Neuromorphic Computing Based Advances in Calorimetry

and the design of the next generation of calorimeters

2nd Jornadas del ICTEA, Oviedo

Pietro Vischia
pietro.vischia@cern.ch
vischia.github.io

Universidad de Oviedo
ICTEA
Ministerio de Ciencia, Innovación y Universidades
Financiado por la Unión Europea NextGenerationEU
Plan de Recuperación, Transformación y Resiliencia
Agencia Estatal de Investigación

If you are reading this as a web page: have fun! If you are reading this as a PDF: please visit

https://www.hep.uniovi.es/vischia/persistent/2025-06-19_NeuromorphicReadoutAtSecond.html

to get the version with working animations.

NeuroMODE

- **New investigation line I opened here in 2023** with my Ramón y Cajal (RYC2021 "NeuroMODE").
  - **11 bachelor, 2 master's, 1 PhD theses so far**
  - Several international internships
- At the intersection of **neurosciences, computer sciences, and physics**
  - Neuromorphic computing
  - Quantum machine learning
  - Experiment design (*if you can build it, we can optimize it.* So far HEP, muon tomography, astrophysics, nuclear physics)
-

Latest publication highlights

- M. Pereira Martínez, X. Cid Vidal, Pietro Vischia, "Automatic Optimization of a Parallel-Plate Avalanche Counter with Optical Readout", Particles 2025, 8(1), 26 (10.3390/particles8010026, 2503.02538)
- P. Vischia, 2025, "AI-assisted design of experiments at the frontiers of computation: methods and new perspectives", Proceedings of Science 476, ICHEP2024, 1022 (2025) (10.22323/1.476.1022)
- S. Sánchez Cruz, M. Kolosova, C. Ramón Álvarez, G. Petrucciani, P. Vischia, 2024, "Equivariant neural networks for robust CP observables", Phys. Rev. D 110, 096023 (10.1103/PhysRevD.110.096023);
- T. Dorigo et al. (P.V.), "Artificial Intelligence in Science and Society: The Vision of USERN", IEEE Access 2025 (13) 15993-16054, 10.1109/ACCESS.2025.3529357

Latest preprints

- E. Lupi et al. (P.V.), "Neuromorphic Readout for Hadron Calorimeters", submitted to Particles, 2502.12693
- A. De Vita et al. (P.V.), "Hadron Identification Prospects With Granular Calorimeters", submitted to Particles, 2502.10817
- K. Schmidt et al. (P.V.), "End-to-End Detector Optimization with Diffusion models: A Case Study in Sampling Calorimeters", submitted to Particles, 2502.02152
- K. G. Barman et al. (P.V.), "Large Physics Models: Towards a collaborative approach with Large Language Models and Foundation Models", submitted to Nature physics roadmap, 2501.05382

NeuroMODE

- **New investigation line I opened here in 2023** with my Ramón y Cajal (RYC2021 "NeuroMODE").
  - **11 bachelor, 2 master's, 1 PhD theses so far**
  - Several international internships
- At the intersection of **neurosciences, computer sciences, and physics**
  - Neuromorphic computing
  - Quantum machine learning
  - Experiment design (*if you can build it, we can optimize it.* So far HEP, muon tomography, astrophysics, nuclear physics)
- **One COST action just funded**: CA24146, "Machine Learning and Quantum Computing for Future Colliders" (MLQC4FC)

Latest publication highlights

- M. Pereira Martínez, X. Cid Vidal, Pietro Vischia, "Automatic Optimization of a Parallel-Plate Avalanche Counter with Optical Readout", Particles 2025, 8(1), 26 (10.3390/particles8010026, 2503.02538)
- P. Vischia, 2025, "AI-assisted design of experiments at the frontiers of computation: methods and new perspectives", Proceedings of Science 476, ICHEP2024, 1022 (2025) (10.22323/1.476.1022)
- S. Sánchez Cruz, M. Kolosova, C. Ramón Álvarez, G. Petrucciani, P. Vischia, 2024, "Equivariant neural networks for robust CP observables", Phys. Rev. D 110, 096023 (10.1103/PhysRevD.110.096023);
- T. Dorigo et al. (P.V.), "Artificial Intelligence in Science and Society: The Vision of USERN", IEEE Access 2025 (13) 15993-16054, 10.1109/ACCESS.2025.3529357

Latest preprints

- E. Lupi et al. (P.V.), "Neuromorphic Readout for Hadron Calorimeters", submitted to Particles, 2502.12693
- A. De Vita et al. (P.V.), "Hadron Identification Prospects With Granular Calorimeters", submitted to Particles, 2502.10817
- K. Schmidt et al. (P.V.), "End-to-End Detector Optimization with Diffusion models: A Case Study in Sampling Calorimeters", submitted to Particles, 2502.02152
- K. G. Barman et al. (P.V.), "Large Physics Models: Towards a collaborative approach with Large Language Models and Foundation Models", submitted to Nature physics roadmap, 2501.05382

Home > Browse Actions > Machine Learning and Quantum Computing for Future Colliders (MLQC4FC)

Description | **Management Committee** | Main Contacts and Leadership | Working

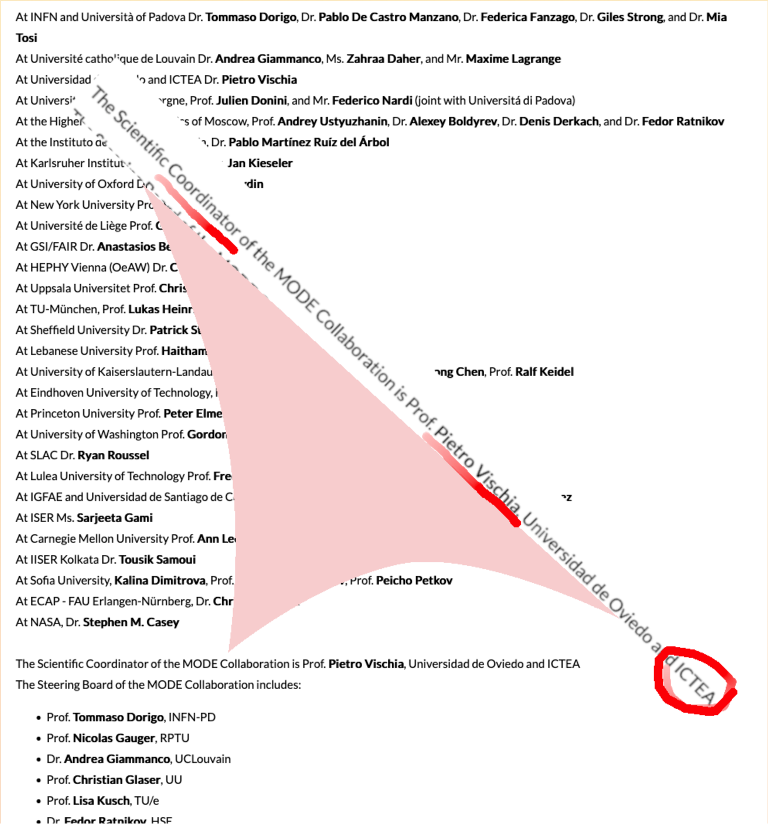
Management Committee

Country	MC Member
Greece	Prof George ZOUPANOS
Spain	Dr German SBORLINI
Spain	Prof Pietro VISCHIA
Switzerland	Prof Tobias GOLLING

The MODE Collaboration

https://mode-collaboration.github.io/

- Machine-Learning Optimized Design of Experiments
- 52 members, 30 institutions, 13 countries (if you are interested, join us!!!)
- Co-founded in 2020, **coordinator since 11-2024**



At INFN and Università of Padova Dr. **Tommaso Dorigo**, Dr. **Pablo De Castro Manzano**, Dr. **Federica Fanzago**, Dr. **Giles Strong**, and Dr. **Mia Tosi**
At Université catholique de Louvain Dr. **Andrea Giammanco**, Ms. **Zahraa Daher**, and Mr. **Maxime Lagrange**
At Universidad ... and ICTEA Dr. **Pietro Vischia**
At Universi... ergne, Prof. **Julien Donini**, and Mr. **Federico Nardi** (joint with Universitá di Padova)
At the Higher ... s of Moscow, Prof. **Andrey Ustyuzhanin**, Dr. **Alexey Boldyrev**, Dr. **Denis Derkach**, and Dr. **Fedor Ratnikov**
At the Instituto de ... Dr. **Pablo Martínez Ruíz del Árbol**
At Karlsruher Institut ... **Jan Kieseler**
At University of Oxford ... **din**
At New York University Pr...
At Université de Liège Prof. ...
At GSI/FAIR Dr. **Anastasios B**...
At HEPHY Vienna (OeAW) Dr. ...
At Uppsala Universitet Prof. **Chris**...
At TU-München, Prof. **Lukas Heinr**...
At Sheffield University Dr. **Patrick S**...
At Lebanese University Prof. **Haitham**...
At University of Kaiserslautern-Landau ... **ng Chen**, Prof. **Ralf Keidel**
At Eindhoven University of Technology, ...
At Princeton University Prof. **Peter Elme**...
At University of Washington Prof. **Gordon**...
At SLAC Dr. **Ryan Roussel**
At Lulea University of Technology Prof. **Fre**...
At IGFAE and Universidad de Santiago de C...
At ISER Ms. **Sarjeeta Gami**
At Carnegie Mellon University Prof. **Ann Le**...
At IISER Kolkata Dr. **Tousik Samoui**
At Sofia University, **Kalina Dimitrova**, Prof. ... Prof. **Peicho Petkov**
At ECAP - FAU Erlangen-Nürnberg, Dr. **Chr**...
At NASA, Dr. **Stephen M. Casey**

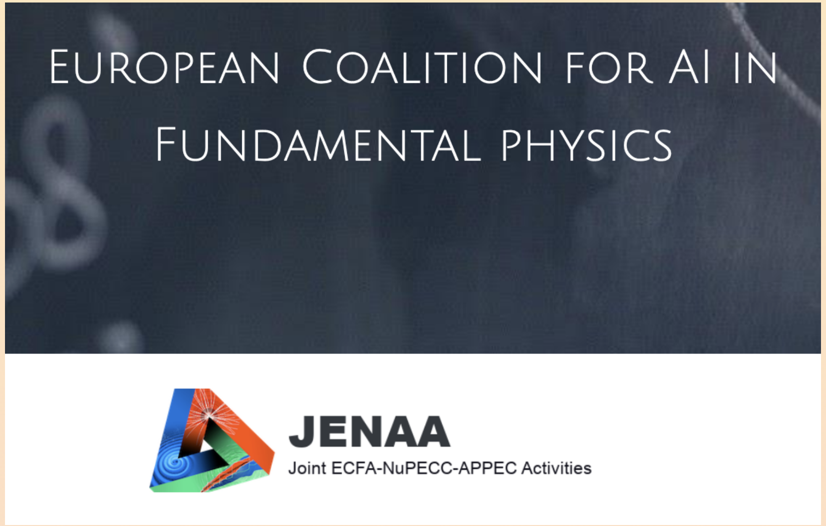
The Scientific Coordinator of the MODE Collaboration is Prof. **Pietro Vischia**, Universidad de Oviedo and ICTEA
The Steering Board of the MODE Collaboration includes:

- Prof. **Tommaso Dorigo**, INFN-PD
- Prof. **Nicolas Gauger**, RPTU
- Dr. **Andrea Giammanco**, UCLouvain
- Prof. **Christian Glaser**, UU
- Prof. **Lisa Kusch**, TU/e
- Dr. **Fedor Ratnikov**, HSE

The Eu Coalition for AI

https://eucaif.org

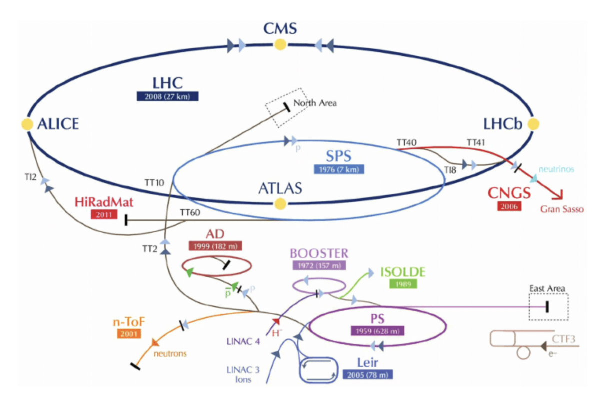
- European initiative for advancing the use of AI in Fundamental Physics
  - Community for developing and promoting AI methods for Physics
  - Bring to AI community advances in AI from physics (see e.g. 2024 Nobel prize to Hopfield and Hinton)
  - Had already two conferences (2024, 2025
  - **Work Package 2: Experiment Design** (leaders: T. Dorigo and **P. Vischia**)



Complex experimental apparata

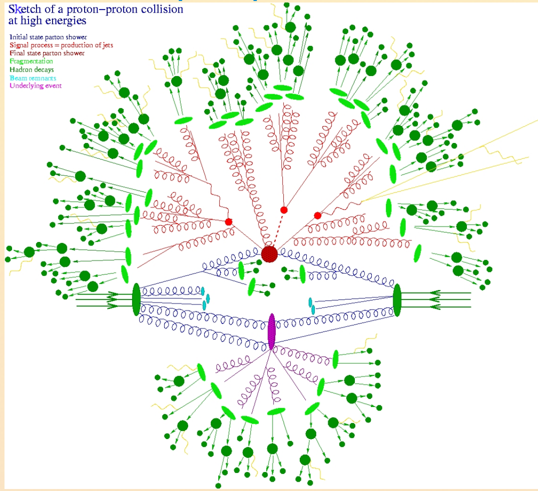
- 2020 European Strategy (EUSUPP): "New large, long-term projects, pushing technological skills to the limit"

Complex accelerators

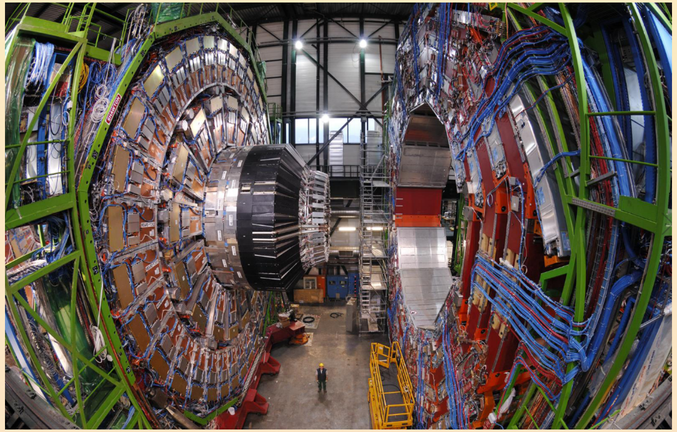


Existing CERN accelerator complex with Large Hadron Collider (LHC), Super Proton Synchrotron (SPS), Proton Synchrotron (PS), Antiproton Decelerator (AD), Low Energy Ion Ring (LEIR), Linear Accelerators (LINAC), CLIC Test Facility (CTF3), CERN to Gran Sasso (CNGS), Isotopes Separation on Line (ISOLDE), and neutrons Time of Flight (n-ToF).

