

Learning the 'Match' Manifold to Accelerate Template Bank Generation

Introducing LearningMatch and TemplateGeNN...

Susanna Green, Andrew Lundgren, Xan Morice-Atkinson

LearningMatch: The Aim

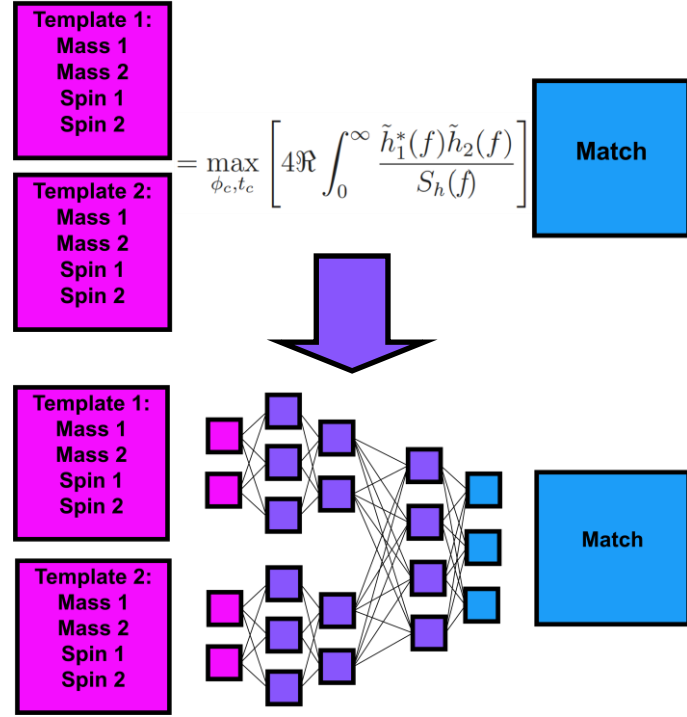


Figure 1: The aim of LearningMatch is to learn the mathematical relationship between the template parameters (specifically the mass and aligned spin of the black holes) and the 'match'.

