AI algorithms and techniques for muon reconstruction in the future ATLAS trigger

SAPIENZA
UNIVERSITÀ DI ROMA

tCSC Student Lightning
Graziella Russo - 04/05/22

INFN

ATLAS
ROMA1

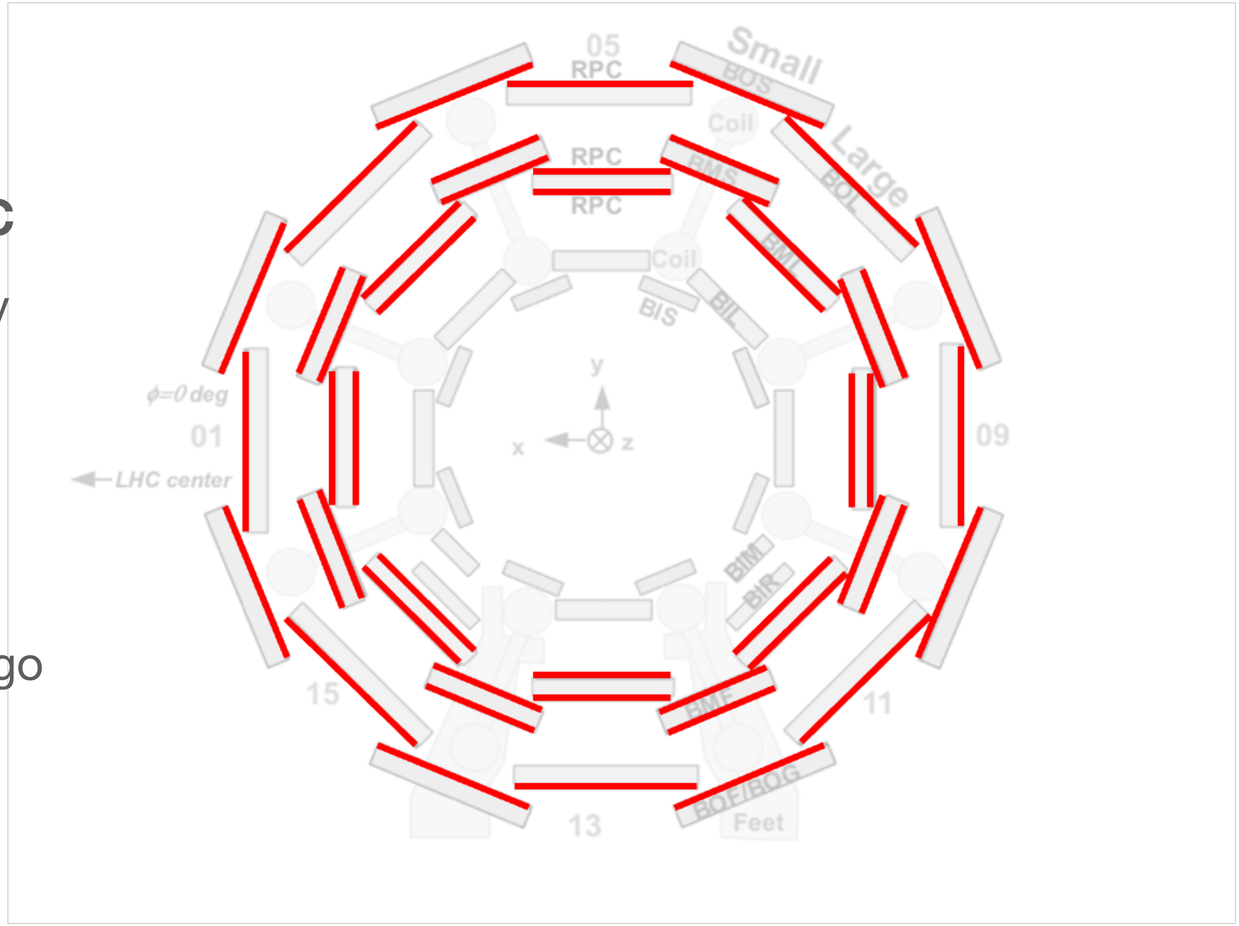
ATLAS Muon Spectrometer Barrel Phase II upgrade

2025-27 **High-Lumi LHC**

Instantaneous luminosity

$\mathscr{L} = 5 \cdot 10^{34}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$

