# Review of CMS HCAL reconstruction performance

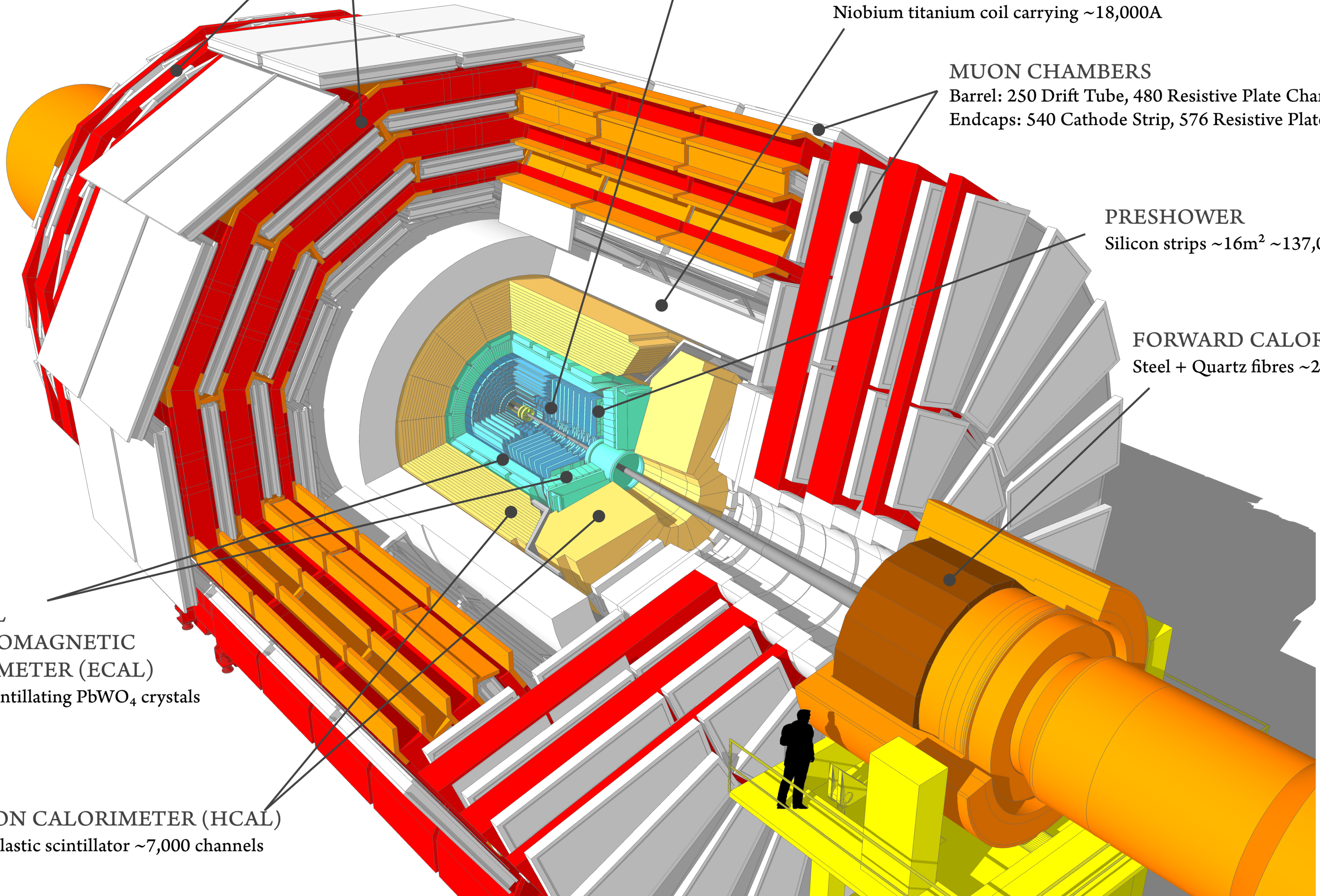
Hui Wang (王徽)
Nanjing Normal University
CLHCP2024, Qingdao

# Outline

- Introduction
- Reconstruction algorithms
- Reconstruction performance
- Reconstruction with ML

# CMS DETECTOR

| | |
|---|---|
| Total weight | : 14,000 tonnes |
| Overall diameter | : 15.0 m |
| Overall length | : 28.7 m |
| Magnetic field | : 3.8 T |



STEEL RETURN YOKE
12,500 tonnes

SILICON TRACKERS
Pixel (100x150 μm) ~1m$^2$ ~66M channels
Microstrips (80x180 μm) ~200m$^2$ ~9.6M channels

SUPERCONDUCTING SOLENOID
Niobium titanium coil carrying ~18,000A

MUON CHAMBERS
Barrel: 250 Drift Tube, 480 Resistive Plate Chambers
Endcaps: 540 Cathode Strip, 576 Resistive Plate Chambers

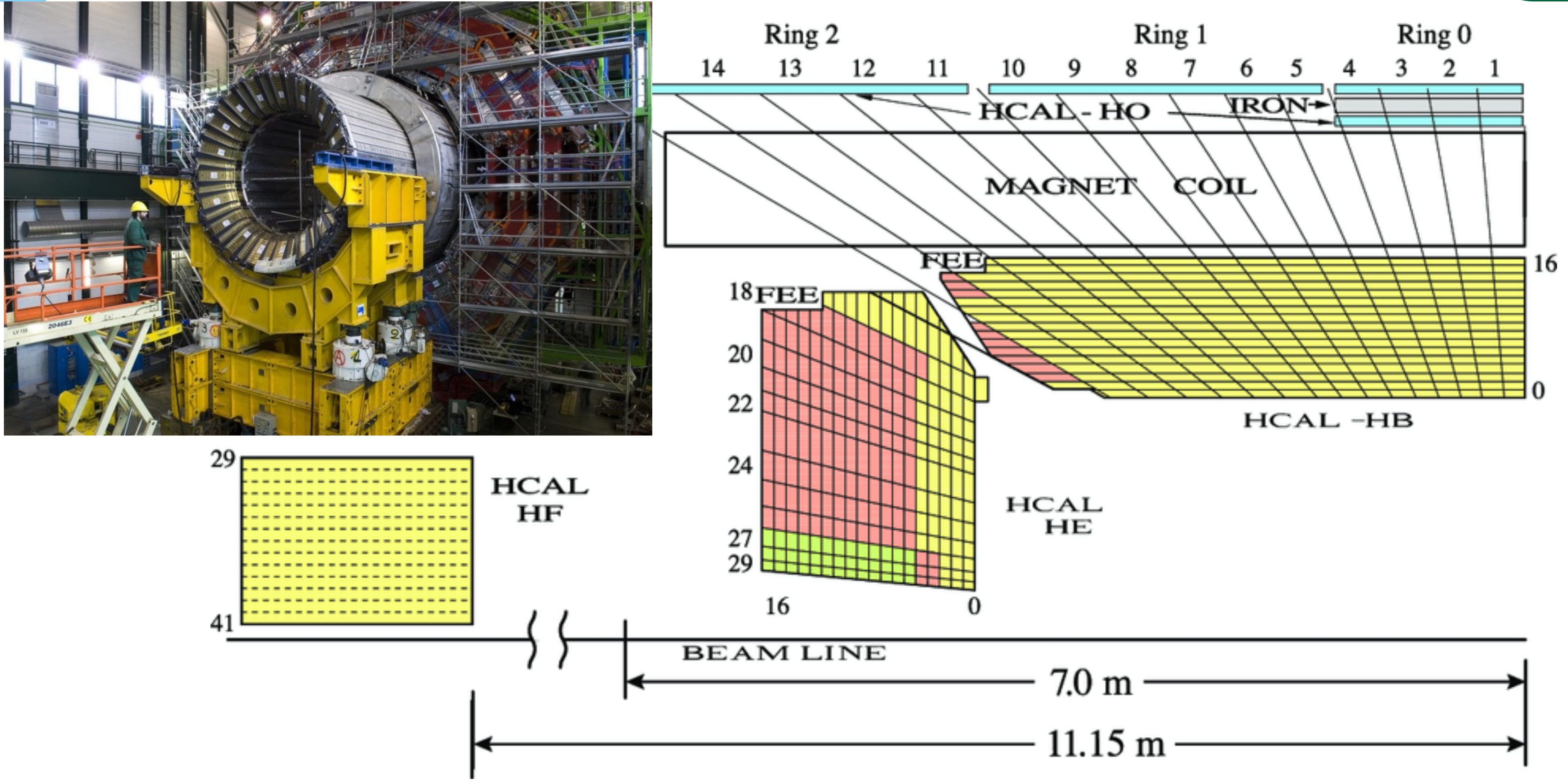
PRESHOWER
Silicon strips ~16m$^2$ ~137,000 channels