TemplateGeNN: The Results

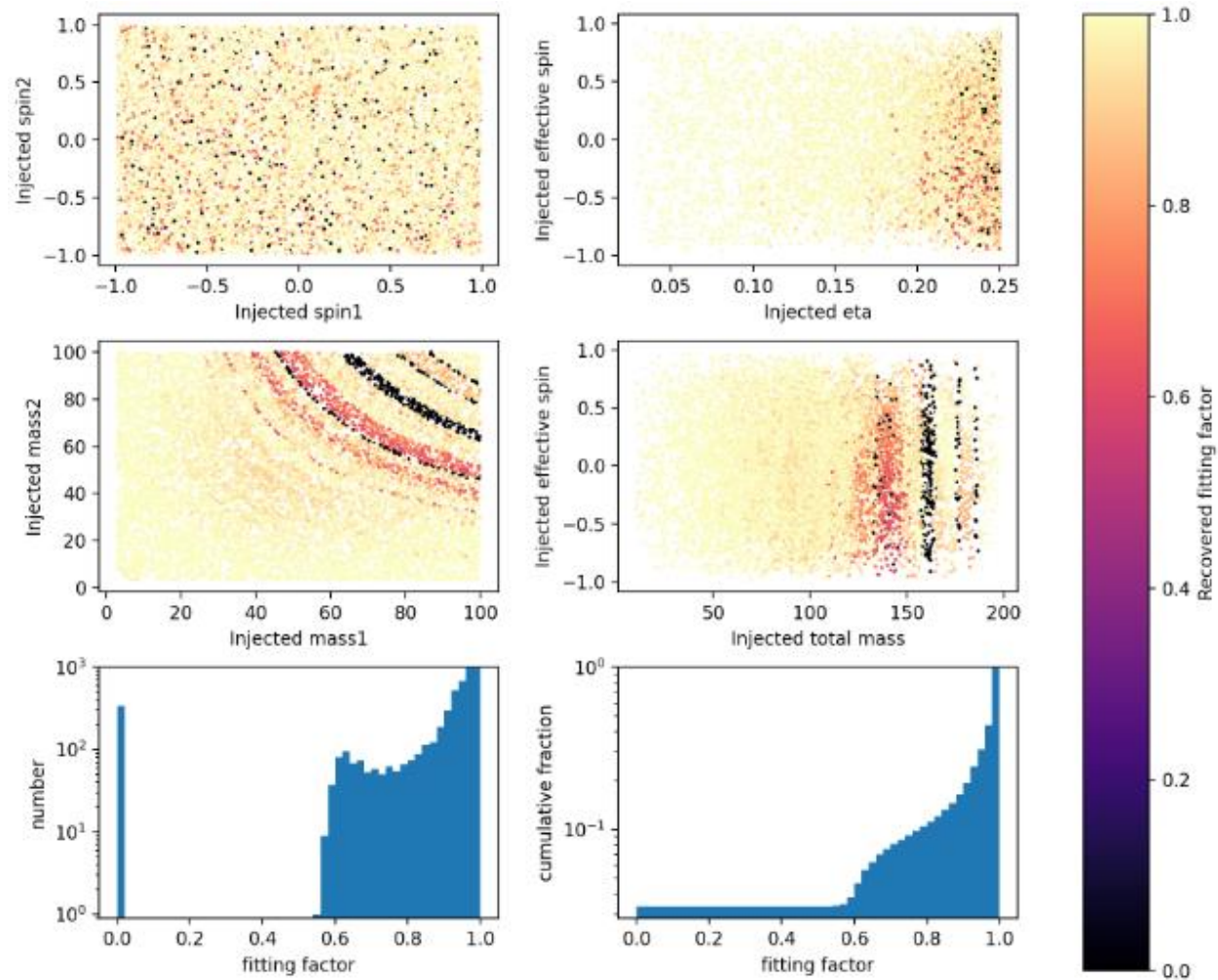


Figure 2: Results from PyCBC template bank verifier.

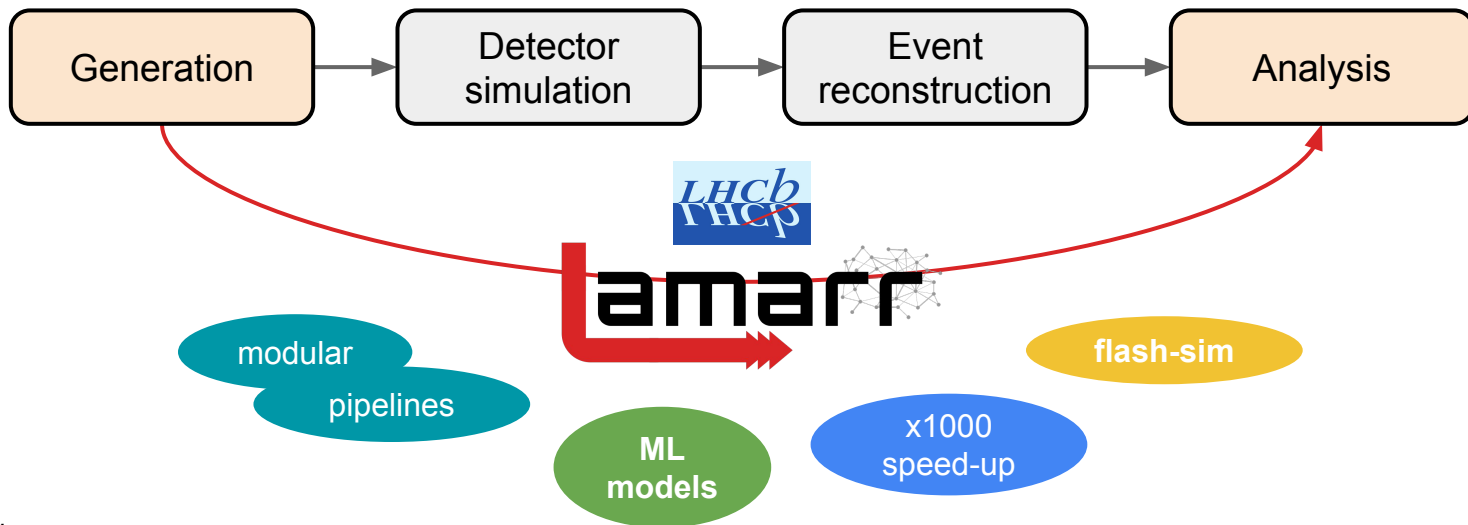


- ✓ Simulations are **crucial** for High Energy Physics experiments
- ✓ Detailed Simulation is computationally **very expensive**
- ✓ Detailed Simulation **not scale** for the future experiment demands

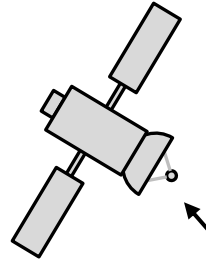


Viable solutions?

