# Status and prospects of DNN-based reconstruction for future Higgs factories

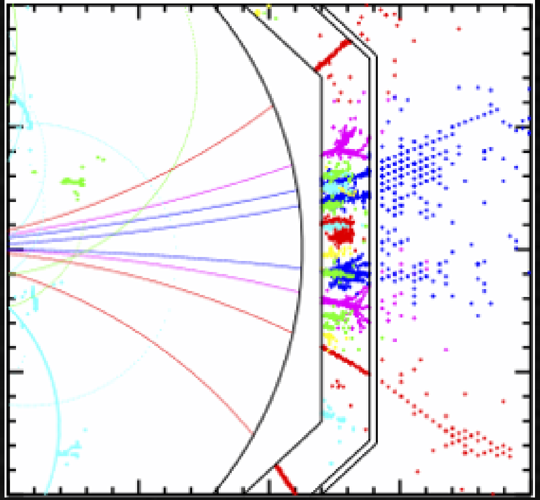
Main results given in
arXiv:2410.08772 (Proc. ICHEP2024)

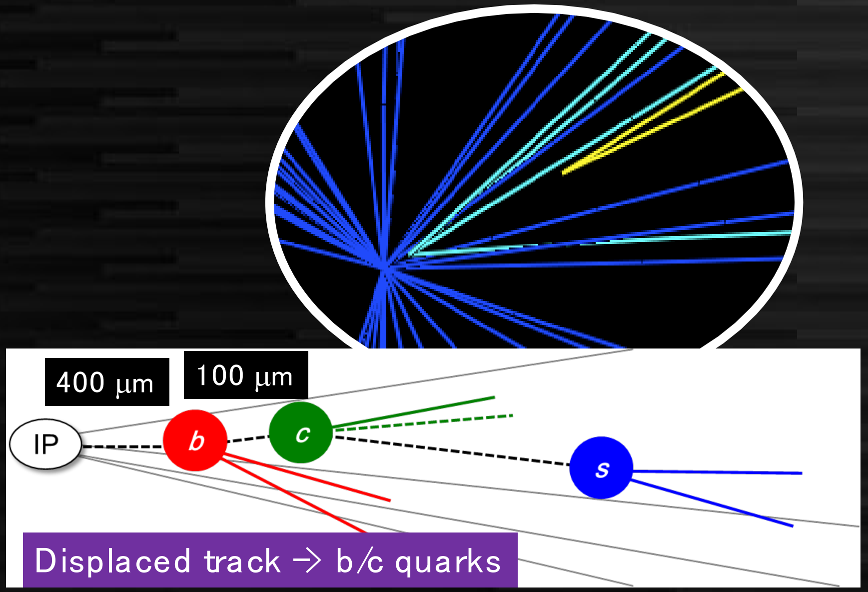
Taikan Suehara / 末原　大幹
(ICEPP, The University of Tokyo)

Collaborators:
R. Tagami, T. Murata, W. Ootani, M. Ishino (ICEPP, UTokyo),
P. Wahlen (ETH/IP Paris/ILANCE UTokyo),
L. Gui (Imperial/Kyushu U.), T. Tanabe (MI-6 Ltd.)

