





# First-Year PhD Progress Report

Xin Xia

Supervisor: Roman Poeschl

Committee: Emi KOU, Dirk ZERWAS

7/24/25

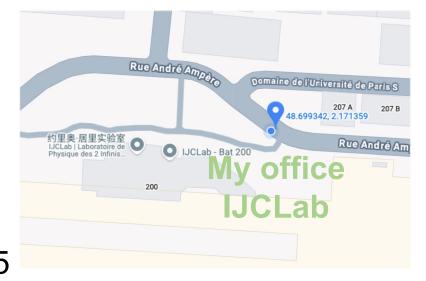


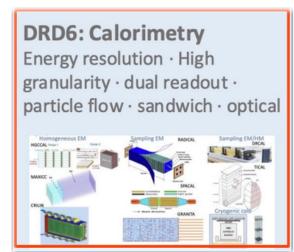
#### Individual information



Now 24 juillet 2025

- 1st year PhD
- Funded by CALO5D project
- 36-month scholarship
- DRD6 group & ILC group







Start point 12 novembre 2024

PhD defense novembre 2027





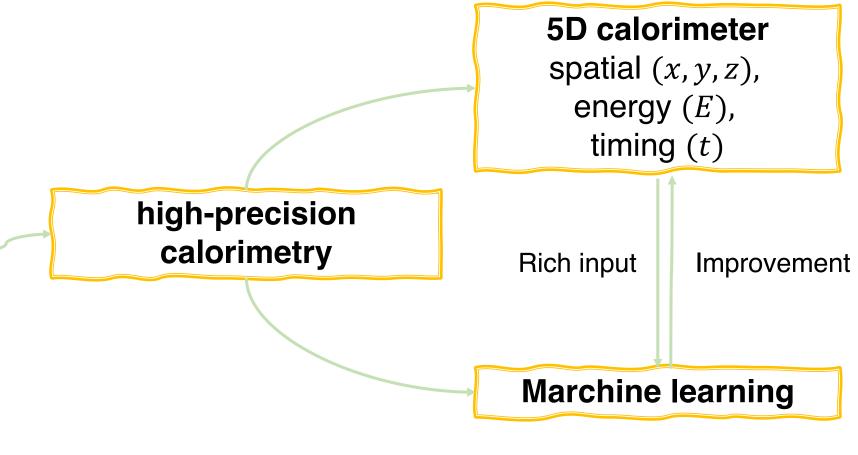
#### Research: Towards 5D Calorimetry for future Higgs factories

- Electroweak Measurements
- Energy reconstruction with Deep learning
- Beam Test experiment (hardware)

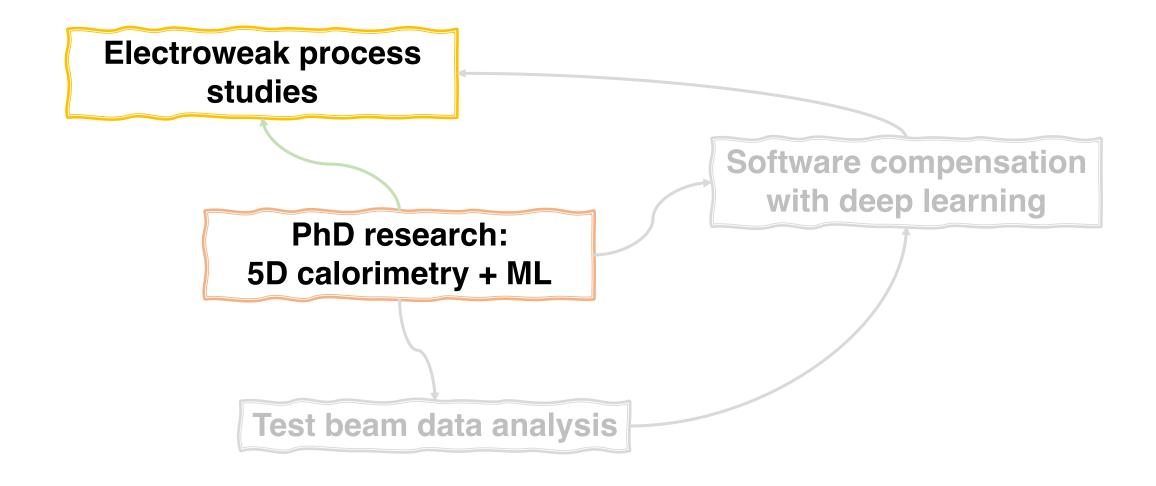


#### **Training courses**

Electroweak
measurements and
new physics searches
At higgs factories









#### **Motivation**

## Electroweak process

#### **Motivation:**

[1]: DOI: <u>10.1016/0550-3213(87)90685-7</u>

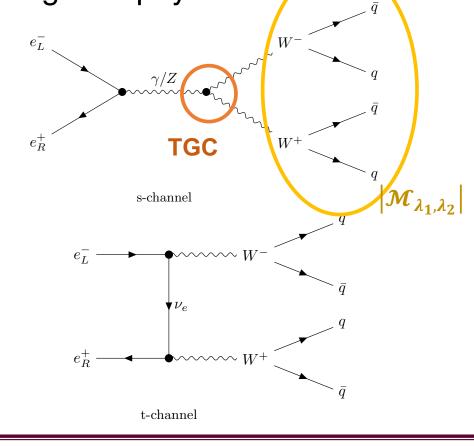
• Electroweak process serves as precision benchmarks for testing the Standard Model, constraining anomalous couplings, and probing new physics.

**Process:**  $e^+e^- \to W^+W^- \to 4q \ (@ 250 \ \text{GeV})$ 

the best channel for Five-fold cross section<sup>[1]</sup>:

$$\frac{d\sigma}{d\Theta \, d\cos\theta \, d\phi \, d\cos\overline{\theta} \, d\overline{\phi}} = \sum_{\lambda_1, \lambda_2} \left| \mathcal{M}_{\lambda_1, \lambda_2} \right| \cdot f$$

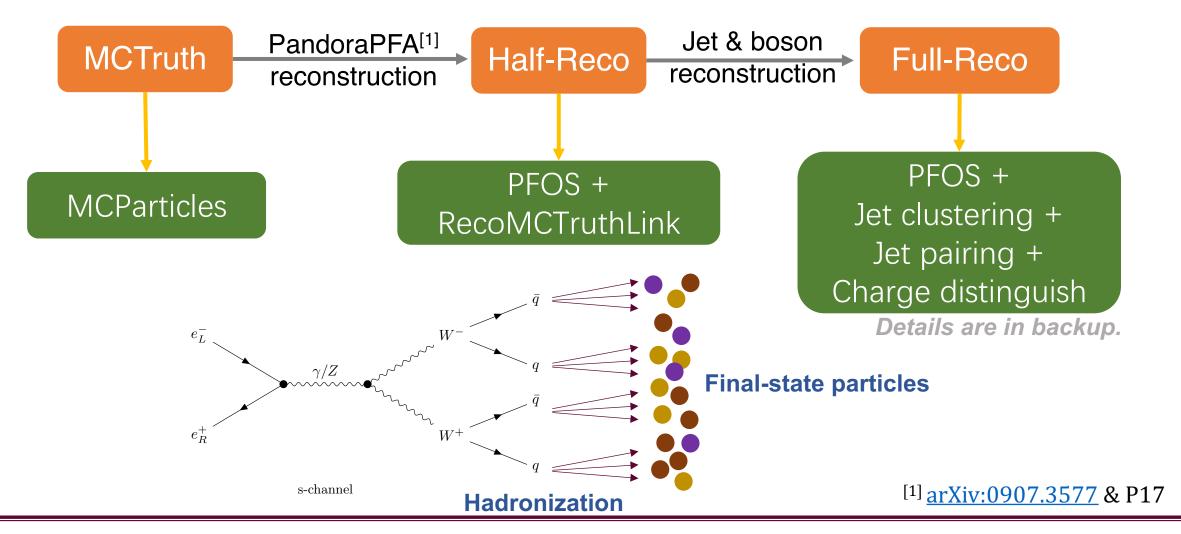
- Sensitive to TGC, CP, polarization, etc
- Give assess to all helicity amplitudes
- Never been examined, the potential is unknown
- First step of the fully hadronic final state of  $t\bar{t}$