FORWARD CALORIMETER
Steel + Quartz fibres ~2,000 Channels

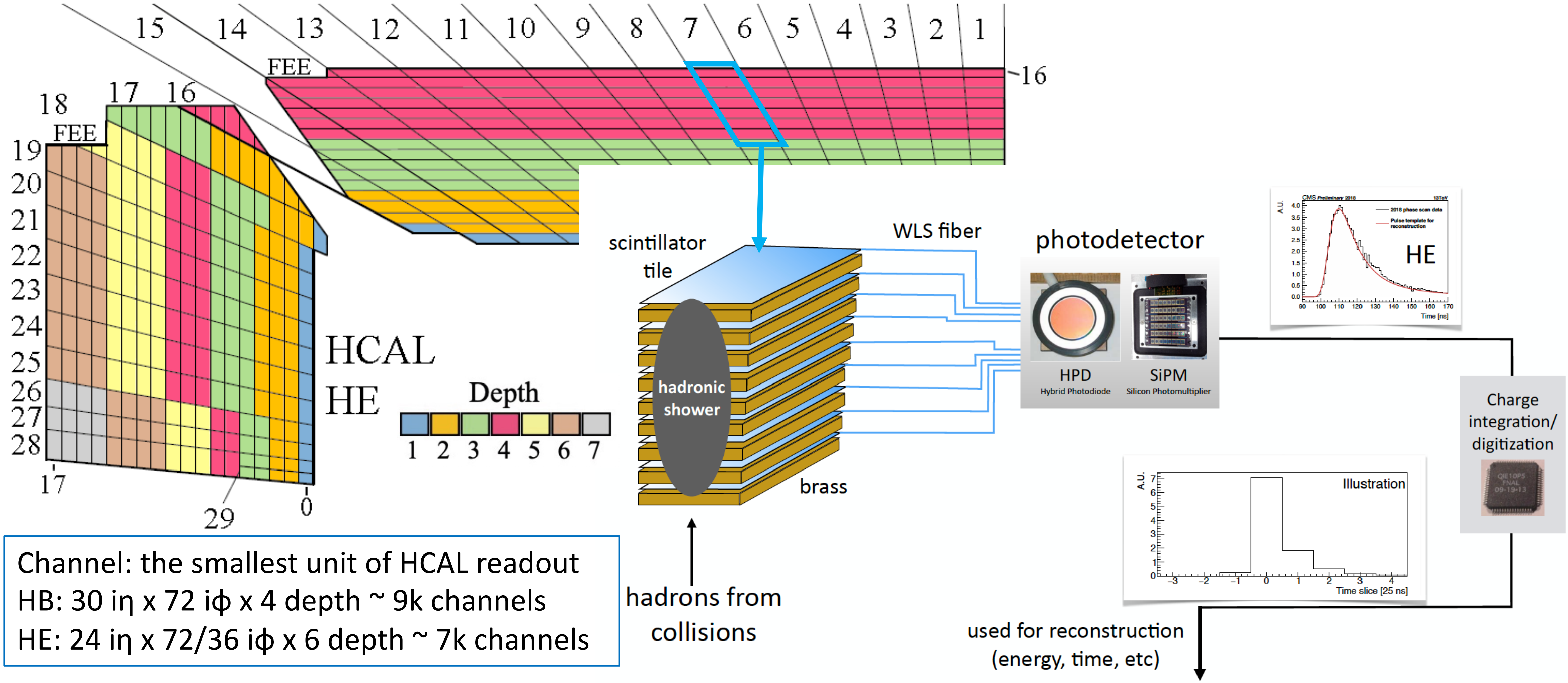
CRYSTAL
ELECTROMAGNETIC
CALORIMETER (ECAL)
~76,000 scintillating $PbWO_4$ crystals

