Highly-Digital Approaches to Front-End Design

Soumyajit Mandal[1], G. W. Deptuch[1], J. Fried[1], D. Gorni[1], S. R. Kharel[1], P. Mukim[1], G. Pinaroli[1], Y. Ren[2], A. Williams[1]

[1]Instrumentation Division, Brookhaven National Laboratory
[2]Computational Science Initiative, Brookhaven National Laboratory

smandal@bnl.gov

XIII Front-End Electronics Workshop
Paris, France — May 18–22, 2026

Brookhaven National Laboratory

Outline

Part 1: Overview & Landscape

- Why go digital early?
- Technology scaling as enabler
- Streaming readout as driver
- State of the field

Part 2: Results from BNL

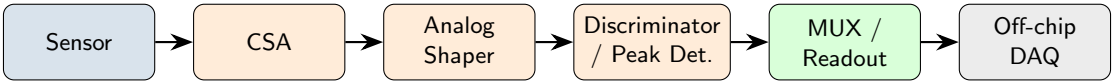
- The AIFE65P1 architecture
- Analog front-end & hybrid ADC
- Model-based DSP design flow
- On-chip neural inference (MLP)
- Measured performance
- Outlook

Update of the invited talk given at the XII FEE Workshop, Torino, June 2023.

The Data Challenge

- Modern and next-generation experiments produce data volumes that are difficult to handle, move, and store.
  - DUNE far-end detector: 9.2 Tbps of waveform data (24k FE ASICs, 12k links).
  - ATLAS LAr calorimeter (HL-LHC): $\sim$345 Tbps across 1524 FEB2 boards.
  - Future colliders (FCC, EIC) will exceed these rates further.
- Key insight: Embedding digital signal processing (DSP) as close as possible to the sensor has high potential for **transformative improvements** in detector readout.
- Detector systems can benefit by converting analog signals to digital *as early as possible* to maintain high SNR and enable flexible, programmable signal processing.

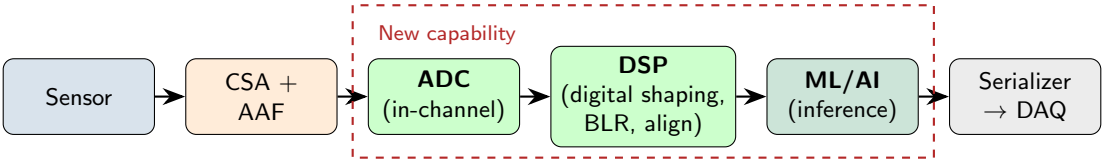
The Traditional Paradigm



All critical signal processing happens in the analog domain

- Classic front-end: CSA → high-order analog shaper → discriminator/peak detector → triggered readout.
- Feature extraction (amplitude, time of arrival) performed using **analog** circuits.
- Well-optimized for specific applications, but:
  - Fixed transfer functions — limited flexibility.
  - Sensitive to PVT variations and mismatch.
  - Cannot be reconfigured post-fabrication.

The Highly-Digital Paradigm



- Minimal analog: CSA + anti-alias filter (AAF) only for charge integration and bandwidth limiting.
- **Early digitization**: In-channel ADC converts the signal to digital as soon as possible.
- All pulse shaping, baseline restoration, feature extraction, and even ML inference done digitally.
- Benefits: programmability, reproducibility, no analog drift, post-fabrication tunability.
- **"Highly digital" also means:** Using digital-like circuit topologies (inverter-based amplifiers, ring amplifiers) and extensive digital calibration/tuning within the analog blocks themselves.

Enabler 1: Technology Scaling

- What was impractical in 250 nm is routine in 65 nm and below.
- Key scaling benefits for digital FE:
  - $\sim$4–5$\times$ reduction in **digital area and power** moving from 65 nm to 28 nm.
  - In-channel SAR ADCs ($\leq$12 bit, 40–90 MS/s) now fit within $\sim$200 μm channel pitch.
  - DSP blocks (FIR, BLR, MLP) become area-feasible.
- ADC architecture trends:
  - Ultra-low-power SAR (e.g., SALT: 6-bit, 40 MS/s, $<$0.5 mW/ch).
  - Hybrid SAR–$\Delta\Sigma$ and SAR–digital-slope for higher resolution.
  - Pipeline-SAR for highest speed.

Implication

The per-channel cost of digitization + DSP is now comparable to that of a traditional high-order analog shaper + discriminator, especially in 65 nm and below.

Key ASICs in the field

SALT	LHCb UT
HGCROC	CMS HGCal
COLUTA	ATLAS LAr
SALSA	EIC MPGDs
VMM3a	ATLAS NSW
Smart Pixels	FNAL

Enabler 2: Streaming Readout

- **Paradigm shift:** From triggered → continuous (streaming) readout.
  - Channels transmit data continuously; no hardware trigger decision.
  - Event building and selection performed in software (back-end).
  - Adopted/planned for EIC (ePIC), DUNE, and many future experiments.
- Streaming readout *demands* highly-digital front-ends because:
  - Continuous data flow ⇒ need on-chip data reduction to stay within bandwidth.
  - Programmable digital processing enables **adaptive** data reduction.
  - Two operational modes become possible:
    1. **Full streaming:** Filtered/decimated waveform samples transmitted.
    2. **Feature streaming:** Only extracted features (amplitude, TOA, particle ID) transmitted.
- This is fundamentally different from traditional triggered readout where the analog chain *is* the data reduction mechanism.