### Electroweak process

#### Simulation and Reconstruction Flow: based on ILC software framework



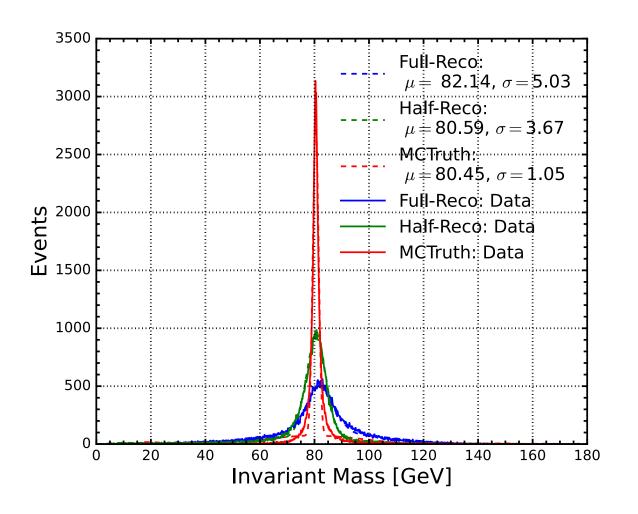


#### W boson Mass

## Electroweak process

#### Mass of W boson

- $m_W = 82.14 \pm 5.03 \, [GeV]$ 
  - Fit function: Chebyshev-gaussian function
  - Not optimize to measure the mass

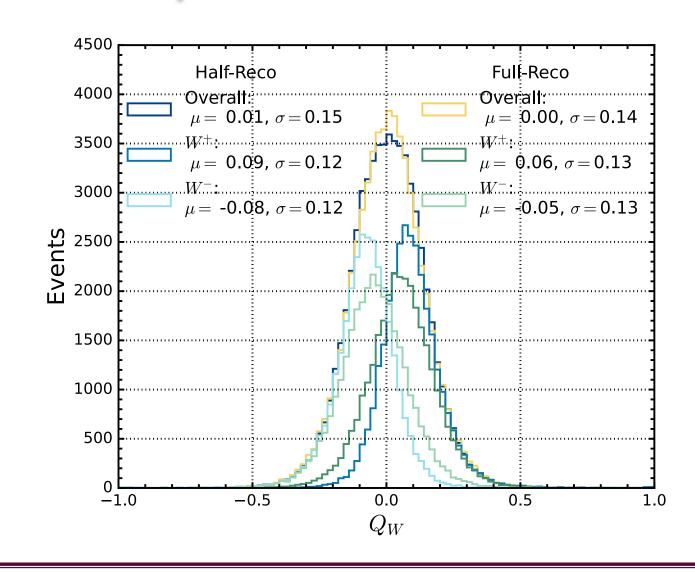




## W boson Charge

## Electroweak process

- For each jet:
  - $Q(\kappa) = \sum_{i \in J} P_i^{\kappa} Q_i / \sum_{i \in J} P_i^{\kappa}$ 
    - $\kappa = 0.25$
    - $P_i$ : PFO momentum
    - Q<sub>i</sub>: PFO charge
    - Q: Jet charge
- For W boson:
  - $Q_W = Q_{I1} + Q_{I2}$

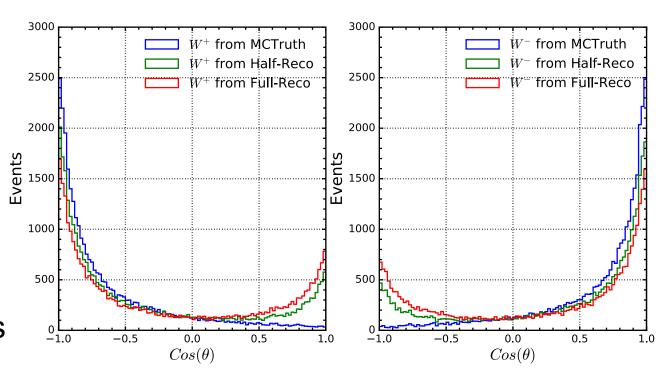




## **Angular distribution**

## Electroweak process

- Charge identification:
  - If  $Q_W > 0$ , this W boson is  $W^+$ .
  - If  $Q_W < 0$ , this W boson is  $W^-$ .
- From MCTruth to Full-Reco, more and more events migrate.
- Can be corrected by Binomial law
  - As was done in 2-fermions analysis





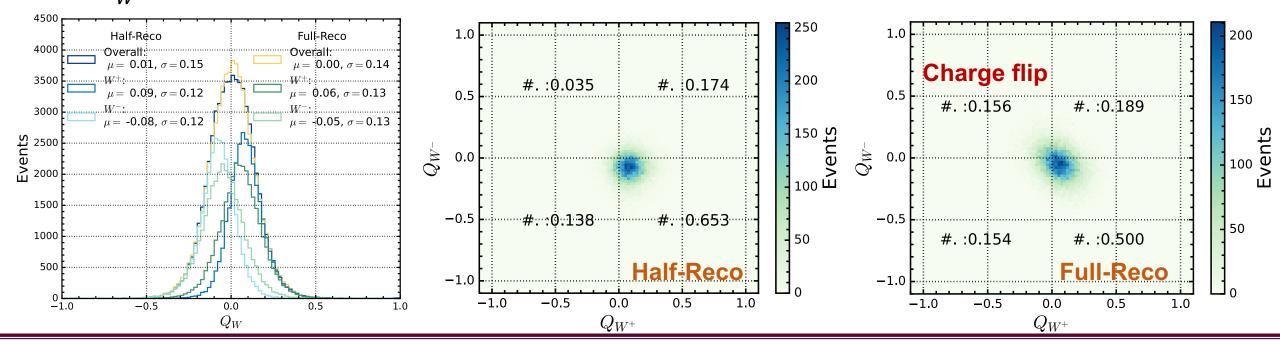
## W boson Charge

## Electroweak process

From the correlation plot of  $Q_{W^+}/Q_{W^-}$ ,

- Not independent between full-reco  $Q_{W^+}/Q_{W^-}$ 
  - Anti-correlation,  $\rho_{X,Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} = -0.314$
  - Jet clustering algorithm

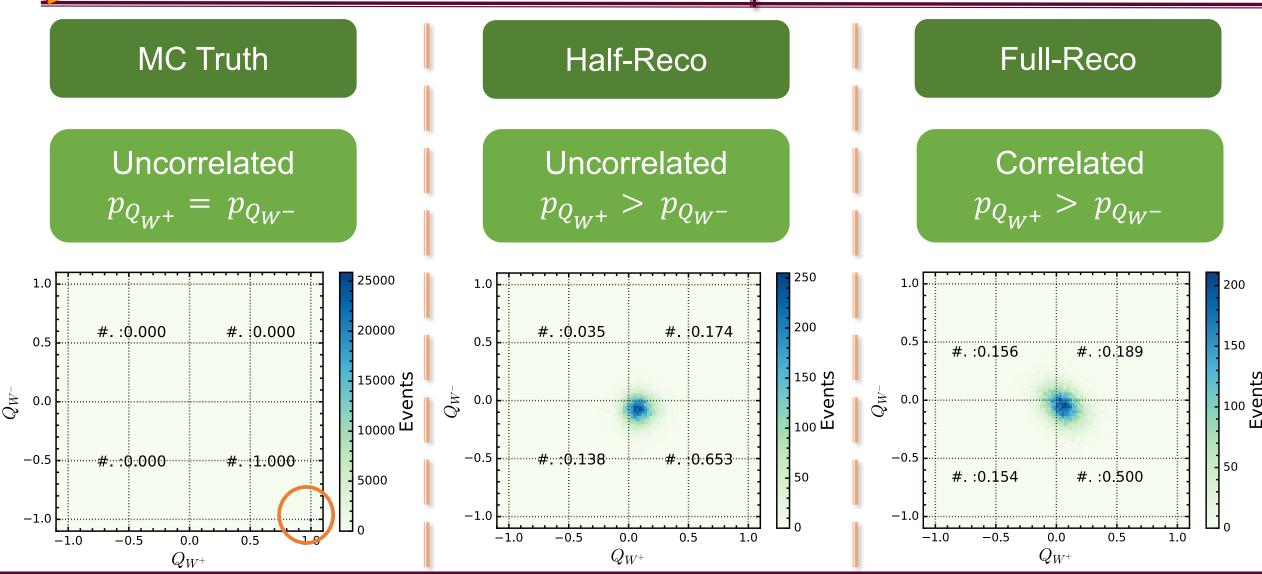
• 
$$p_{Q_{W^+}} > p_{Q_{W^-}}$$





#### W boson Charge

### Electroweak process





## **Summary & Plans**

## Electroweak process

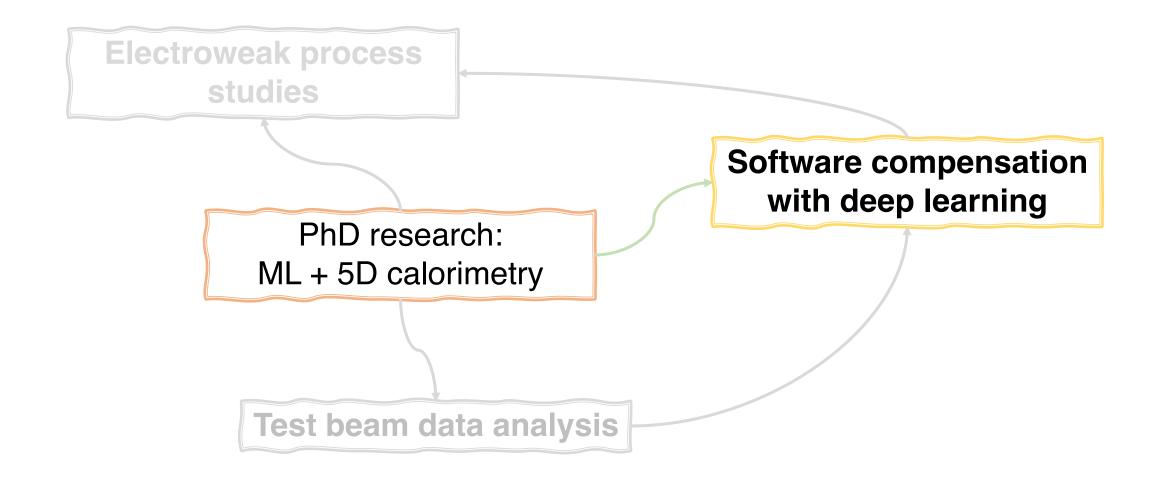
#### Summary:

