4$ electrons (set by detector thickness and experiment features), with a threshold of 600 to 1000 electrons
- **Dead-time**: 400 ns max (to minimize pile up)
- **Amplitude measurement**: ≥ 5 bits
- **Capability** to comply with relatively large currents, up to 10 nA, from the detector (radiation-induced)
- **Area**: 50 um x 50 um (or 25 um x 100 um) per pixel; less than a half for the analog section
- **Current consumption**: 4 uA max (@1.2 V) for the analog section (~half of the available power budget per pixel, the overall power budget being 0.4 W/cm$^2$)
- **Noise and threshold dispersion**: root of the quadratic sum should not exceed 130 electrons (for a noise hit rate below $10^{-6}$ Hz)
RD53A, a large scale prototype

- On the end of August 2017, the RD53 collaboration submitted the RD53A chip

- **Goal:** demonstrate in a large format IC
  - suitability of the 65 nm technology (including radiation tolerance)
  - high hit rate capabilities: 3 GHz/cm²
  - trigger rate: 1 MHz
  - low threshold operation with the chosen isolation strategy and power distribution

- RD53A is not intended to be a production chip
  - contains design variations for testing purposes (with three different versions of the analog front-end)
23
L. Ratti, Pixel-embedded signal processing for the next generation detectors at the LHC and FELs
5th INFIERI International Summer School, Wuhan, China, May 12-26 2019
Serial powering scheme

RD53A is designed to operate with **Serial Powering** → constant current to power chips/modules in series

Based on ShuntLDO

Designed having in mind the production chip

**Three operating modes:**

- **Shunt-LDO:** constant input current $I_{in} \rightarrow$ locally regulated VDD
- **LDO (Shunt is OFF):** external un-regulated voltage $\rightarrow$ locally regulated VDD
- **Externally regulated VDD** (Shunt-LDO bypassed)
Pixel region

- 2 x 2 analog island in deep N-well, digital sea all around
- Routing in lower metals
- Shielding of bias lines
- Top three metals for power grid
- Digital place & route at 67%, some space still available
- Total power ~92 uW \rightarrow 5.7 \text{ uW} per pixel

elementary cell
The RD53A chip

Sync FE
128 columns

Lin FE
136 columns

Diff FE
136 columns

Chip size: 20.066 x 11.538 mm²
400x192

Aug. 31, 2017: Submission
Dec. 6, 2017: First chip test
Mar. 15, 2018: 25 wafers ordered
Apr. 13, 2018: First bump-bonded chip test
Pixel front-end: an asynchronous approach

- ToT architecture for charge digitization
- No shaper to minimize area and power (shaper-less channel)
- 4 bit DAC for threshold correction

65 nm LP CMOS technology

Developed by L. Gaioni, M. Manghisoni, G. Traversi, V. Re (University of Bergamo and INFN Pavia), F. De Canio and L. Ratti (University of Pavia and INFN Pavia)
Krummenacher feedback network: leakage compensation

- Provides $I_L$ current required by the detector (leakage), only limited by operating conditions of $M_2$ and $M_3$
- Response is a falling exponential for a small input signal, triangular for large signals (Krummenacher stage no longer linear)
- Also sets the DC voltage at the preamplifier output
Time over threshold calculation

- The response to a large signal (larger than 10% of the max signal) has a triangular shape.
- If $t_p$ is very small, $\text{ToT}$ is linear with $Q$.

\[ \text{ToT} = \frac{V_p - V_{\text{TH}}}{I_K} \approx 2Q - \frac{V_{\text{TH}}C_F}{I_K} \]

\[ \text{ToT}_{\text{MAX}} = 2 \frac{Q_{\text{MAX}} - V_{\text{TH}}C_F}{I_K} \approx 2 \frac{Q_{\text{MAX}}}{I_K} \quad I_K \approx 2 \frac{Q_{\text{MAX}}}{\text{ToT}_{\text{MAX}}} \]

- The maximum $\text{ToT}$ (set to 400 ns) value is the one obtained for $Q_{\text{MAX}}$. 