HADRON CALORIMETER (HCAL)
Brass + Plastic scintillator ~7,000 channels

# HCAL structure



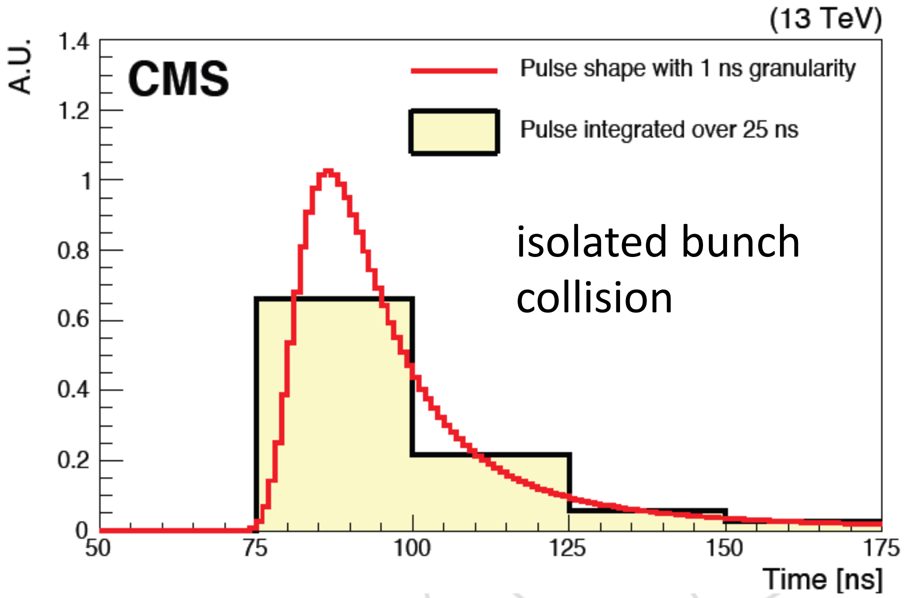
# HCAL Readout Chain



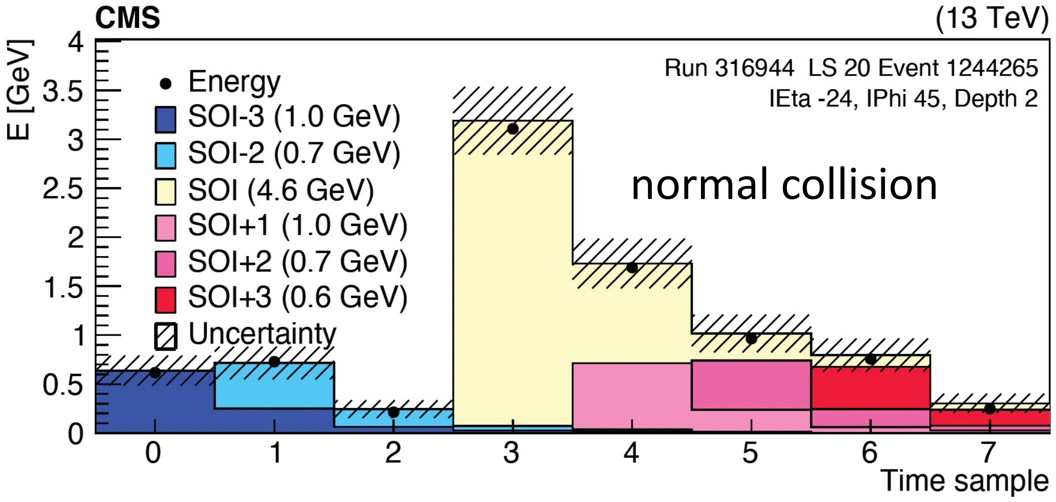
Channel: the smallest unit of HCAL readout
HB: 30 iη x 72 iϕ x 4 depth ~ 9k channels
HE: 24 iη x 72/36 iϕ x 6 depth ~ 7k channels

# Reconstruction algorithms

# HCAL Energy Reconstruction



isolated bunch collision

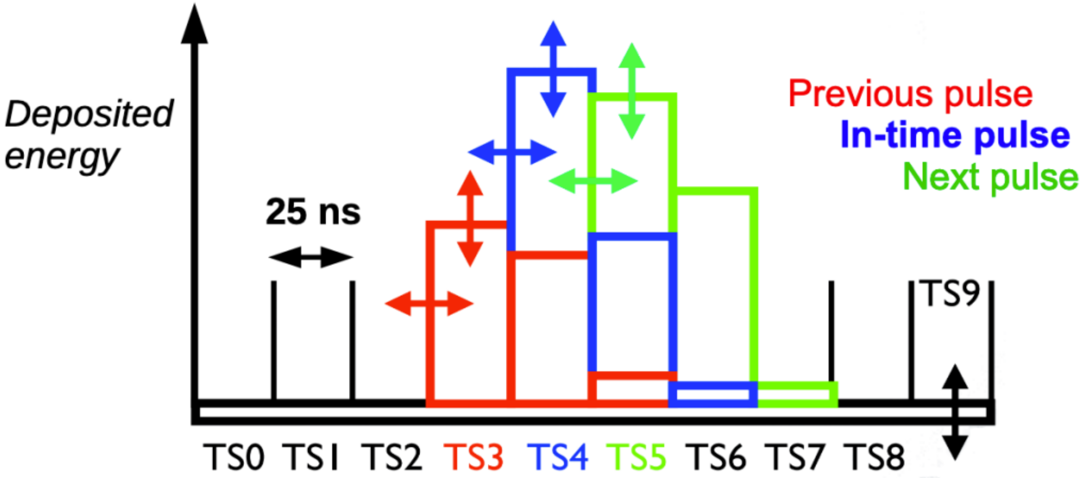


normal collision

- Reco input: digitized charge in 8 LHC bunch crossings (BX) in buffer, called time samples
  - Current BX (BX0): 75-100 ns (Time sample 3) ~60% total charge
  - BX+1: ~20% total charge
- First reco algorithm: Method 0
  - Used in Run1 (50 ns bunch spacing)
  - OOT PU almost negligible
  - ($Q_{BX0}$ + $Q_{BX+1}$) x scale factors
- Pulse fitting algorithms
  - In use since Run2 (25 ns bunch spacing)
  - 2016-2017: Method 2 (3) offline (HLT)
  - from 2018: MAHI both offline and HLT

# Method 2

- M2 estimates the energy of SOI pulse by minimizing $\chi^2$ using MIGRAD algorithm in Minuit
- Fits up to 3 pulses (SOI - 1, SOI and SOI + 1) to QIE digis in 10 TS
- Starts with fitting 1 pulse. If $\chi^2$ > 15 and charge < 100 fC for HPD or 25000 fC for SiPM (both correspond to ~20 GeV), then switches to 3 pulses



$$\chi^2 = \sum_{i=0}^{9} \frac{(A_i - \mu_i)^2}{\sigma^2_{\mathrm{p},i}} + \sum_{j=0}^{2} \frac{(t_j - \langle t \rangle)^2}{\sigma^2_{\mathrm{t}}} + \frac{(\mathrm{ped} - \langle \mathrm{ped} \rangle)^2}{\sigma^2_{\mathrm{ped}}}$$

$A_i$: QIE digi in ith TS

$\mu_i$: sum of fitted amplitudes in ith TS

$\sigma^2_{p,i}$: quadratic sum of uncertainties

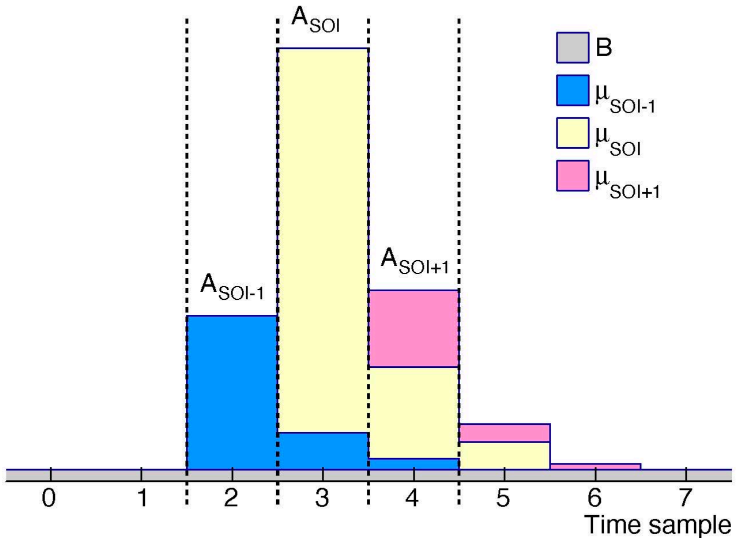
(pedestals, QIE granularity, and photostatistics)

$t_j$: pulse arrival time

ped: floating baseline

# Method 3

- M3 was developed to meet HLT timing requirment
- Compared to M2, M3:
  - Fits 3 pulses (SOI - 1, SOI and SOI + 1) to only 3 TS
  - Drops the arrival time term
  - Uses constant baseline term
  - Fitting → solving linear equations



$$\begin{bmatrix} A_{\mathrm{SOI}-1} \\ A_{\mathrm{SOI}} \\ A_{\mathrm{SOI}+1} \end{bmatrix} = \begin{bmatrix} f_0 & 0 & 0 \\ f_1 & f_0 & 0 \\ f_2 & f_1 & f_0 \end{bmatrix} \begin{bmatrix} \mu_{\mathrm{SOI}-1} \\ \mu_{\mathrm{SOI}} \\ \mu_{\mathrm{SOI}+1} \end{bmatrix} + \begin{bmatrix} B \\ B \\ B \end{bmatrix}$$