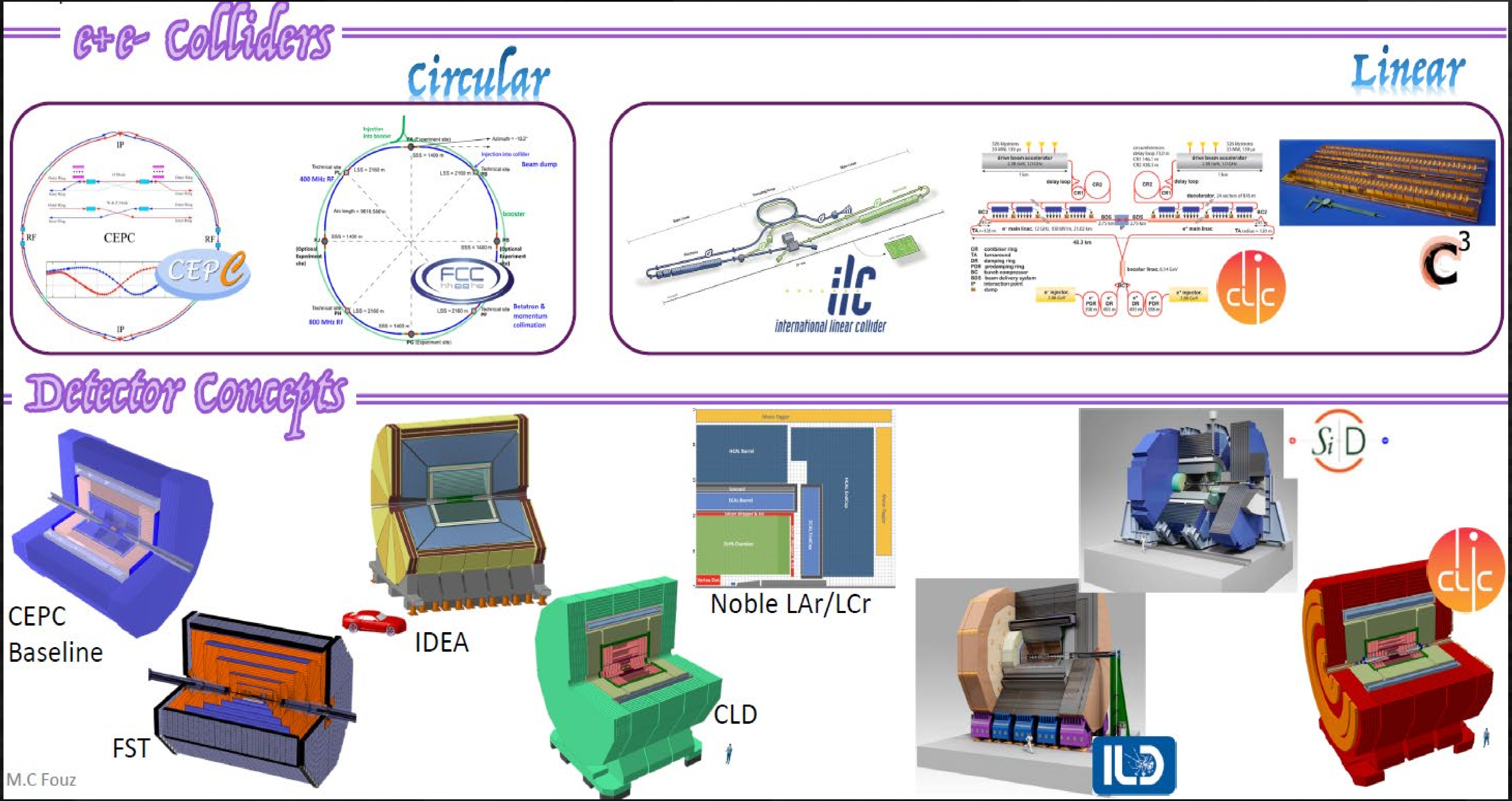
# Today's topics

- Introduction on Higgs factories – including situations
- Detectors and Reconstruction strategies
- ML topics (status and prospects)
  - Particle flow with GNN / Transformer
  - Flavor tagging with ParT





# Higgs factories and detectors



e+e- Colliders

Circular

Linear

Detector Concepts

# $e^+e^-$ collider projects

- Linear colliders
  - ILC (Japan) 250 GeV (initial) → multi-TeV
    Superconducting LC to be started in end of 2030s. The most mature project.
  - CLIC (CERN) 380 GeV → 3 TeV
    Normal conducting (X-band) LC. The alternative option to FCC in EPPSU. Affordable for CERN.
  - CCC (US) 250 GeV → multi-TeV
    Cooled normal conducting (C-band) LC. Currently at Pre-CDR. Realization in > 2040.
- Circular colliders
  - FCCee (CERN) 91 GeV → 240 GeV → 365 GeV
    Coupled with 100 TeV hadron collider. Need non-CERN contribution. Operation start at 2048 (at Z-pole?)
  - CEPC (China)
    Slightly conservative than FCCee. TDR just published. To be upgraded to SppC (hadron collider)