L. Ratti, Pixel-embedded signal processing for the next generation detectors at the LHC and FELs
5th INFIERI International Summer School, Wuhan, China, May 12-26 2019
Time walk

- Delay of the signal at the discriminator output depends on the amplitude of the signal at the input → time walk
- An event may fail to be assigned to the right bunch crossing (40 MHz rate) if delays are too large

Response to the largest possible signal (30000 electrons)

Response to the smallest detectable signal (900 electrons, 300 electrons over the minimum threshold)
Threshold discriminator

- Indicate a significant event in the detector with the rising edge of the output signal
- Freeze the content of a local time stamp → provide a time label to the event
- Enable a local counter for direct digitization of ToT

Transimpedance amplifier: provides a low impedance path and a small time constant for fast switching

\[
p = C_p R_p = \frac{C_p}{G_m} \frac{1}{1 - G_{\text{loop}}}
\]

\[-(g_{m,21} + g_{m,22})(r_{o,21} // r_{o,22})\]

Series of \(g_{m,i}\) and \(g_{m,i+1}\),
i=19 on \(0 \rightarrow 1\) transitions \((i_{\text{tr}} < 0)\),
i=17 on \(1 \rightarrow 0\) transitions \((i_{\text{tr}} > 0)\)

Minimize crowbar current in the output stage
Measurements at the preamplifier output and ToT

Good ToT linearity for $Q \geq 2000 \text{ e}^-$ (and well in excess of $30000 \text{ e}^-$)

In-time overdrive smaller than $300 \text{ e}^-$ for a threshold of $600 \text{ e}^- @ C_D^* = 80 \text{ fF}$
Noise in MOSFET transistors

- Noise is the ultimate limit to the accuracy with which a measurement can be performed.
- Noise in a MOSFET can be circuitually described by means of a voltage source $e_n$ in series to the gate terminal of the device; this source, formally represented by means of its power spectral density, includes two terms:
  - A frequency independent one ($S_W$), dominated by the channel thermal noise, which originates from thermal agitation of carriers in the device channel.
  - A term which is inversely proportional to the frequency, also called $1/f$ or flicker noise, which arises from continuous, random capture and release of carriers by border (very close to the Si/SiO$_2$ interface) traps in the oxide.

\[ \frac{d^2e_n}{df} = S_w + A_f f \]

- Noise is the ultimate limit to the accuracy with which a measurement can be performed.

\[ S_w = \frac{4k_B T}{g_m} \]

\[ A_f = \frac{K_f}{C_{OX} W L} \]

- $k_B$ = Boltzmann’s constant
- $T$ = absolute temperature
- $\Gamma$ = channel thermal noise coefficient
- $g_m$ = channel transconductance
- $K_f$ = $1/f$ noise parameter
- $W, L$ = channel width and length
- $f$ = frequency
- $\alpha_f$ = flicker noise slope coefficient
Equivalent noise charge

Figure of merit for a charge measuring system: charge to be injected at the preamplifier input to have a unit signal-to-noise ratio at the output

\[
ENC = \sqrt{\frac{v_n^2}{G_Q^2}}
\]

\(v_n^2\): mean square value of the noise at the channel output

\(G_Q\): charge sensitivity
Total ionizing dose effects in MOSFET devices

An MOS device exposed to ionizing radiation typically suffers degradation in one or more of its parameters (threshold voltage, gate voltage to drain current gain, or transconductance, channel leakage, noise); changes may not be constant with time after irradiation and may depend on the dose rate, bias conditions during irradiation, temperature during and after irradiation.

After being irradiated, an integrated CMOS circuit may slow down, show higher leakage (parasitic) currents, exhibit degraded noise performance or even cease functioning properly (catastrophic failure).

Damage responsible for these total dose effects occurs in the insulating layers ($\text{SiO}_2$) of the device structures and at the interface between the silicon substrate of the device and the oxide, and consists of three main components:

1. buildup of (positive) charge trapped in the oxide (the gate oxide and/or the field oxides, used to isolate devices from each other)
2. increase in the density of traps at the Si/$\text{SiO}_2$ interface and in $\text{SiO}_2$ close to the interface (border traps)
3. increase in the density of traps in the oxide bulk
ENC measurements, including radiation effects

Increase in the slope compatible with an increase in the series noise contribution (flicker noise in particular)
Threshold setting

Threshold chosen in such a way to optimize detection efficiency (also depending on the chip readout architecture, prompt readout of detected hits or delayed readout with on-chip buffering)

\[ f_n = f_{n0} \exp \left( - \frac{V_{TH}^2}{2 \sigma_n^2} \right) \]

noise hit rate at zero threshold

- noise hits (false events)
- missing true hits

\[ \text{noise hit rate} \]

\[ f_n = f_{n0} \exp \left( - \frac{V_{TH}^2}{2 \sigma_n^2} \right) \]

mean square noise
Threshold setting in the presence of noise

• Suppose noise distribution is a Dirac delta (all the channels have the same noise)
• Maximum noise hit rate is set to $f_{n,\text{max}}$

$$f_{n,\text{max}} = f_{n0} \exp\left(-\frac{V_{\text{TH}}^2}{2\sigma_n^2}\right)$$

$$V_{\text{TH}} = \sqrt{2\ln\left(\frac{f_{n0}}{f_{n,\text{max}}^n}\right)} = (f_{n,\text{max}})^n$$