- Developed a solid understanding of  $W^+/W^-$  reconstruction in the fully hadronic channel  $e^+e^- \rightarrow W^+W^- \rightarrow 4q$  (@ 250 GeV).
- Studied and implemented the charge-weighted PFOs to measure the W charge.
- Identified a key limitation: jet clustering algorithms significantly affect charge reconstruction

#### Plans:

- Try other jet clustering algorithms, to assess their impact on W charge reconstruction performance.
- Explore decay-level information, including individual track charges and angular correlations.
- Develop machine learning approaches (e.g., BDTs or deep neural networks) to:
  - Refine W/jet charge determination
  - Reduce misassignments from jet clustering
- Aim to perform a full five-fold differential cross section measurement, using the reconstructed angular variables to access W helicity amplitudes and probe Triple Gauge Couplings.



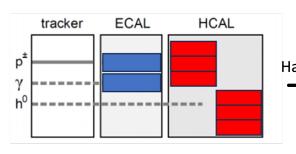


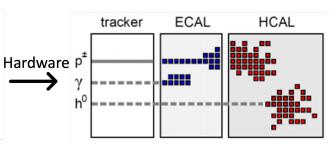


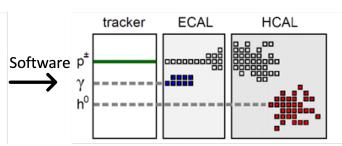
#### **Motivation**

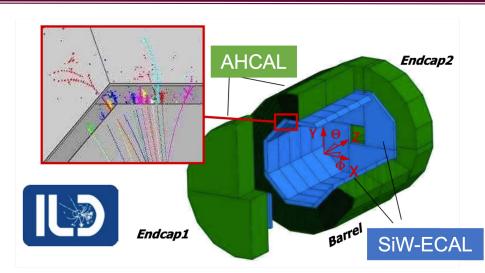
### Software compensation

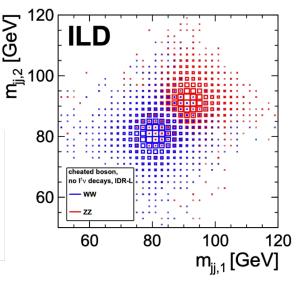
- High-precision energy measurements is essential for W/H/Z studies
- Target: Jet energy resolution  $< 30\%/\sqrt{E}$
- A major option: Particle Flow Algorithm (PFA) + PFA-oriented detector system
  - PFA:  $E_{jet} = E_{tracker} + E_{ECAL} + E_{HCAL}$
  - High granularity calorimeter (imaging):
    - SiW-ECAL + AHCAL
- $E_{ECAL} + E_{HCAL}$  as part of PFA & intrinsic performance of calorimeter













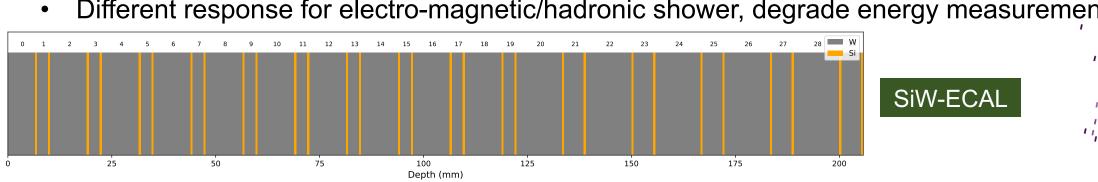
#### **Detector introduction**

#### Software compensation

Non-Compensating

#### **Detector introduction:**

- Configuration: shown in table; Energy resolution of hadron < 60%/  $\sqrt{E}$
- Different response for electro-magnetic/hadronic shower, degrade energy measurement,



 $3 \times 3 \text{ cm}^2$ 

**AHCAL** 

 $\sim 5\lambda_I$ 

#. Layers

30 in 20 cm

48 in 1 m

SiW-ECAL

AHCAL



Steel

Scintillator

2025/7/24

## **SC** Algorithm

# Software compensation

#### **Software Compensation Algorithm:**

- For a hadron in a non-compensating calorimeter
  - Have electromagnetic components (e) and hadronic components (h), e/h >1

• 
$$\pi = f_{em} \cdot e + (1 - f_{em}) \cdot h$$

• 
$$E_{rec} = \frac{e}{\pi} \cdot E_{dep} = \frac{e}{f_{em} \cdot e + (1 - f_{em}) \cdot h} \cdot E_{dep} = \frac{e/h}{1 + f_{em}(e/h - 1)} \cdot E_{dep}$$

- e/h: a constant value which depends on calorimeter
- $f_{em}$ : generated by  $\pi^0$  in hadronic shower, fluctuates strongly from event to event, Measured as the ratio of the energy deposited by electronmagnetic components (i.e., photons and  $e^{\pm}$  from  $\pi^0 \to \gamma \gamma$  decays) to the total deposited energy.

$$f_{em} = \frac{\sum_{i}^{em} e_i}{\sum_{i}^{all} e_i} (em \ is \ \gamma, e^{\pm}, \pi^0)$$

 $f_h$ 

 $f_e$ 

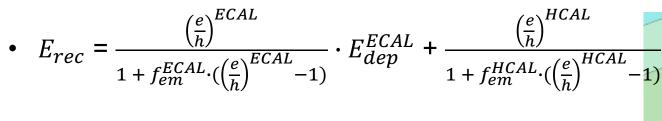


## SC Algorithm

# Software compensation

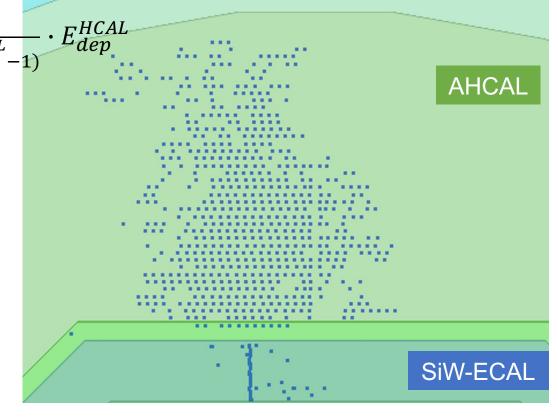
#### **Software Compensation Algorithm:**

• In my case, a hadron in a system (SiW-ECAL + AHCAL):





- fem
- fHCAL





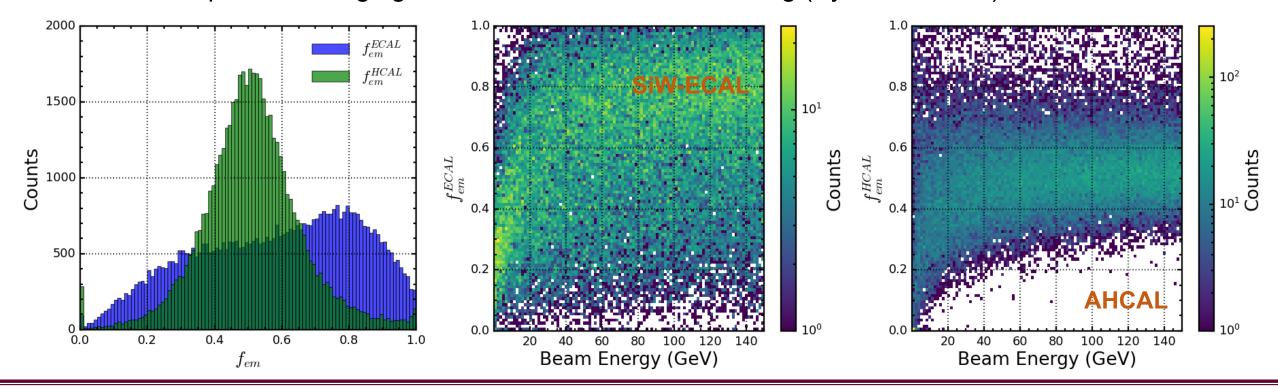
# f<sub>em</sub> behaviour

#### Software compensation

- $f_{em}$  fluctuates strongly from event to event
- How to determine  $f_{em}$  event-by-event basis?
  - Dual-readout techniques