# Target Energies of e+e- colliders

91~250 GeV
Oblique parameters, W/Z mass, b/τ rare decays

250 GeV
Higgs couplings (~1%), Higgs rare decay (light BSM)
(TeV BSM indirect search)

350 GeV
Top mass → vacuum stability

500–550 GeV
Higgs self coupling (20–30%), ttH coupling

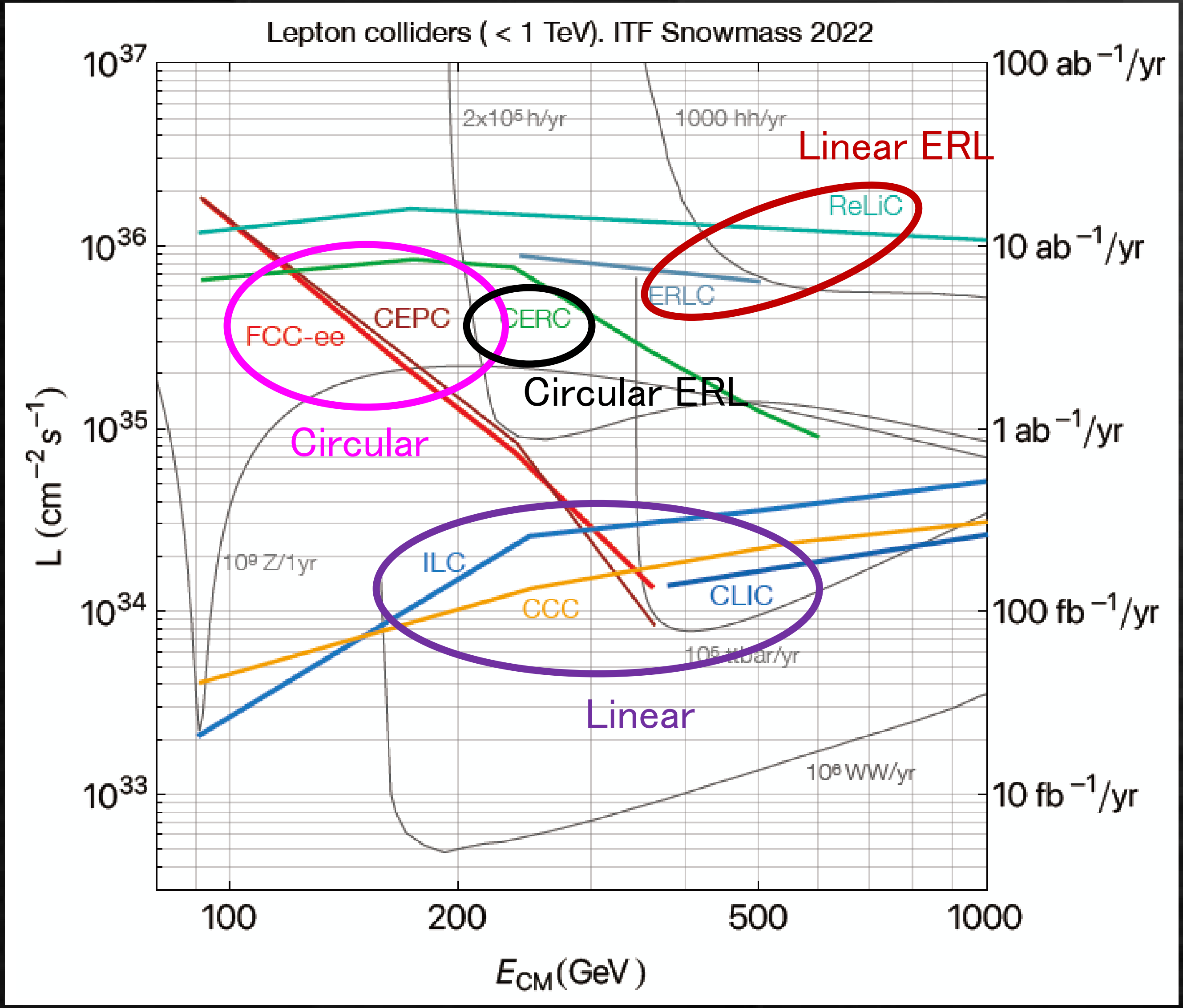
1 TeV
Higgs self coupling (10%) → baryogenesis

250 GeV – a few TeV
TeV BSM direct search

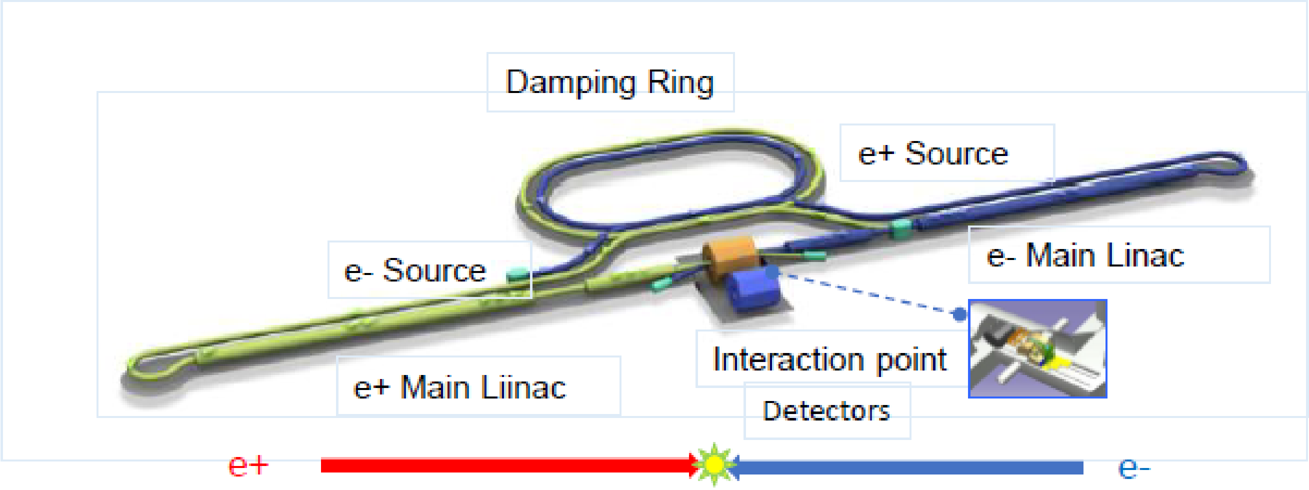
Natural SUSY (250 GeV – 1 TeV) 1 TeV Higgsino 3 TeV Wino

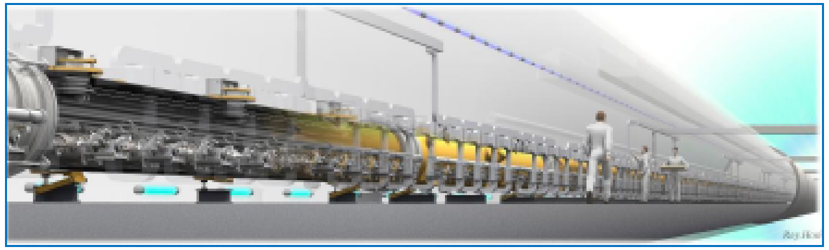
# Circular and Linear collider?