$A_i$: QIE digi in ith TS

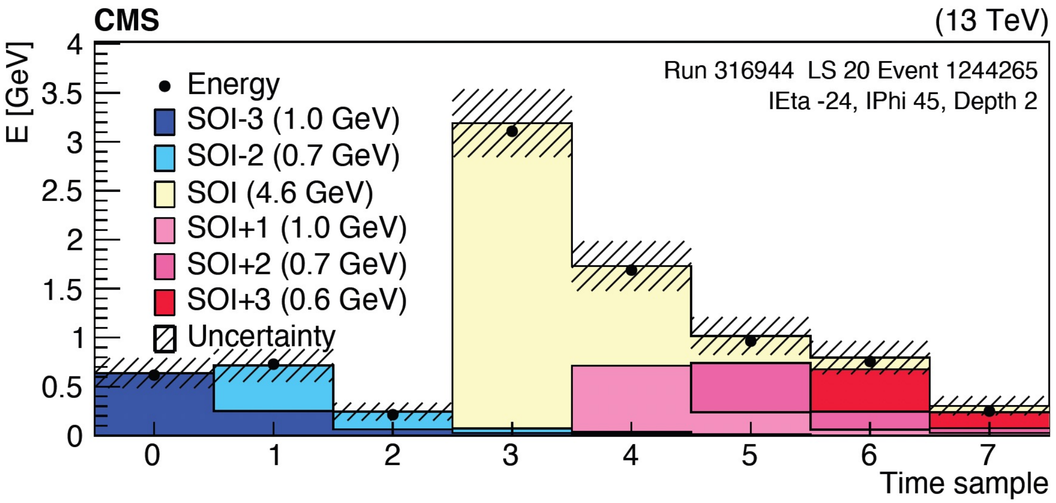
$f_{0,1,2}$: pre-measured fractions of the pulse template in +0, +1 and +2 TS, respectively

$\mu_i$: amplitudes of ith pulse

B: constant baseline (average of pedestals in all TS except SOI and SOI+1)

# MAHI

- MAHI (Minimization At HCAL, Iteratively) estimates the energy of SOI pulse by minimizing $\chi^2$ in an iterative approach, using Non-Negative Least Square (NNLS) algorithm instead of MIGRAD in M2
- Reconstruction speed: MAHI is O(10) faster than M2 and O(10) slower than M3



$$\mathbf{V} = \sum_{j=0}^{7} \mu_j^2 \mathbf{D}_j^{\text{pulse}} + \mathbf{D}^{\text{noise}}$$

$\mu_j$: amplitudes of jth pulse

$D_j^{pulse}$: pulse shape uncertainty

$D^{noise}$: total noise (pedestals, QIE granularity, and photostatistics)

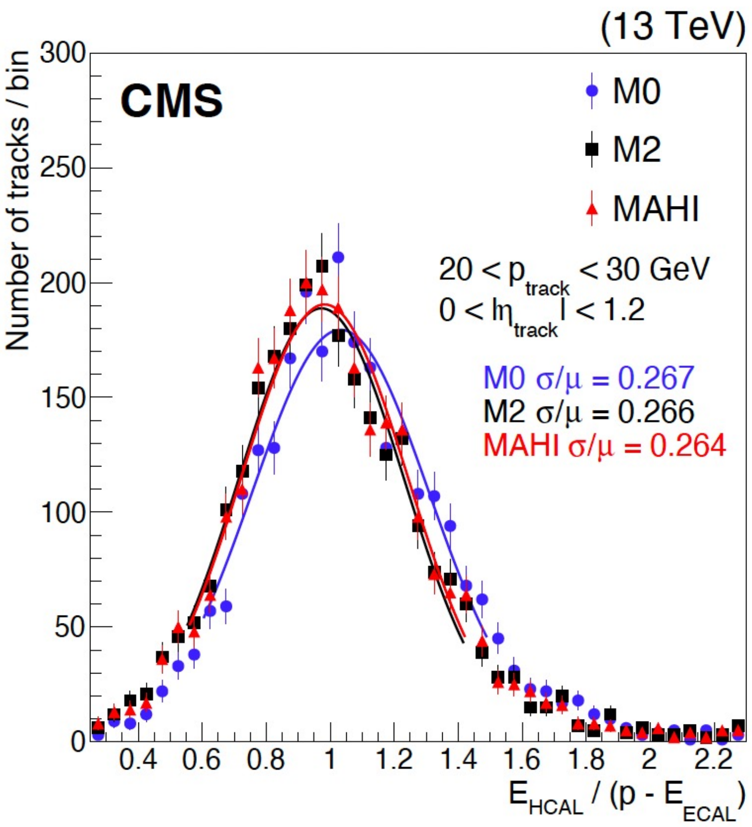
$$\chi^2 = \left[\sum_j \vec{P}_j \mu_j - \vec{d}\right]^T \mathbf{V}^{-1} \left[\sum_j \vec{P}_j \mu_j - \vec{d}\right]$$

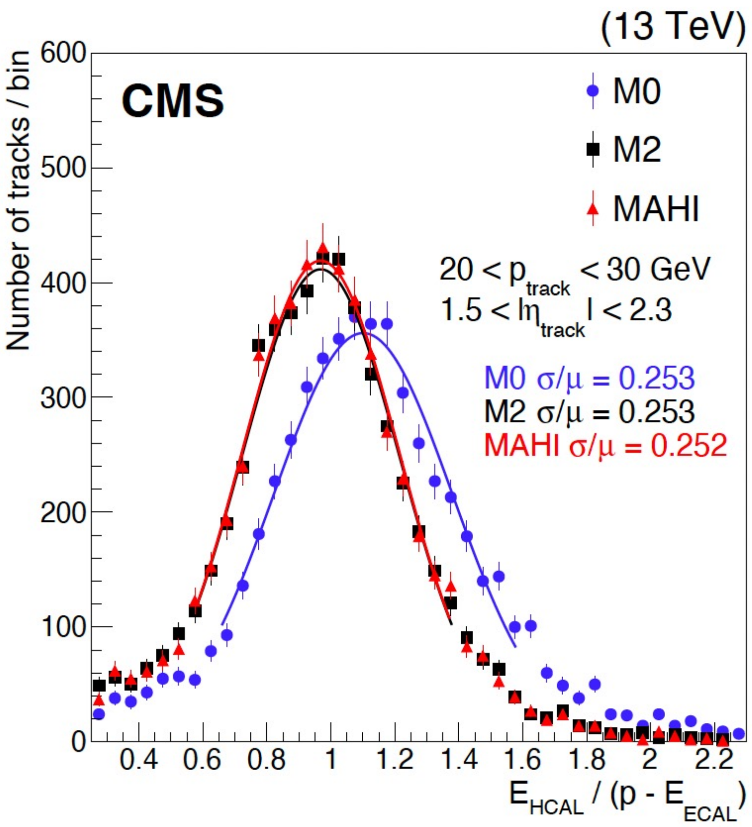
$\overrightarrow{P_j}$: 8x8 matrix contains pulse template