Complex phenomena



Complex experiments and reconstruction



$$P(\mathbf{x}|\alpha) = \boxed{\frac{1}{A_\alpha \sigma_\alpha}} \int \boxed{d\Phi(y) \frac{dx_1 dx_2}{x_1 x_2 s} f(x_1) f(x_2)} \boxed{|\mathcal{M}_\alpha(y, x_1, x_2)|^2} \boxed{W(\mathbf{x}|y)\, \epsilon_\alpha(y)}$$

Normalization factor

We collide protons and that is a mess

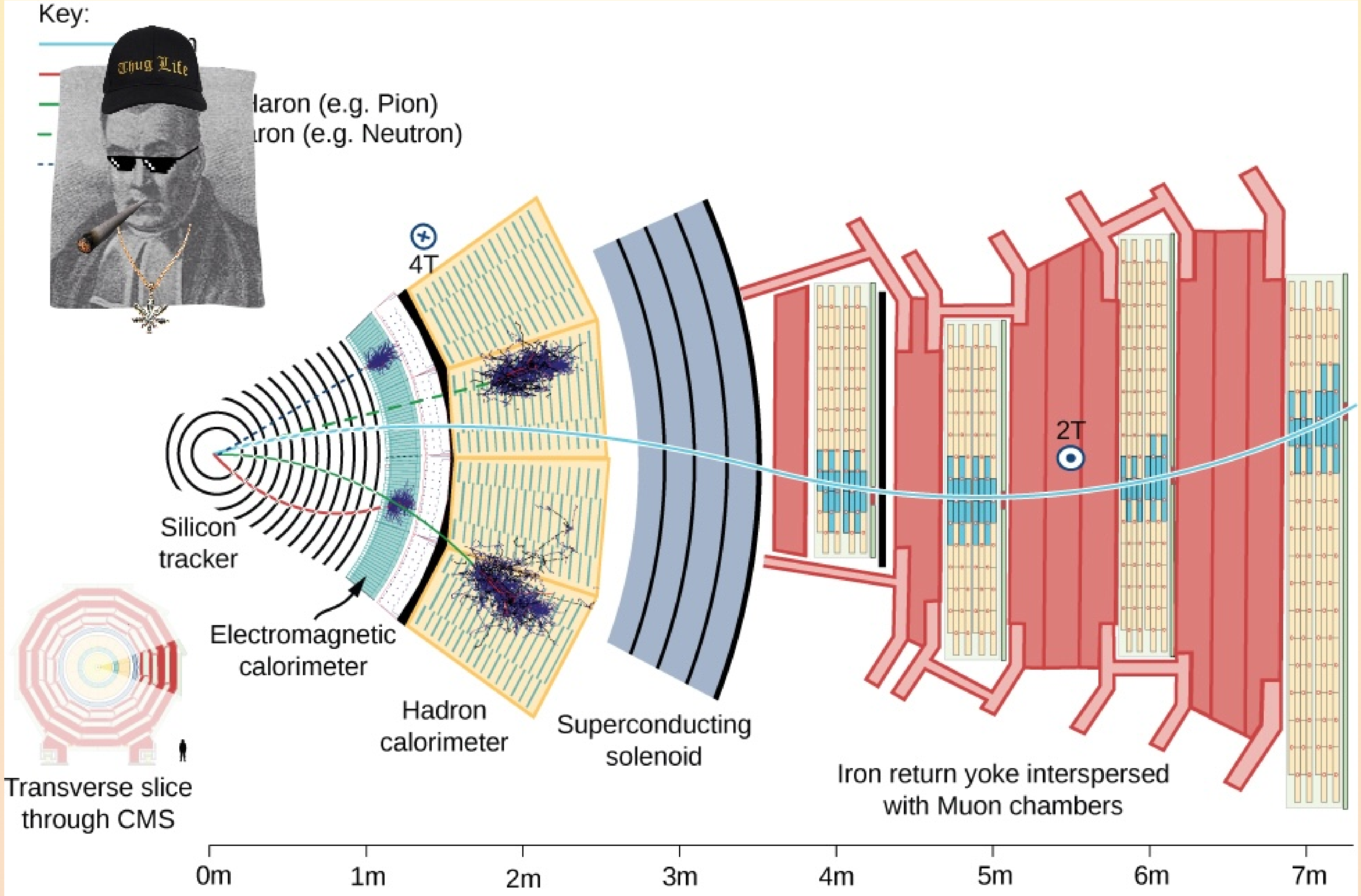
Linear operator in Hamiltonian formalism, describes the physics process

Detector response, experimental efficiencies

- Stochastic processes $\longrightarrow$ intractable likelihood (matrix element, parton shower, detector simulation... result in latent variables)
- Costly MonteCarlo simulators to generate $x \sim p(x|\theta)$ (for each event, several thousand randomized choices)

Image from 10.1103/PhysRevSTAB.16.054801 and from the CMS Collaboration

What is a calorimeter?

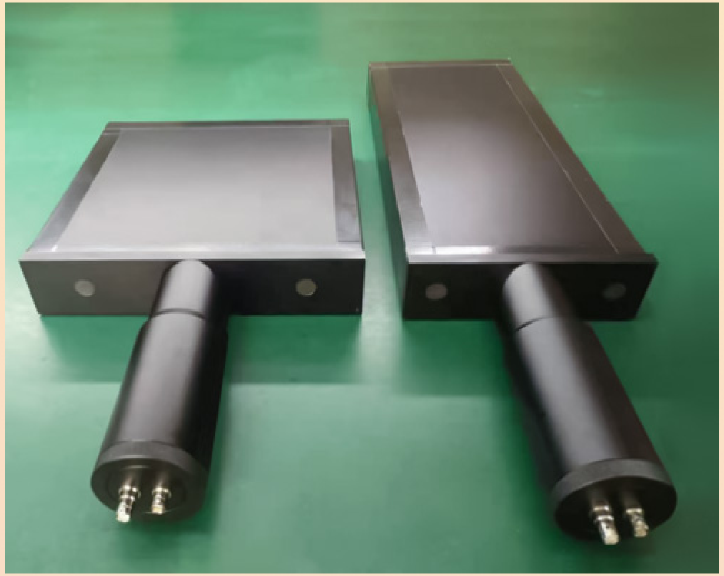


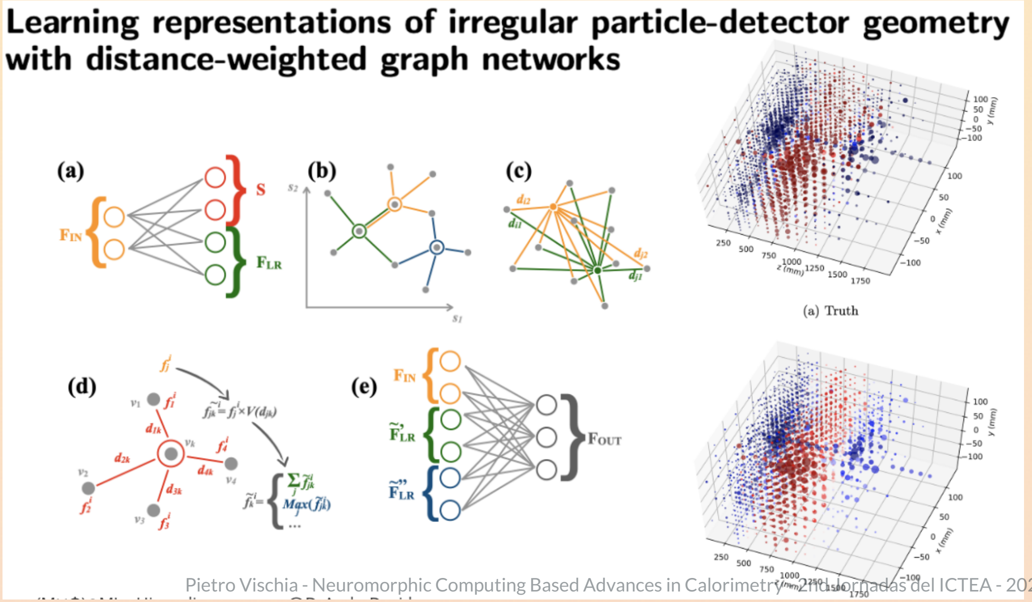
—
Images by the CMS Collaboration

What is a calorimeter?

- Scintillating material (e.g. polystirene, $PbWO_4$) producing light
- Readout (e.g. silicon photomultipliers)
- Absorber material (Iron, Lead)
- **Size of the scintillating elements drives resolution**

- **Increasing the granularity is costly**
  - E.g. **CMS spent** 67M CHF **to upgrade** in 2026-2030 the endcap calorimeter
  - High Granularity Calorimeter (HGCal): 6 million Si channels, 250000 scintillator tiles readout by SiPMs,
  - First imaging calorimeter at a high-PU hadron collider
  - Precision particle flow reconstruction (sensitivity to VBF and VBS improves, more precise jet substructure studies)
  - Timing extends reach for long-lived particles

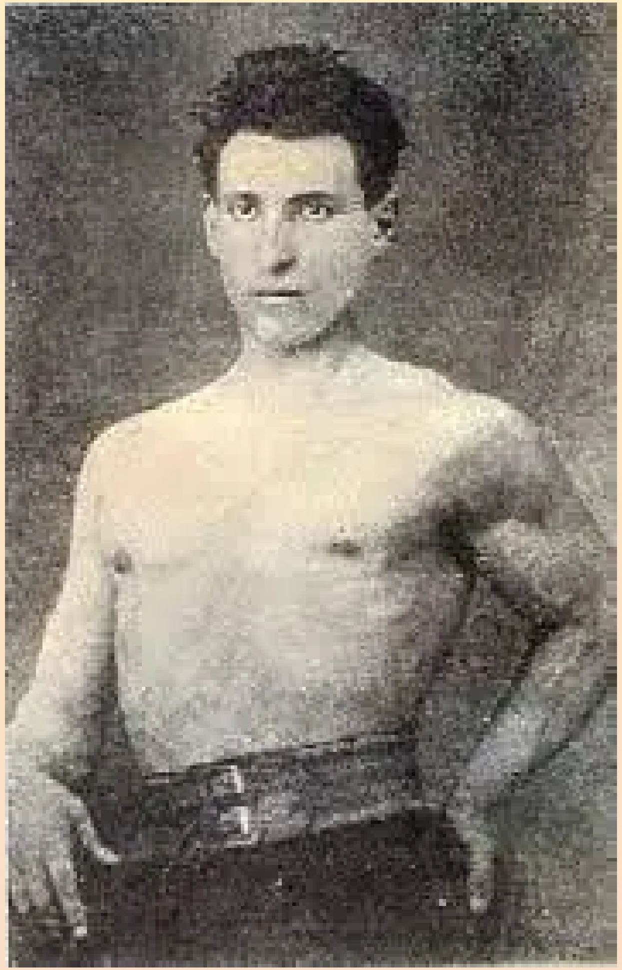




Neuromorphic Computing can help!

Santiago Ramón y Cajal First become a camorrista and a bodybuilder...