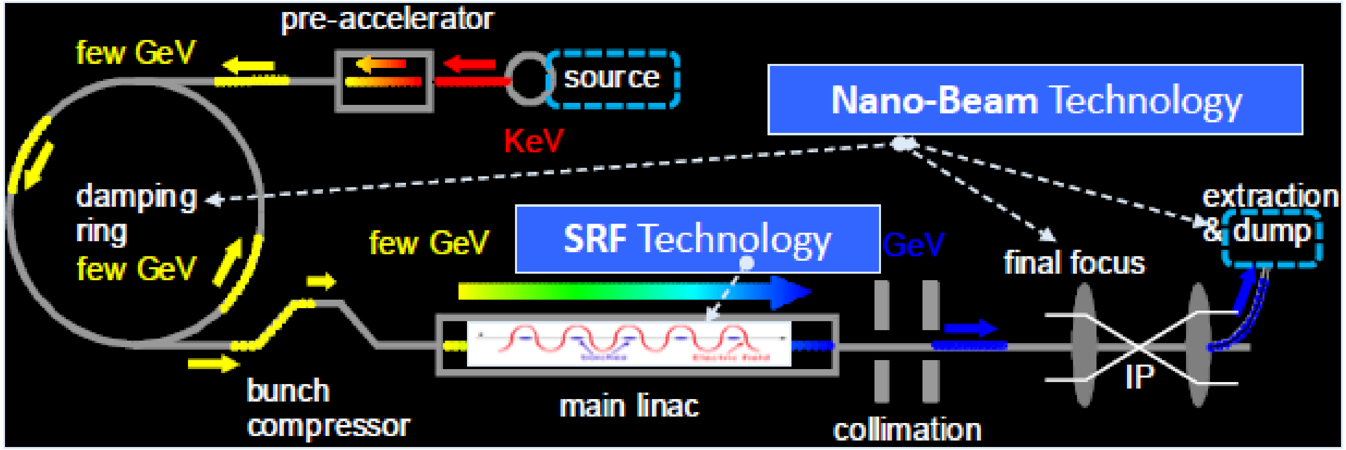
- Luminosity @ 240/250 GeV
  - – A few times higher at circular colliders
- Luminosity @ 350 GeV
  - – Less efficient with circular
- Polarization
  - – Obvious in LC
  - – Not excluded but not guaranteed in circular
- Self coupling, ttH
  - – Indirect only in circular