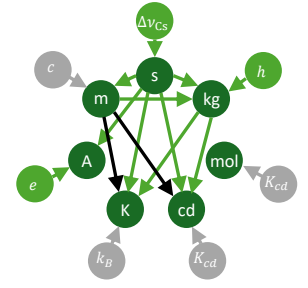
- ❌ Renouncing to increase statistics → **lower demand**
- ✅ Developing **faster options** for simulation → at LHCb, **Lamarr**



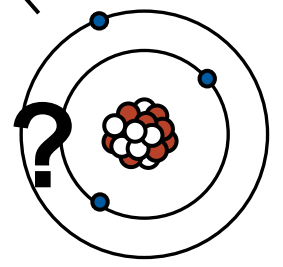
Astrophysics



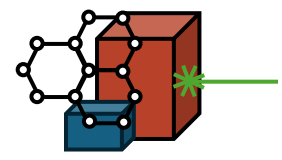
SI-Units



Structure
of Ions and
Atoms

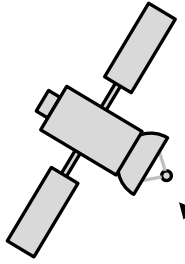


Material science
Plasma diagnostics

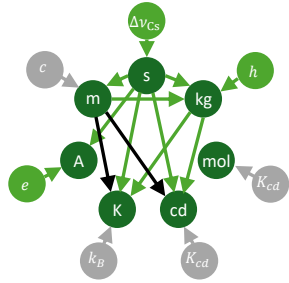


...

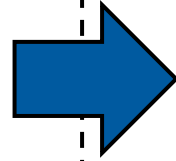
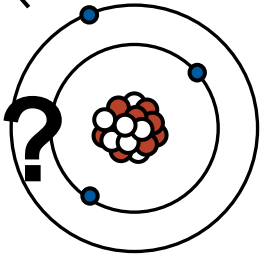
Astrophysics



SI-Units

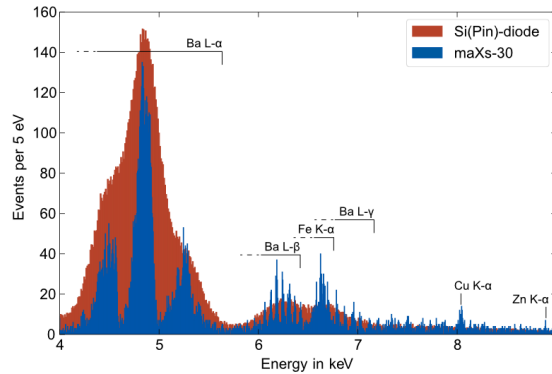


Structure of Ions and Atoms



Bound-state
Quantum Electro Dynamics

$$\propto Z^4$$

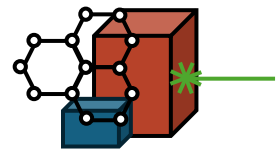


Close to Schwinger-limit

Transition ≈ 100 keV

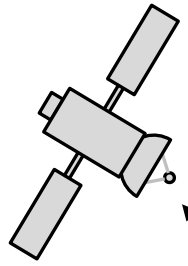
2nd order QED ≈ 1 eV

\Rightarrow Precision

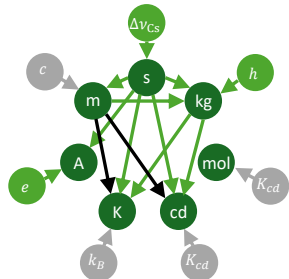


Material science
Plasma diagnostics

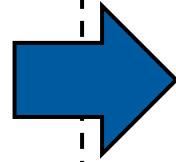
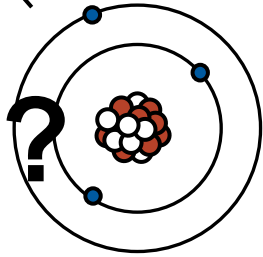
Astrophysics