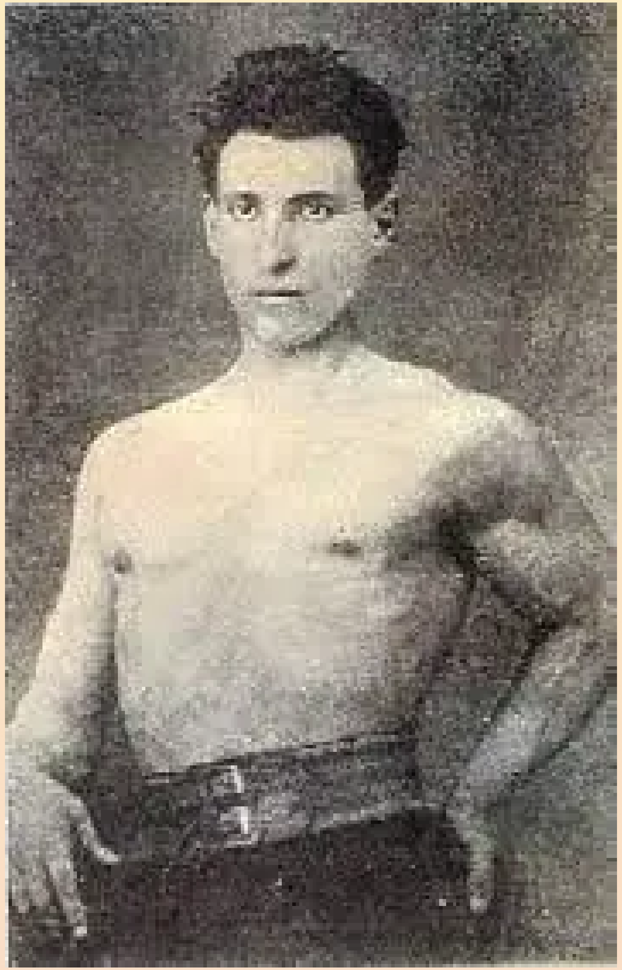
"El prurito de lucir el esfuerzo de mi brazo me arrastró más de una vez, contra mi temperamento nativamente bonachón, a parecer camorrista y hasta agresivo."[1]

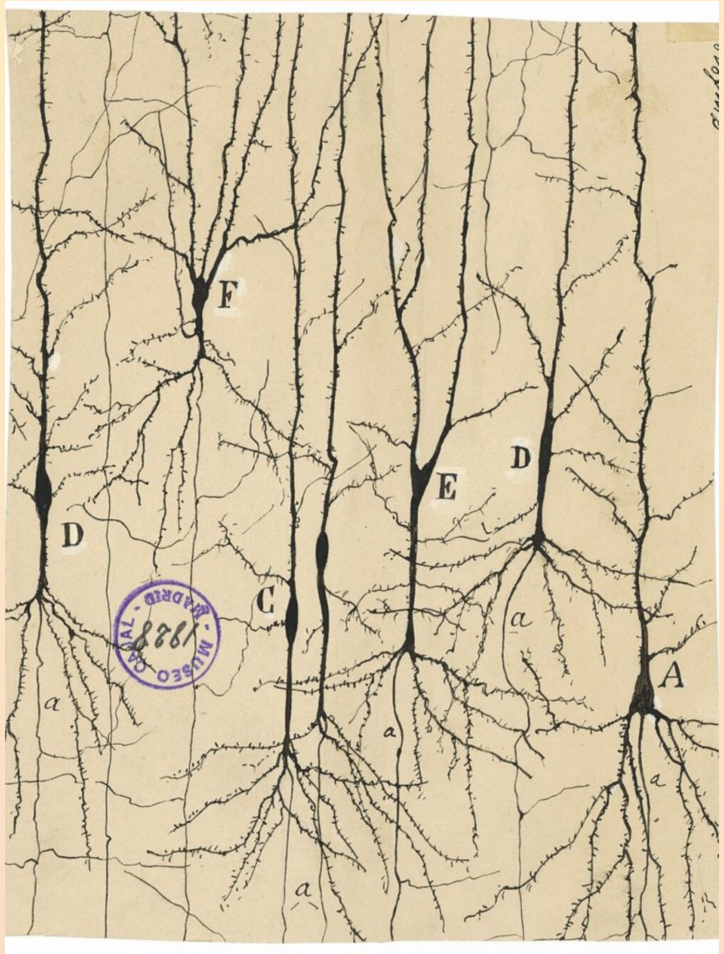


—
1: Santiago Ramón y Cajal, "Recuerdos de mi vida", Capítulo XXI. Image from La Razon

Santiago Ramón y Cajal ... then get the Nobel Prize (1906, in Physiology or Medicine)

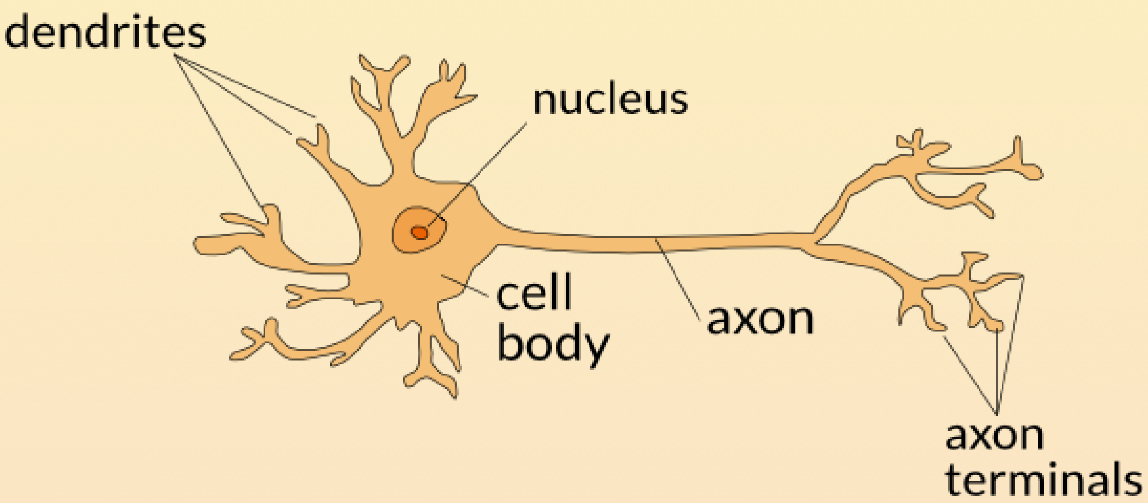
"El prurito de lucir el esfuerzo de mi brazo me arrastró más de una vez, contra mi temperamento nativamente bonachón, a parecer camorrista y hasta agresivo."[1]



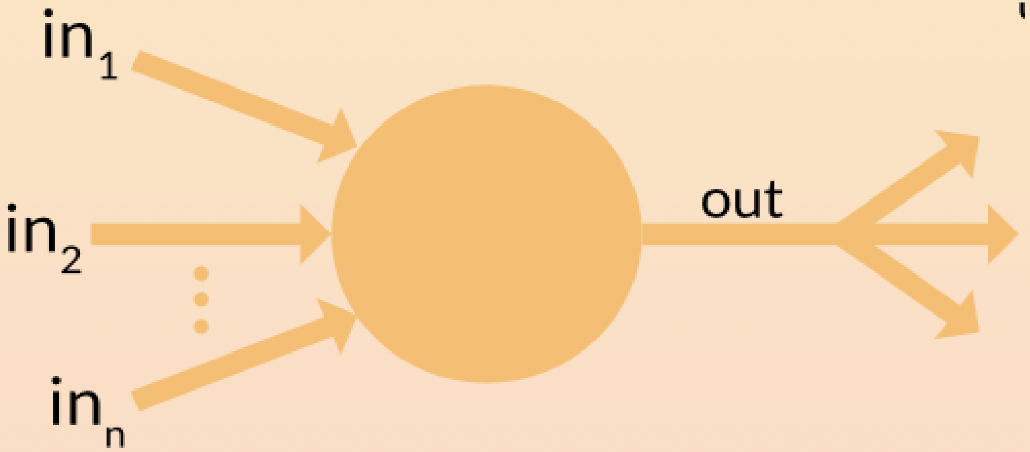


1: Santiago Ramón y Cajal, "Recuerdos de mi vida", Capítulo XXI. Images from La Razon and The Nobel Prize

From Neurons to Perceptrons...



$$
\begin{aligned}
I(t) = \quad & C\frac{dV}{dt} \\
+ \quad & G_{Na}m^3h(V - V_{Na}) \\
+ \quad & G_K n^4(V - V_K) \\
+ \quad & G_L(V - V_L)
\end{aligned}
$$

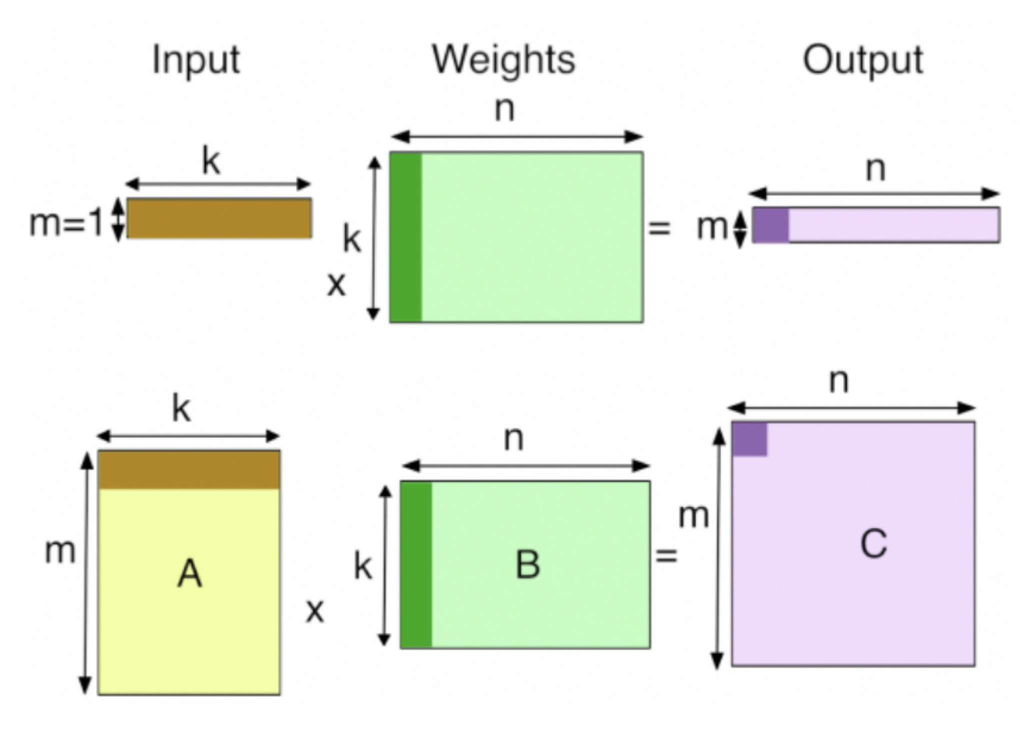


$$y = f\left(b_i + \sum w_i x_i\right)$$

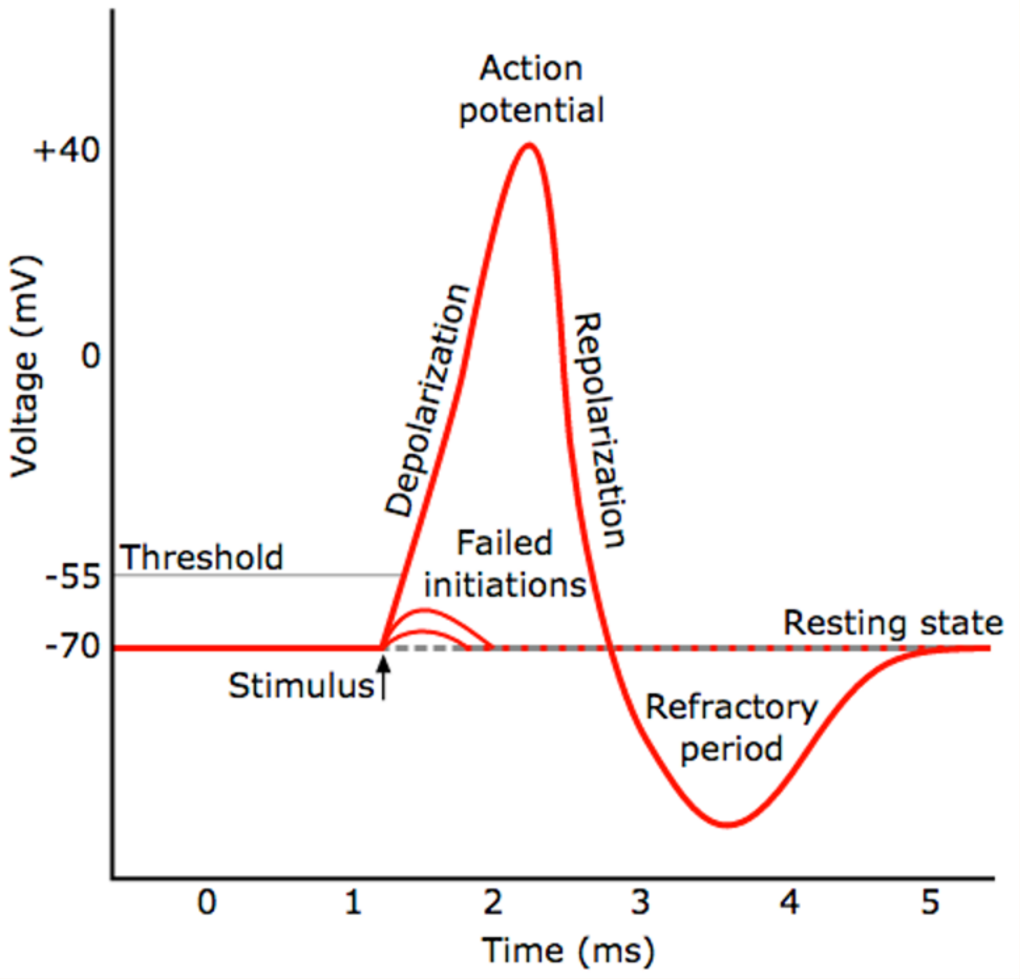
Image from https://appliedgo.net/perceptron/

...and back: a paradigm shift

From perceptrons and matrices...