$\vec{d}$: vector contains QIE digis of 8TS

# Reconstruction performance

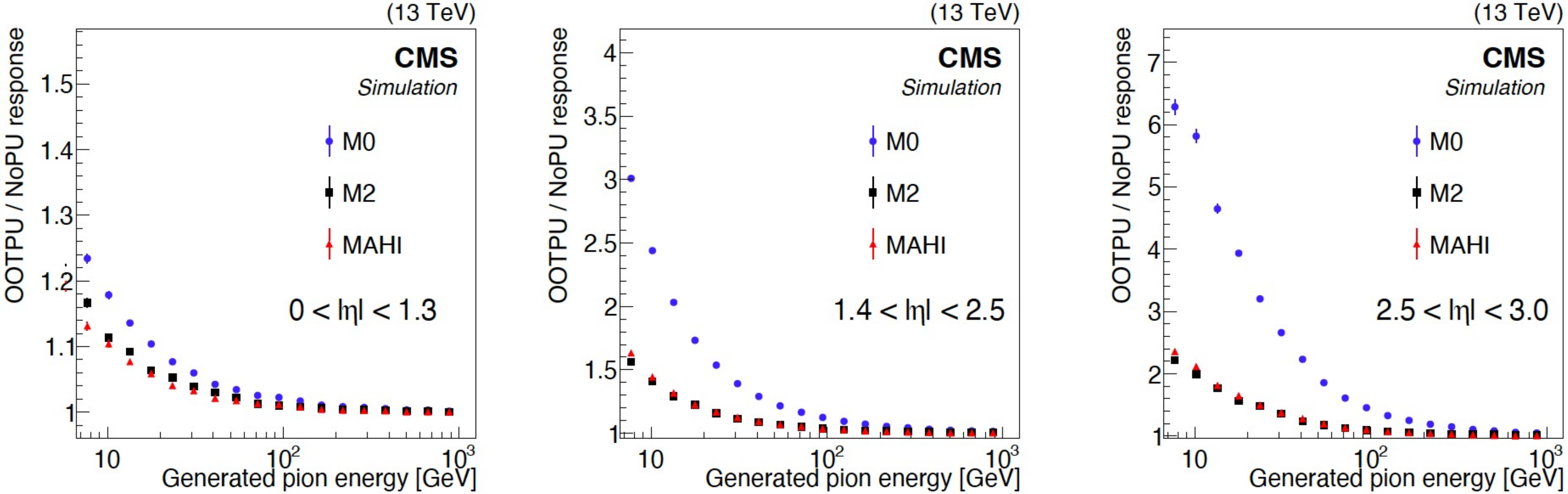
# Charged pion resolution in data





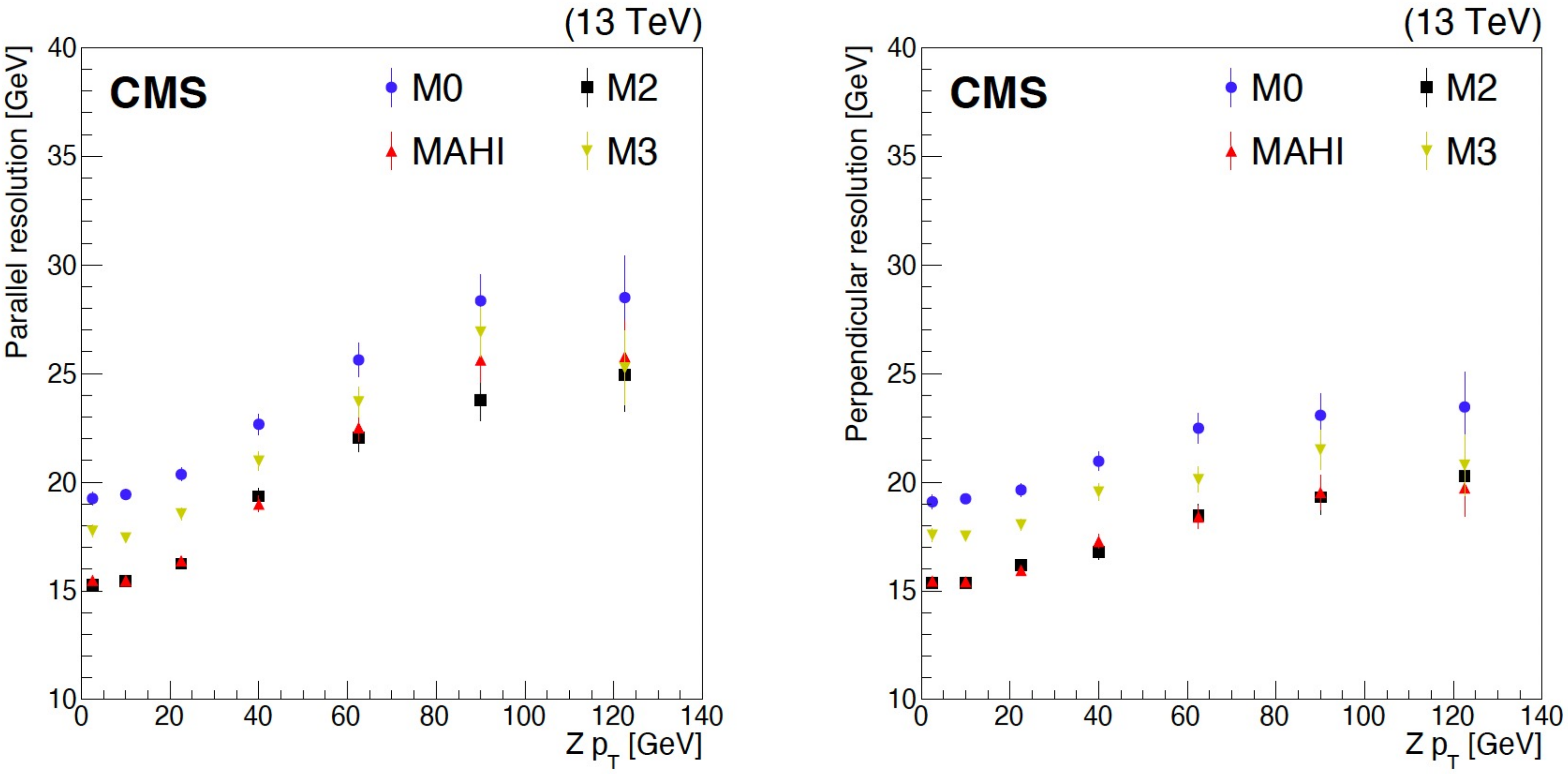
- Extrapolate isolated tracks to calorimeter and match to a cone
- Use track momentum - ECAL energy in that cone as “truth” HCAL energy
- M0, M2 and MAHI have similar resolutions, but M0 has high response because of OOT-PU

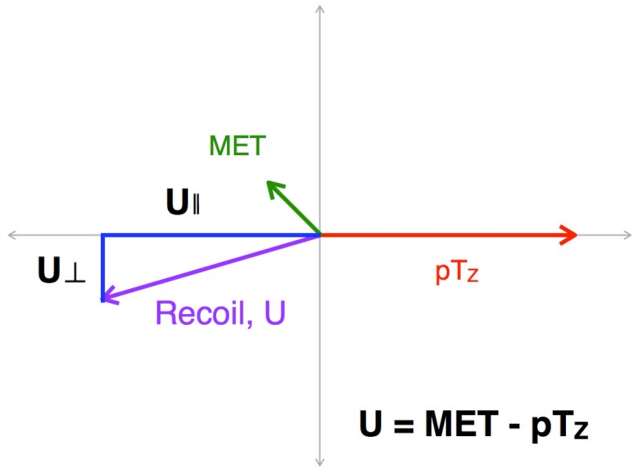
# Response of pions in MC



- Two MC samples from the same GEN step pion gun
  - One has only OOT-PU
  - The other has no PU
- Extrapolate GEN pion tracks to calorimeter and match to a cone
- Response = cone energy / GEN pion energy
- Plot ratios of responses in OOT-PU sample and no PU sample
- Performance: M2/MAHI better than M0, especially in low energy / high eta regions, because of OOT-PU subtractions