ATLAS detectors will undergo an upgrade too



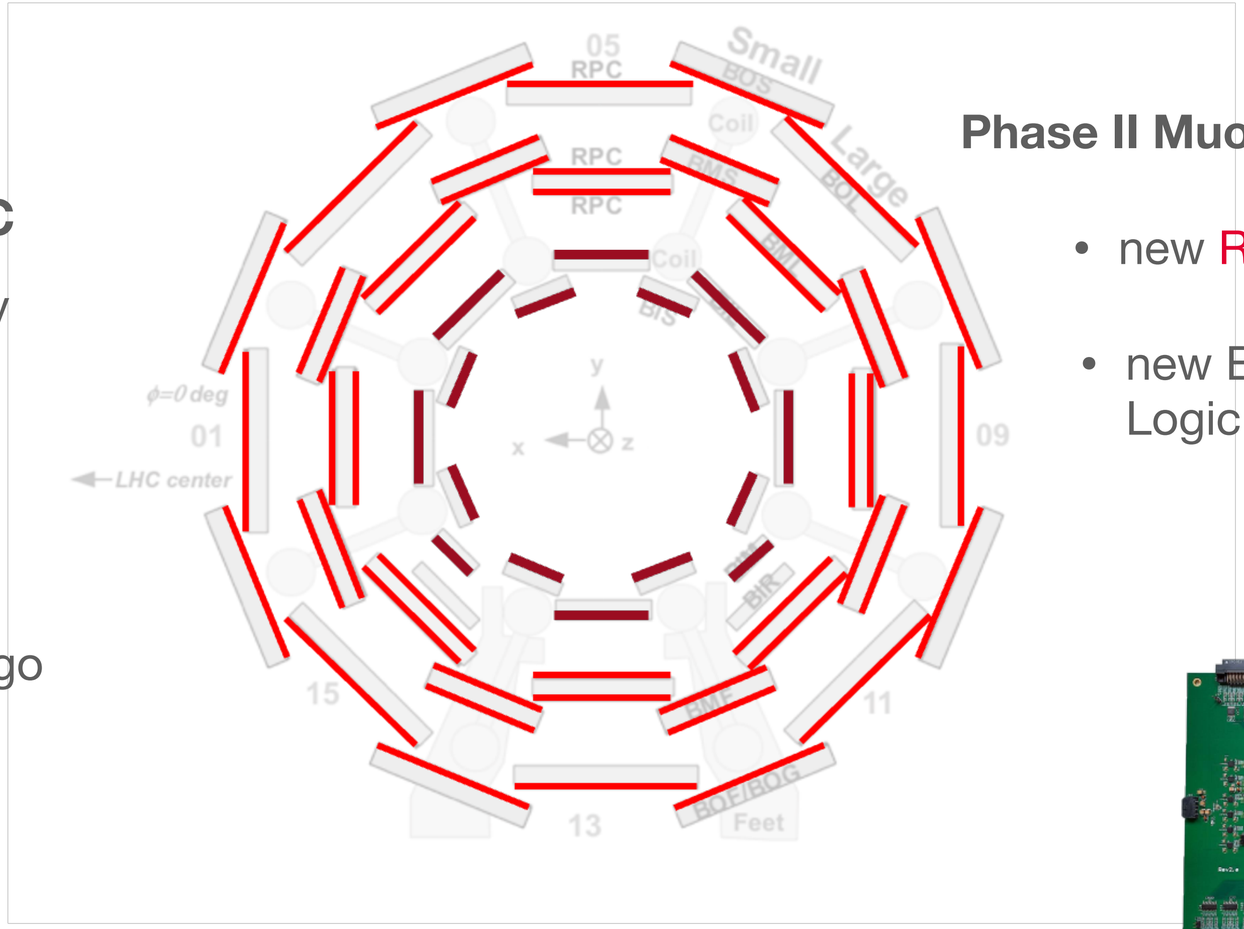
ATLAS Muon Spectrometer Barrel Phase II upgrade

2025-27 **High-Lumi LHC**

Instantaneous luminosity

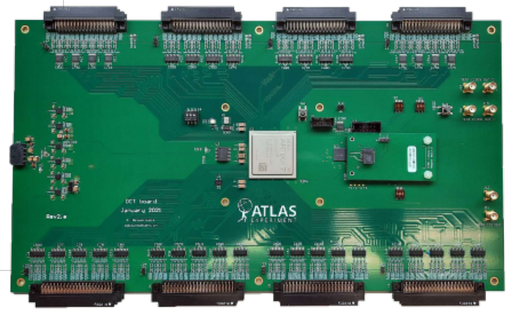
$\mathscr{L} = 5 \cdot 10^{34}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$