# ILC: accelerator overview



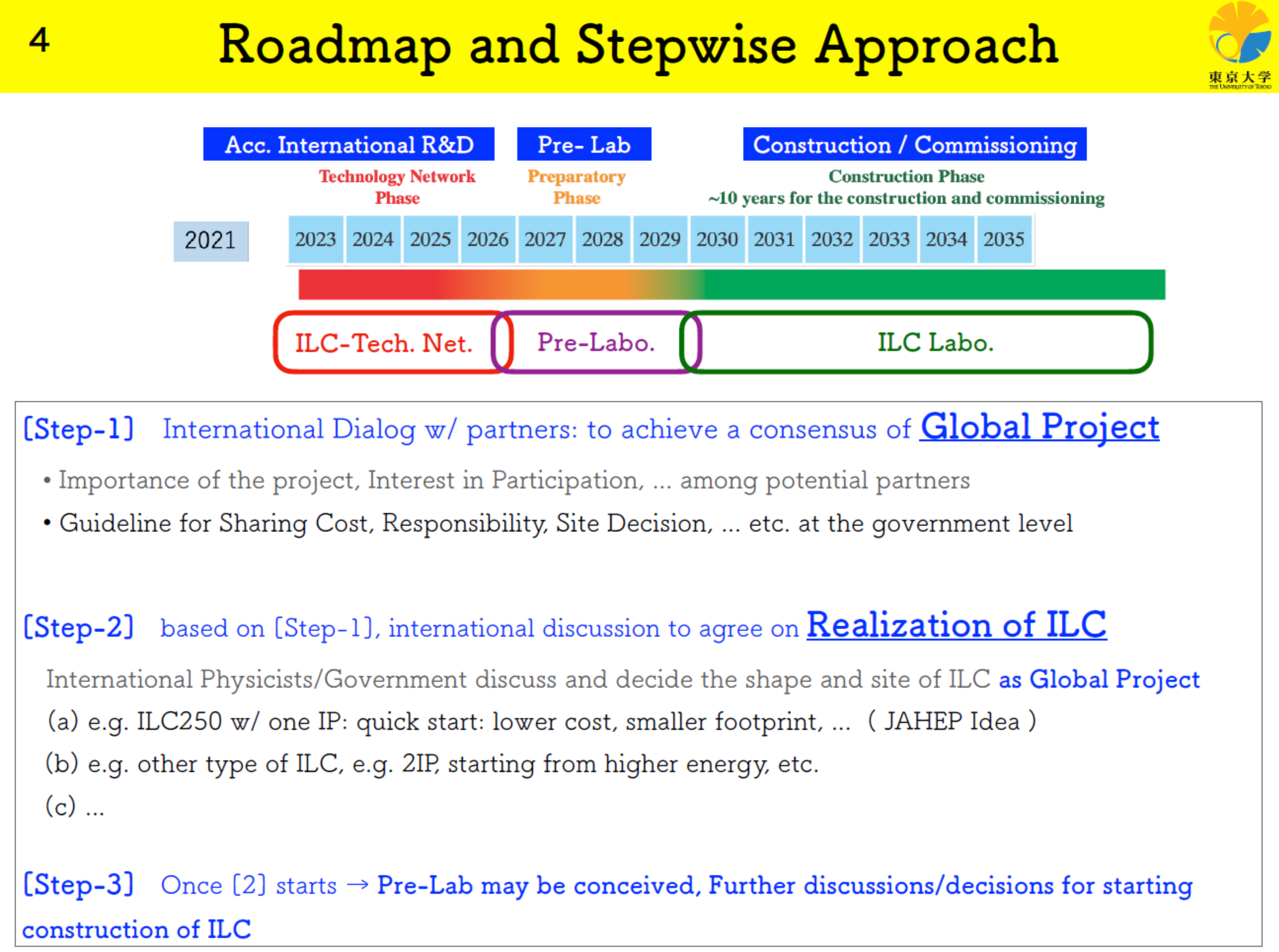




| Parameters | Value |
|---|---|
| Beam Energy | 125 + 125 GeV |
| Luminosity | 1.35 / 2.7 x $10^{10}$ $cm^2/s$ |
| Beam rep. rate | 5 Hz |
| Pulse duration | 0.73 / 0.961 ms |
| # bunch / pulse | 1312 / 2625 |
| Beam Current | 5.8 / 8.8 mA |
| Beam size (y) at FF | **7.7** nm |
| SRF Field gradient | < 31.5 > MV/m (+/-20%)<br>$Q_0$ = $1x10^{10}$ |
| #SRF 9-cell cavities (CM) | ~ 8,000 (~ 900) |
| AC-plug Power | 111 / 138 MW |

S. Michizono, LCWS2023

M. Ishino



# Roadmap and Stepwise Approach

**[Step-1]** International Dialog w/ partners: to achieve a consensus of **Global Project**

- Importance of the project, Interest in Participation, … among potential partners
- Guideline for Sharing Cost, Responsibility, Site Decision, … etc. at the government level

**[Step-2]** based on [Step-1], international discussion to agree on **Realization of ILC**

International Physicists/Government discuss and decide the shape and site of ILC **as Global Project**

(a) e.g. ILC250 w/ one IP: quick start: lower cost, smaller footprint, … ( JAHEP Idea )

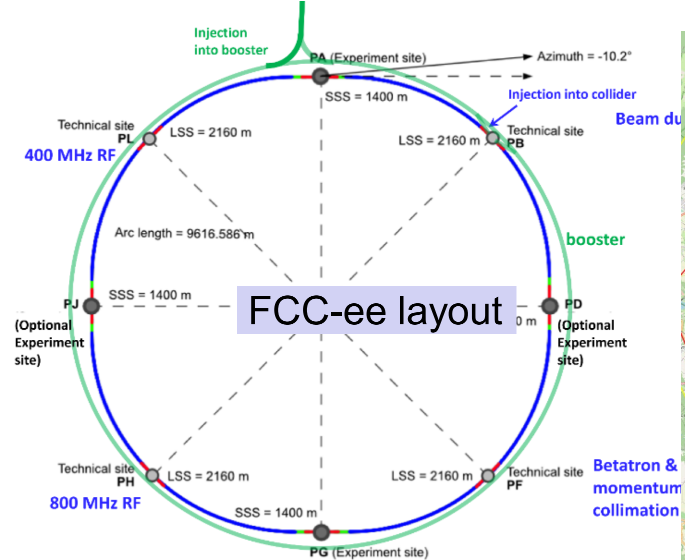
(b) e.g. other type of ILC, e.g. 2IP, starting from higher energy, etc.