Noise is not the only constraint
Threshold dispersion

- Random and systematic variations of process (doping) and geometrical (device dimensions) parameters → non uniformities in the parallel path followed by the signals → two nominally identical channels (including nominally identical discriminators) with the same threshold may provide different responses.

\[ V_{TH} \]

\[ V_{th} \]

\[ V_{th} + \Delta V_{th} \]

\[ V_{TH} + \Delta V_{TH} \]
Effect of threshold dispersion

\[ \rho(f_{n,\text{max}}) \sigma_n \]

\[ \sigma_n \]

\[ V_{th} \]

\[ \text{noise distribution} \]

\[ \text{threshold voltage distribution} \]

\[ \text{half of the channels has noise hit rate } > f_{n,\text{max}} \]
$V_{TH}$ setting in the presence of threshold dispersion

If the probability density function of the threshold voltage is Gaussian

$$n_{hc,max} = \sqrt{2} \cdot Erfc^{-1}(2n_{hc,max})$$

Not only noise, but also threshold dispersion should be minimized

If $\rho(f_{n,max}) \sigma_n + \lambda(n_{hc,max}) \sigma_{th}$

maximum fraction of "hot" channels

L. Ratti, Pixel-embedded signal processing for the next generation detectors at the LHC and FELs
5th INFIERI International Summer School, Wuhan, China, May 12-26 2019
Threshold correction

Threshold tuning procedure

- With threshold scan techniques the threshold voltage is measured for each channel.
- The average threshold voltage value is calculated.
- For each channel, the difference between its threshold and the average value defines the bit sequence.
- The bit correction sequence is loaded during the programming phase.

Signal from the shaper stage

$V_{TH}$

Discriminator

IN-PIXEL LOGIC

Decoder

AVDD

$I_{BIAS}$

$I_{DAC}$

$I_{DAC}$

$I_{DAC}$

$I_{DAC}$

$B_0$ $B_1$ $B_2$ $B_3$

15

$I_{bias} = 350 \text{ nA}$

$I_{cell} = 10 \text{ nA}$

Before correction

After correction

Count

1000

800

600

400

200

0

Iout DAC [V]

0

1.06

1.065

1.07

1.075

1.08

L. Ratti, Pixel-embedded signal processing for the next generation detectors at the LHC and FELs
5th INFIERI International Summer School, Wuhan, China, May 12-26 2019

42
Threshold correction vs DAC range

\[ \frac{\sigma_{th,c}}{\sigma_{th}} \]

\[ \theta \text{ (i.e., DAC output range/} \sigma_{th} \text{)} \]

DAC resolution

- 3-bit
- 4-bit
- 5-bit
- 6-bit
- 7-bit
- 8-bit
Optimum configuration for threshold correction

\[ J_{\text{opt}}(n) = a + b \times n \]

\[ a \approx 2.96, \quad b \approx 0.63 \]

\[ J_{\text{sth}, c}(n) = c + e^{-d \times n} \]

\[ a \approx 1.26, \quad b \approx 0.61 \]
Pixel front-end: a synchronous approach

- ToT architecture (5 bit) for charge digitization
- No shaper to minimize area and power (shaper-less channel)
- Offset compensation and local threshold adjustment based on auto-zero technique
- Local fast oscillator for high resolution (8 bit) measurements

Developed by N. Demaria, E. Monteil, L. Pacher, A. Rivetti, M. Da Rocha Rolo, University of Torino and INFN Torino, Italy

65 nm LP CMOS technology
Synchronous hit discrimination

Track & latch comparator architecture

- voltage difference at the input is amplified by a chain of low-gain/high bandwidth amplifiers (track)
- voltage difference at the latch input is further amplified through a fast, regenerative mechanism (latch)

Comparator operation is synchronous with bunch crossing clock \(\rightarrow\) for a fast enough preamplifier, no ambiguity in the time stamp assignment of the event

Latch based comparators have the potential for very low power operation
Offset cancellation and threshold adjustment