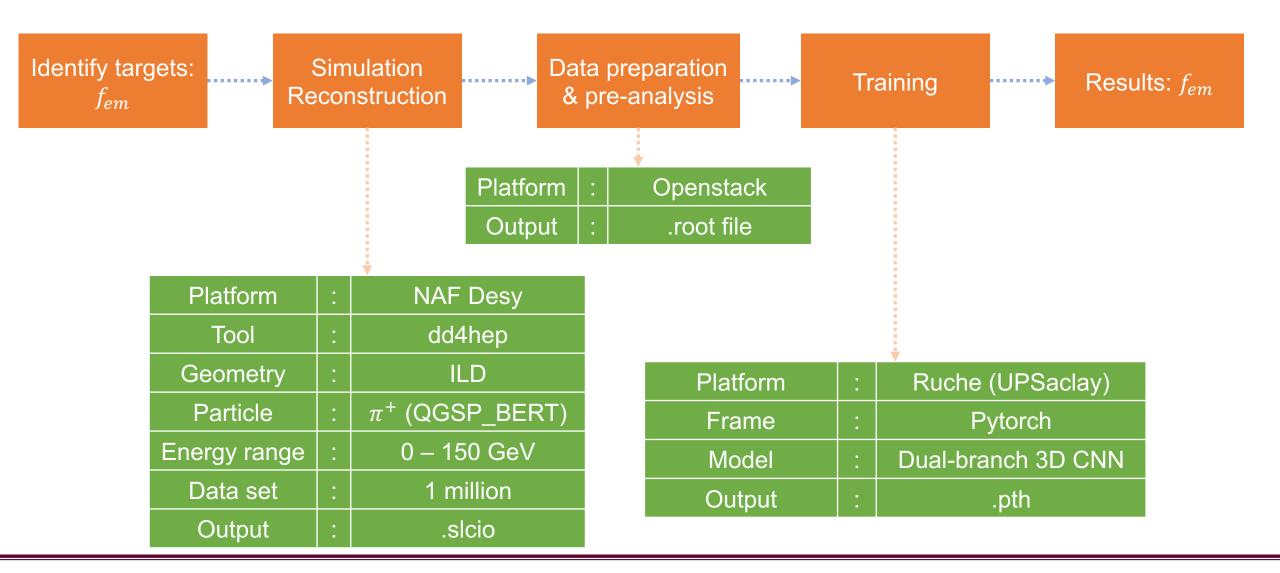
EM Narrow, compact, Higher component shorter Lower, spread out components

Rich input from imaging calorimeter + Marchine Learning (Pytorch + CNN)





#### Software compensation

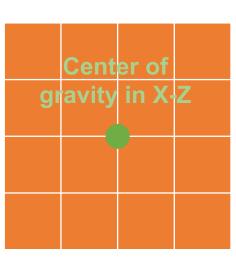




## Challenges

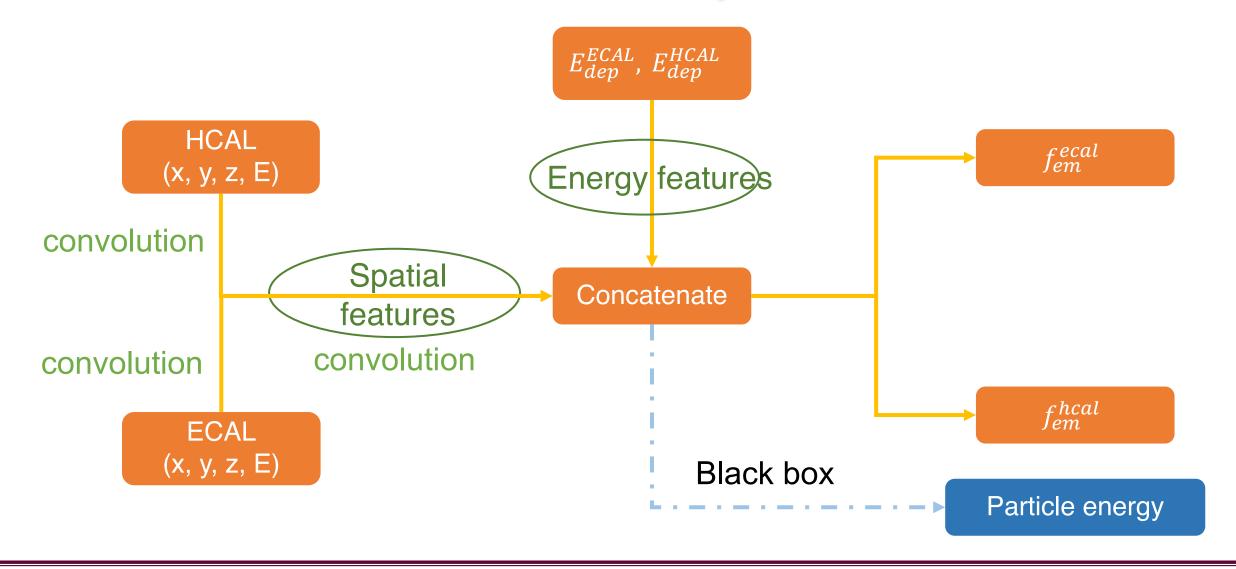
## Software compensation

- Environment Setup
  - Issue: PyTorch GPU training requires CUDA; incompatible versions cause installation conflicts
  - Solution: Try multiple versions to match system & PyTorch
- Input Shape & Model Design
  - Issue:
    - Voxel input per event extremely large (~100M channels)
    - Total dataset size too big for disk/memory; I/O very slow
  - Solution:
    - Center voxel grid using shower COG to reduce input size
- Data Handling & Shuffling
  - Issue:
    - 1M events (1k energies × 1k events); train/val/test split should reflect full spectrum, not block by energy
    - random access in .h5 is slow
    - Full shuffling causes OOM
  - Solution:
    - Use pre-chunked .h5 shards, mimic shuffle while avoiding OOM





### Software compensation





## Results - $f_{em}$

### Software compensation

Data

0.0

Gaussian Fit

 $= 0.004, \sigma = 0.059$ 

0.8

1.0

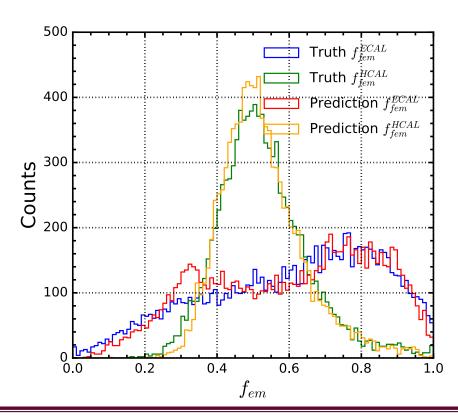
0.6

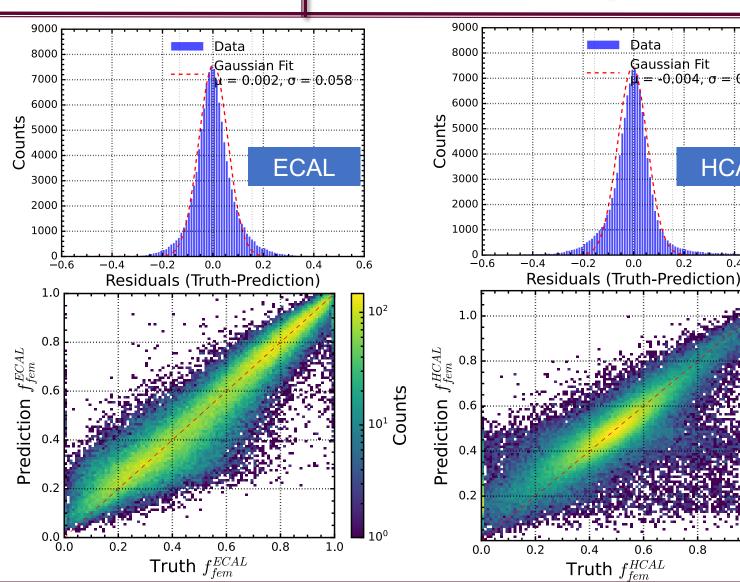
**HCAL** 

10<sup>2</sup>

Results for  $f_{em}^{ECAL}$  and  $f_{em}^{HCAL}$ :