# Resolution of MET in Z→μμ data



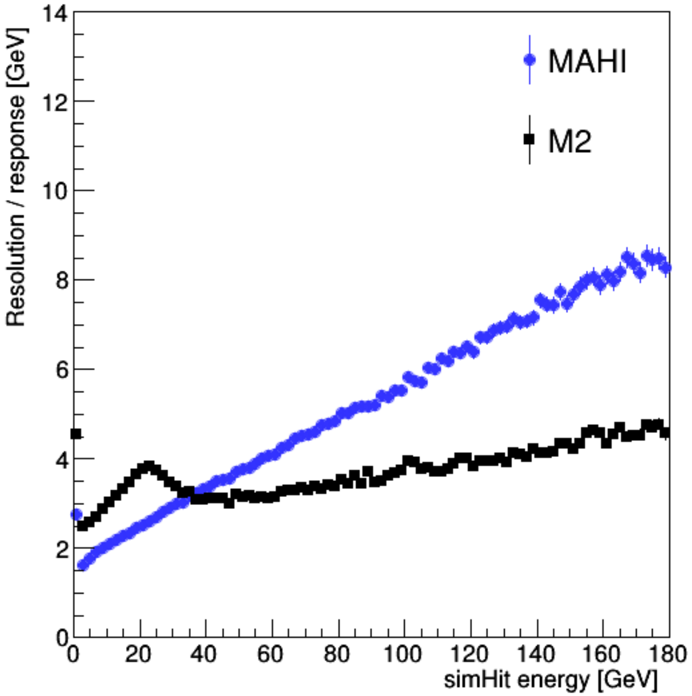


**U = MET - pTz**

- Select events with a well reconstructed Z boson
- Project MET to Z pT, and measure the resolutions of its parallel and perpendicular components
- Performance: M2/MAHI better than M3 than M0

# Reconstruction with ML

# Limitation of analytical algorithms

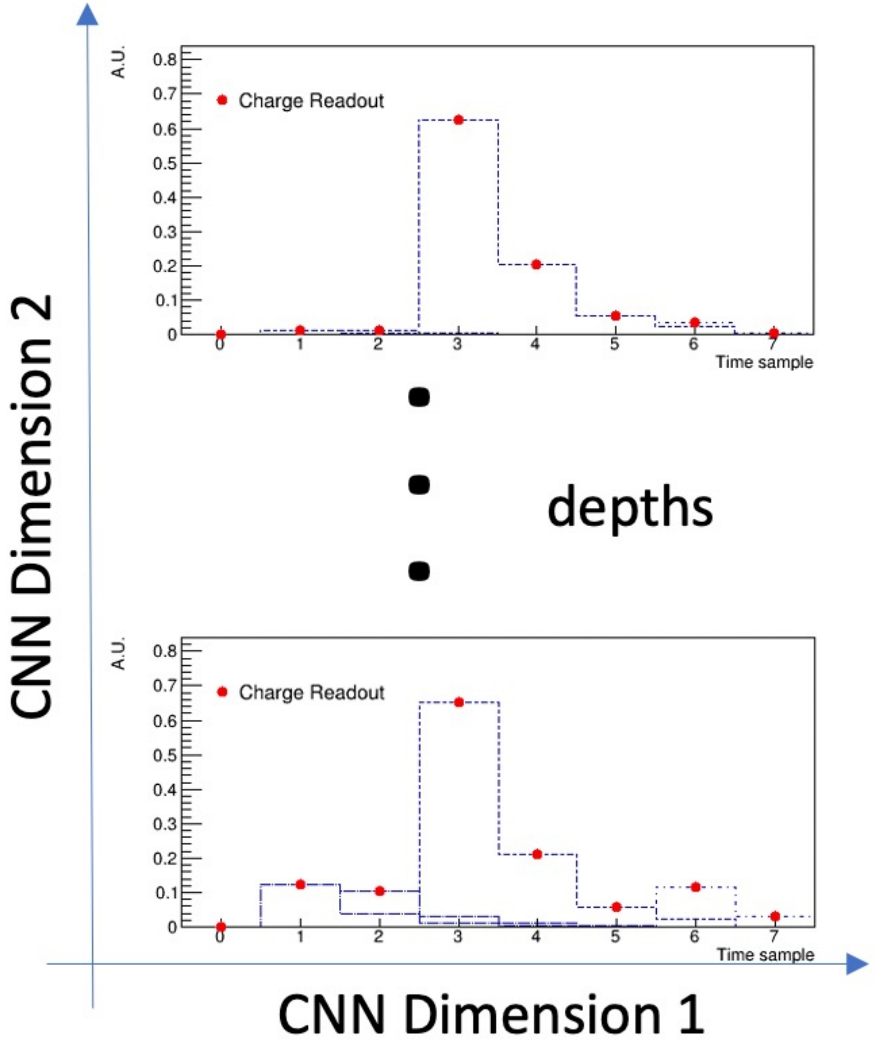


- Reconstructed energy resolution in each channel
  - MAHI: not fitting pulse arrival time

    Bad performance at high energy
  - M2: too slow - only fits up to 3 pulses

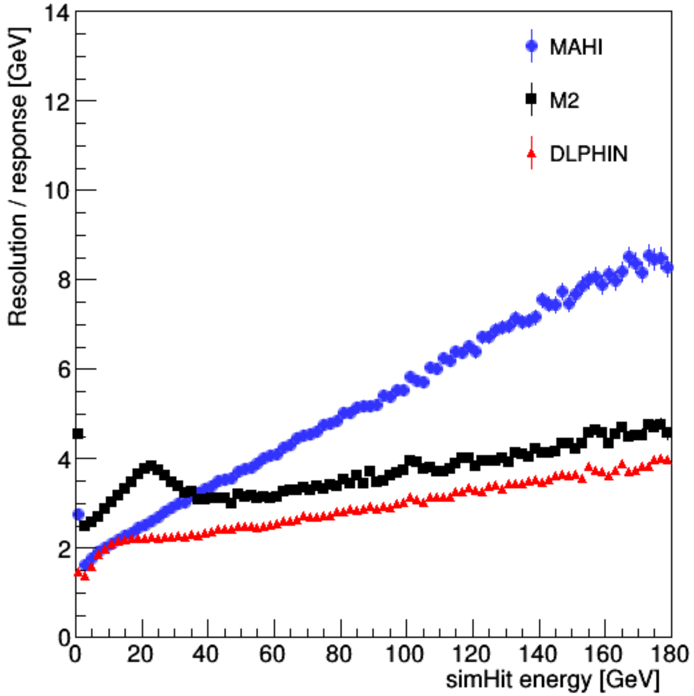
    Bad performance at low energy
- Is there an algorithm that has better resolution at both low and high energy?
- Machine learning can achieve this!