Auto-zeroed comparator with output offset storage
- effect of the input offset at the preamplifier output is sensed, sampled and stored
- stored value is added to the preamplifier output → offset effect is removed

One cycle to be periodically used for offset compensation → minimum offset sampling frequency set by storing capacitance leakage ~10 kHz (<< 40 MHz HL-LHC bunch crossing frequency)

Offset efficiently stored for 100 us

L. Ratti, Pixel-embedded signal processing for the next generation detectors at the LHC and FELs
5th INFIERI International Summer School, Wuhan, China, May 12-26 2019
Fast ToT encoding

Asynchronous logic network is used to turn the discrete time comparator into an oscillator

- high frequency, self generated clock for fast ToT digitization
- oscillation frequency depending on size of latch transistors and on delays set by NAND and NOR gates

Oscillation frequencies in the order of 4 GHz - high resolution amplitude measurements, 8-9 bit ToT

L. Ratti, Pixel-embedded signal processing for the next generation detectors at the LHC and FELs 5th INFIERI International Summer School, Wuhan, China, May 12-26 2019
On-pixel calibration circuits play a fundamental role in verifying proper front-end functionality when no detector is connected:

- Externally generated pulses may be subjected to timing issues due to RC effects.
- On-chip pulser, generally in the periphery, does not solve the problem, while increasing power dissipation.

- Power-efficient solution: local generation of test pulse from two DC voltage levels (12 bit on-chip DAC) distributed to all pixels - a digital signal is used to switch from one level to the other.

- Two operation modes which allow to generate two consecutive signals of the same polarity or to inject a different amount of charge in neighboring pixels at the same time.

L. Ratti, Pixel-embedded signal processing for the next generation detectors at the LHC and FELs
5th INFIERI International Summer School, Wuhan, China, May 12-26 2019
Measurement results

Output signal with triangular shape (when output not saturated), discharge time (corresponding to dead time) can be set by changing the current in the Krummenacher network

ENC of about 80 electrons at $C_D=50$ fF

Response to 100 ke$^-$

- Ifeed=40 nA → falling time=90 ns → <0.3% inefficiency
- Ifeed=10 nA → falling time=360 ns → ~1% inefficiency
Differential analog front-end

- detector leakage compensation circuit (LCC)
- first stage: continuously reset charge preamplifier DC coupled with the following pre-comparator
- two-stage open loop comparator with fully differential input
- threshold adjustment with global 8bit DAC and two per pixel 4bit DACs

A. Mekkaoui (FNAL), D. Gnani, A. Krieger, T. Heim, M. Garcia-Sciveres (LBNL) C.A. Gottardo (U. Bonn)
Preamplifier response

- Straight regulated cascode architecture with NMOS input transistor in weak inversion
  - operated with a single bias at a 2 uA current
  - gain and LCC selectable
Leakage compensation circuit

**Differential OTA**
- Negligible power (0.05-1nA)
- A small bandwidth loop provides the DC leakage current for the detector
- Bias DAC sets BW and DC sensitivity
Comparator

Two-stage comparator w/ hysteresis

- Fully differential input
- Limited output slew-rate, relies on nearby digital buffer (local buffer planned for RD53B)
- Static power: 1.2V x 500nA nom.
Noise and threshold distribution

- Bug in the A/D interface: missing P&R constraint on the Diff. FE hit output → Varying load capacitance on comparator output → systematic variation of delay and ToT
- This bug did not prevent the Diff FE full characterization → Non default parameters to minimize the effect of load capacitance
- Low threshold achieved with 35 e^- rms threshold dispersion in non-default configuration → (slower wrt nominal)
RD53A bump bonded to a detector

- Image of a nut placed on the sensor backside, illuminated with Am241 source
- Hit-OR-trigger scan, LIN and DIFF FE, both set to 3 ke- threshold, un-tuned
- Need some more FW/SW development to implement auto-zero sequence for SYNC FE
Processing the signal from pixel detectors in photon science
X-ray sources as probing tools

Since their discovery by W. Röntgen, X-rays have been used as a powerful tool to probe matter.

The use of large accelerometer-driven X-ray sources such as synchrotron light sources and X-ray free-electron lasers (FEL) is continually growing and expanding to many scientific disciplines worldwide.

These facilities are now driving the state of the art of X-ray science, therefore shaping the requirements for many types of X-ray detectors.

X-ray FELs in particular can offer unprecedented capabilities in penetrating the microscopic structure of organic and inorganic systems, new materials and matter under extreme conditions and in recording and understanding the time evolution of fast biochemical phenomena at the nano-scale.
Free electron laser operation

- In free electron lasers, a beam of relativistic electrons moves through the magnetic field generated by a periodic structure (wiggler or undulator).