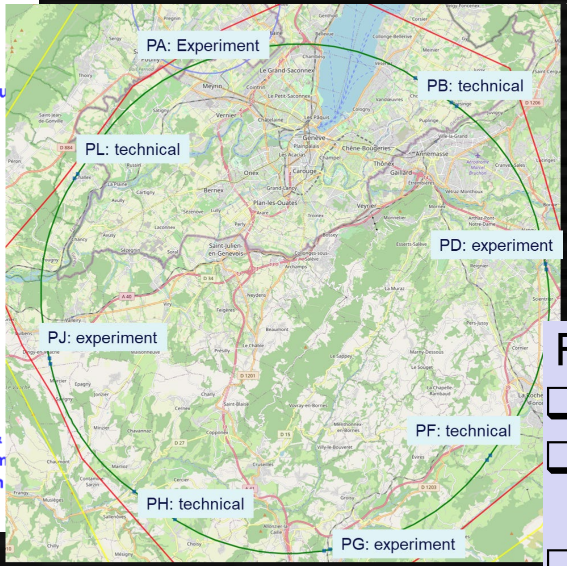
(c) …

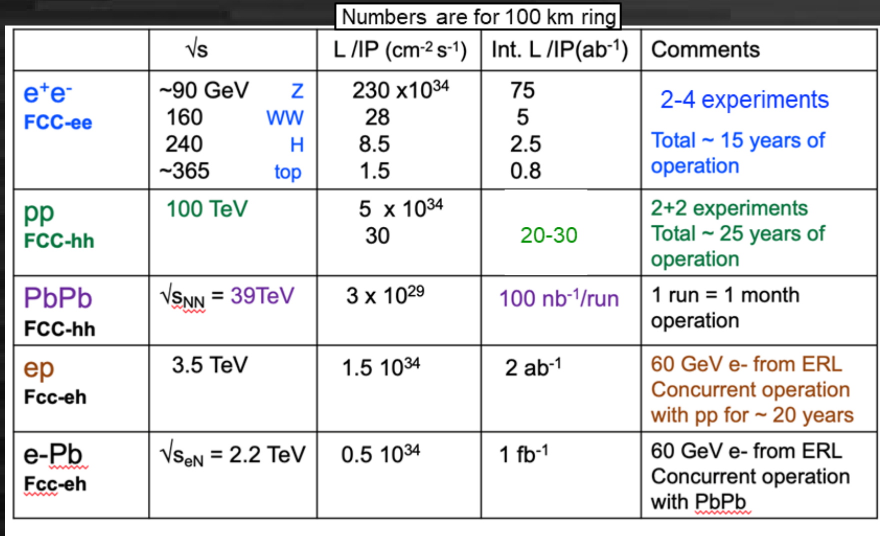
**[Step-3]** Once [2] starts → **Pre-Lab may be conceived, Further discussions/decisions for starting construction of ILC**

# FCCee

91.1 km tunnel, 2 IP, CM energy 91-365 GeV





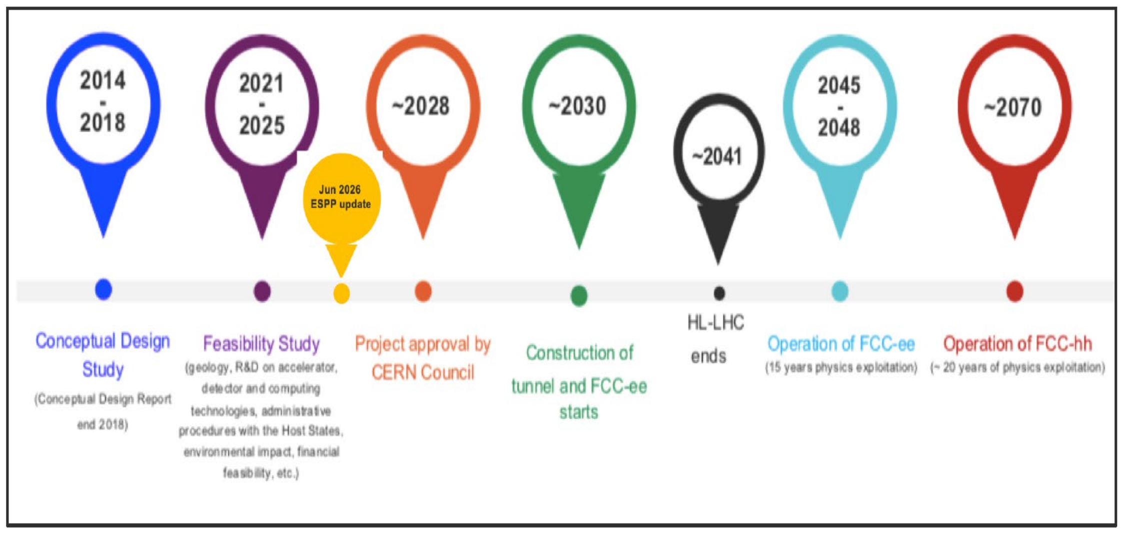


Numbers are for 100 km ring

| | $\sqrt{s}$ | L /IP ($cm^{-2} s^{-1}$) | Int. L /IP($ab^{-1}$) | Comments |
|---|---|---|---|---|
| $e^+e^-$<br>**FCC-ee** | ~90 GeV Z<br>160 WW<br>240 H<br>~365 top | 230 x$10^{34}$<br>28<br>8.5<br>1.5 | 75<br>5<br>2.5<br>0.8 | 2-4 experiments<br>Total ~ 15 years of operation |
| pp<br>**FCC-hh** | 100 TeV | 5 x $10^{34}$<br>30 | 20-30 | 2+2 experiments<br>Total ~ 25 years of operation |
| PbPb<br>**FCC-hh** | $\sqrt{s_{NN}}$ = 39TeV | 3 x $10^{29}$ | 100 $nb^{-1}$/run | 1 run = 1 month operation |
| ep<br>**Fcc-eh** | 3.5 TeV | 1.5 $10^{34}$ | 2 $ab^{-1}$ | 60 GeV e- from ERL<br>Concurrent operation with pp for ~ 20 years |
| e-Pb<br>**Fcc-eh** | $\sqrt{s_{eN}}$ = 2.2 TeV | 0.5 $10^{34}$ | 1 $fb^{-1}$ | 60 GeV e- from ERL<br>Concurrent operation with PbPb |

Feasibility Study:

- ❑ Focus is on FCC-ee and magnet R&D
- ❑ ~ 40 MCHF/year from CERN budget (half for magnet R&D) Additional funding from EU and collaborating institutes (e.g. CHA
- ❑ Results will be summarised in Feasibility Study Report in 2025