 Shows good consistency between truth and prediction





CSI | XIN XIA 23 2025/7/24



## **Summary & Plans**

## Software compensation

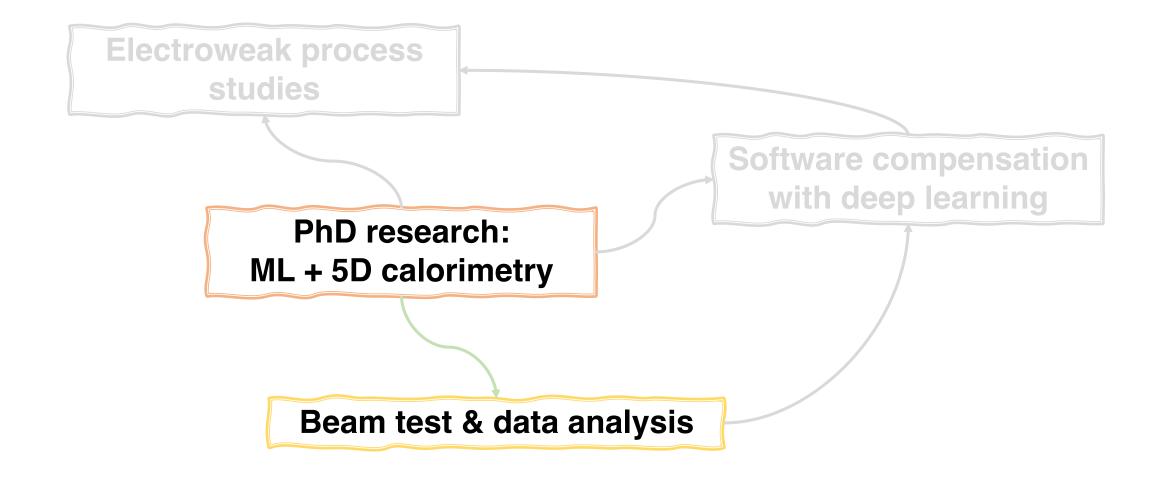
#### **Summary:**

- Full ML-based compensation chain established
- $f_{em}$  (ECAL & HCAL) regression model shows accurate predictions

#### Plans:

- Determine the energy resolution based on fem (first-step results in back-up)
- Expolit timing information to enhance performance
- Test alternative physics lists for better hadronic modeling
- Explore other advanced architectures: GNNs, Transformers, etc.

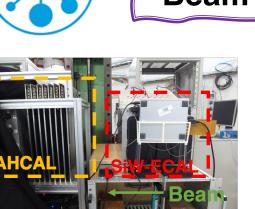








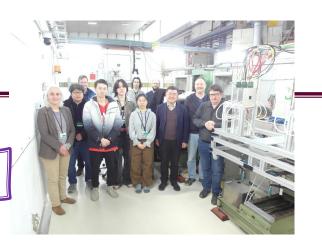
#### **Beam test experiment**

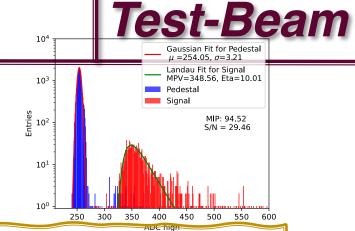


My work



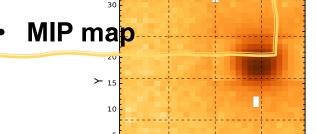
- Bunch crossing ID match
- Event display for ECAL+HCAL combination





**ECAL** alone analysis

- S/N ratio
- Pedestal map
- Hit map





## Beam test & data analysis

#### Test-Beam

The result was presented on DRD Calo 2025 Collaboration Meeting







# **Training courses**

Category	Courses	Points
Reinforcing scientific culture	Understanding basic principles of particle accelerators	5
Methods, tools, techniques and concepts for conducting doctoral research	Paris-Saclay Computing Center Howto	1
	The training for new users of the Mesocenter	1
Communication skills	French course A1	6
Prepare professional future	the 10th edition of the Workshop Perspectives	2
Sustainable development	Enjeux de la transition écologique (SPOC)	1
Research ethics and scientific integrity	MOOC Ethique de la recherche et intégrité scientifique	1
Leading a professional team	Veiller à la qualité de vie au travail	1
In total		18 / 25



#### **PhD Fest**



cnrs

#### **Towards 5D Calorimetry for future Higgs factories:**

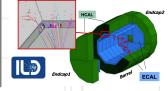
#### Electroweak Measurements, Energy reconstruction and Beam Test

Xin XIA, IJCLab-Orsay xiax@ijclab.in2p3.fr Supervisor: Roman Pöschl





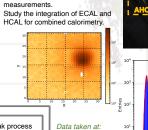
- Future electron-positron Higgs factories require high-precision calorimetry for accurate electroweak measurements and new physics searches.
- **5D calorimeters** with precise spatial (x, y, z), energy (E), and timing (t) information, provide unprecedented resolution and rich input for advanced reconstruction algorithms.
- My PhD research explores machine learning techniques to fully exploit this 5D information, focusing on three main directions:
- · Electroweak process studies at Higgs factories
- · Software compensation with deep learning for improved energy reconstruction
- · Test beam data analysis of high-granularity SiW-ECAL and AHCAL prototypes



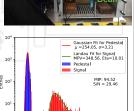
#### Electroweak process studies Software compensation with deep learning Motivation: Motivation: · Electroweak processes serve as precision benchmarks Energy measurements is essential for W/H/Z studies for testing the Standard Model, constraining anomalous Non-compensating calorimetry (SiW-ECAL + AHCAL) respond couplings, and probing new physics. differently to electromagnetic and hadronic shower components, Requires accurate reconstruction of jets and $W^{\pm}$ charge degrading energy resolution. to reduce systematic uncertainties. Particle: $\pi$ Process: $e^+e^- \rightarrow W^+W^- \rightarrow 4q$ $\frac{e}{h}$ : calorimeter-dependent constant response ratio $f_{em} = \frac{\sum_{i}^{em} E_{i}}{\sum_{i}^{all} E_{i}}$ : fluctuates event by event Approach: a dual-branches 3D CNN model (based on Ruche and Openstack) Simulation and Reconstruction: PandoraPFA Spatial features PFOS + 0.4 0.6 Truth f... Jet pairing + Charge distinguis Energy Jet clustering: FastJet Jet pairing: Future work: mass constrain Check energy linearity and resolution Charge distinguish: Incorporate timing information into the polar angle constrain model for better performance. Validate and refine the model using real beam test data. Test beam data analysis

#### Validate the performance of the

- SiW-ECAL prototype under realistic beam conditions.
- Assess detector stability and data quality to ensure reliable energy measurements.



DESY.



250 300 350 400 450 500 550 60 ADC high

PHENIICS FEST

#### Summary

- studies, with plans to apply ML solutions.
- Achieved promising results in software compensation with deep learning on highgranularity calorimeters.

Problem: charge flipped

 $\frac{d\sigma}{d\cos\theta_{w}}$  distribution with MC \_1 to get TGC parameter

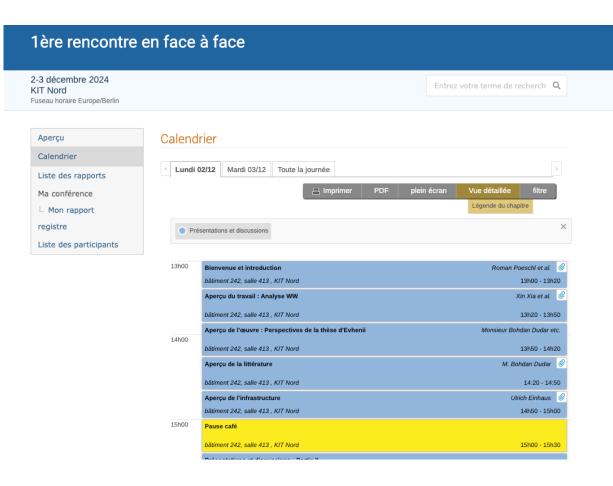
between W+/W Main limiting factor: Jet clustering algorithm Future work: · Use ML to optimize charge reconstruction Compare reconstructed

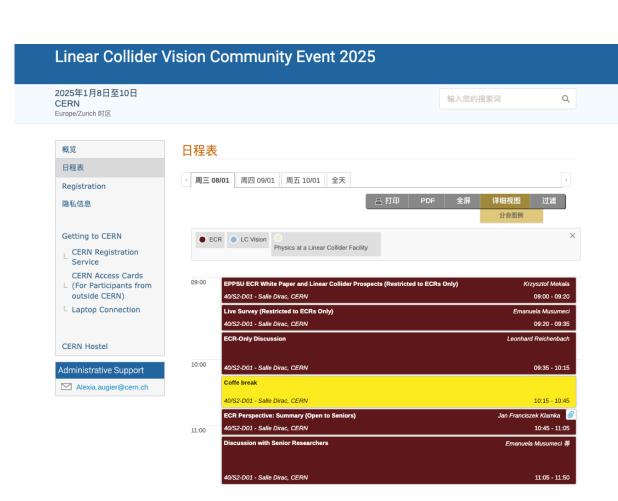
#### Identified and analyzed charge reconstruction challenges in electroweak process

Validated SiW-ECAL and AHCAL performance through test beam data analysis.