Electrons acquire an undulatory motion in the plane orthogonal to the magnetic field → acceleration → longitudinal emission of the synchrotron radiation type.

- Stimulated emission process comes about through the interaction between the e.m. field of the laser beam with the relativistic electrons.

- The wavelength of the emitted radiation depends on undulator geometry, magnetic field intensity, and the electron energy.

Free electrons → not bound to a single atom or molecule or confined along a chain of atoms or in a crystal lattice.
Free electron laser facilities

FELs provide high intensity beam of ultrafast X-rays
• energy range: 100 eV to >10 keV ($\lambda$ from 10 nm to 0.1 nm)
• pulse duration: femtoseconds to picoseconds
• repetition rate: 10 Hz (continuous mode) to 5 MHz (burst mode)
• peak brightness may exceed $10^{33}$ ph s$^{-1}$ mm$^{-2}$ mrad$^{-2}$

<table>
<thead>
<tr>
<th>Project</th>
<th>Start of operation</th>
<th>Electron beam energy [GeV]</th>
<th>Photon energy [keV]</th>
<th>Repetition rate [Hz]</th>
<th>Number of X-ray pulses/burst @inter-pulse period</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLASH@DESY</td>
<td>2005</td>
<td>1.25</td>
<td>0.03-0.3</td>
<td>5</td>
<td>800@1 us</td>
</tr>
<tr>
<td>LCLS@SLAC</td>
<td>2009</td>
<td>14.5</td>
<td>0.3-10</td>
<td>120</td>
<td>1</td>
</tr>
<tr>
<td>SACLA@RIKEN</td>
<td>2010</td>
<td>8</td>
<td>4.5-15</td>
<td>60</td>
<td>1</td>
</tr>
<tr>
<td>Fermi@ELETTRA</td>
<td>2010</td>
<td>2.4</td>
<td>0.01-0.06</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>SwissFEL</td>
<td>2016</td>
<td>5.8</td>
<td>12</td>
<td>100</td>
<td>2@50 ns</td>
</tr>
<tr>
<td>Eu-XFEL</td>
<td>2017</td>
<td>17.5</td>
<td>0.4-20</td>
<td>10</td>
<td>2700@220 ns</td>
</tr>
<tr>
<td>LCLSII</td>
<td>&gt;2020</td>
<td>4-14.5</td>
<td>0.2-25</td>
<td>120-10$^6$</td>
<td>1</td>
</tr>
</tbody>
</table>
Beam-line and beam-time structure

Beam lines with different photon energies available at each facility

Very different beam structure from one FEL facility to the other - some pose very challenging requirements on the instrumentation

**Eu-XFEL**

- ~2700 pulses/600μs
- 99.4ms
- 220ns
- X-ray photons
- FEL process

**Today’s x-ray laser sources**

- ~ millijoule
- ~ milliseconds

**LCLS**

- Intense pulses at low rep rate

**LCLSII**

- Intense pulses at high rep rate

**Tomorrow’s x-ray laser sources**

- ~100 microjoule
- ~ microseconds
- ~ attoseconds to femtoseconds
What is so special about FELs

Broad science base accessible: structural biology, chemistry, material science, atomic and molecular science

Unprecedented features of X-ray FELs can make a number of new measurements possible

- In X-ray crystallography, the sample is a periodic structure \( \Rightarrow \) scattered amplitude at the Bragg peaks amplified by the square of the number \( N \) of unit cells
- small and/or non-periodic samples (single biomolecules or small cells) \( \Rightarrow \) increased source brightness to compensate for the lack of Bragg peak amplification

- In some applications (e.g., X-ray absorption spectroscopy) a pump probe technique is used - 1) a dynamic process is initiated, 2) after a pre-set delay \( \tau \), the excited sample is probed by a synchronized X-ray pulse, 3) measurement is repeated for different values of \( \tau \)
- process evolving while the pulse is passing through the sample (which also undergoes Coulomb explosion) \( \Rightarrow \) very short pulses to avoid “blurred” pictures

What is so special about FELs

STRUCTURAL BIOLOGY

CHEMISTRY

L. Ratti, Pixel-embedded signal processing for the next generation detectors at the LHC and FELs
5th INFIERI International Summer School, Wuhan, China, May 12-26 2019
2D X-ray imaging challenges

Many measurements based on scattering of coherent X-ray pulses and detection of diffraction patterns with a large pixel camera.