ATLAS detectors will undergo an upgrade too

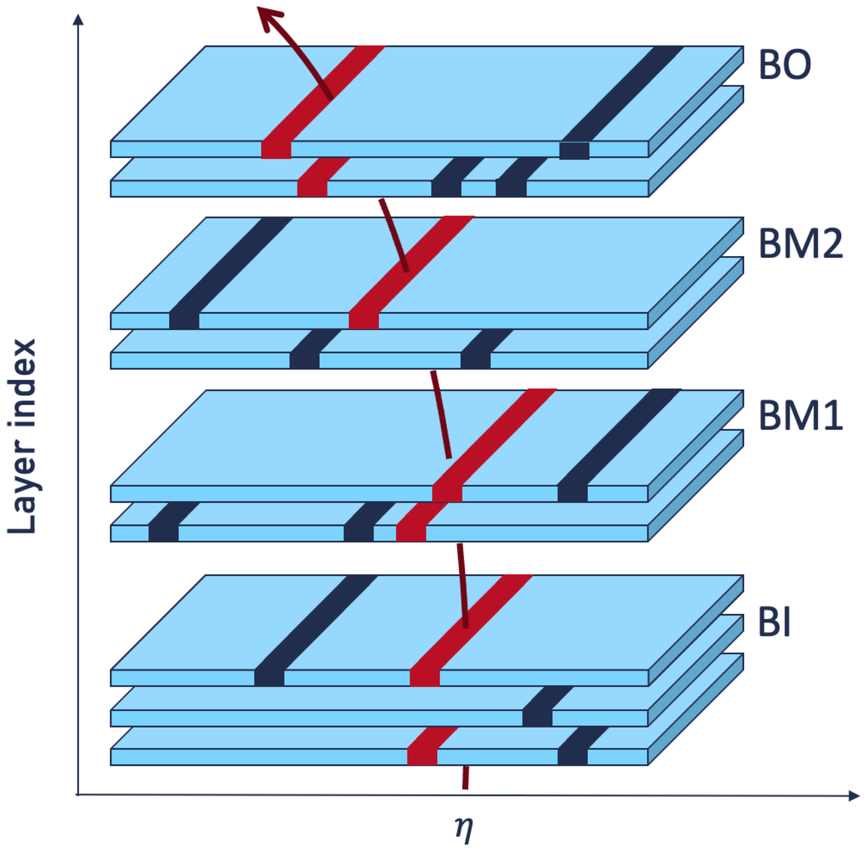


Phase II Muon Spectrometer Barrel

- new RPC triplet
- new Barrel readout and Sector Logic on FPGAs



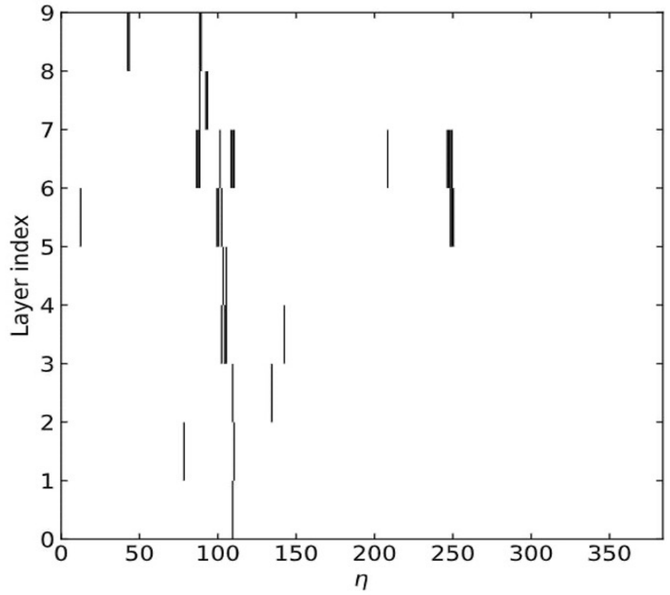
From the data to the ML dataset



Dataset of ~200k images with:

- random noise-only hits
- single muon tracks ($p_T \in [2.5, 20]$ GeV) + noise hits