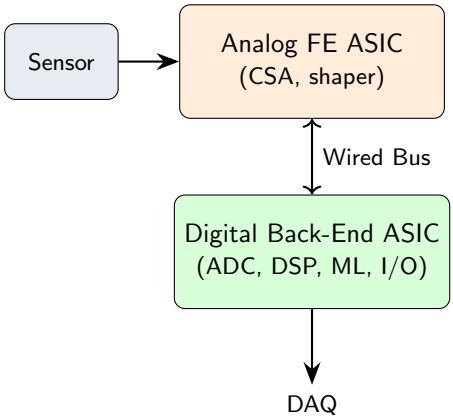
State of the Field: Common Architectures

Architecture	Digitization	DSP	Examples
Binary / ToT	Comparator only	None (analog)	Medipix/Timepix, VMM3a
Analog sampling + off-chip ADC	External ADC	Off-chip (FPGA)	SCA-based (SAMPA, DRS4)
In-channel ADC + triggered digital output	In-channel SAR/Pipeline	Minimal (MUX, zero-suppression)	SALT (LHCb), COLUTA (ATLAS LAr), HGCROC (CMS)
In-channel ADC + DSP + ML	**In-channel hybrid ADC**	**FIR, BLR, waveform alignment, MLP**	**This work (AIFE65P1)**

- The field is progressively moving toward earlier digitization and richer on-chip processing.
- Open question: How much intelligence can we embed per-channel under realistic power/area budgets?

Advantages of a Two-Chip Approach

- Separate the high-sensitivity **analog front-end ASIC** from the **mixed-signal/digital back-end ASIC**.
- Advantages:
  - Optimal allocation of functionality, fewer risks.
  - Independent, parallel development paths → faster development.
  - Analog ASIC can be simpler, smaller, and tuned for noise.
  - Digital ASIC benefits from process scaling.
  - Managed complexity including packaging.
- Digital back-end ASIC is a **programmable** design that can be paired with different analog front-ends for different applications (e.g., X-ray spectroscopy, noble-liquid TPCs).



A Multi-Functional Streaming Readout ASIC with Channel-Embedded DSP and Neural Inference

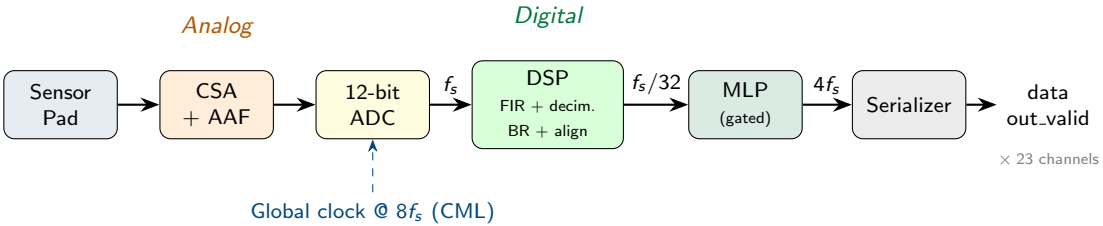
Context:

- Evolution of the full-field fluorescence (3FI) project for X-ray spectroscopy (Si drift detectors, Ge strip detectors).
- Also potentially applicable to noble-liquid TPCs (DUNE) and calorimeters (FCC-ee).

What's new since FEE 2023:

- Fabricated 23-channel ASIC in 65 nm CMOS technology.
- Full measurement results: AFE, ADC, DSP, and **on-chip MLP inference**.
- Paper submitted to IEEE TCAS-I.

Chip Architecture: Per-Channel Signal Chain



- 23 independent channels, 200 μm pitch (direct wirebonding to strip sensors).
- Two operating modes: **full streaming** (filtered waveforms) and **feature streaming** (MLP outputs).
- Deterministic one-sample-per-cycle streaming constraint at ADC rate.
- MLP is **clock-gated** (activated only when a peak is detected) $\rightarrow$ saves dynamic power.

Analog Front-End (AFE)

- **Charge-sensitive amplifier (CSA):**
  - Two-stage design with self-cascoded FET (SCFET)-based pole–zero cancellation.
  - Configurable polarity (electrons/holes), baseline trimming.
  - On-chip charge injector for calibration and testing.
- **Anti-aliasing filter (AAF):**
  - Third-order OTA-based active filter.
  - Limits bandwidth prior to ADC sampling.
  - *Final pulse shaping is done digitally* (programmable FIR).
- **Single-ended to differential converter (SEDC):**
  - Drives the fully-differential ADC input buffer.
  - Buffer input cap. $\approx$300 fF (10$\times$ less than the ADC).
- All bias currents generated locally via constant-G_m references and DACs — fully self-contained.