New detectors are needed to comply with extremely challenging requirements:

- **pulse structure**: high rate single shot imaging (4.5 MHz), frame storage of a complete bunch train (2700 pulses)
- **dynamic range**: single photon resolution, integration of up to $10^4$ ph/pixel/pulse
- **energy range**: 0.25 keV to 25 keV
- **radiation hardness**: 10 MGy to 1 GGy over three years of operation
- **angular resolution**: 7 mrad to 4 urad ⇒ pixel size from 700 um to 16 um (also depending on the distance from the sample)
- **angular coverage**: diffraction experiments require a 0.1 nm resolution ⇒ scattering angles of 120°
- **dead area**: as small as possible
Beam structure at the Eu-XFEL

- X-ray pulses at a repetition rate of 4.5 MHz with a time interval of ~100 ms between two subsequent bursts ➔ in-pixel storage needed
- Burst duration of about 600 us with 2700 pulses per train
Dynamic range

- In diffraction imaging applications, the dynamic range of the signal to be processed can be as large as 80 dB (1 to $10^4$ photons)
- Single photon resolution is a desirable feature of the readout electronics - 1 ADC count for each additional photon

Assuming $V_{\text{max}}=1$ V, $n_{\text{max}}=10^4$, $E_{\text{ph}}=1$ keV
- LSB=100 uV, no. of bits ≥ 14 ⇒ large area, power dissipation, data rate
- $G=100$ uV/ph≈0.4 uV/el≈2.5 mV/fC ⇒ poor SNR, or large power dissipation
Signal compression

Possible solution: cover the dynamic range using a piecewise linear gain (signal compression) - the simplest choice may be a bilinear characteristic

- implement single photon resolution for \( n_{ph} \leq n_H \)
- chose the gain accordingly for the remaining part of the characteristic

\[
G_H = \frac{V_H}{n_H}, \quad G_L \approx \frac{V_{max} - V_H}{n_{max}}
\]

Assuming \( V_{max} = 1 \text{ V}, n_{max} = 10^4, n_H = 100, V_H = 200 \text{ mV} \)

- \( G_H = 2 \text{ mV/ph}, G_L \approx 80 \text{ uV/ph} \)
- LSB = 2 mV, no. of bits \( \geq 9 \)
- in the low gain region, \( n = 25 \text{ ph/ADC count} \rightarrow \) quantization noise

\[
q = \sqrt{\frac{n^2 - 1}{12}}
\]
**Shot or Poisson noise**

The number of photons hitting a pixel during diffraction imaging experiments is subject to fluctuations according to a random process described by Bose Einstein statistics \( \Rightarrow \) photon shot noise with variance \( \sigma^2_p \)

\[
\sigma^2_p (n_{ph}) = n_{ph} \exp \left( \frac{\hbar c}{k_B T} \right) \left[ \exp \left( \frac{\hbar c}{k_B T} \right) - 1 \right]
\]

- \( h = \text{Plank's constant} \)
- \( k_B = \text{Boltzmann's constant} \)
- \( T = \text{absolute temperature} \)
- \( \lambda = \text{photon wavelength} \)

Actually, for \( \frac{\hbar c}{\lambda} >> k_B T \), Poisson statistics is a good approximation of Bose-Einstein statistics and

\[
\sigma^2_p (n_{ph}) \approx n_{ph}
\]

Shot noise increases with the square root of the number of photons \( \Rightarrow \) no use in having single photon resolution when the number of detected photons gets large
Quantization vs Poisson noise

- Choose $n_H$ (i.e., the range in which single photon resolution is implemented) and the number of bits in such a way that

\[ q^2 \ll p^2 \]

- Once the number of bits $n_{\text{bit}}$ is given

\[ q = \sqrt{\frac{1}{12} \left( \frac{n_{\text{max}} - n_H}{2^n_{\text{bit}} - n_H} \right)^2 - 1} \]

- For a bilinear, compressed characteristic, at least 9 bits are needed for quantization noise to be always smaller than Poisson noise
Overview of Eu-XFEL imaging detectors

Modular detectors based on hybrid pixel sensor technology

DSSC - DEPFET Sensor with Signal Compression
- energy range: 0.5-6 keV
- dyn. range: $10^4$ ph@1 keV
- single photon sensitivity
- storage cells (digital): ~640
- pixel pitch: 236 um
- sensor with compression features