## Message from CERN council, FCC week June 2024

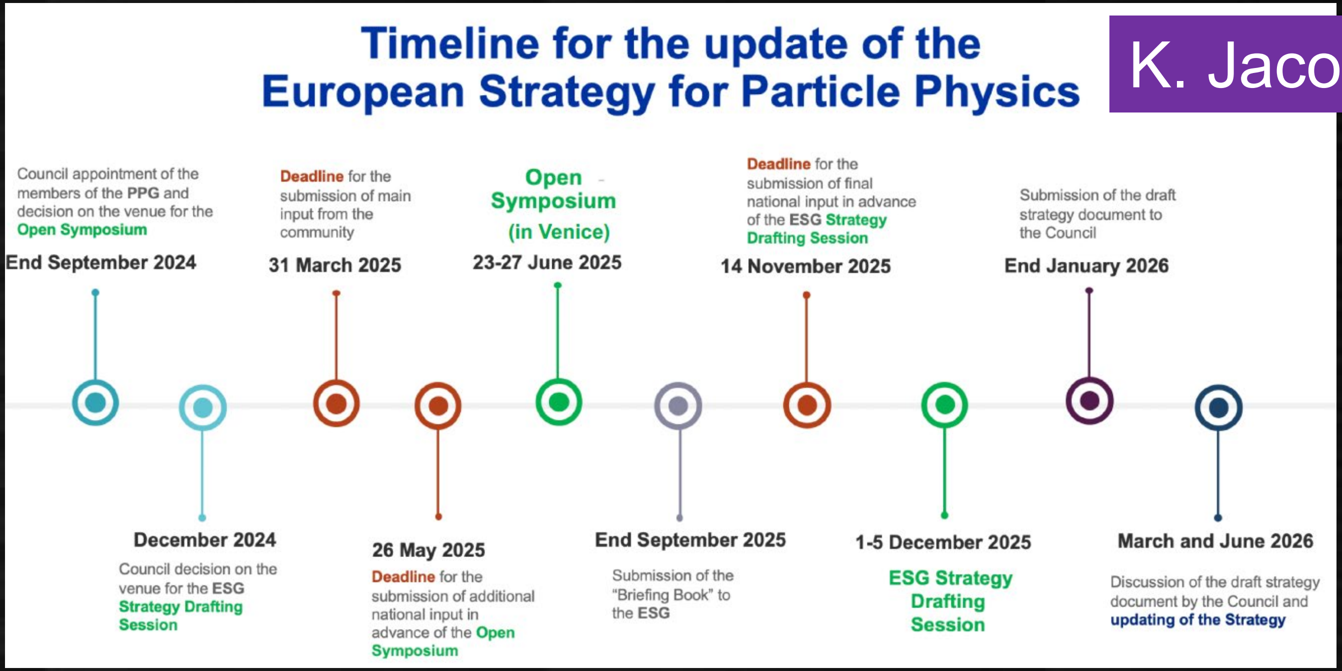
- CERN COUNCIL IS UNITED IN THE VISION OF WANTING TO MAKE SURE CERN CONTINUES TO PROVIDE THE MOST INTERERSTING AND SCIENTIFIC POSSIBILITIES AND THE MOST ADVANCED TECHNOLOGICAL TOOLS TO DO THE PHYSICS.
- CERN COUNCIL DOES NOT, YET, HAVE A CONSENSUS ON HOW THE ACTUALIZE THIS VISION.

# European strategy update

*In June 2024, the CERN Council established and approved the* ***remit of the European Strategy Group***

*"The aim of the Strategy update should be to develop a* ***visionary and concrete plan*** *that greatly advances human knowledge in fundamental physics through the* ***realisation of the next flagship project at CERN****. This plan should attract and value* ***international collaboration*** *and should* ***allow Europe to continue to play a leading role in the field."***

*The Strategy update should include the* ***preferred option*** *for the next collider at CERN and* ***prioritised alternative options*** *to be pursued if the chosen preferred plan turns out not to be feasible or competitive.*



**Timeline for the update of the European Strategy for Particle Physics**

**End September 2024**
Council appointment of the members of the **PPG** and decision on the venue for the **Open Symposium**

**December 2024**
Council decision on the venue for the **ESG Strategy Drafting Session**

**31 March 2025**
**Deadline** for the submission of main input from the community

**26 May 2025**
**Deadline** for the submission of additional national input in advance of the **Open Symposium**

**23-27 June 2025**
**Open Symposium (in Venice)**

**End September 2025**
Submission of the "Briefing Book" to the ESG

**14 November 2025**
**Deadline** for the submission of final national input in advance of the **ESG Strategy Drafting Session**