Design Philosophy

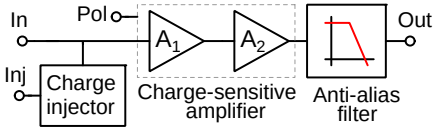
Minimize analog complexity:
CSA integrates charge,
AAF limits bandwidth.

Everything else is digital.

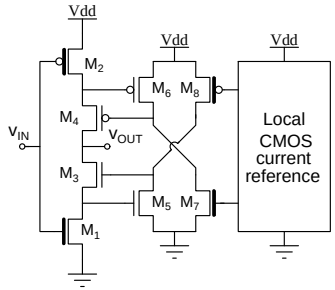
Key Specifications

Gate cap.	$\approx$35 fF
Amp. bias	$\approx$50 μA
ENC (no det.)	$\approx$17 e^-
Peaking time	$\sim$140 ns
Total power	$\approx$0.27 mW/ch

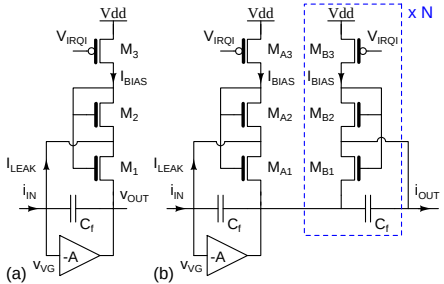
AFE Circuit Details: Digital-Friendly Analog Design



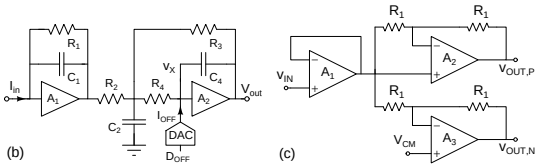
Block diagram of the analog front-end (AFE).



CMOS inverter-based gain stage with active cascodes.
Thick-oxide transistors are shown with heavy gate lines.



Full CSA with SCFET-based feedback, (a) basic circuit, and (b) including pole–zero cancellation.



(a) Anti-aliasing filter. (b) SEDC.

12-bit Hybrid SAR–Digital-Slope ADC

- **Architecture:** 8-bit SAR (coarse) + 5-bit DS (fine), 1 redundant bit → 12-bit output.
- **Successive Approximation (SAR) stage:**
  - Split capacitive DAC ($C_{\text{unit}} = 20\,\text{fF}$).
  - Merge-and-split switching for energy efficiency (6× better than binary).
  - Asynchronous state machine (no high-freq. clock, up to 2× faster settling).
- **Digital-slope (DS) stage:**
  - Self-timed delay line + 32-step cap array.
  - Embedded within SAR DAC ($C = C_{\text{unit}}/8$) → no inter-stage gain error.
  - Low-power continuous-time comparator.
- Charge neutralization reduces V_{REF} ripple.
- Foreground calibration of 4 MSB caps + offset cancellation of both comparators.

Why Hybrid SAR–DS?

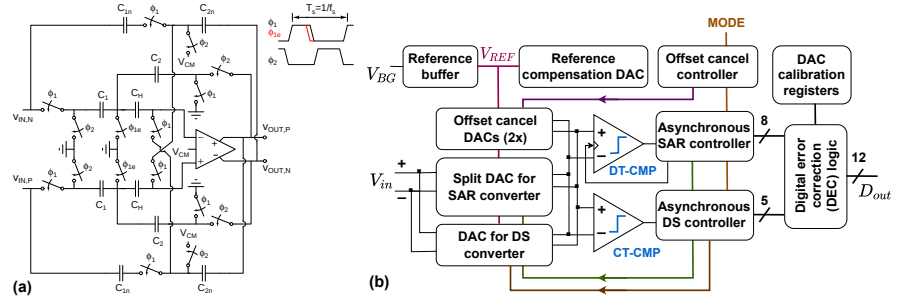
Pure SAR at 12-bit resolution ⇒ comparator noise and DAC settling challenges.

DS stage resolves fine bits in the *time domain* ⇒ relaxes comparator noise and reduces power.

ADC Performance

Resolution	12 bits
ENOB (sim.)	≈11 bits @ 50 MS/s
Core power	≈0.34 mW
Ref. buffer	≈1.15 mW
FoM (Walden)	7.8 fJ/conv-step
Area	550×200 μm^2

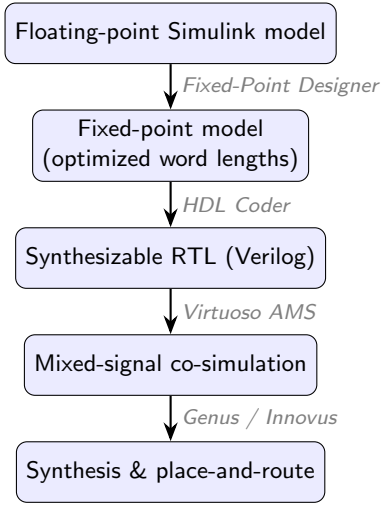
ADC Circuit Details: Digitally-Assisted Architecture



(a) Ring amplifier-based ADC driver. (b) Block diagram of the hybrid SAR–digital-slope ADC. (c) Discrete-time comparator for the SAR with MOS-based input capacitance neutralization.