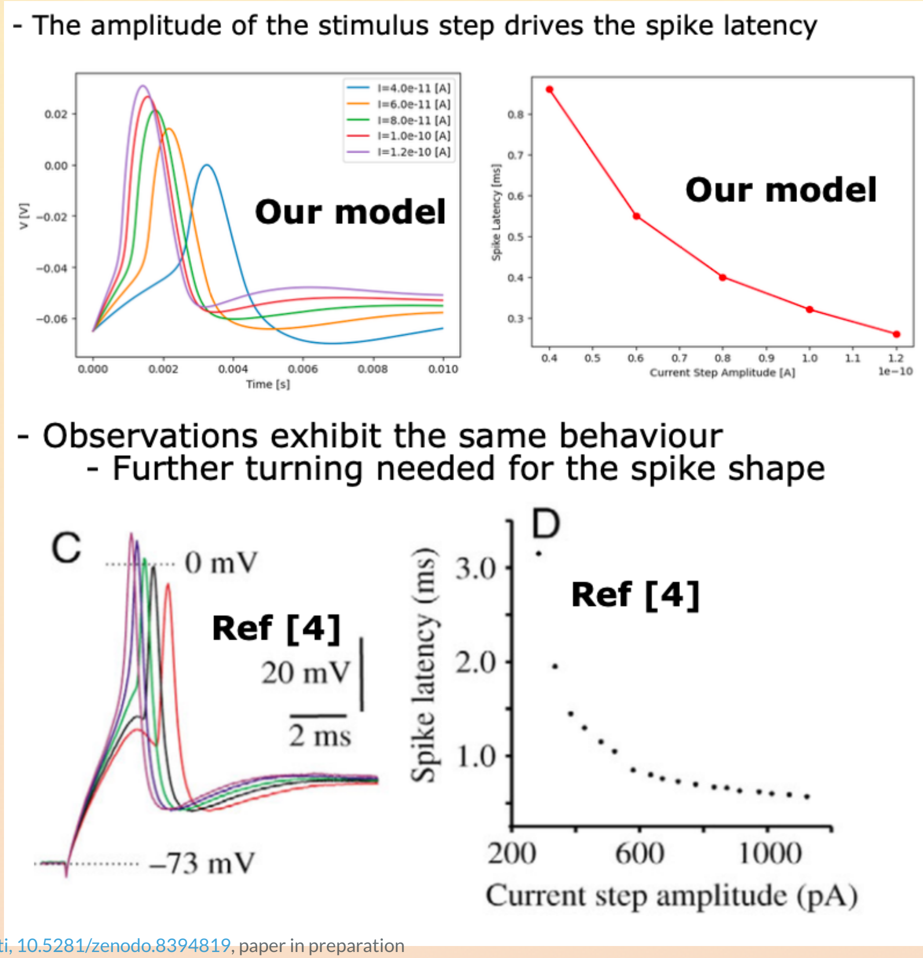


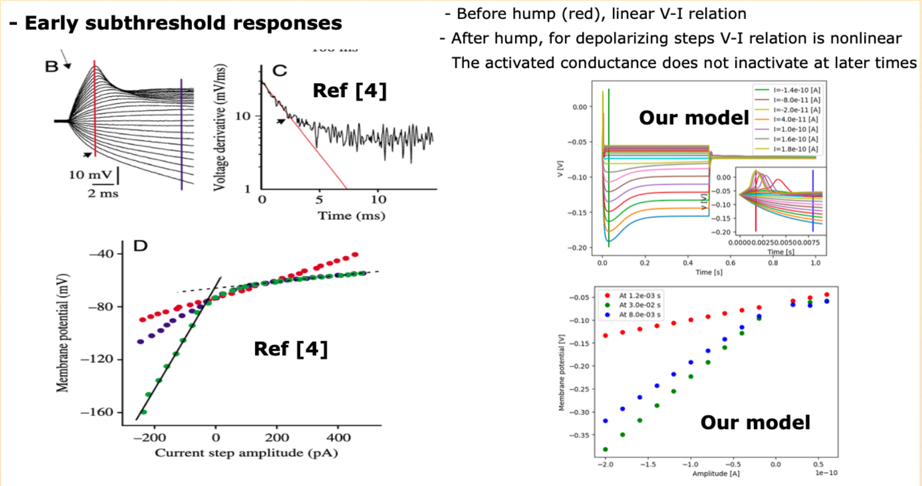
to spiking neurons modulating spatiotemporal patterns

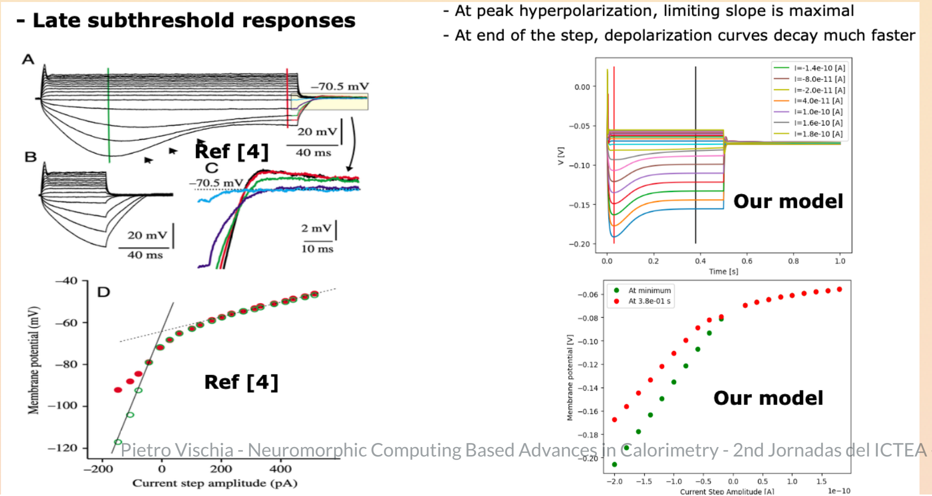


Realistic neuronal models P. Vischia, A. Caputi, 10.5281/zenodo.8394819

- Spherical neuron with four channels (different thresholds and time constants) for *Gymnotus Omarorum*
  - Vischia, Caputi 2023: computational model compared with data from "[4]" (J Exp Biol (2006) 209 (6): 1122–1134.)





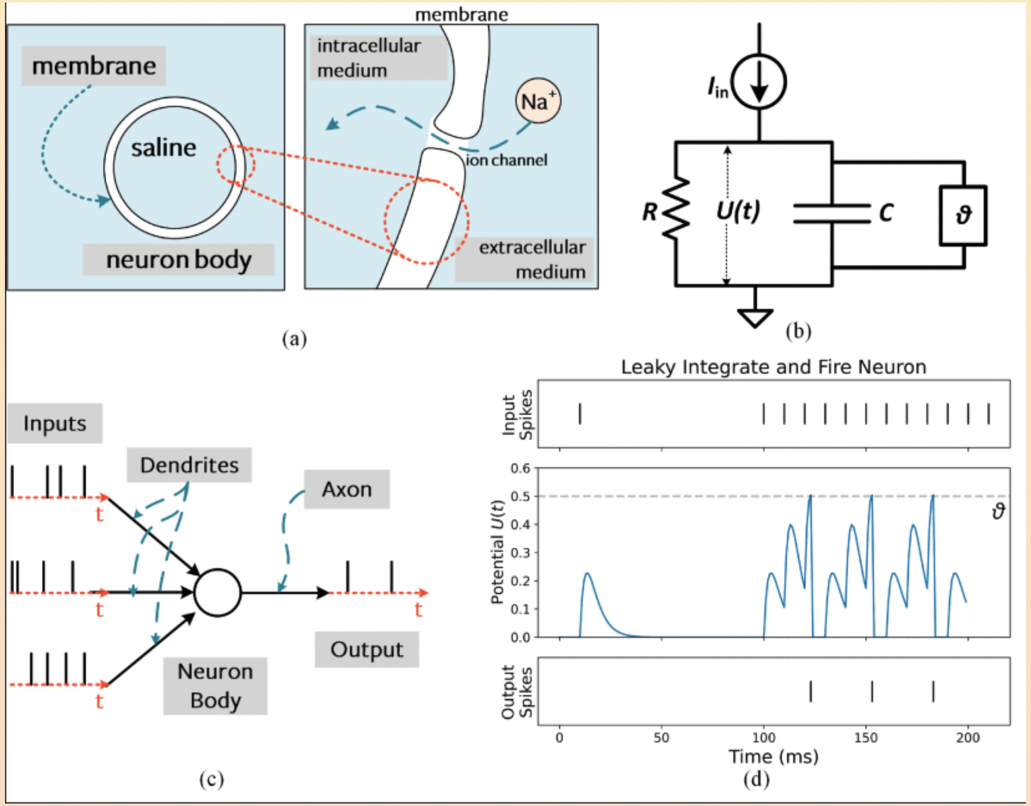


P. Vischia, A. Caputi, 10.5281/zenodo.8394819, paper in preparation

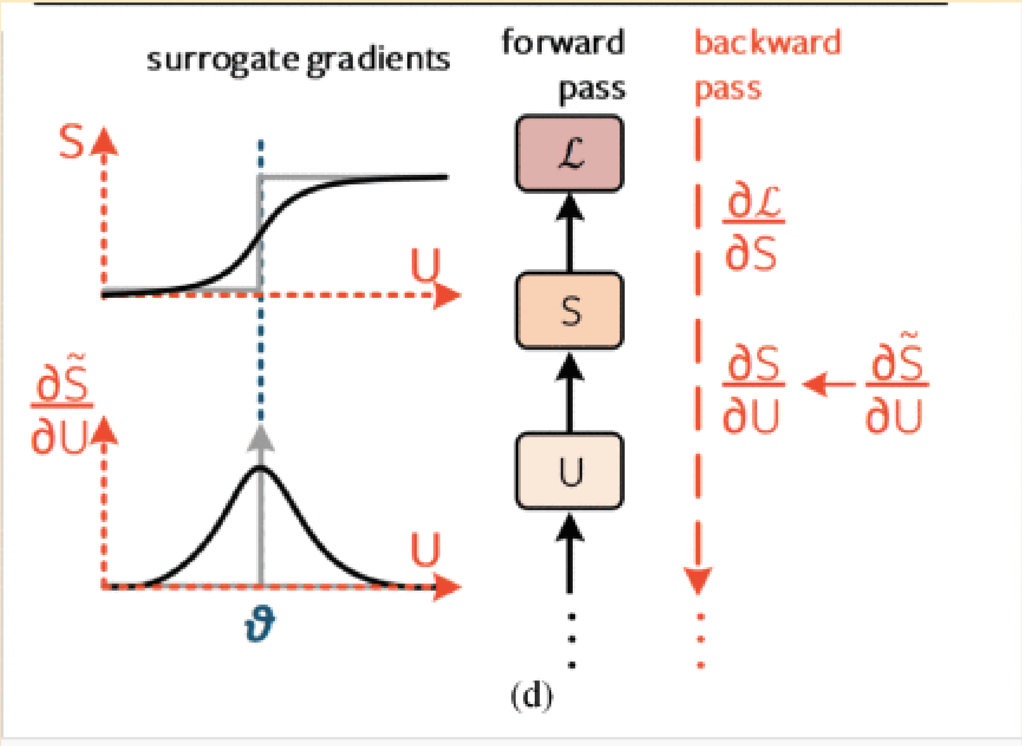
Artificial Spiking Networks

- Neuronal model vastly simplified: (Leaky) Integrate-and-fire Model. Other training schemes are available (not illustrated here)

$$U[t] = \underbrace{\beta U[t-1]}_{\text{decay}} + \underbrace{WX[t]}_{\text{input}} - \underbrace{S_{\text{out}}[t-1]\theta}_{\text{reset}}.$$



- Surrogate gradient to mimick the usual gradient descent
  - Use automatic differentiation for training!

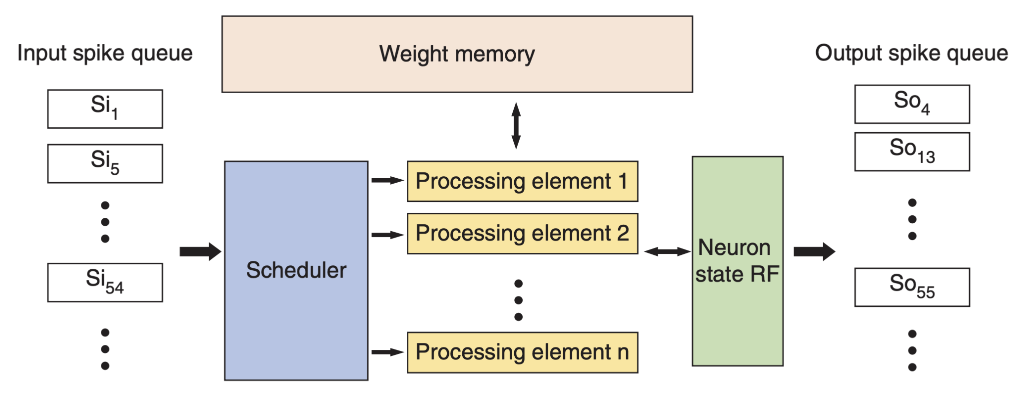


Images from J. K. Eshraghian et al. 10.1109/JPROC.2023.3308088

Energy-efficient architectures

Sparser inputs → less time and energy

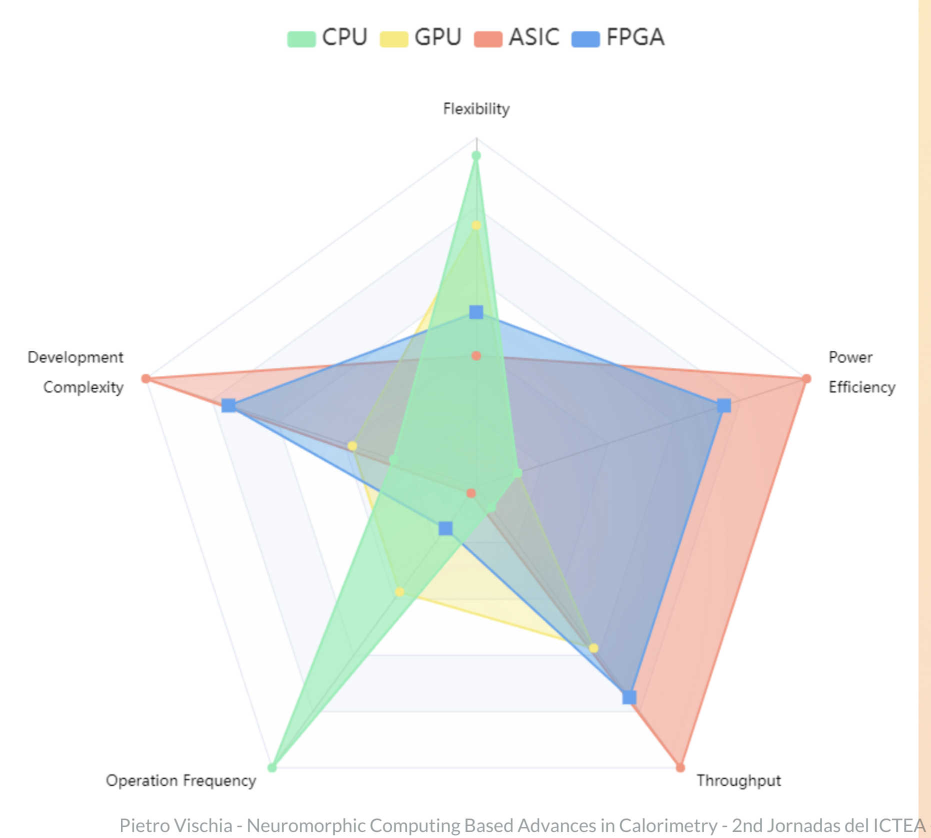
- Event-driven computation
- Efficient, threshold-based encoding
  - GPU (RTX3090): 40 GW simulation on discrete states
  - Human brain equivalent: 20 W on dynamic states
- Energy consumption drastically limited in sparse settings



	Brain	Spikey	SpiNNaker	R2600X	Intel mobile	RTX2070
Housekeeping	4.75E-11	1.37E-06	1.66E-04	4.49E-04	1.23E-04	**9.76E-07**
Resting potential	5.77E-11	**3.83E-08**	8.99E-05	4.77E-05	4.25E-05	3.63E-06
Action potential	1.96E-11	**4.39E-10**	1.04E-08	3.04E-08	4.46E-09	4.71E-09
Transmission	8.17E-15	**1.08E-11**	9.59E-09	5.82E-08	2.14E-08	3.40E-09
Single neuron	2.49E-10	**1.49E-06**	3.33E-04	9.62E-04	3.37E-04	3.18E-05
Full brain	2.15E+01	**1.29E+05**	2.87E+07	8.29E+07	2.90E+07	2.74E+06

Values for the simulation of 1s of model time are reported in Joule. The single neuron and full brain estimates assume a fan-out of 2,000 synapses and a spike rate of 4Hz. R2600X: AMD Ryzen 2600X. Intel mobile: Intel Core i7-4710MQ. RTX2070: NVIDIA RTX 2070. Both CPUs are measured using a PeakTech power meter. The lowest values from simulators/emulators are highlighted in bold.