#.:0.500







# Merci beaucoup!

# Backup

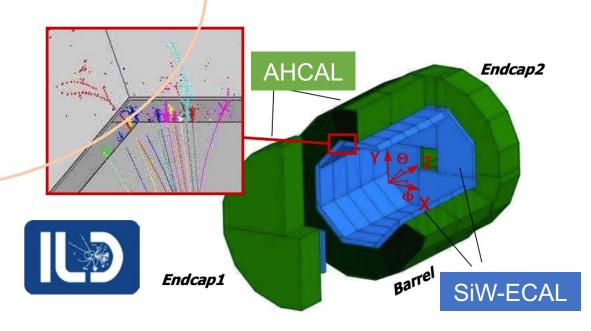


Electroweak process studies

PhD research: ML + 5D calorimetry

**Beam-test experiment** 

Software compensation with deep learning





## Electroweak process

Jet clustering

FastJet algorithm

Jet pairing

$$|(Jet_i + Jet_j)$$
.Mass-80.4 $|$  + $|(Jet_k + Jet_l)$ .Mass-80.4 $|$   $W1 = Jet_i + Jet_j$ ,  $W2 = Jet_k + Jet_l$ 

Charge distinguish

$$\sqrt{(\theta_{W_1} - \theta_{W_+})^2 + (\varphi_{W_1} - \varphi_{W_+})^2} + \sqrt{(\theta_{W_2} - \theta_{W_-})^2 + (\varphi_{W_2} - \varphi_{W_-})^2} < \sqrt{(\theta_{W_1} - \theta_{W_-})^2 + (\varphi_{W_1} - \varphi_{W_-})^2} + \sqrt{(\theta_{W_2} - \theta_{W_+})^2 + (\varphi_{W_2} - \varphi_{W_+})^2}),$$

$$W1 \text{ is } W^+, W2 \text{ is } W^-$$



#### Electroweak process studies

p,q method to quantify the charge identification power.

• 
$$N_{acc} = N \cdot p^2 + N \cdot q^2$$

• 
$$N_{rej} = 2 \cdot N \cdot p \cdot q$$

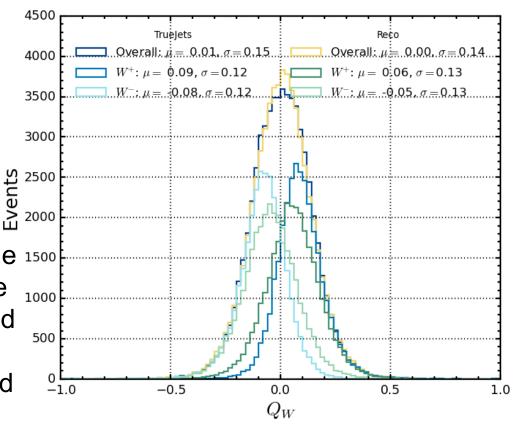
• 
$$p + q = 1$$

• 
$$N_{acc} + N_{rej} = N$$

#### Here the:

- *N*: total number of events
- p : probability of correctly identifying the generated W charge
- q: probability of wrongly identifying the generated W charge
- $N \cdot p^2$ : number of events that both  $W^+$  and  $W^-$  are identified
- correctly
- $N \cdot q^2$ : number of events that both  $W^+$  and  $W^-$  are identified
- wrongly
- $2 \cdot N \cdot p \cdot q$ : one is identified correctly, one is identified wrongly

So p = 0.5 + 0.5 
$$\sqrt{2 \cdot \frac{N_{acc}}{N} - 1}$$



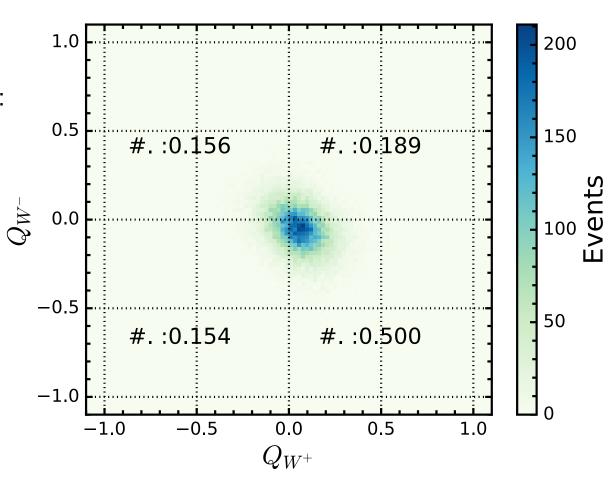
35



#### Electroweak process studies

- For the  $W^+$  identified by MCtruth information, fill their reconstructed charge in x axis
- For the  $W^-$  identified by MCtruth information, fill their reconstructed charge in y axis
- From the right plot, We can observe that:
  - Anti-correlation, lots of events are charge-flipped, :
  - Ratio of  $(Q_W^+ < 0 \text{ and } Q_{W^-} > 0)$ : 0.156, these events are charge flipped.
    - $q^2 = 0.156$ , q = 0.395
  - Ratio of  $(Q_W^+ < 0 \text{ and } Q_W^- > 0)$ : 0.500, these events are what we wanted.
    - $p^2 = 0.500$ , p = 0.707
  - Very clear that  $p + q \neq 1$
- Probabilities of correctly identifying the generated  $W^+/W^-$  charge are not equal.
  - $p_{W^+} = 0.689$
  - $p_{W^-} = 0.654$
  - Bias =  $(N_W^+ N_{W^-})/\sqrt{(N_W^+ N_{W^-})}$  = 5.3 $\sigma$

#### Refined4Jets





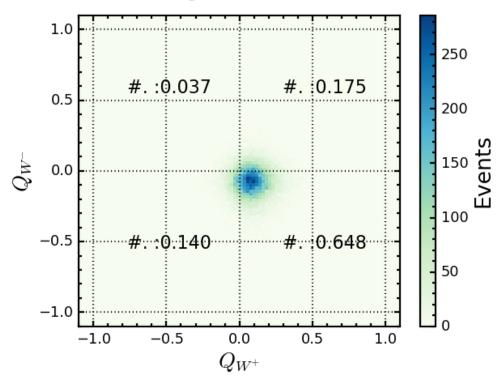
#### Electroweak process studies

#### From the right plot, We can observe that:

- p-q method works :
  - Ratio of  $(Q_{W^+} < 0$  and  $Q_{W^-} > 0$ ): 0.037, these events are charge flipped.
    - $q^2 = 0.037$ , q = 0.192
  - Ratio of  $(Q_{W^+} < 0 \text{ and } Q_{W^-} < 0)$ : 0.648, these events are what we wanted.
    - $p^2 = 0.648$ , p = 0.805
  - $p + q \approx 1$
- Probabilities of correctly identifying the generated  $W^+/W^-$  charge are not equal.
  - $p_{W^+} = 0.823$
  - $p_{W^-} = 0.788$
  - Bais =  $(N_W^+ N_{W^-})/\sqrt{(N_W^+ N_{W^-})} = 4.8\sigma$

#### **TrueJets**

Total:30596
Bias between W+/W-:4.792  $\sigma$ Reconstruction Correlation:0.003
Identification Agreement Correlation:-0.003





#### **Electroweak process studies**

#### The reason of anti-correlation:

- Jet clustering algorithm
  - The reconstructed of W<sup>+</sup> and W<sup>-</sup> are not independent.
  - If  $W^+$  loses a charged PFO,  $W^-$  will gain a charged PFO.

And vice versa

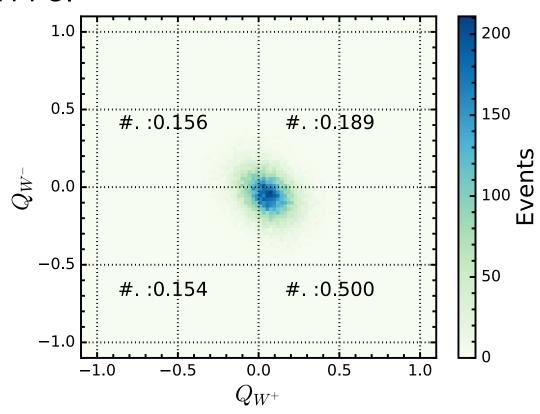
• If one of  $PFO^+$  in  $W^+$  is classified into  $W^-$ 

• 
$$W^+ \to W^-$$
,  $W^- \to W^+$ 

- If one of PFO<sup>-</sup> in W<sup>-</sup> is classified into W<sup>+</sup>
  - $W^+ \to W^-$ ,  $W^- \to W^+$
- If one of PFO<sup>-</sup> in W<sup>+</sup> is classified into W<sup>-</sup>
  - $W^+ \to W^+$ ,  $W^- \to W^-$
- If one of  $PFO^+$  in  $W^-$  is classified into  $W^+$