- **Ring amplifier (RA) ADC driver:** Fully-differential, self-biased — *digital-like* topology.
  - CMOS inverter-based gain stages provide low noise and rail-to-rail output.
  - 5-bit DAC adjusts bias current $I_{B,RA}$ for PVT compensation.
- **Digital calibration and tuning:**
  - Two-bit time delay control within asynchronous SAR state machine.
  - Coarse + fine reference V_{REF} trimming registers.

Model-Based HLS Design Flow for the DSP



- **Key idea:** Automated word-length optimization + RTL generation from MATLAB/Simulink.
- Constrained optimization: minimize total bit width subject to accuracy constraints.
- Single-clock architecture (no time-multiplexing) $\Rightarrow$ simplified timing closure.
- Co-simulation with transistor-level analog blocks for end-to-end verification.

Per-Channel DSP Pipeline

1. Programmable FIR Shaping Filter

- Linear-phase, $N = 17$ taps, 9 distinct coefficients (symmetric).
- Replaces analog shaper $\rightarrow$ reconfigurable.
- Operates at full ADC rate f_s.

2. Decimation ($\times$4)

- Polyphase filter ($K = 32$, $A_{\text{stop}} = 40\,\text{dB}$).
- Reduces rate to $f_s/4 \rightarrow$ lowers DSP power.

3. Baseline Restorer

- EMA-based baseline estimator with separate learning rates for mean and variance.
- Suppresses low-frequency drift without broadening noise floor.

4. Waveform Aligner

- Peak-centering algorithm on sliding 16-sample windows.
- Extracts 9-sample snippet centered on detected peak.
- Amplitude gate: signal must exceed $\gamma \cdot \hat{\sigma}$.
- Increases peak-to-min σ ratio.

Data Rate Reduction

$$f_s \xrightarrow{\text{FIR}} f_s \xrightarrow{\div 4} f_s/4 \xrightarrow{\text{window}} f_s/32$$

Output inference rate $= f_s/32$
($0.625\,\text{MS/s}$ @ $f_s = 20\,\text{MS/s}$)

On-Chip Neural Inference: The MLP

- **Architecture:** Fully-connected two-layer MLP, two outputs ($9 \to 16 \to 1$ or $9 \to 16 \to 2$).
- Either **regression** (amplitude estimation) or **classification** (pulse-shape discrimination).
- **Quantization:** Q1.6 weights (8-bit), 12-bit activations, stochastic rounding.
- **Training procedure:**
  1. Float pre-training (400 epochs, Adam optimizer).
  2. Quantization-Aware Training (QAT, 400 epochs).
- **Clock gating:** MLP activates only when `pkFound` is asserted $\to$ dynamic power savings.
- **Dual-output encoding** for regression: coarse + fine outputs to increase dynamic range over Q1.6.

Why MLP over classic algorithms?

Matched filter (MF): requires FFT/IFFT — complex hardware.

Non-linear recursive filter (NRF): requires LUTs for exp/log.

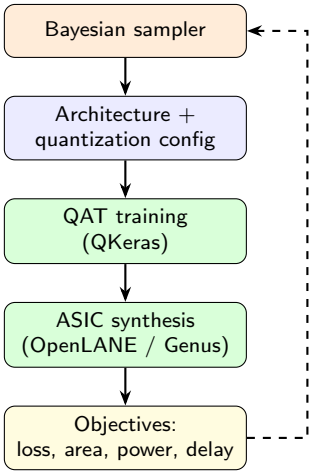
MLP: Only multiply-accumulate + ReLU. Simple, flexible, retrainable.

Network Parameters

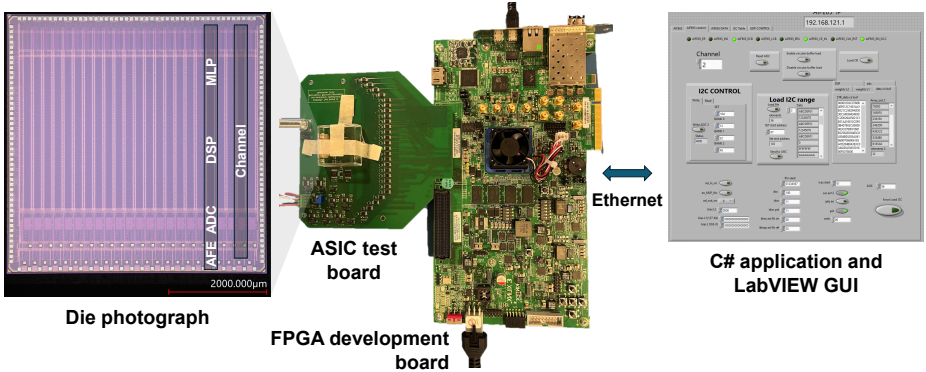
Input width	9 samples
Hidden layer	16 neurons
Output	1 (class) or 2 (regr.)
Total params	194
Weight format	Q1.6 (8-bit)
Activation	ReLU