- Neuromorphic hardware has maximal energy efficiency
  - **Memristors are optimal**, otherwise CMOS/ASIC
  - FPGA implementations also available (e.g. 2502.20415, 2411.01628, 2307.03910, 2305.11252



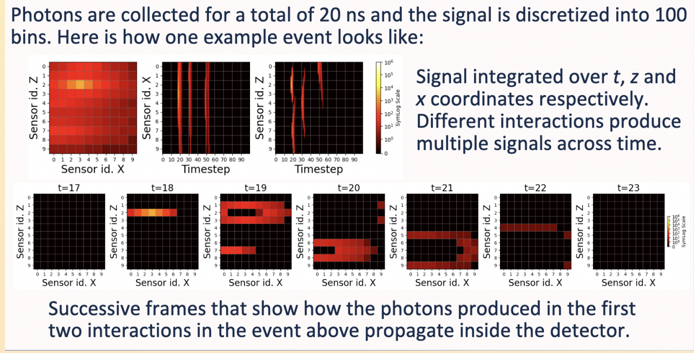
—
Figure from 10.1038/s41586-021-04362-w and 2307.03910

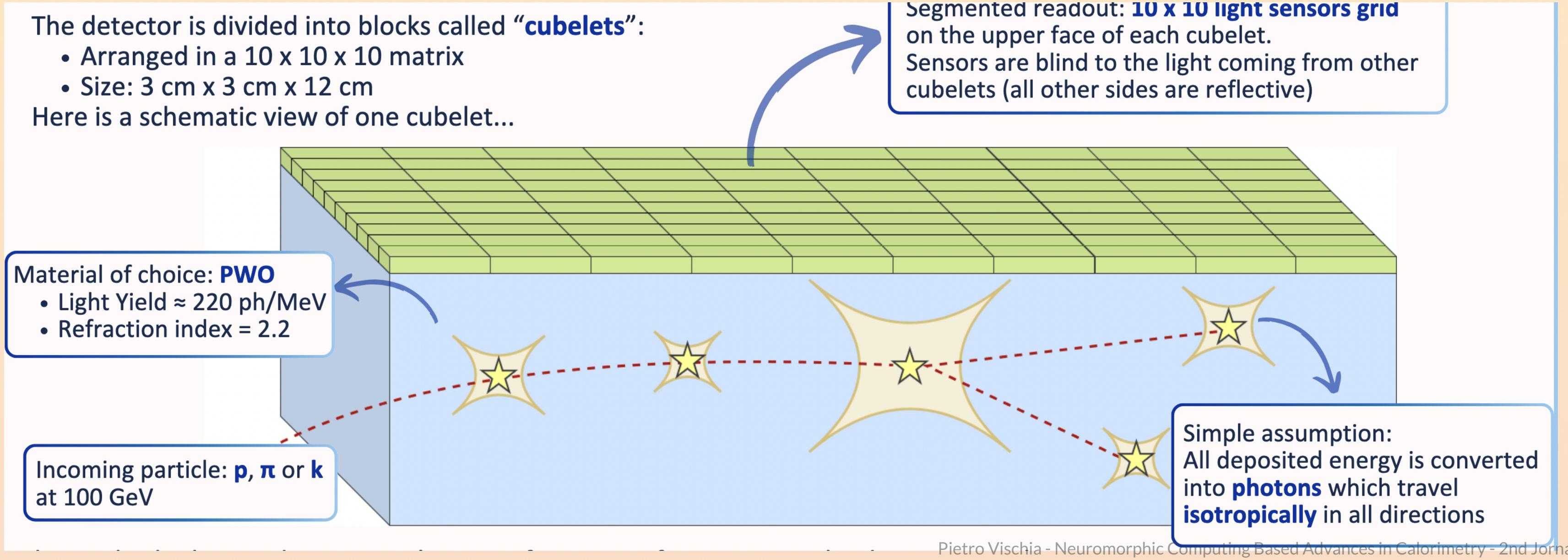
Can we apply Neuromorphic Computing to HEP apparata?

Neuromorphic Readout for HCals E. Lupi et al. (P.V.) Particles 2025 8(2) 52

- Hadron Calorimeter readout via light-sensitive nanowires
  - Fast, energy-efficient local computation
  - Generate informative high-level primitives
- GEANT4 simulation of 100-GeV $p/K/\pi$
- Simulated calorimeter: 1000 cubelets, 1M nanowire cells
  - $300 \times 300 \times 1200 mm^3$: lateral/longitudinal containment of 100% (87%)
  - Record every 0.2ns for 20ns (90% of the shower is deposited)

- Time evolution accessible via spiking networks
  - **Bypass the need for high-granularity calorimeters**



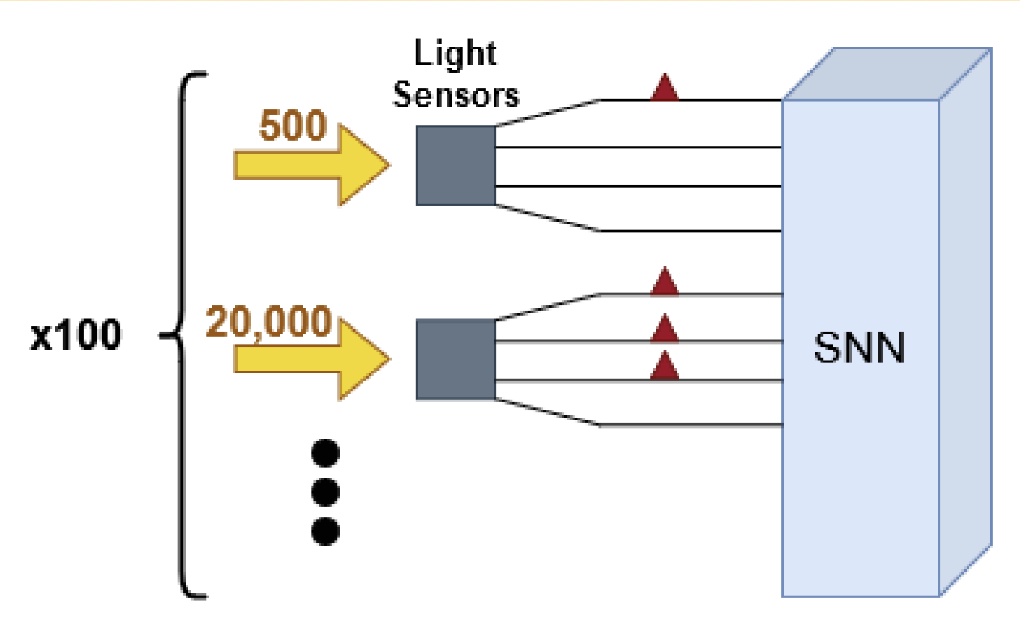


Encoding sensor data E. Lupi et al. (P.V.) Particles 2025 8(2) 52

- Each sensor feeds four channels to a spiking network $\longrightarrow$ have to convert incoming data into a spike train
- For each time step t, each channel can either be activated (carry a spike) or inactivated
  - Define threshold on incoming light that activates the channel
- Each channel has a different activation threshold $\longrightarrow$ spikes carry both timing and intensity information
- For the i-th channel and t-th timestep is the following:

$$S^i[t] = \begin{cases} 1 & \text{if } N_{ph}[t] \geq 10^{i+2} \\ 0 & \text{otherwise} \end{cases}$$

$$i = 1, 2, 3$$



Decoding scheme E. Lupi et al. (P.V.) Particles 2025 8(2) 52

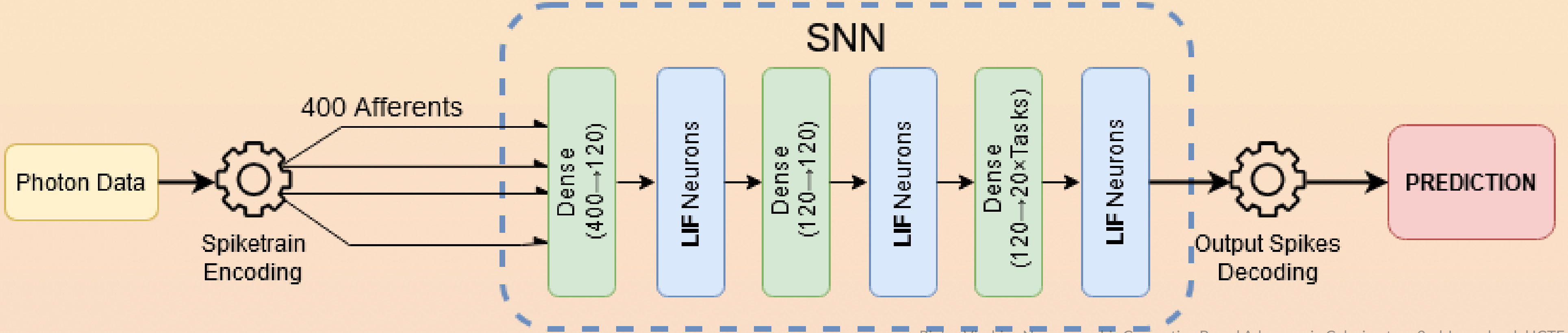
- Decode spike train via fully connected SNN out of LIF neurons
  - LIF neurons with learnable U_{thr} and β, and `arctan` function as surrogate gradient.
- Last layer of SNNs natural choice as a regressor (either outpus spikes or membrane potential)

Membrane potential: higher precision, higher latency and computational complexity

$$\hat{X} = \frac{1}{N_{pop}} \sum_{i=1}^{N_{pop}} U^{(i)}[T]$$

Output spikes: requires careful information encoding/decoding schemes (e.g. rate- or latency-based), but better performance

$$\hat{X} = \frac{1}{N_{pop}} \sum_{i=1}^{N_{pop}} \sum_{t=1}^{T} U^{(i)}[t]$$

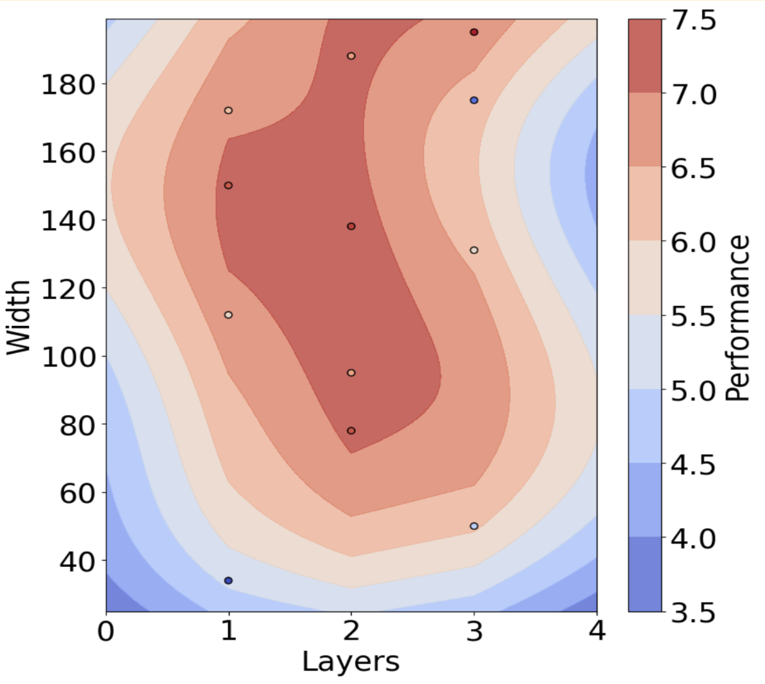


Network optimization E. Lupi et al. (P.V.) Particles 2025 8(2) 52

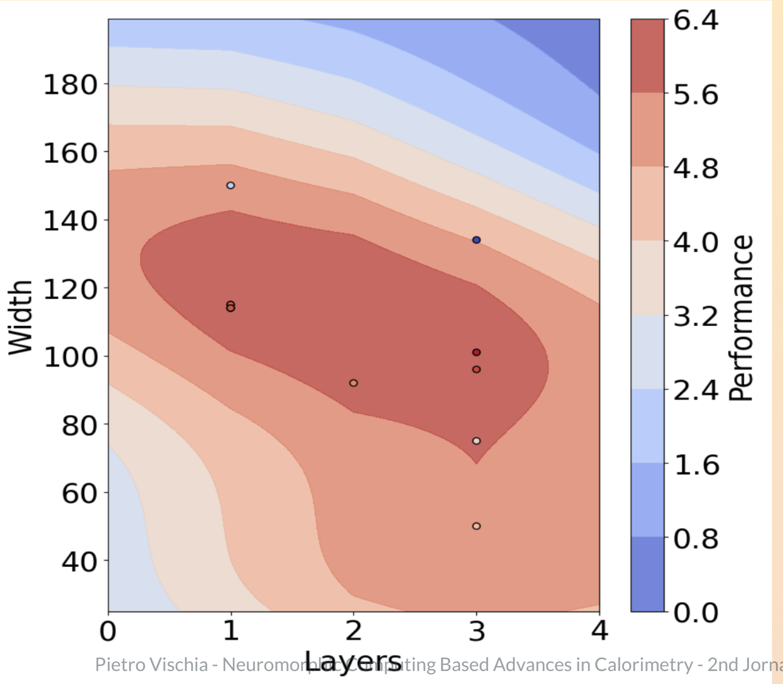
- Architecture optimized via Bayesian optimization, separately for each multi-target regression
  - Prior> gaussian surrogate model of the objective function
  - New observations generated maximizing Expected Improvement $\rightarrow$ used to update the prior in the function space
  - Sampling directed towards the maximum of the surrogate model (Conditioning on sampling towards maxima.
  - Access to full posterior over parameters

$$P = \frac{1}{N_{targets}} \sum_{i}^{N_{targets}} \frac{1}{\epsilon_i}$$

Regress log(E/MeV) and centroid



Regress log(E/MeV) and dispersion



Regression E. Lupi et al. (P.V.) Particles 2025 8(2) 52

- Single- and multi-target regression give **consistent results**
- Centroid of energy distribution determined to **less-than-one-cell accuracy**
- Large energy range results in nonoptimal error
- z dispersion more precise than x and y