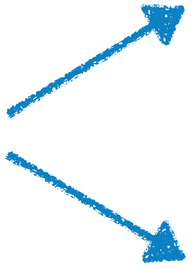
Target label:
$p_T, \eta, (q)$



Extreme experimental requirements

- Fit within the Virtex UltraScale+ 13 FPGA resources;
- Maximum latency (= time interval of algorithmic response) allowed of ∼400 ns;
- Fake efficiency (= trigger efficiency on noisy events) < 2‰

Train a CNN model (VGG-like)
+ help from compression techniques



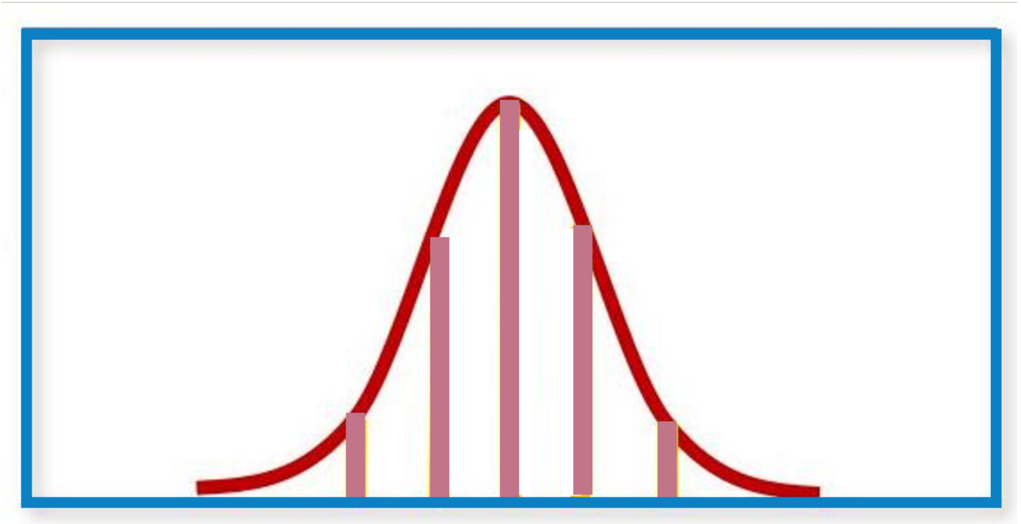
Quantization

Knowledge Distillation

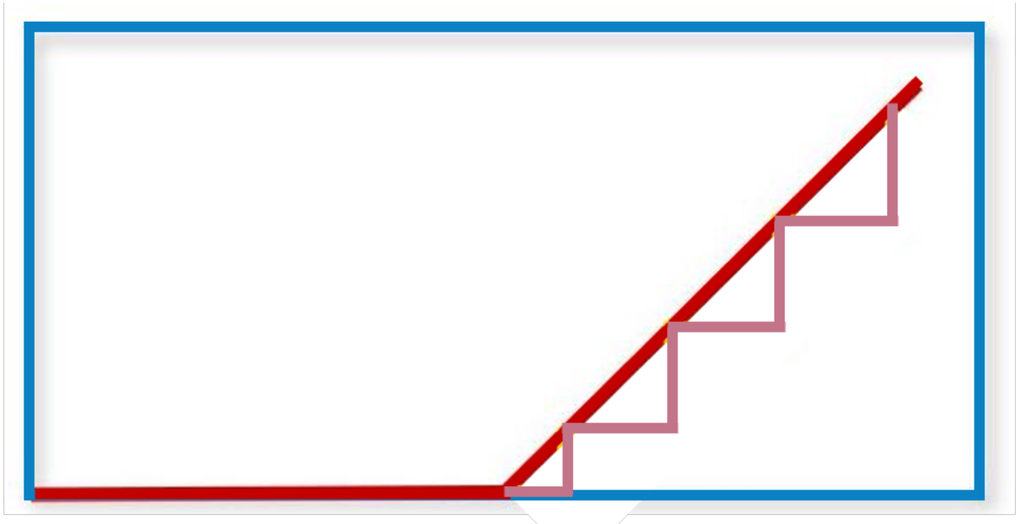
Quantization

The model weights and activations (except the output layer) are not trained in the usual 32 or 64 bit precision floating point arithmetic, but by fixing a lower number of bit $n_{bits} = 3, 4$