Bayesian Co-Design of Neural Network and ASIC

- **Problem:** How to choose MLP architecture (depth, width) and quantization jointly with ASIC synthesis strategy?
  - Co-design space exceeds **2 billion** configurations.
  - Competing objectives: accuracy vs. area vs. power vs. delay.
- **Our approach** [Kharel et al., arXiv:2407.14560]:
  - Multi-objective Bayesian optimization (using Optuna) with logic synthesis in the loop.
  - Each trial: QAT training → automatic RTL generation → logic synthesis → report area, power, delay.
  - Uses *actual synthesis metrics*, not theoretical proxies.
- **Key finding:** Theory-guided optimization saturates early with poor Pareto diversity; synthesis-guided search finds significantly better solutions.
- Selected design (9→16→2, Q1.6) is Pareto-optimal for our area/power/accuracy constraints.



Evaluation Platform and Die Photograph



- 65 nm CMOS, $5 \times 5\,\mathrm{mm}^2$ die, 23 channels at 200 μm pitch.
- CML clock input at $8f_s$; all internal clocks derived on-chip by frequency division.
- FPGA captures LVDS data $\rightarrow$ UDP packets $\rightarrow$ real-time display + offline analysis.
- Dedicated pad and HV bias for direct wirebonding to silicon strip sensors.

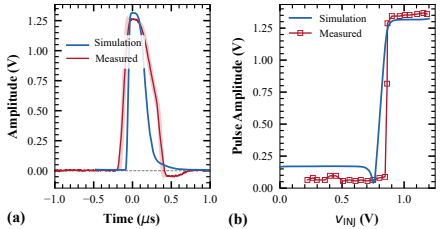
Measured Results: AFE and Noise

AFE Pulse Shape

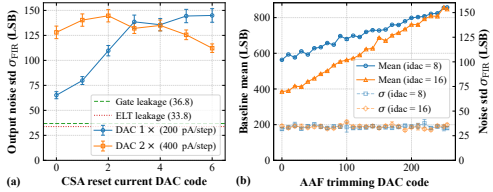
- Measured pulse shape closely tracks post-layout simulation.
- FWHM: 418 ns (measured) vs. 226 ns (sim.) — difference attributed to AAF bandwidth/slew-rate limitations.
- Peak amplitude within 3.6% of simulation.

AFE Noise and Dynamic Range

- No detector connected: $C_D \approx 85$ fF.
- **Meas. ENC** $\approx$ **17 e$^-$** at low reset bias.
- Consistent with series-noise-limited CSA.
- Max. signal $\approx$2,900 e$^-$ (10.4 keV in Si).
- Dynamic range $\approx$ 45 dB.



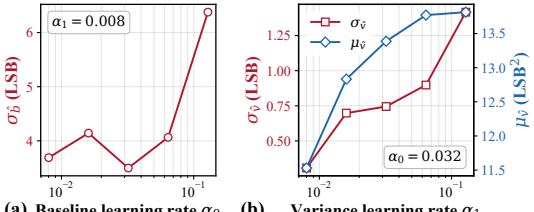
(a) AFE pulse shape: sim. vs. measured
(b) Injected pulse amplitude: sim. vs. measured



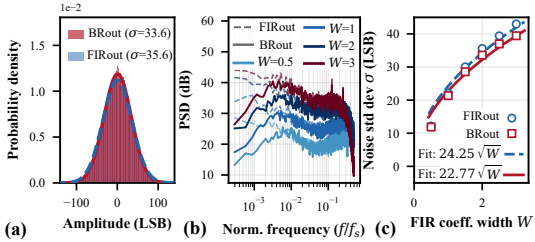
FIR output: (a) quiescent noise and (b) baseline tunability (Gaussian coefficients: $W = 2$, peak $= 1$)

Measured Results: Digital Baseline Restorer

- Exponential moving average (EMA)-based digital baseline restorer (BR) with separate learning-rate parameters for baseline and variance estimators.
- **Key results:**
  - Suppresses low-frequency drift without broadening noise floor.
  - Learning-rate parameters independently control baseline and variance estimators as predicted by EMA model.
- Operates in streaming mode at a rate of $f_s/4$ after decimation.



BR output vs. learning rate parameters

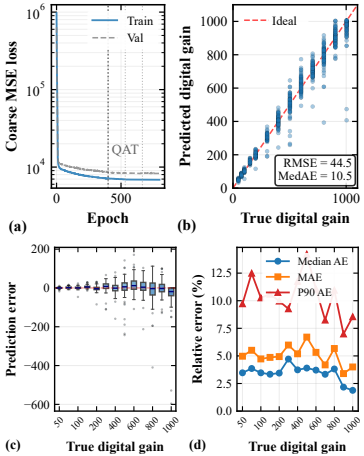


Noise characterization of FIR and BR outputs

Measured Results: MLP Regression

Amplitude estimation (pulse width $W = 1$):

- 9→16→2 network, Q1.6 weights.
- Training on data from on-chip charge injector (12 gain levels × 10 streams).
- Dual-output (coarse + fine) encoding.
- **Results (quantized, on held-out set):**
  - RMSE = 44.5 (digital gain units)
  - **Median absolute error ≈ 1.4% of full scale**
  - Negligible degradation after QAT.
- Consistent performance across pulse shapes (Gaussian coefficients, $W = 1, 2, 3$): median absolute error (AE) ≈ 2–4% in all cases.