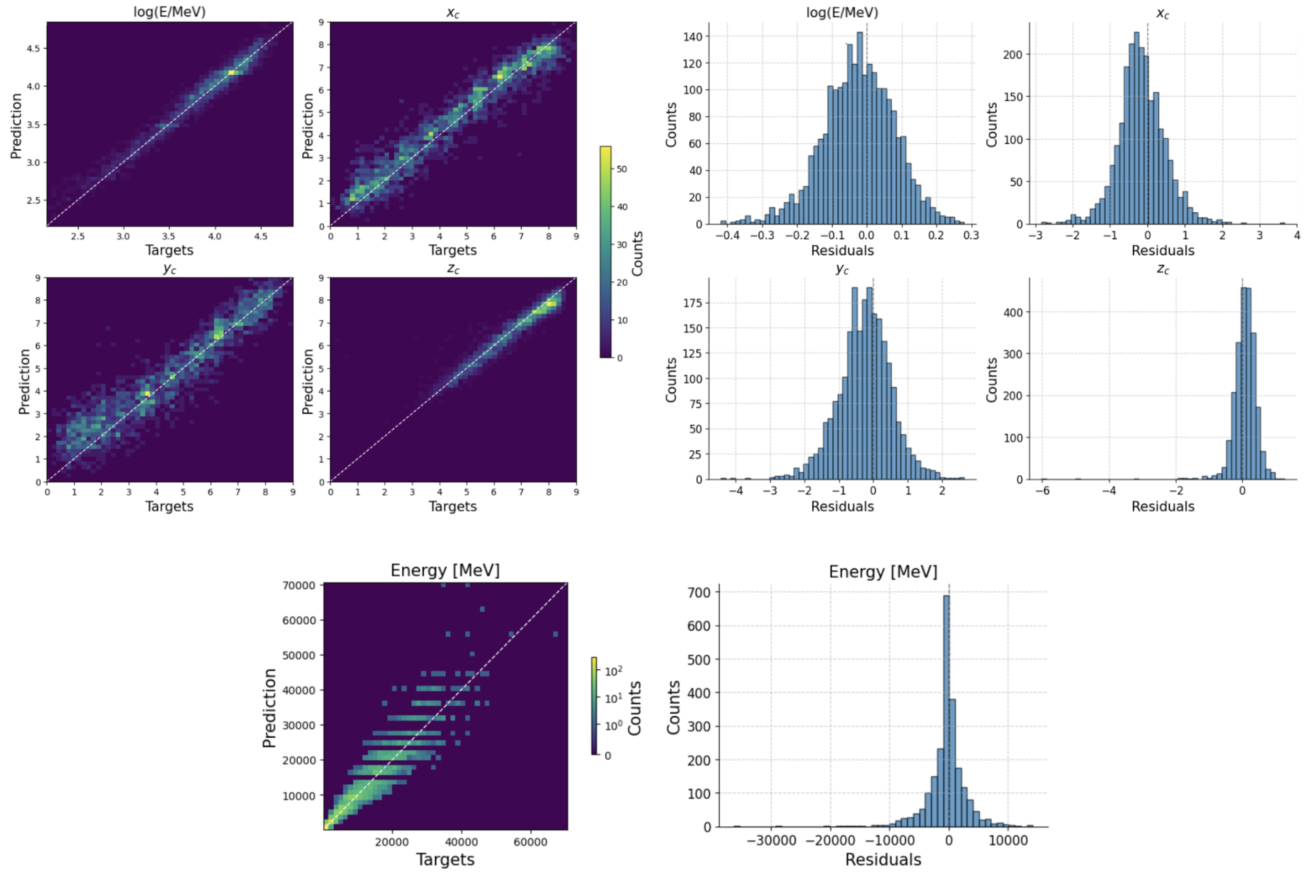


Figure 8. Output of the energy deposition and centroid network. The plots show the correlation between targets and predictions (left) and the residuals (right).

Nanophotonics implementation E. Lupi et al. (P.V.) Particles 2025 8(2) 52

- Most NC systems operate at 200 MHz (5ns cycle)
  - Required sub-ns (0.2ns) sampling implies challenges
- III-V semiconductor nanowire: detect and emit light, work as electronics
  - Spiking photonic neurons in the picoseconds
  - InP can detect light in the 450-550 nm bandwidth of $PbWO_4$
- Large energy range results in nonoptimal error
- z dispersion more precise than x and y

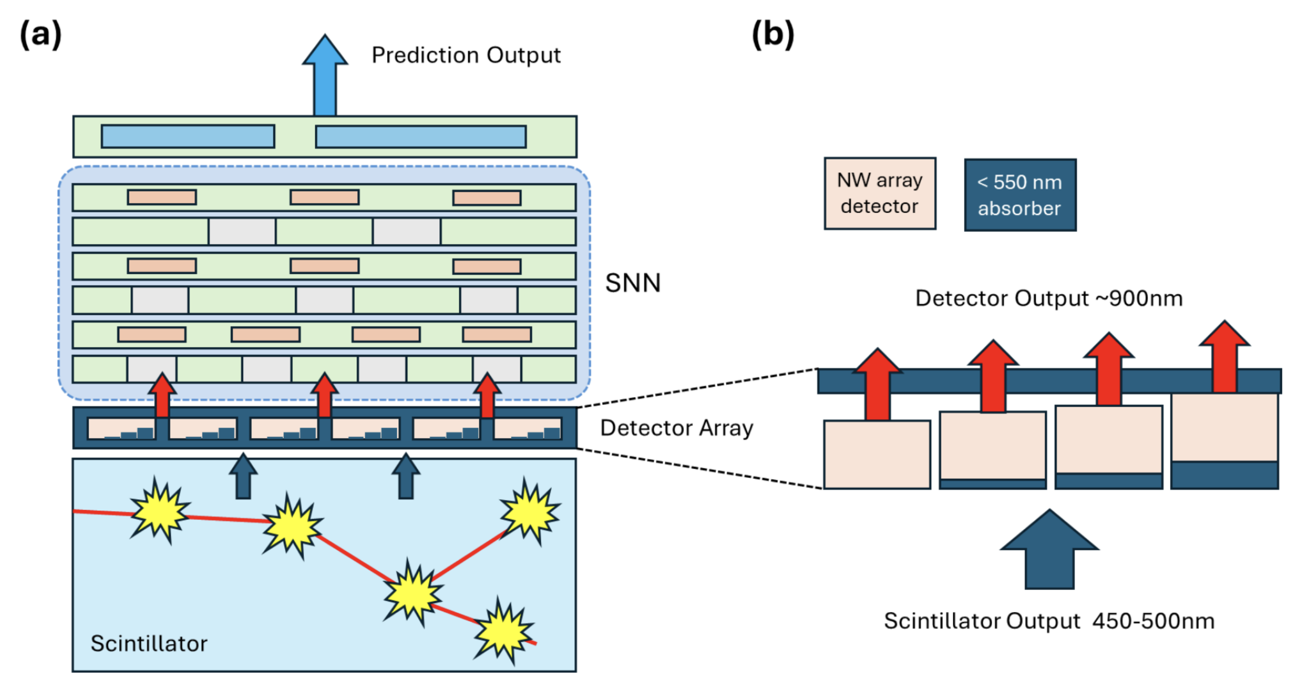
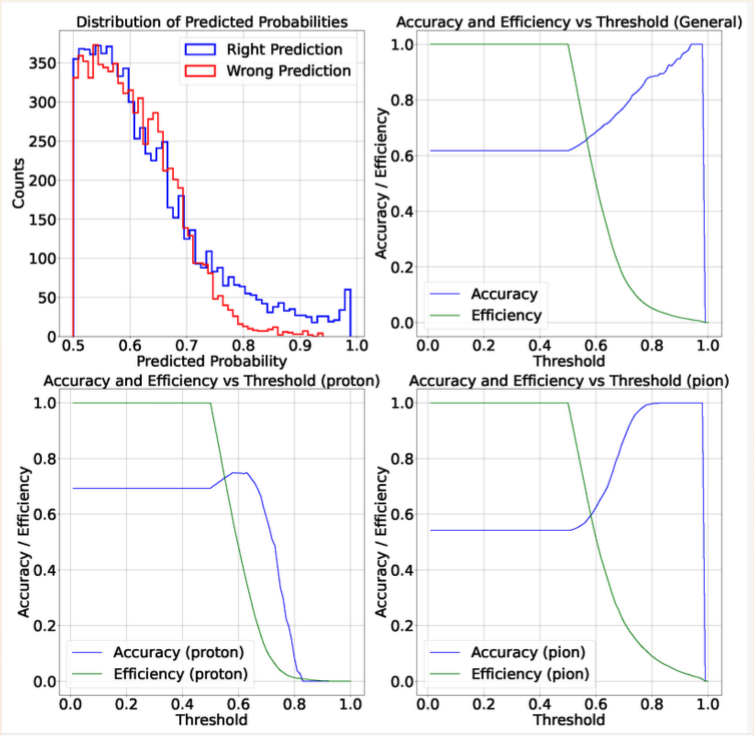
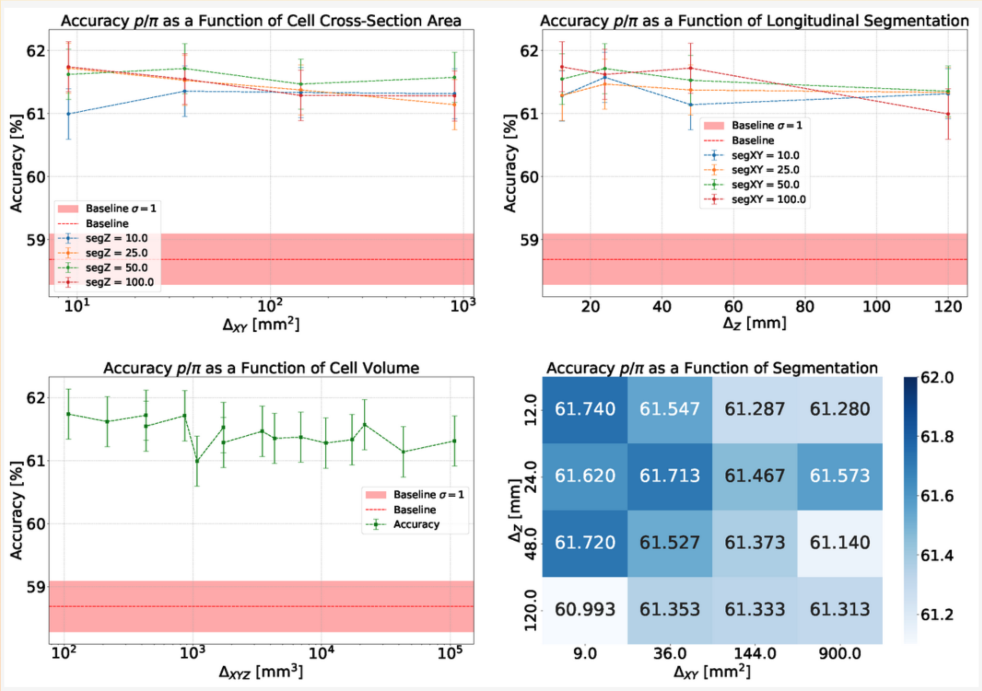


Figure 12. (a) Possible implementation of a readout system for a calorimeter block (bottom left) – light pulses are received by the layer of detector arrays and passed onto the SNN, which comprises of multiple layers of metalens/waveguide broadcasting and LIF neurons. Finally, the prediction output from the SNN is decoded and read out. (b) 4 channel detector superpixel, with individual detectors receiving progressively attenuated signals due to their respective thicker absorbers.