SI-Units

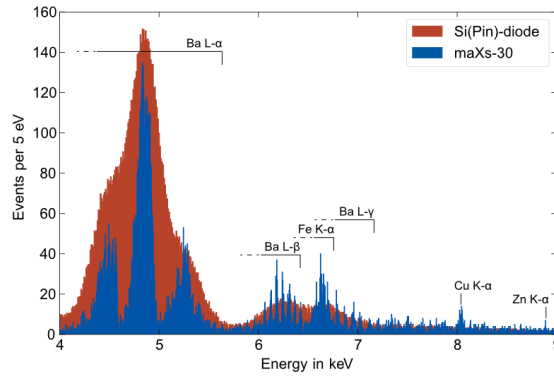


Structure of Ions and Atoms



Bound-state Quantum Electro Dynamics

$$\propto Z^4$$



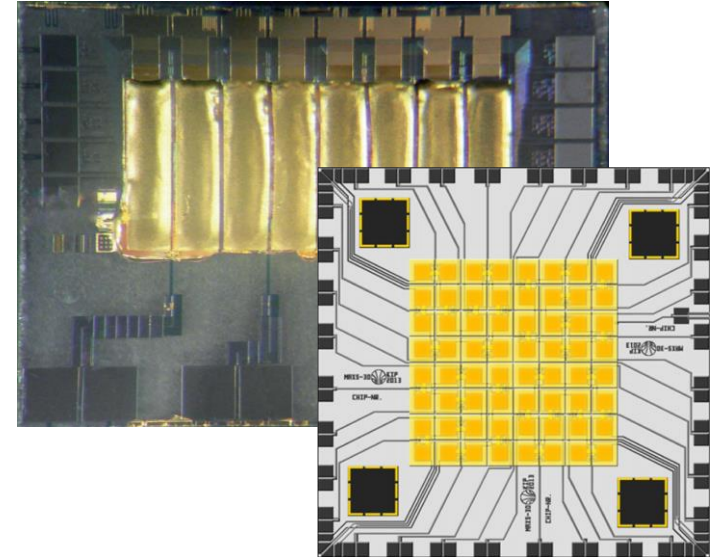
Close to Schwinger-limit
Transition ≈ 100 keV
2nd order QED ≈ 1 eV

\Rightarrow Precision

Micro-Calorimeters

» Small thermometers for measuring single particle energies «

maXs-200 (8 pixels)



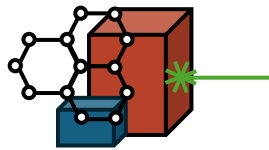
maXs-30 (64 px)

High energy resolution

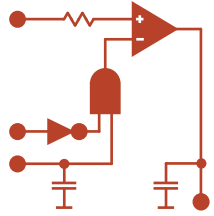


High quantum efficiency

Material science
Plasma diagnostics



High sensitivity...
... also to **noise**



Signal processing:

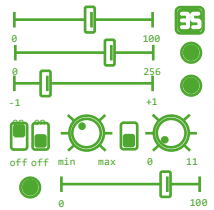
Analog



```
01 float get_e
02 {
03     uint32_t
04     for ( uin
05     {
06         float v
07         mwd_maf
```

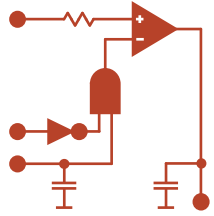
Digital

Compensate
artifacts



Optimization of many
parameters / pixel
⇒ **Future: More Pixels**
> 1000 px

High sensitivity...
... also to **noise**



Signal processing:

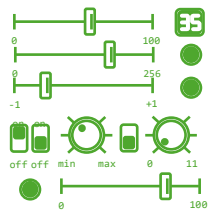
Analog



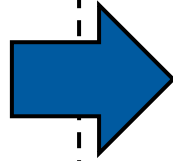
```
01 float get_e
02 {
03   uint32_t
04   for ( uin
05   {
06     float v
07     mwd_maf
```

Digital

Compensate
artifacts



Optimization of many
parameters / pixel
⇒ **Future: More Pixels**
> 1000 px



Solution:
Artificial Intelligence?



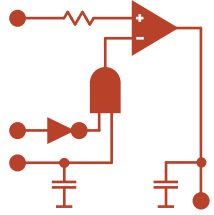
Parameter optimization
Signal characterization
1D temporal analysis

...