Regression for $W = 1$ with quantized weights.
(a) Training dynamics. (b) Predicted vs. true gain.
(c) Error distribution. (d) Relative error vs. gain.

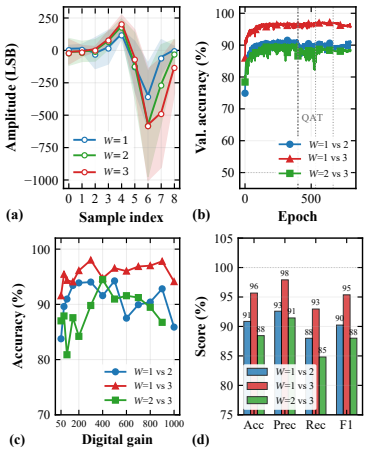
Measured Results: MLP Classification

Pairwise pulse-width classification:

- Three binary classifiers: $W = 1$ vs. 2, $W = 1$ vs. 3, $W = 2$ vs. 3.
- Same 9→16→1 architecture, Q1.6 weights.
- **Results (quantized):**

Pair	Acc.	Prec.	Rec.	F1
$W = 1$ vs. 2	90.1%	93%	88%	89.3%
$W = 1$ vs. 3	96.2%	98%	93%	96.0%
$W = 2$ vs. 3	88.5%	91%	85%	87.9%

- Quantization penalty $\leq 1.6\%$.
- Gain-invariant training prevents amplitude-induced bias.



Pairwise classification results. (a) Pulse shapes ($W = 1, 2, 3$) (b) Accuracy vs. epoch (c) Per-gain accuracy (d) Summary metrics.

Performance Summary

Chip Overview	
Technology / die area / channels	65 nm / 5×5 mm^2 / 23
Sampling and Throughput	
ADC sampling rate (f_s)	20 MS/s (nominal)
ADC resolution / ENOB (sim.)	12-bit / 11 bits @ 50 MS/s
Post-decimation rate	$f_s/4 = 5$ MS/s
MLP inference rate	$f_s/32 = 0.625$ MS/s
Noise Performance (No Detector)	
Gate capacitance / peaking time	≈35 fF / ≈140 ns
Measured ENC ($C_d = 0$)	**≈17 e$^-$**
Dynamic range	≈45 dB
Power Consumption per Channel	
AFE / ADC core / ref. buffer	0.27 / 0.34 / 1.15 mW
DSP (incl. FIR + MLP)	≈3.6 mW
Total per channel	**≈5.4 mW**

Key takeaways:

- Full signal chain — charge integration to neural inference — implemented **per channel** in 65 nm CMOS.
- ADC Walden FoM of **7.8 fJ/conv-step** (core only) competitive with state-of-the-art in-channel converters.
- ENC of 17 e$^-$ at $C_d = 0$ demonstrates low-noise operation of inverter-based CSA with SCFET reset.
- DSP + MLP dominate power/area (≈67% of core) — optimize by process scaling, SRAM-based weight storage.
- **All parameters programmable** via I^{2}C: FIR coefficients, BR learning rates, MLP weights, AFE/ADC bias trims.

Conclusions

- Highly-digital front-end architectures with **early digitization, programmable DSP, and on-chip ML** are now practical in 65 nm CMOS.
- The benefits will continue to increase with process scaling (e.g., $\sim$4–5$\times$ area/power reduction moving to 28 nm).
- Our 23-channel streaming readout ASIC demonstrates:
  - Low-noise AFE: ENC $\approx 17\,e^-$ (no detector).
  - Energy-efficient 12-bit hybrid SAR–digital-slope ADC (7.8 fJ/conv-step).
  - Full DSP chain: FIR shaping, decimation, baseline restoration, waveform alignment.
  - **First demonstration of per-channel MLP inference** in a multi-channel strip readout ASIC.
  - Amplitude regression: 2–4% median absolute error. Classification: $>$88% accuracy.
- Total power: $\approx$5.4 mW/channel at 20 MS/s.
- Two-chip solutions (analog front-end + digital back-end) are favored long-term for flexibility and risk management.

Outlook and Future Work

- **Detector characterization:**
  - Direct wirebonding to silicon strip sensors — evaluate noise, linearity, and MLP inference under realistic signal/background conditions.
- **Architecture improvements:**
  - Deeper/wider MLP topologies for multi-class or multi-parameter inference enabled by denser memory (SRAM macros: $\sim$10$\times$ denser than D-latches).
  - Exploration of convolutional layers for improved feature extraction.
  - Event-driven spiking neural networks (SNNs) with multi-channel inputs and compressive encoders for spatiotemporal pattern recognition.
- **Process scaling:**
  - Port to 28 nm for $\sim$4–5$\times$ improvement in digital area/power.
  - Enables richer on-chip intelligence within same power envelope.
- **Application broadening:**
  - Noble-liquid TPCs (DUNE FD3/FD4): waveform denoising, pile-up resolution.
  - X-ray spectroscopy (synchrotron microprobes): high-rate amplitude estimation.