Why is segmentation useful? A. De Vita et al. (P.V.) Particles 2025, 8(2), 58

- High granularity opens the possibility of separating $p/K/pi$
- Can define binary decisions with $\mathcal{O}(60\%)$ accuracy at 100% efficiency
- Neuromorphic implementation
  - **can do this within the chips themselves!!!**
  - **can do that without increasing granularity of the calorimeter!!!** (only the readout)





Traditional experiment/detector design

- Manual or sampled design process is inefficient and sometimes even unfeasible

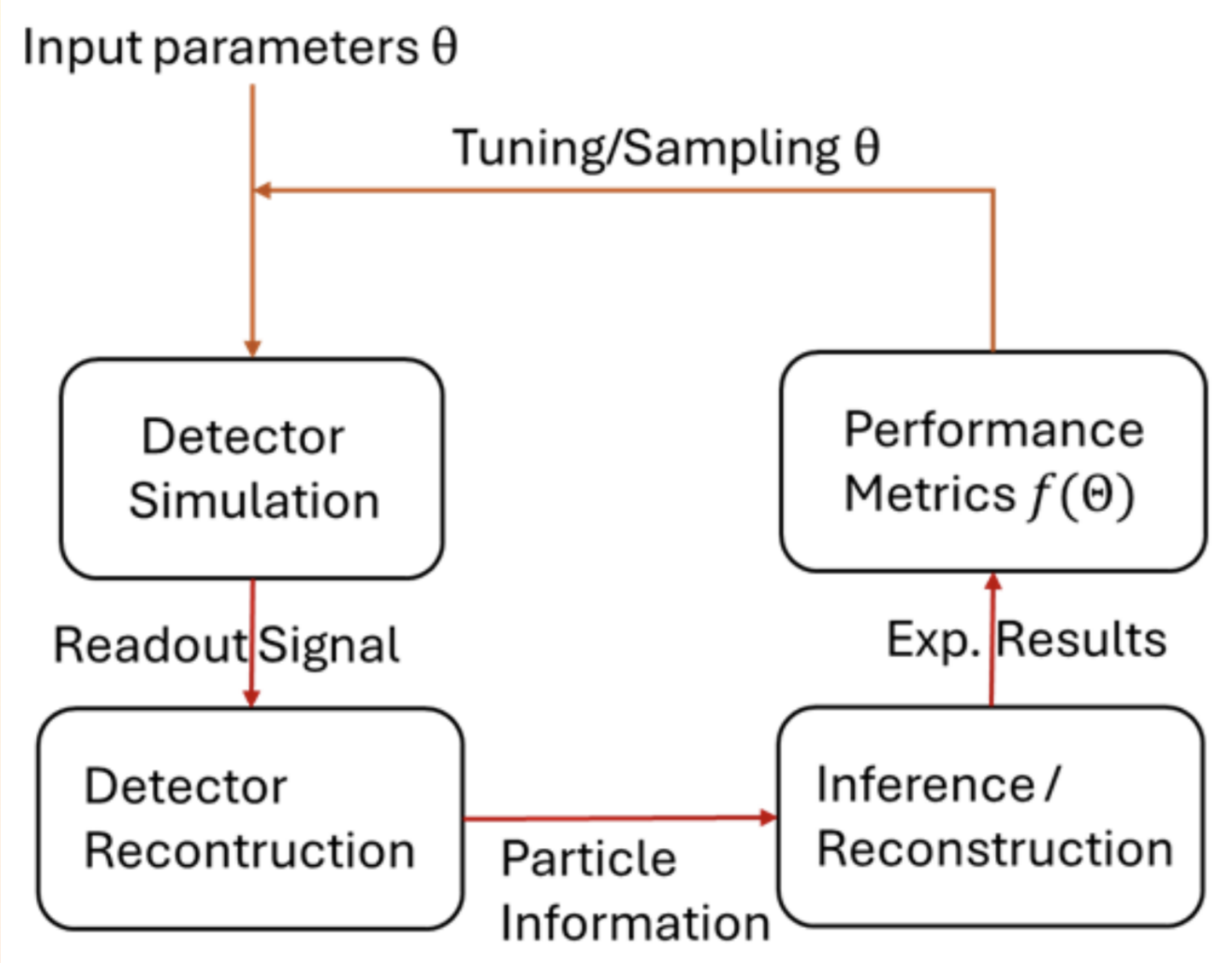
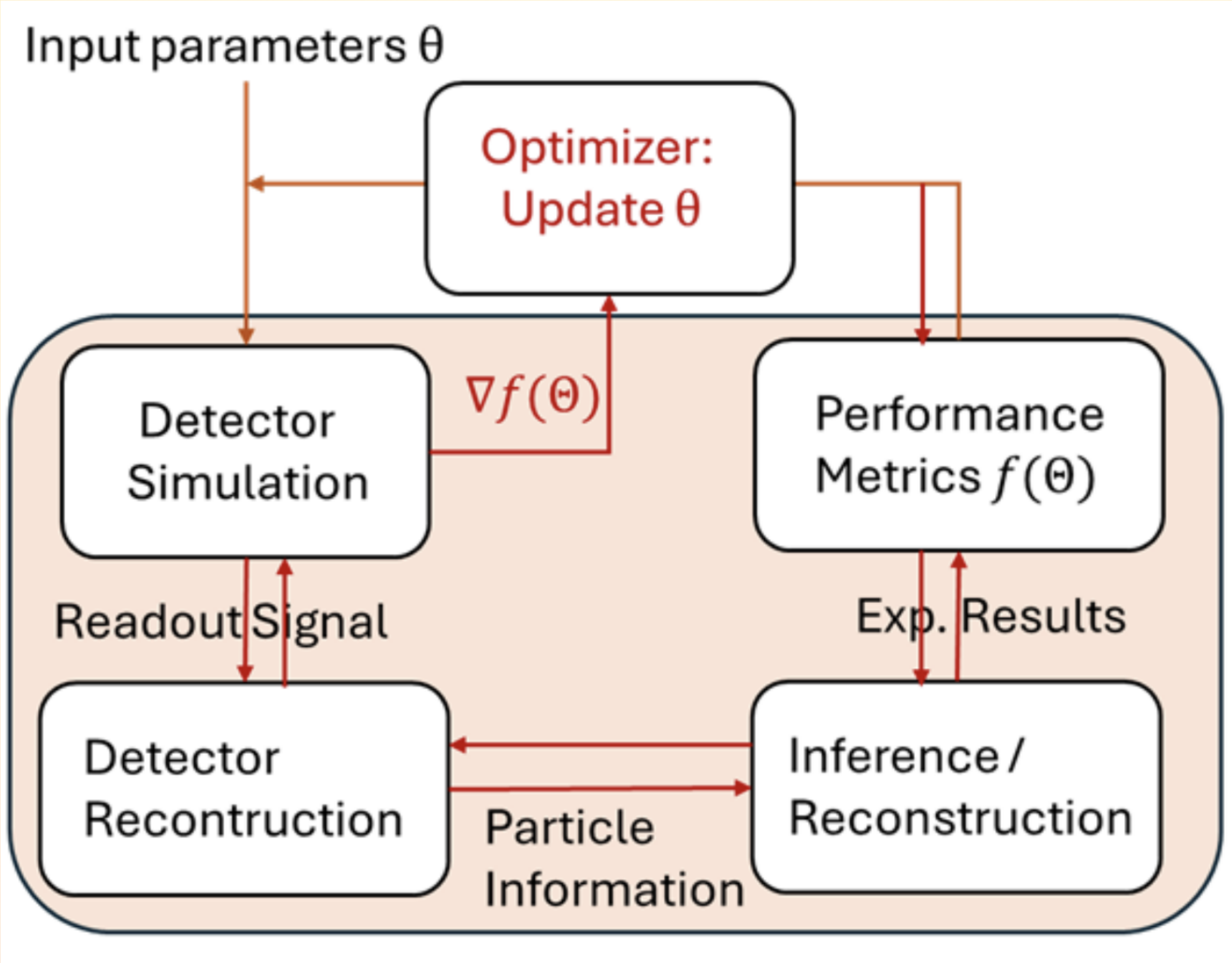


Diagram by Lisa Kusch

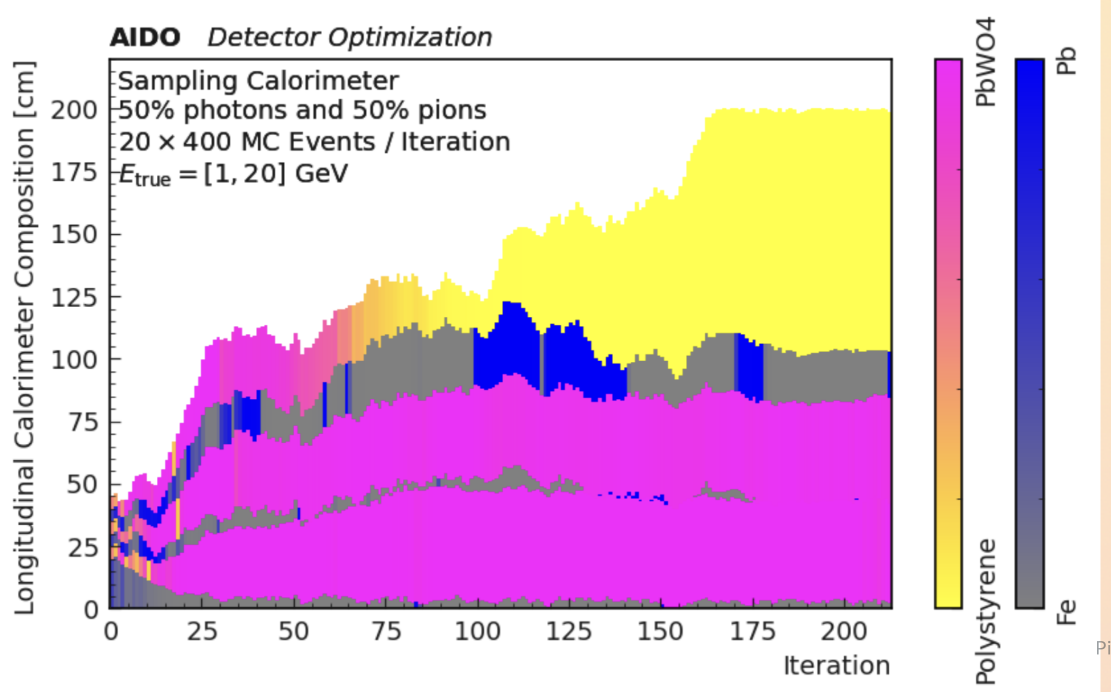
AI-assisted experiment/detector design

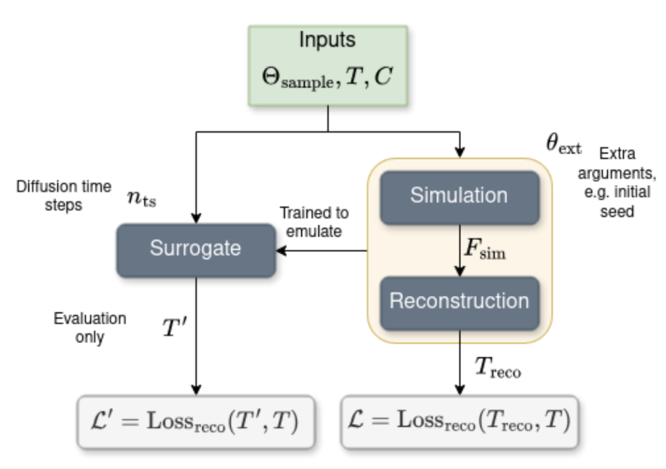
- Optimize a cost function's multidimensional landscape via gradient or reinforcement learning methods

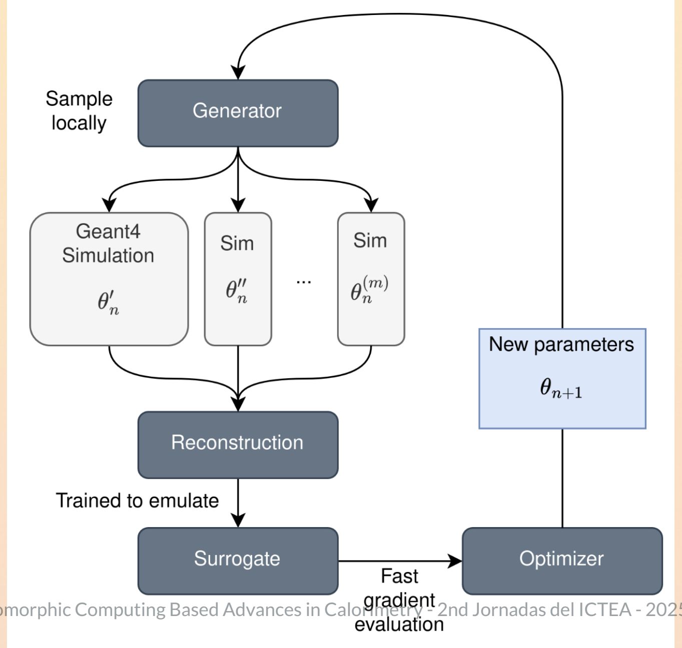


E2E detector optimization K. Schmidt et al. (P.V.) Particles 2025 8(2) 47

- AIDO: Diffusion model replaces the full simulation and reconstruction chain
  - This work: focus on sampling calorimeter
  - General approach extendable to full detector
- Learn directly output of reconstruction step
  - Surrogate via parallelized GEANT4 simulations
- Constrained optimization: 200cm length, 200k EUR budget
  - Better materials (lead for absorber, $PbWO_4$ for scintillator) cost more

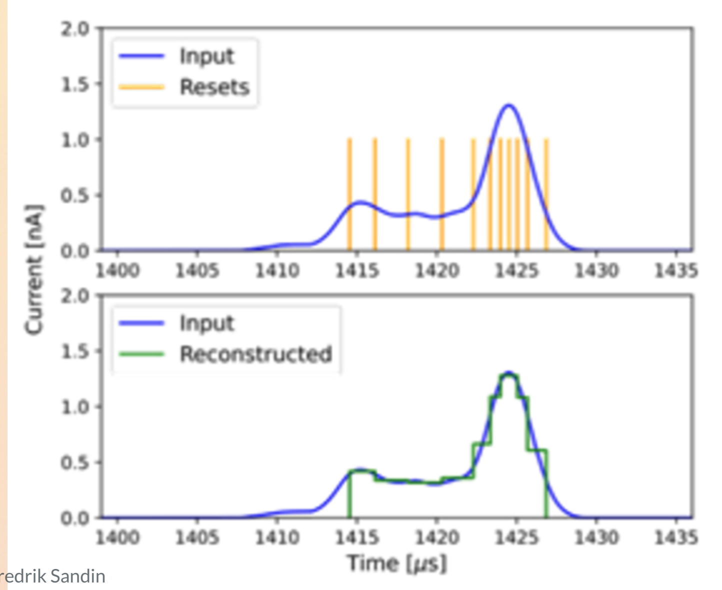


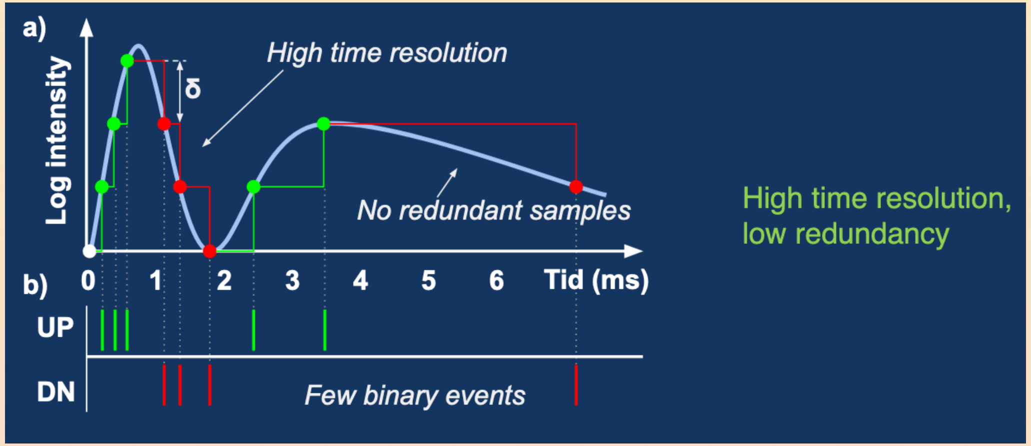




The next developments P.V. 10.22323/1.476.1022

- Further studies for hadron calorimetry ongoing
  - Funding requested for demonstrator
- Scheme can be adapted to CMS-style ECAL and other subsystems, some studies ongoing/starting
- Funding requested for NC-based trigger systems, studies ongoing
- Charge readout in LArTPC for neutrino physics
  - The Q-Pix consortium demonstrated pixel-based readout in ASIC, equivalent to an integrator
  - Neuromorphic systems are well suited to in-chip signal processing
- Large-scale optimization of experiment design