⇒ **Lots of potential!**

High sensitivity...
... also to noise



Signal processing:

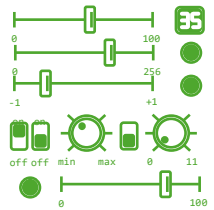
Analog



```
01 float get_e
02 {
03   uint32_t
04   for (uin
05   {
06     float v
07     mwd_maf
```

Digital

Compensate artifacts



Optimization of many parameters / pixel

⇒ Future: More Pixels
> 1000 px

Solution:
Artificial Intelligence?

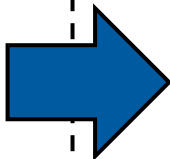


Parameter optimization
Signal characterization
1D temporal analysis

...



⇒ Lots of potential!



Utilizing Artificial Intelligence Technologies for the Enhancement of X-ray Spectroscopy with Metallic-Magnetic Calorimeters (MMC) 110

M. O. Herdrich^{1,2,3,4}, Ph. Pfäffli^{1,2,3}, G. Weber^{1,2,3}, D. Hengstler^{1,2}, A. Fleischmann^{1,2}, C. Ernst^{1,2}, and Th. Söhler^{1,2,3,4}

¹Helmholtz-Institute Jena, Jena, Germany
²Faculty for Physics and Quantum Electronics, Friedrich-Schiller-Universität Jena, Germany
³DFG Collaborative Center for Matter at Research, Jena, Germany
⁴DFG Collaborative Center for Matter at Research, Jena, Germany

What are MMCs?

Small thermometer for measuring single particle energies

max: 30 (MMC)

Developed within SPARC collaboration: max8

Why use MMCs?

Outstanding properties associated with MMCs (1,2,3):

- Fast signal rise time up to $\tau_{90} \approx 100$ ns
- High energy resolution $\Delta E_{FWHM} = 1.6$ eV @ 6 keV
- Excellent linearity $\Delta E / E < 5.9\%$ @ 60 keV

Several successful benchmark campaigns on ion accelerators and an ion trap [4,5,6]

Measurement of K α line-splitting in U⁹² at CRYRING of FAIR [7]

Challenges of using MMCs?

However: Best performance is only achievable ... with a transition Analog ⇒ Digital signal processing

MMC is susceptible to environmental changes... vibrations, magnetic flux, etc. ⇒ corrections needed

Development of a complex signal analysis framework

- Test and improvement through experiments [8]
- Requires multitude of numerical values and hardware settings to be optimized
- Partial automation but several manual steps involved for individual pixels
- Future: From few to many pixels. ⇒ Setup 'by hand' is no longer feasible

How could AI be involved?

Hardware

- Auto-tuning of read-out SQUID electronics
- Goal: Use Feedback between operation parameters and signal quality to improve the performance ⇒ Reinforcement Learning (RL)

Software

- Simple signal characterization
- Goal: Automate the rejection of false-positives and improvement of trigger timing capabilities ⇒ Classification
- Numerical parameter optimization
- Goal: Automate and improve the optimization of the finite response filter (FIR) used for signal analysis ⇒ Reinforcement Learning
- Full signal analysis
- Train a neural network on the specific MMC detector pulse characteristics
- Goal: Extract relevant measurements directly
- Bonus: Use NN to synthesis MMC pulses for testing ⇒ 1D temporal analysis

First steps: Full signal analysis ⇒ Extraction of pulse amplitude

Setup: Synthesize MMC detector pulses with known parameters Perform RL on 1D temporal signal processing network [9]:

Use n layers of a convolutional neural network (CNN) followed by m layers of a simple perceptron ⇒ Amplitude

Under optimal conditions: FIR-filter outperforms AI in reading energy from pulses However: AI is less effected by jitter of pulse timings Next: Optimize the model

(Lower is better)

[1] S. Karpel et al., Supercond. Sci. Technol. 28 (2015)
[2] C. Ernst et al., J. Low Temp. Phys. 187 (2015)
[3] M. Herdrich et al., Thin. Appl. Phys. 98 (2005)
[4] M. O. Herdrich et al., Phys. Rev. Lett. 102 (2009)
[5] M. O. Herdrich et al., Eur. Phys. J. D 97 (2012)
[6] M. O. Herdrich et al., Phys. Rev. D 85 (2012)
[7] Ph. Pfäffli et al., Phys. Rev. Lett. 109 (2012)
[8] M. O. Herdrich et al., Thin. Appl. Phys. 98 (2005)
[9] S. Karpel et al., J. Low Temp. Phys. 187 (2015)