# DLPHIN

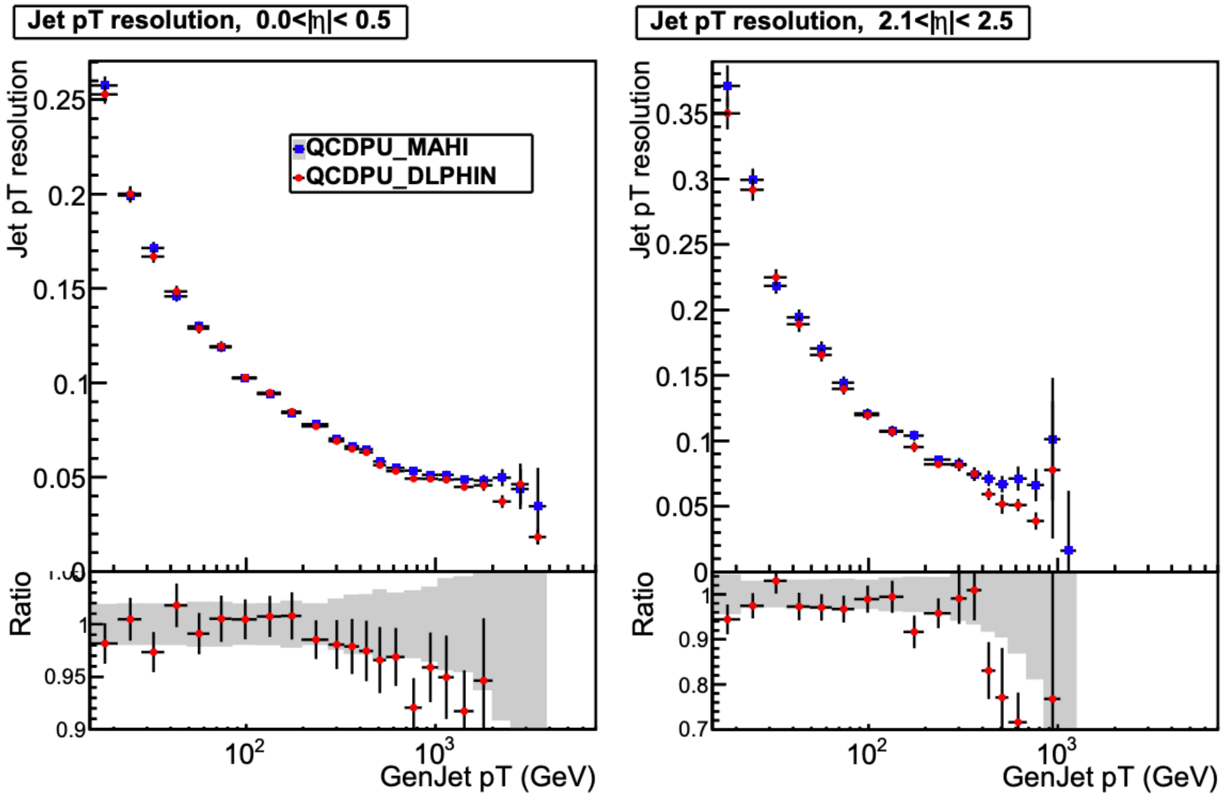


- Deep Learning Processes for HCAL INtegration
- Novel architecture based on 2D CNN
  - Dim. 1: digitized charge in 8 BX
  - Dim. 2: depth → exploit correlations among channels in an HCAL tower
- More than 3 times faster than MAHI
- Better perform from upstream to downstream Channel-level → single particle-level → jet-level
- Will benefit almost all physics analyses

# DLPHIN performance



Channel-level performance
DLPHIN up to 50% better resolution than MAHI



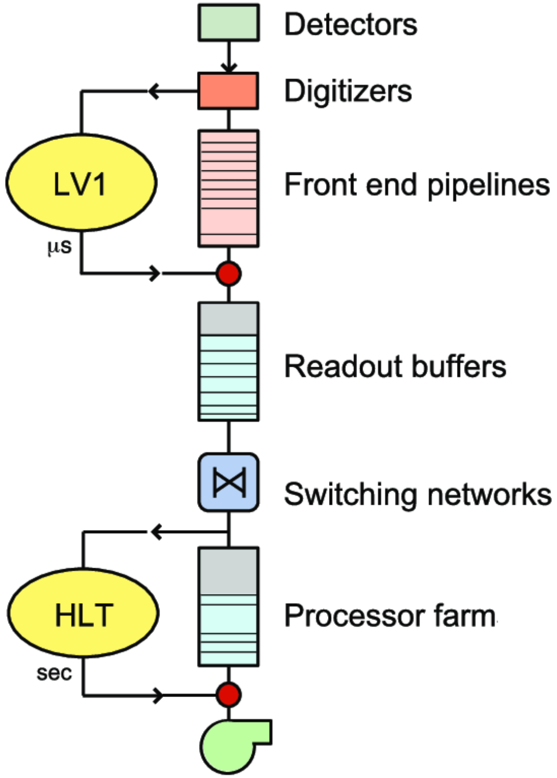
PF jet-level performance
5% / 10% better resolution in HB / HE at high energy

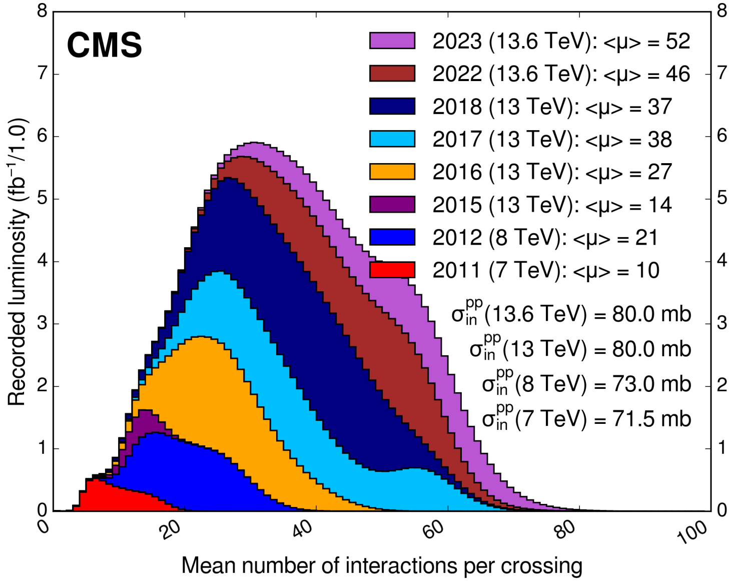
Also better lepton/photon isolation and MET resolution

# Backup Slides

# Trigger System and Pileup

- Two-level trigger system
  - Reduce the event rates from 40 MHz to ~1kHz
  - While keeping most of the interesting events
- Level-1 trigger (L1T)
  - Custom ASIC, FPGA, etc
  - Reduce rate to 100 kHz
- High-level trigger (HLT)
  - Commercial CPU + GPU
  - Rate reduce to ~1k Hz