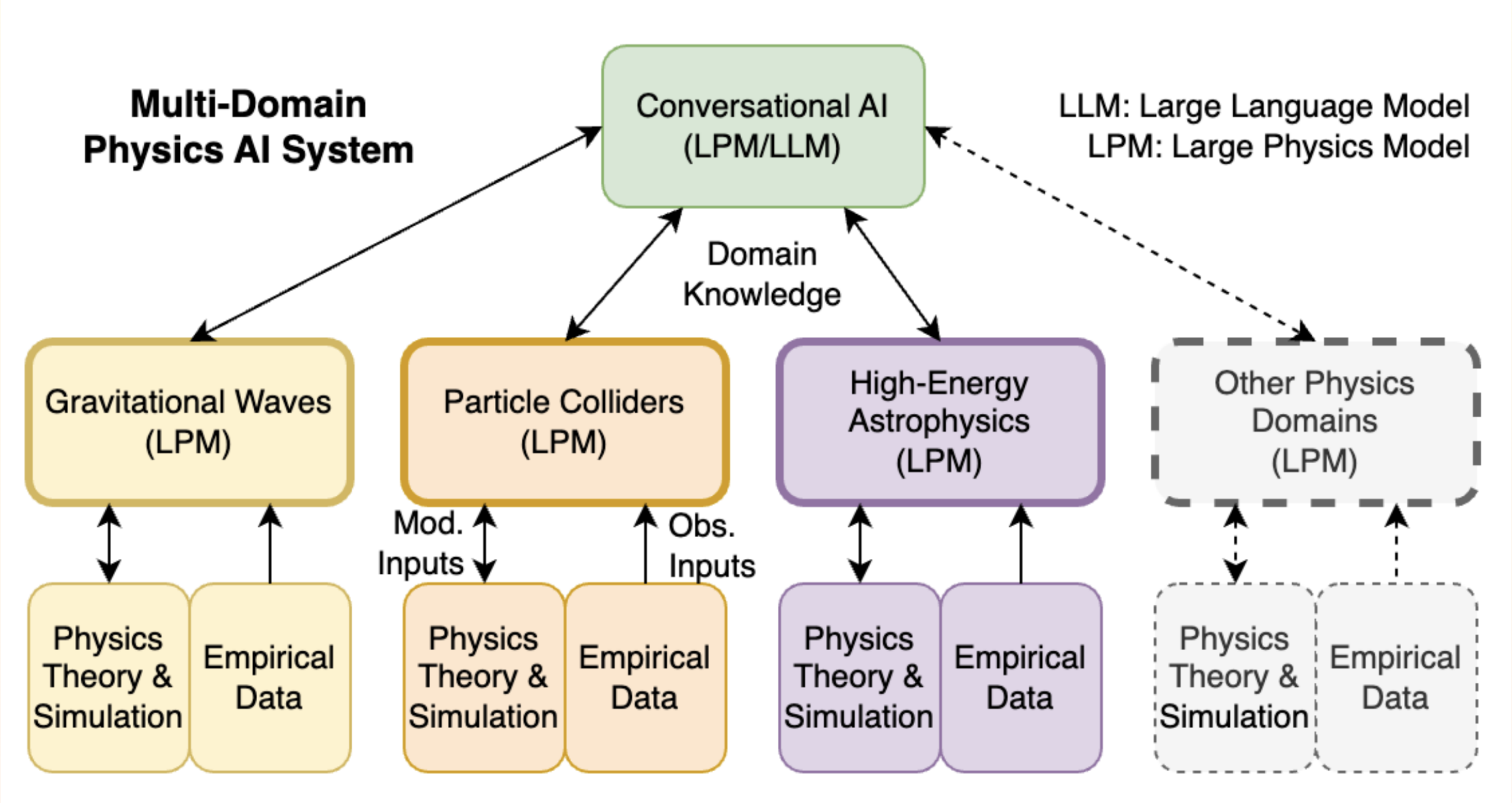




Figures by Shion Kubota and Fredrik Sandin

Component of Large Physics Models? K. G. Barman et al. (P.V.) 2501.05382

- Tailored model for detector optimization may inform future conversational AI physics model



Thank you!

Backup

Autodiff powers most of modern ML

- By design, simple in software

```python
import torch, math
x0 = torch.tensor(1., requires_grad=True)
x1 = torch.tensor(2., requires_grad=True)
p = 2*x0 + x0*torch.sin(x1) + x1**3
print(p)
p.backward()
print(x0.grad, x1.grad)
```

yielding

```
Primal: tensor(10.9093, grad_fn=<AddBackward0>)
Adjoint: tensor(2.9093) tensor(11.5839)
```

- Computational cost of calculating $\mathbf{J}_f(\mathbf{x})$ in $\mathbb{R}^n \times \mathbb{R}^m$ for $f : \mathbb{R}^n \to \mathbb{R}^m$
  - $\mathcal{O}(n \text{ time}(f))$
  - $\mathcal{O}(m \text{ time}(f))$

$$y(\mathbf{x}) = 2x_0 + x_0\, sin(x_1) + x_1^3$$

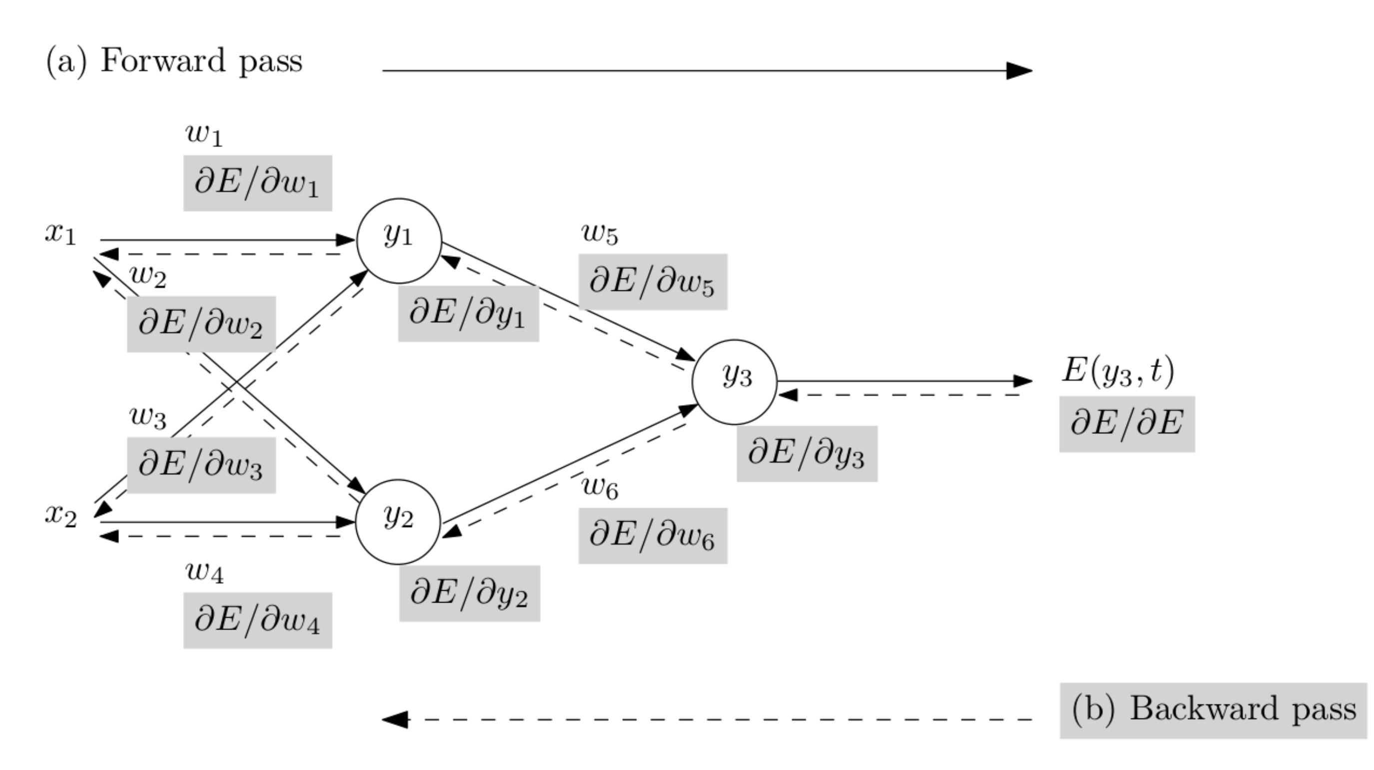
Fwd Primal Trace **Atomic operation**	**Value in** $(1, 2)$	***Fwd Tangent Trace (set*** $\dot{x}_0 = 1$ ***to compute*** $\frac{\partial y}{\partial x_0}$***)*** **Atomic operation**	**Value in** $(1, 2)$
$v_0 = x_0$ $v_1 = x_1$	1 2	$\dot{v}_0 = \dot{x}_0$ $\dot{v}_1 = \dot{x}_1$	1 0
$v_2 = 2v_0$ $v_3 = sin(v_1)$ $v_4 = v_0 v_3$ $v_5 = v_1^3$ $v_6 = v_2 + v_4 + v_5$	2 0.9093 0.9093 8 10.9093	$\dot{v}_2 = 2\dot{v}_0$ $\dot{v}_3 = \dot{v}_1 cos(v_1)$ $\dot{v}_4 = \dot{v}_0 v_3 + v_0 \dot{v}_3$ $\dot{v}_5 = 3\dot{v}_1 v_1^2$ $\dot{v}_6 = \dot{v}_2 + \dot{v}_4 + \dot{v}_5$	2×1 0×-0.41 $1 \times 0.9093 + 1 \times 0$ $3 \times 0 \times 4$ $2 + 0.9093 + 0$
$y = v_6$	10.9093	$\dot{y} = \dot{v}_6$	2.9093

Fwd Primal Trace **Atomic operation**	**Value in** $(1, 2)$	***Rev Adjoint Trace (set*** $\bar{y} = 1$ ***to compute*** $\frac{\partial v}{\partial y}$***)*** **Atomic operation**	**Value in** $(1, 2)$
$v_0 = x_0$ $v_1 = x_1$	1 2	$\bar{x}_0 = \bar{v}_0$ $\bar{x}_1 = \bar{v}_1$	2.9093 11.5839
$v_2 = 2v_0$ $v_3 = sin(v_1)$ $v_4 = v_0 v_3$ $v_5 = v_1^3$ $v_6 = v_2 + v_4 + v_5$	2 0.9093 0.9093 8 10.9093	$\bar{v}_0 = \bar{v}_0 + \bar{v}_2 \partial v_2 / \partial v_0$ $\bar{v}_0 = \bar{v}_4 \partial v_4 / \partial v_0$ $\bar{v}_1 = \bar{v}_1 + \bar{v}_3 \partial v_3 / \partial v_1$ $\bar{v}_1 = \bar{v}_5 \partial v_5 / \partial v_1$ $\bar{v}_2 = \bar{v}_6 \partial v_6 / \partial v_2$ $\bar{v}_3 = \bar{v}_4 \partial v_4 / \partial v_3$ $\bar{v}_4 = \bar{v}_6 \partial v_6 / \partial v_4$ $\bar{v}_5 = \bar{v}_6 \partial v_6 / \partial v_5$	$\bar{v}_0 + \bar{v}_2 \times 2 = 2.9093$ $\bar{v}_4 \times v_3 = 0.9093$ $\bar{v}_1 + \bar{v}_3 \times cos(v_1) = 11.5839$ $\bar{v}_5 \times 3v_1^2 = 12$ $\bar{v}_6 \times 1 = 1$ $\bar{v}_4 \times v_0 = 1$ $\bar{v}_6 \times 1 = 1$ $\bar{v}_6 \times 1 = 1$
$y = v_6$	10.9093	$\bar{v}_6 = \bar{y}$	1

Image by P. Vischia, book in preparation

Empirical Risk Minimization Vapnik, '90s

$$\mathbf{J}(\mathbf{W}) = \frac{1}{n}\sum_{i=1}^{n}\mathcal{L}(f(x^{(i)};\mathbf{W}), y^{*(i)}), \qquad \mathbf{W}^0 = argmin_{\mathcal{W}}\mathbf{J}(\mathbf{W}), \qquad \mathbf{W} \leftarrow \mathbf{W} + \eta\frac{\partial \mathbf{J}(\mathbf{W})}{\partial \mathbf{W}}$$



- Efficient matrix multiplication in dedicated hardware (GPUs, FPGAs)