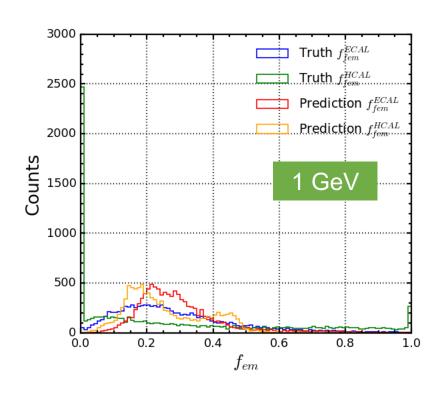
• 
$$W^+ \to W^+$$
,  $W^- \to W^-$ 

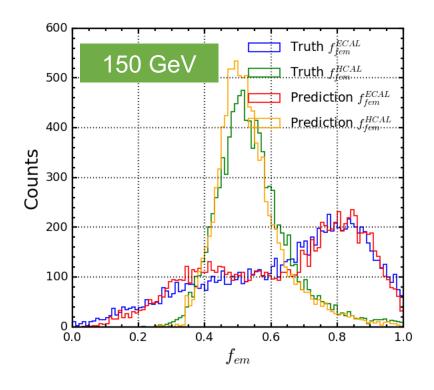
• 
$$\rho_{X,Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} = -0.314$$





#### Software compensation

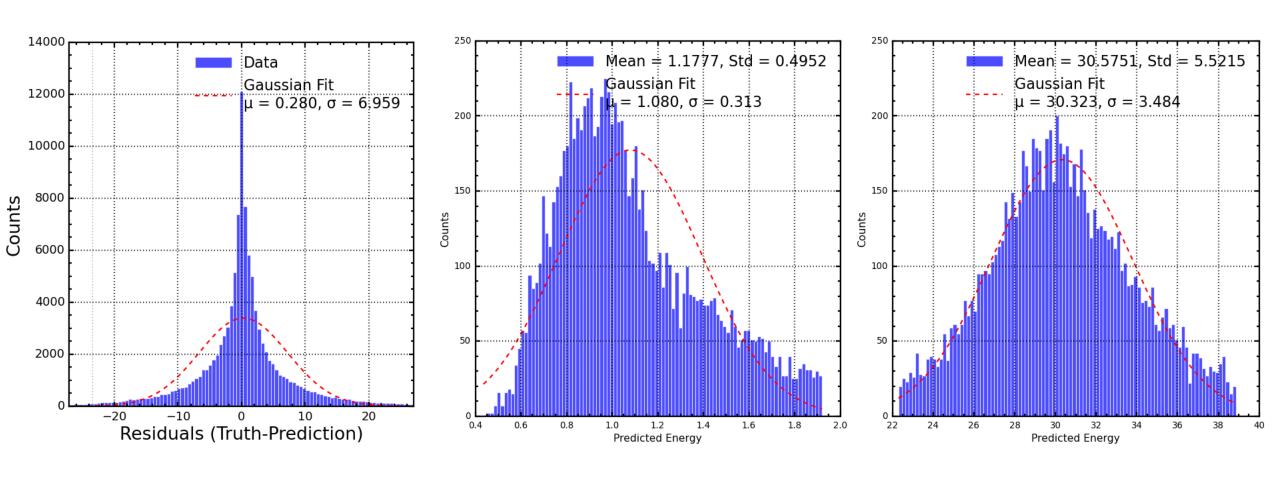






# **Results - Energy**

### Software compensation



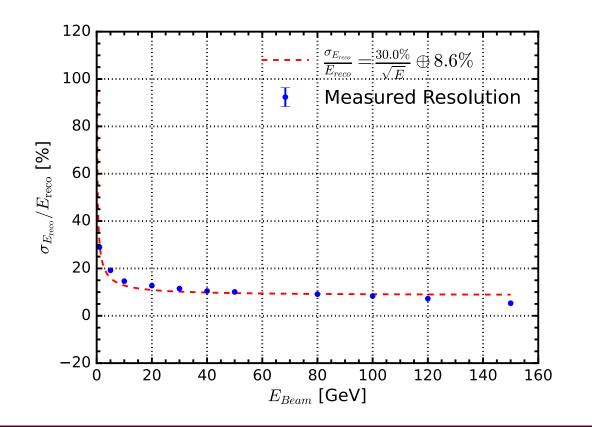


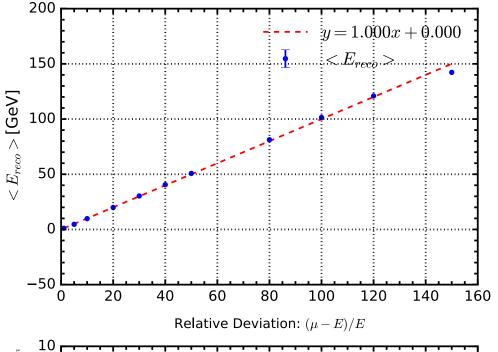
## **Results - Energy**

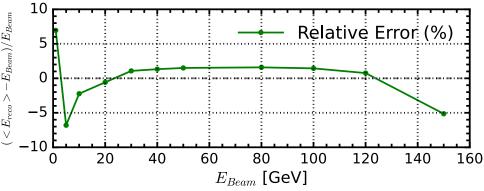
#### Software compensation

Results for particle energy:

- Energy resolution: S =  $30.0\% \sqrt{GeV}$ ; C = 8.6%
- Energy linearity: in 7 %









# **Experiment**

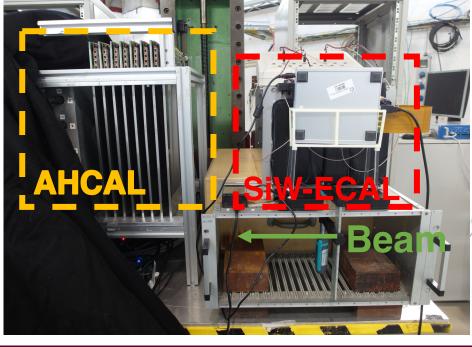
#### Test-Beam

#### **Beam test at DESY:**

- Take part in the beam-test experiment from 02/03/2025 to 10/03/2025
- Prototype: SiW-ECAL and AHCAL









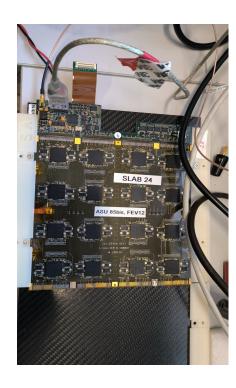


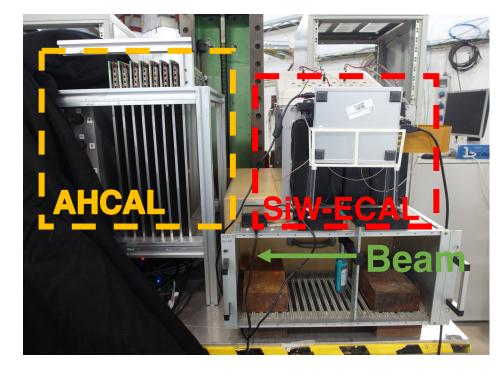
#### **ECAL** geometry

- 3 layers
  - 1 chip on board, 2 produced recently
  - 16 SKIROC chips
  - 32 $\times$ 32 cells: 5.5 $\times$ 5.5  $mm^2$
  - No absorbers

#### **Data Taking**

- ECAL:Mar 4th Mar 6th
  - Configuration and calibration
  - Position scan
- ECAL+HCAL: Mar 7<sup>th</sup> Mar 8<sup>th</sup>
  - Configuration and calibration
  - Position scan
  - TDC test