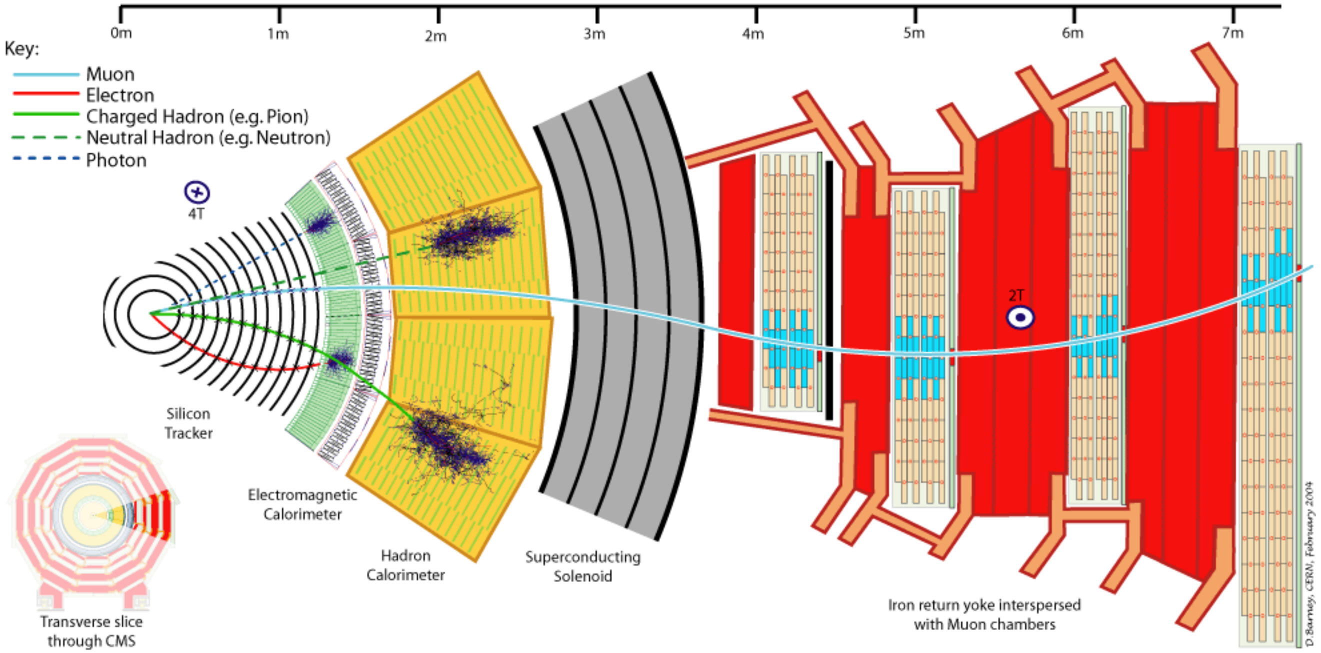




pileup (PU)

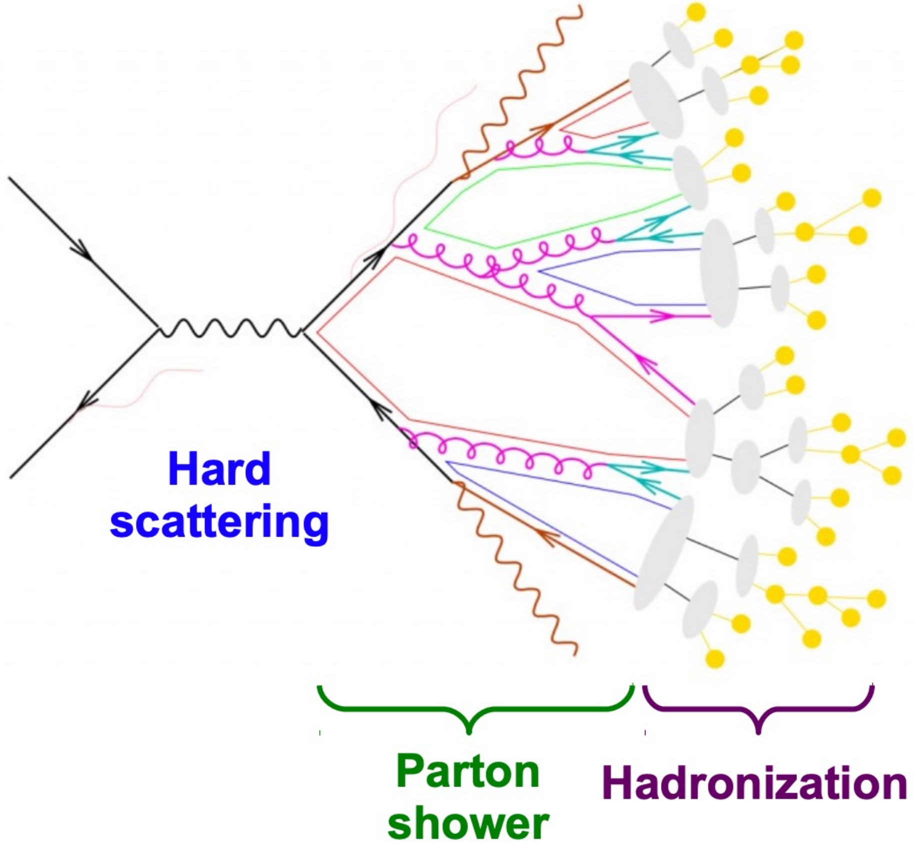
- In-time PU: current bunch crossing (BX)
- Out-of-time PU: other BX, very important for calorimeter reconstruction

# Event Reconstruction

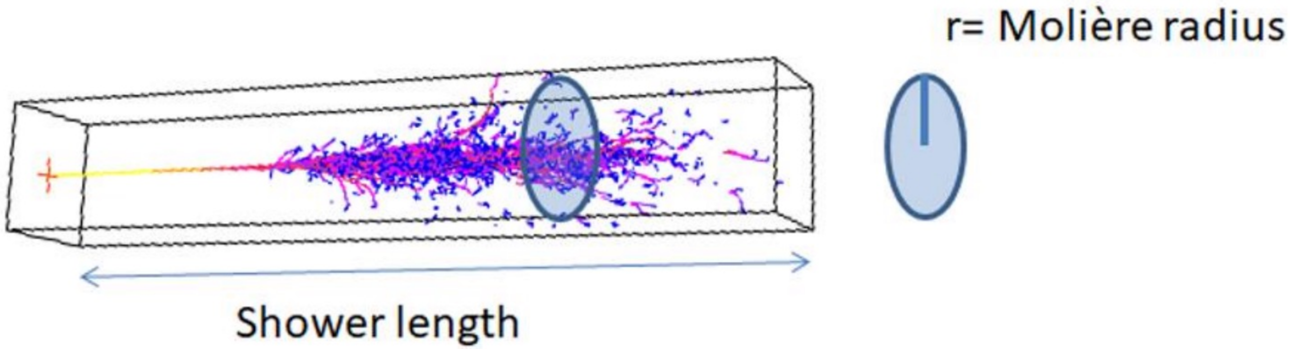


- Particle Flow (PF) Algorithm
  - Runs on HLT and offline reconstruction
  - Synthesizes information from all sub-detectors and reconstructs particles based on their signatures
    1. Muon
    2. Electron and Photon
    3. Charged and Neutral Hadron
- Then PF particles are clustered as jets
  - Usually anti-$k_T$ algorithm in CMS
- Last global quantities of an event
  - e.g. missing transverse momentum $p_T^{miss}$, aka MET usually a manifest of neutrinos, but may also from BSM :P

# Parton Shower vs Hadron Shower



Parton shower (+ hadronization that form a jet) typically in a cone



Hadron shower (interacting with detector material)

typically in a cylinder

- longitudinal development: radiation or interaction length
- lateral development: Molière radius (a cylinder containing on average 90% of the shower's energy deposition)
- Typical Molière radius for a pion is an HCAL tower (0.087 x 0.087 rad.)