Quark/Gluon Discrimination and Top Tagging with Dual Attention Transformer

— EuCAIFCon24 —

Daohan Wang

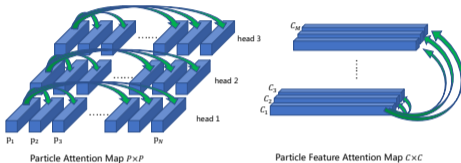
Institute of High Energy Physics (HEPHY), Austrian Academy of Sciences (OeAW)

April 30, 2024

Dual Attention Mechanism

$$\mathcal{A}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Concat}(\text{head}_1, \dots, \text{head}_{N_h})$$

$$\text{where head}_i = \text{softmax} \left[\frac{\mathbf{Q}_i (\mathbf{K}_i)^T}{\sqrt{C_h}} + \mathbf{U}_1 \right] \mathbf{V}_i$$



$$\mathcal{A}(\mathbf{Q}_i, \mathbf{K}_i, \mathbf{V}_i) = \text{softmax} \left[\frac{\mathbf{Q}_i^T \mathbf{K}_i}{\sqrt{C}} + \mathbf{U}_2 \right] \mathbf{V}_i^T$$

Particle interaction matrix \mathbf{U}_1 :

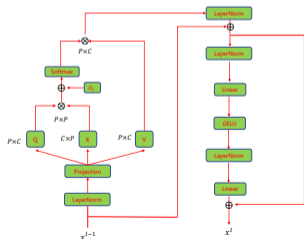
$$\Delta R = \sqrt{(y_a - y_b)^2 + (\phi_a - \phi_b)^2},$$

$$k_T = \min(p_{T,a}, p_{T,b}) \Delta,$$

$$z = \min(p_{T,a}, p_{T,b}) / (p_{T,a} + p_{T,b}),$$

$$m^2 = (E_a + E_b)^2 - \|\mathbf{p}_a + \mathbf{p}_b\|^2,$$

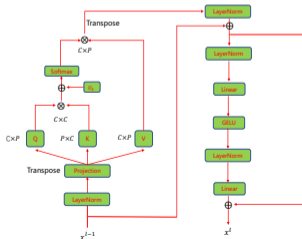
$$\Delta p_T = |p_{T,a} - p_{T,b}|$$



Channel interaction matrix \mathbf{U}_2 :

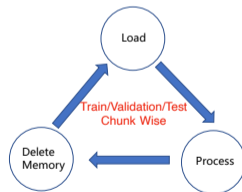
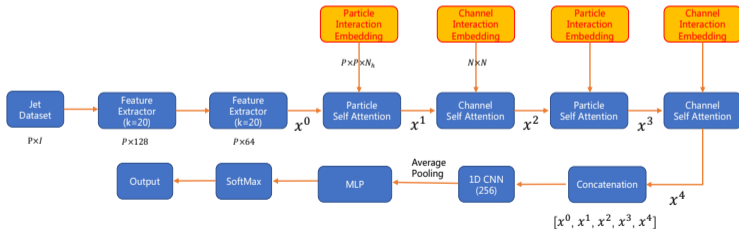
$$\{E_J, p_{T_J}, \sum p_{T_f}, \sum E_f, \Delta \eta, \Delta \phi, \Delta \bar{R}, \text{PID}\}$$

where $\Delta \eta$, $\Delta \phi$ and $\Delta \bar{R}$ correspond to the transverse momentum weighted sum of the $\Delta \eta$, $\Delta \phi$, ΔR of all the constituent particles inside the input jet, respectively. Here $\Delta \eta$, $\Delta \phi$ and ΔR refer to the distances in the $\eta - \phi$ space between each constituent particle and the input jet.



Model Architecture

- Input features: $\log E$, $\log p_T$, $\frac{p_T}{E}$, $\frac{E}{E_j}$, $\Delta\eta$, $\Delta\phi$, ΔR , PID of leading 100 particles.
- The particle attention module ($P \times P$ attention map) and the channel attention module ($C \times C$ attention map) are stacked while maintaining a consistent feature dimension of $N = 64$ and they can complement each other.
- Particle - Dual Attention Transformer: 2 Feature Extractor (1 EdgeConv + 3 Conv2D + 1 AvgPool) + 2 Particle Attention modules + 2 Channel Attention modules + 1D CNN + MLP.



Chunk Loading Strategy