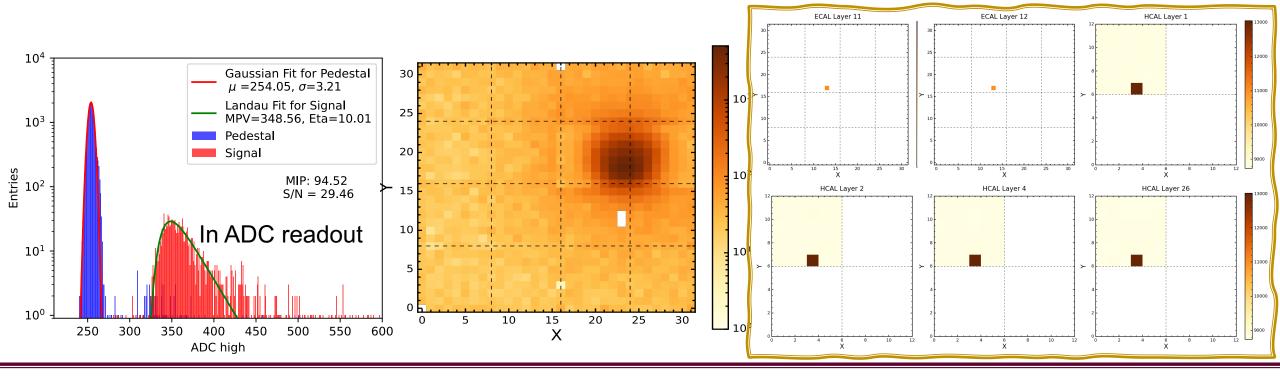


## data analysis

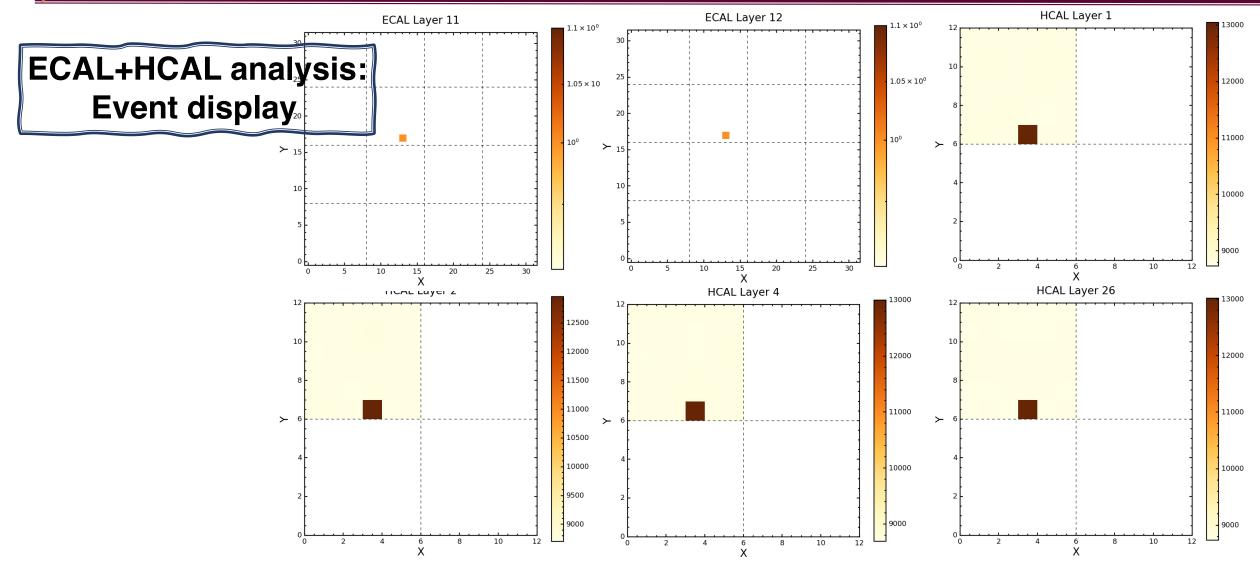
#### Test-Beam

#### **ECAL** response & **ECAL-HCAL** matching

- Validate the performance of the SiW-ECAL prototype under realistic beam conditions.
- Assess detector stability and data quality to ensure reliable energy measurements.
- Study the integration of ECAL and HCAL for combined calorimetry.



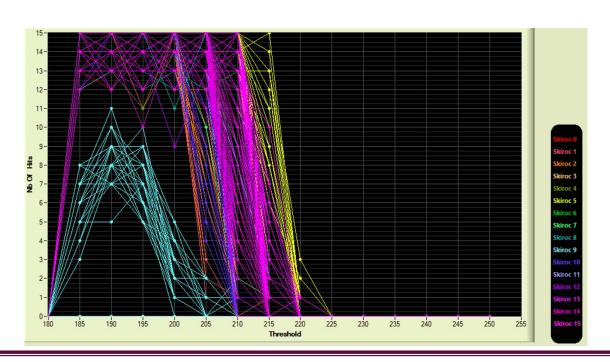






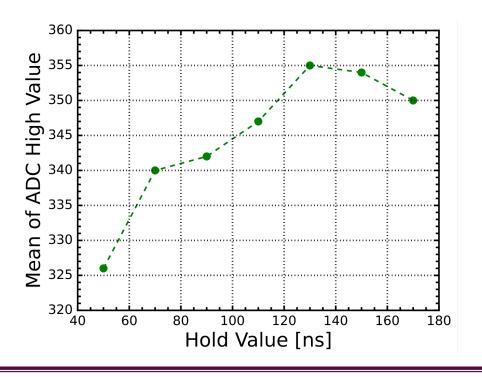
#### Threshold scan

- Scanning was processed separately for each channel without signal
- A threshold of 230 ADC was applied during the beam test, with exceptions for certain chips in specific runs.

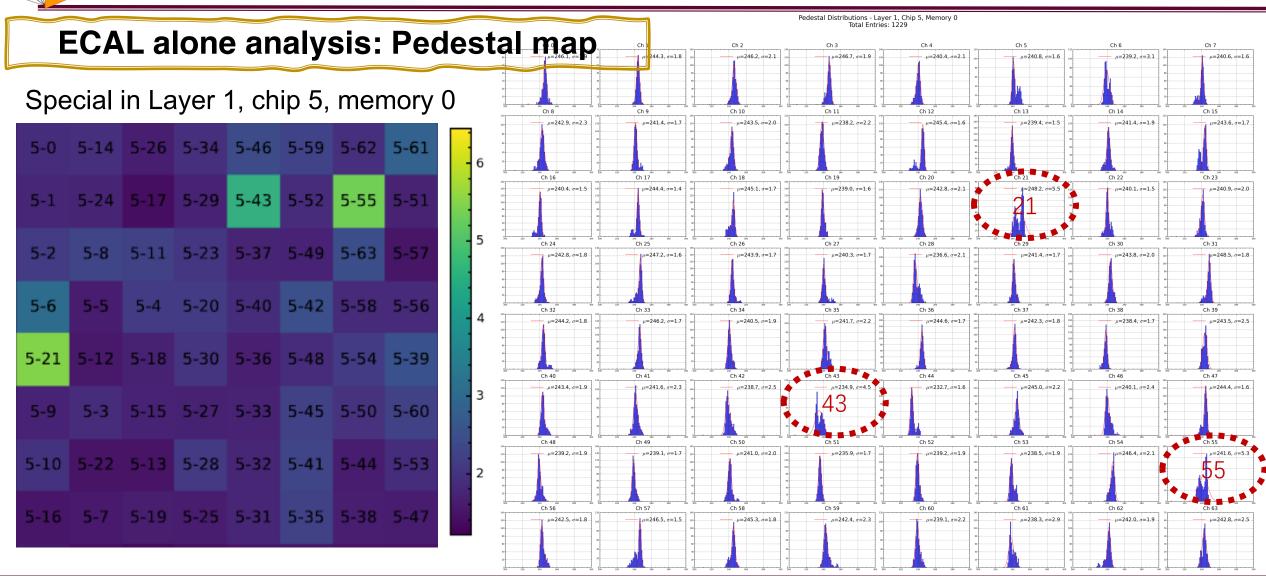


#### Hold value scan

- Hold value was scanned with beam on
- The hold value was set up as 130 DAC(1 DAC is approximately 1 ns)







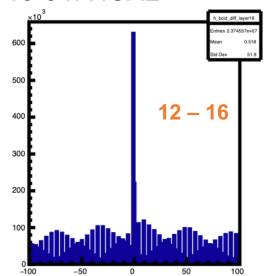


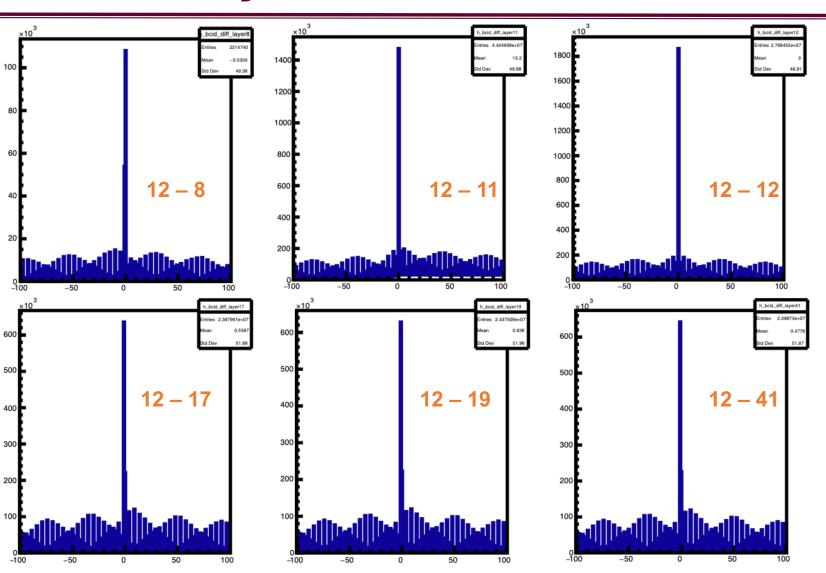
#### ECAL+HCAL analysis: Bunch crossing ID match

#### Layer:

• 0-14: ECAL

• 15-54: HCAL







#### **ECAL** alone analysis: Pedestal map

- In each chip, the pedestals are relatively consistent.
- But in some chips, the pedestal sigma is higher than 4 because of two peaks.
- Need further check.