References

1. S. Mandal et al., "A 23-Channel Streaming Readout ASIC...," *IEEE Trans. Circuits Syst. I*, submitted, 2026.
2. G. Deptuch et al. "Pixel detectors with built-in signal processing and bandwidth-efficient data transmission." *10th International Workshop on Semiconductor Pixel Detectors for Particles and Imaging (PIXEL)*, 2023.
3. S. R. Kharel et al., "Automated and Holistic Co-design of Neural Networks and ASICs...," arXiv:2407.14560, 2024.
4. S. Mandal et al., "Highly-Digital Approach Towards FE Design," invited talk, XII FEE Workshop, Torino, 2023.
5. S. Miryala et al., "Waveform processing using neural network algorithms...," *JINST*, vol. 17, C01039, 2022.
6. D. Gorni et al., "Integration of EDWARD readout architecture...," *JINST*, vol. 19, C04035, 2024.
7. S. Mandal et al., "Real-time feature extraction for readout ASICs...," IEEE NSS/MIC/RTSD, 2023.
8. C. Abellan Beteta et al. (SALT), *Sensors*, vol. 22, no. 1, p. 107, 2021.
9. D. Neyret et al. (SALSA), *JINST*, vol. 20, C06057, 2025.
10. G. De Geronimo et al. (VMM3a), *IEEE Trans. Nucl. Sci.*, vol. 69, no. 4, pp. 976–985, 2022.
11. F. Bouyjou et al. (HGCROC3), *JINST*, vol. 17, C03015, 2022.
12. G. Pinheiro Matos (COLUTA), *JINST*, vol. 18, C02018, 2023.
13. J. Yoo et al. (Smart Pixels), *Mach. Learn.: Sci. Technol.*, vol. 5, no. 3, 035047, 2024.
14. H. Hernandez et al. (SAMPA), *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, pp. 2686–2697, 2019.

Thank You!

Questions?

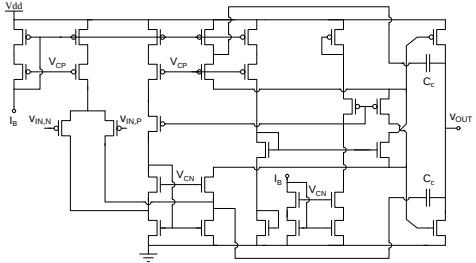
Contact: smandal@bnl.gov

This work was supported by the U.S. Department of Energy, Office of Science.
Brookhaven National Laboratory is operated under Contract No. DE-SC0012704.

Backup Slides

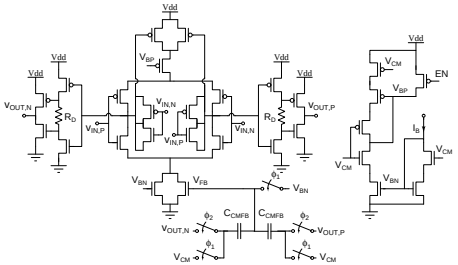
Amplifier Designs

Op-amp used in the SEDC



- Two-stage design: PMOS-input folded-cascode 1st stage + class-AB output stage.
- Operates at 1.8 V for rail-to-rail input range in the 1.5 V AFE domain.
- Indirect feedback compensation to reduce power → GBW = 550 MHz at $I_B = 60\,\mu$A.
- 5-bit DAC adjusts I_B for PVT tuning; total output noise ≈200 μV_{rms} (0.28× LSB).

Ring amplifier used in the ADC buffer



- Three inverter-based gain stages → high gain, rail-to-rail swing, compact layout.
- Differential input (1st stage) for CMRR/PSRR; cross-coupled MOS caps for input neutralization.
- Series resistor R_D in 2nd stage sets adaptive dead zone V_{DZ} for large-signal stability.
- Swited-capacitor CMFB on output stage.
- $I_{B,RA} = 10\,\mu$A (5-bit DAC tunable); output noise = 510 μV_{rms} (0.70× LSB) at 40 MS/s.