The QKeras library does the job

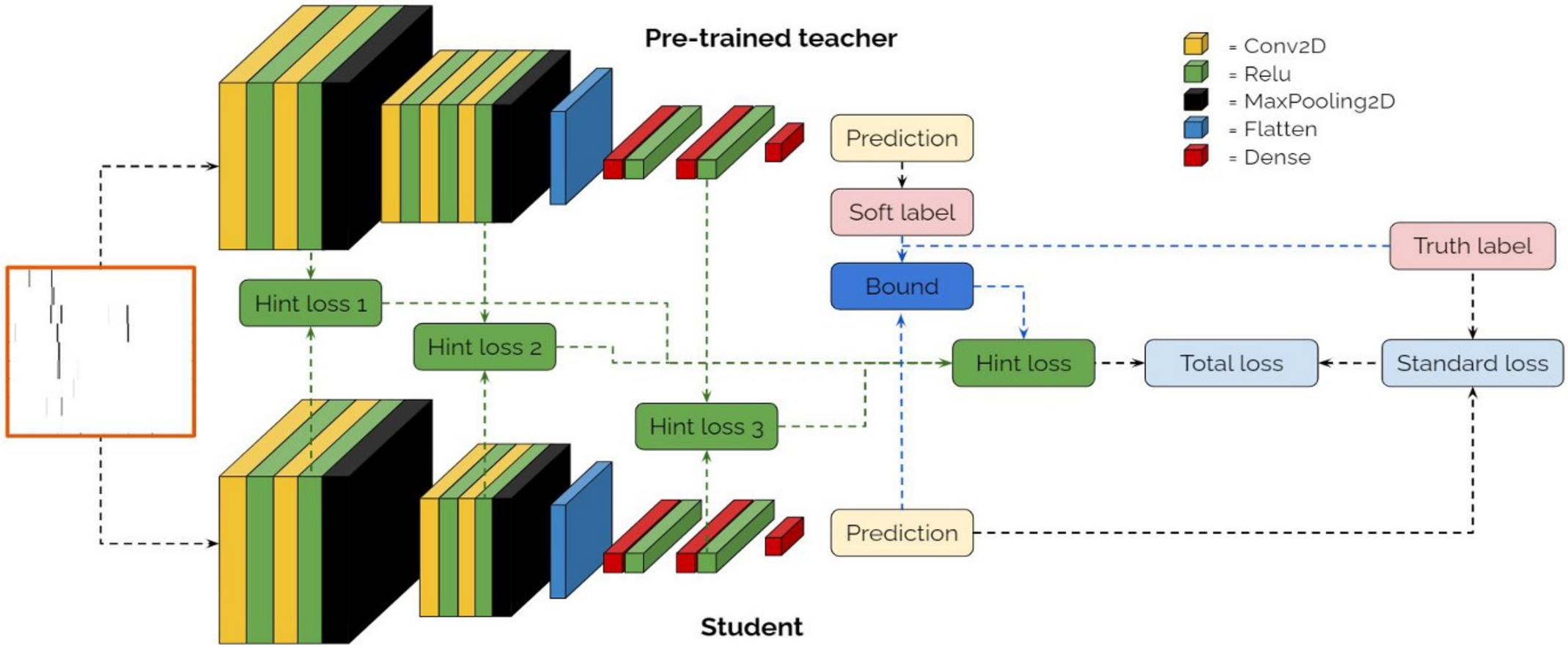


Weight value



Activation function

Knowledge Distillation (KD)



Models and performance

Layer type	**Teacher** 9×16 Output shape	Weights	**Student** 9×16 Output shape	Weights
Input	(9, 16, 1)	0	(9, 16, 1)	0
Conv2D	(9, 16, 10)	100	(7, 14, 1)	10
Conv2D	(9, 16, 10)	910	(5, 12, 1)	10
MaxPooling2D	(9, 8, 10)	0		
	Activation: ReLU, padding: same		Activation: ReLU, padding: valid	
Conv2D	(9, 8, 17)	1547	(3, 10, 6)	60
Conv2D	(9, 8, 17)	2618	(1, 8, 6)	330
Conv2D	(9, 8, 17)	2618		
MaxPooling2D	(4, 4, 17)	0		
	Activation: ReLU, padding: same		Activation: ReLU, padding: valid	
Flatten	272	0	48	0
Dense	10	2730	10	490
Dense	10	110	10	110
	Activation: ReLU		Activation: ReLU	
Dense	2	22	2	22
Model total		**10655**		**732**