AGIPD - Adaptive Gain Integrating Pixel Detector
- energy range: 3-13 keV
- dyn. range: $10^4$ ph@12 keV
- single photon sensitivity
- storage cells (analog): ~352
- pixel pitch: 200 um
- compression obtained through dynamically switched gain

LPD - Large Pixel Detector
- energy range: 1-25 keV
- dyn. range: $10^5$ ph@12 keV
- single photon sensitivity
- storage cells (analog): ~512
- pixel pitch: 500 um
- no real compression, three channels with different gain
Long term goal of the PixFEL project

Develop a four-side buttable, multi-layer module for the assembly of large area detectors with minimum dead area

- Good efficiency from 1 keV (optimized entrance window) up to 10 keV (450 um thickness)
- 9 bit resolution (effective), 5 MHz sampling rate
- 1 kframe
- Wide dynamic range (1 to 10000 photons), single photon sensitivity
- Burst and continuous mode operation

L. Ratti, Pixel-embedded signal processing for the next generation detectors at the LHC and FELs
5th INFIERI International Summer School, Wuhan, China, May 12-26 2019
Readout channel

- **Charge sensitive amplifier** - dynamic signal compression
- **Time invariant filter** - gain and integration time selection options
- **Analog-to-digital conversion** - 10 bit SAR ADC
- **Pitch**: 100 um $\xrightarrow{65}$ nm CMOS and 3D integration (high density TSVs and interconnect)

- **dead area**: as small as possible, 2% seems feasible $\xrightarrow{\text{active edge technology and low density TSV interconnects}}$
- **wide dynamic range**: single photon resolution for small number of photons, full dynamic range of $10^4$ ph $\xrightarrow{\text{1 keV or 10 keV}}$ low noise charge preamplifier with dynamic compression
- **readout**: 10 bit @ 5 MS/s
Charge preamplifier: inversion mode MOS cap

**Capacitor terminals:** one is source shorted to drain, gate is the second one (sub to GND)

- $0 < V_{G,SD} < V_{Th}$ → $C_{G,SD}$ is set to its minimum value, the sum of the overlap gate-to-source and gate-to-drain capacitances, $C_{gs,ov}$ and $C_{gd,ov}$ respectively

$$C_{\text{min}} = C_{gs,ov} + C_{gd,ov} = 2WL C_{OX}$$

- $V_{G,SD} > V_{Th}$ → $C_{G,SD}$ is set to its maximum value, mostly the gate-to-channel $C_{gc}$ capacitance

$$C_{\text{max}} = C_{gc} = 2WL C_{OX}$$

**Basic idea:** take advantage of the non-linear feature of the MOS capacitor to dynamically change the gain of the charge sensitive amplifier with the input signal amplitude
Dynamic compression

Based on the non-linear feature of a MOSFET operated in inversion mode

- $|\Delta v_{out}| \ll V_{Th} \Rightarrow C_f = C_{min}, \text{Gain} = G_{he}$
- $|\Delta v_{out}| \gg V_{Th} \Rightarrow C_f = C_{max}, \text{Gain} = G_{le}$

Appropriate choice of $W$ and $L$ to configure the gain in the low and high energy regime, under the constraint set by the preamplifier output range

The MOS capacitance in inversion mode can be empirically modeled with the following equation with two fitting parameters \( \alpha \) and \( \beta \)

\[
C_{\text{MOS}}(V_{GS}) = \frac{C_{\text{max}} + C_{\text{min}}}{2} + \frac{C_{\text{max}} - C_{\text{min}}}{2} \tanh\left( \frac{V_{GS} - \beta}{\alpha} \right)
\]

\( \alpha \) and \( \beta \) depend on the device polarity and threshold voltage.

An incremental change \( dv_0 \) as a response to an infinitesimal injected charge \( dq \) will be given by

\[
dv_0 = \frac{dq}{C_{\text{MOS}}(v_0)} \Rightarrow \int dq = Q = \int C_{\text{MOS}}(v_0) dv_0
\]

By replacing \( C_{\text{MOS}} \) with its expression above we get

\[
Q = \frac{C_{\text{max}} + C_{\text{min}}}{2} V + \frac{C_{\text{max}} - C_{\text{min}}}{2} \ln\left( \cosh\left( \frac{V}{2} \right) \right) \Rightarrow V = f^{-1}(Q) = f^{-1}\left( n_{ph} E_{ph} \right)
\]

\( n_{ph} \) is the number of photons, \( E_{ph} \) is the photon energy, and the photon energy is 3.6 eV.
Modeling the inversion mode MOS cap

different MOS geometries

about 2 orders of magnitude in capacitance

almost 3 orders of magnitude in $V_{GS}$
CSA architecture

Reset network - for large input charge provides the current needed to restore voltage in A and B

Active folded cascode

Feedback MOS cap - gain can be adjusted for two different photon energies, 1 and 10 keV

Based on a White follower configuration

- Open loop DC gain: 60 dB
- GBP: 140 MHz
- Phase margin ($C_f=10$ pF): 52 deg
- Power consumption: 100 uW
CSA time response

[Diagram showing time response of CSA output voltage with labeled time intervals: ~20 ns and ~30 ns.]