Algorithm 1: Noise Estimation and Digital Baseline Removal

Algorithm 1: Noise estimation and digital baseline removal

Require: Input window $\mathbf{x}$, threshold coefficient γ, update rates α_b, α_v, enable flag

Compute window mean $\mu = \frac{1}{N}\sum x[n]$
Compute variance $v = \frac{1}{N}\sum (x[n] - \mu)^2$
Update raw variance estimate $v_{\text{raw}} \leftarrow (1 - \alpha_v)\, v_{\text{raw}} + \alpha_v\, v$
Compute peak-to-peak metric $v_{\text{pp}} \leftarrow (\max \mathbf{x} - \min \mathbf{x})^2$
$\texttt{sig\_present} \leftarrow (v_{\text{pp}} > \gamma\, v_{\text{raw}})$
if not `sig_present` **then**
  Update baseline estimate $b \leftarrow (1 - \alpha_b)\, b + \alpha_b\, \mu$
  Update filtered variance estimate $v_{\text{est}} \leftarrow (1 - \alpha_v)\, v_{\text{est}} + \alpha_v\, v$
end if
if enable **then**
  Output $\mathbf{y} = \mathbf{x} - b$
else
  Output $\mathbf{y} = \mathbf{x}$
end if

- **Key idea:** Exponential moving average (EMA)-based feedforward baseline tracker with variance-gated updates.
- Baseline recursion is *frozen* when signal is detected ($v_{\text{pp}} > \gamma\, v_{\text{raw}}$), preventing pulse energy from corrupting the baseline estimate.
- Both α_b (baseline learning rate) and γ (detection threshold) are programmable per channel.

Algorithm 2: Waveform Alignment by Peak-Centering

Algorithm 2: Waveform alignment by peak-centering

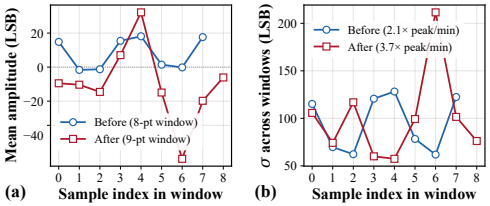
Require: Half-windows $\mathbf{x}_1$, $\mathbf{x}_2$ (8 samples each), polarity flag `pol`, noise estimate v, threshold coefficient γ, enable flag

Set $N \leftarrow 16$, $N/2 \leftarrow 8$, $N/4 \leftarrow 4$; define search indices $k \in \{5, \ldots, 12\}$
Form buffer $\mathbf{x}[1\!:\!N] \leftarrow [\mathbf{x}_1 \;\; \mathbf{x}_2]$
if `pol` indicates negative pulses **then** $\mathbf{x} \leftarrow -\mathbf{x}$ **end if**
Find candidate maximum in central region: $(p,\, \ell_c) \leftarrow \max_{k \in \{5,\ldots,12\}} \mathbf{x}[k]$
Amplitude gate: $\texttt{sig\_found} \leftarrow (p^2 > \gamma\, v)$
if enable **then**
  Convert to absolute index: $\ell \leftarrow \ell_c + N/4$
  Extract 9-sample sub-window: $w[m] \leftarrow \mathbf{x}[\ell_c + m],\ m = 0, \ldots, 8$
  Local-maximum test: $\texttt{is\_peak} \leftarrow (p > \mathbf{x}[\ell - 1]) \;\wedge\; (p > \mathbf{x}[\ell + 1])$
  $\texttt{pkFound} \leftarrow \texttt{sig\_found} \wedge \texttt{is\_peak}$
else
  Output unaligned central segment: $\mathbf{w} \leftarrow \mathbf{x}[5\!:\!13]$; $\texttt{pkFound} \leftarrow \texttt{false}$
end if

- **Key idea:** Center pulses within windows to minimize variability due to asynchronous arrivals.
- Two half-windows concatenated into a 16-sample buffer ($N = 16$, fixed by data rates).
- Peak search restricted to central 50% to avoid edge effects.
- Noise estimate v and threshold γ *reused* from baseline-removal block $\rightarrow$ no extra hardware.
- `pkFound` flag gates the MLP clock $\rightarrow$ inference only on candidate events.

Waveform Aligner Validation

- Pulse waveforms recorded **before** the aligner (BR output, 8-sample windows) and **after** it (DSP output, 9-sample windows).
- **Assessment method:** Average all windows at each sample position.
  - If pulses always arrive at the same position ⇒ average preserves the sharp pulse shape.
  - If they jitter ⇒ average approaches zero.
- **Experimental results:**
  (a) Before alignment: mean amplitude nearly flat ($\approx \pm 20$ LSB). After: well-resolved doublet at positions 4 and 6, amplitude $\sim 2\times$ larger.
  (b) Per-position σ: peak-to-min ratio increases from **2.1**$\times$ (before) to **3.7**$\times$ (after).
- The doublet shape arises from the two edges of each 1 μs charge-injection pulse.



(a) Mean amplitude at each window position, averaged over all windows. (b) Per-position standard deviation.
Red = after alignment (9-pt window);
Blue = before alignment (8-pt window).

Evolution: FEE 2023 → FEE 2026

	FEE XII (2023)	**FEE XIII (2026)**
Chip status	Concept + standalone prototypes (CSA, ADC, CML driver)	**Fabricated and measured the 23-channel ASIC**
ADC	Standalone test chip (SINAD limited by clock feedthrough)	Integrated in-channel, self-contained with on-chip ref/bias
DSP	Verilog blocks synthesized separately (FIR, aligner, BLH)	**Full streaming pipeline verified in silicon**
ML inference	Synthesis results only (area, power estimates)	**Measured regression and classification on-chip**
Training methodology	Simulated data	**On-chip data** with quantization-aware training