L. Ratti, Pixel-embedded signal processing for the next generation detectors at the LHC and FELs
5th INFIERI International Summer School, Wuhan, China, May 12-26 2019
 CSA compressed trans-characteristic

- Low energy gain: $G_{he}=1 \text{ mv/phot}$
- High energy gain: $G_{le}=25 \text{ uV/phot}$

- Compression factor: $G_{he}/G_{le}=40$
- Full dynamic range of $10^4$ photons is covered
Flip capacitor filter

- events with a known repetition rate → time-variant shaping
- trapezoidal weighting function by feedback capacitor flipping
- performs correlated double sampling (CDS)
- 45 µW power dissipation


L. Ratti, Pixel-embedded signal processing for the next generation detectors at the LHC and FELs
5th INFIERI International Summer School, Wuhan, China, May 12-26 2019
Differential integrator

Can replace the transconductor + FCF

Baseline integration

\[ V_o(t) = -\frac{t}{RC} V_A \]

\[ V_o(\cdot) = -\frac{1}{RC} V_A^- \]

Signal integration

\[ V_o(t) = V_o(2) + \frac{t}{RC} V_A^+ \]

\[ V_o(3) = \frac{1}{RC} (V_A^+ - V_A^-) \]

Charge sensitivity

\[ G_Q = \frac{V_{out}}{Q} = \frac{1}{C_fRC} \]
10 bit interleaved SAR ADC

Two split capacitive DACs in a time interleaved structure; for each DAC

- Pre-charge during one sampling period
- Conversion during the subsequent period

No need for a dedicated stage to charge the DAC (~2.3 pF)

Avoid large current peaks due to fast capacitance charge

Sampling rate of 5 MHz, 10 bit conversion in 11 clock periods (each period ~18 ns)
Larger ENC in the case of NMOS feedback capacitor

- larger stray capacitance at the preamplifier input
- the PMOS capacitance, being integrated in an N-well, is less sensitive to noise propagating through the substrate
Full channel trans-characteristic

![Graph showing full channel trans-characteristic with ADC count on the y-axis and number of photons on the x-axis. The graph includes two curves: one for 50 fF injection capacitance and another for 500 fF injection capacitance.]

- 50 fF injection capacitance
- 500 fF injection capacitance

1 keV mode operation
1 μs processing period
(τ=250 ns)

large injection capacitance to test the channel response with large amounts of charge at the input
Analog layer in monolithic technology

- Pixel array
- Top guard ring
- Peripheral electronics
- Wells with readout electronics
- Charge collection electrodes
- High resistivity substrate
- Backside contact
- Entrance window
- Bottom guard rings

**Standard CMOS technology with process modifications**

- High-resistivity substrate
- Custom implantations
- Backside processing
A two-layer X-ray sensor

- **Two-tier detector**
  - analog layer - sensor, front-end (110 nm CMOS)
  - digital layer - ADC, memory and fast readout (65 nm CMOS)

- **Low noise + high frame readout rate**

- **Small pitch bump bonding with IZM**

- **Development required on sensor (trenches for slim edge design, backside processing for low energy photons)**
#### Conclusion

Pixel detector requirements for next generation experiments at high luminosity colliders and X-ray sources are extremely challenging:

- **High granularity →** small room for electronic circuits
- **Large amount of data and high hit rate →** high speed, on-chip memory, data reduction
- **Large area chip →** integration issues, analog-digital cohabitation problems
- **Radiation hardness**

More functionalities (data conversion and storage, channel calibration, parameter programmability) need to be built into the readout chip to satisfy the specifications.

Classical analog problems (low noise signal amplification and shaping, noise, threshold dispersion) require new solutions to comply with new demanding specifications and provide an exciting challenge for integrated circuit designers.

The evolution of microelectronic technologies makes it possible to follow the trend and, actually, has a role in enhancing it, leaving room for creativity and continuous innovation.