**1-5 December 2025**
**ESG Strategy Drafting Session**

**End January 2026**
Submission of the draft strategy document to the Council

**March and June 2026**
Discussion of the draft strategy document by the Council and **updating of the Strategy**

K. Jacobs

Higgs factory is one of primary focus on European strategy

- FCCee
- LC@CERN

As well as

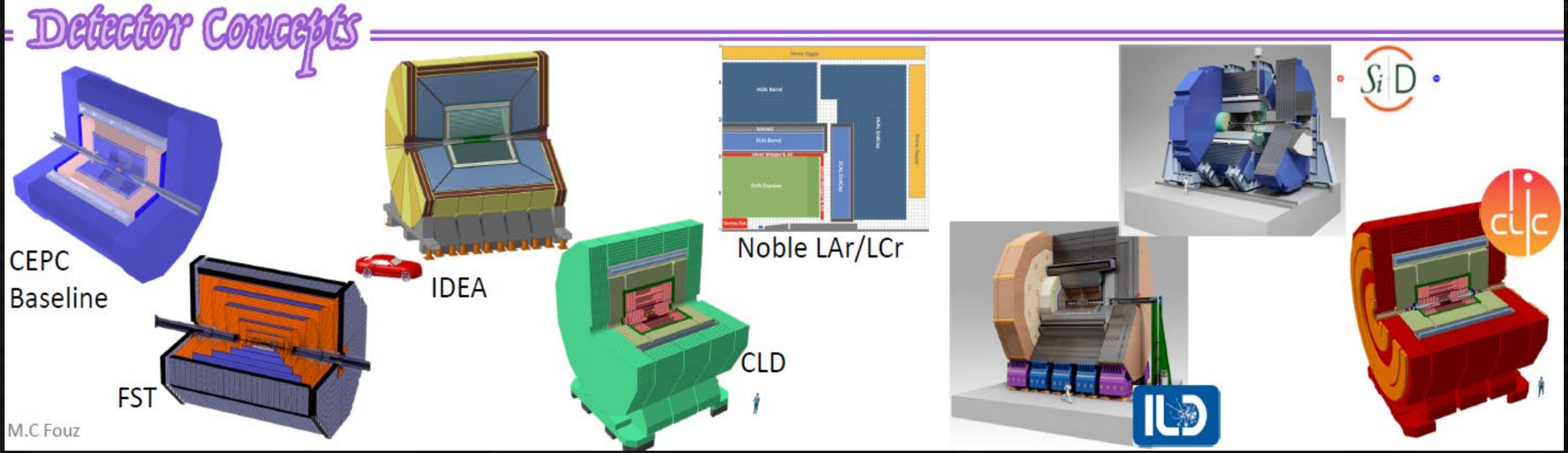
- Low energy hadron collider
- ep collider

# Personal observations

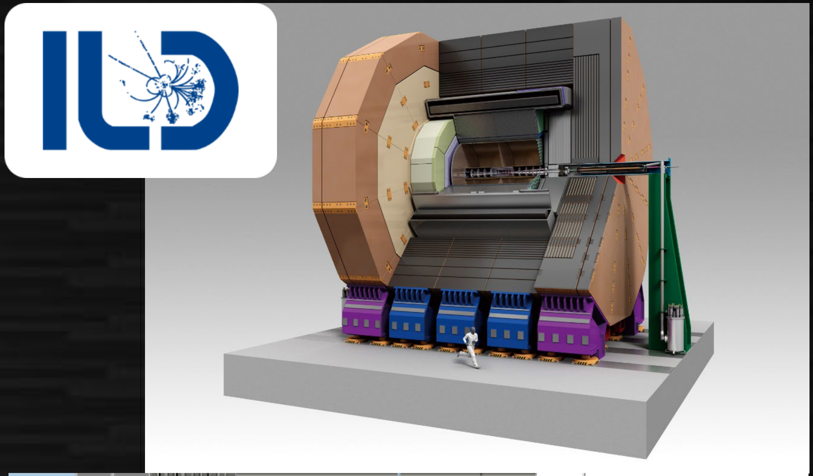
- Higgs factory projects are recognized as a strong candidate for the next high-energy collider
  - With solid case of physics
  - With upgrade paths
- Uncertainty of the realization is still high
  - Cost of FCCee is (~2x) beyond CERN's affordable budget
  - Uncertainty of global projects (and weak JPY for ILC in Japan)
  - CEPC (ask Manqi for details)
- We will see some changes (hopefully progress) in a few years
  - By European strategy update and CEPC progress

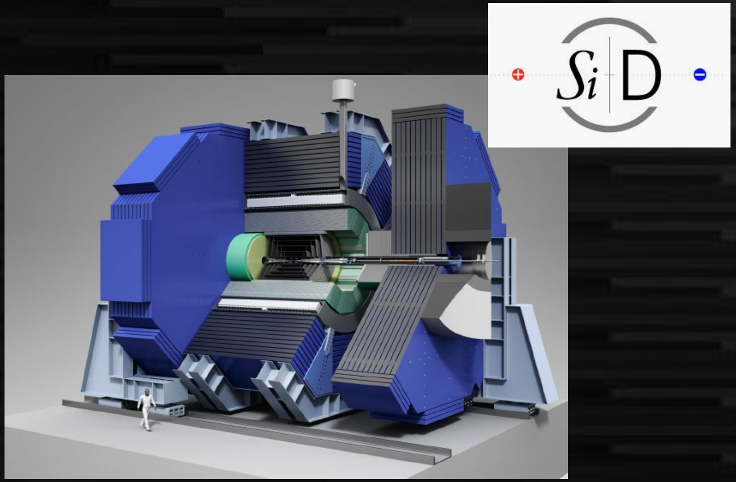
# Higgs factories detector concepts

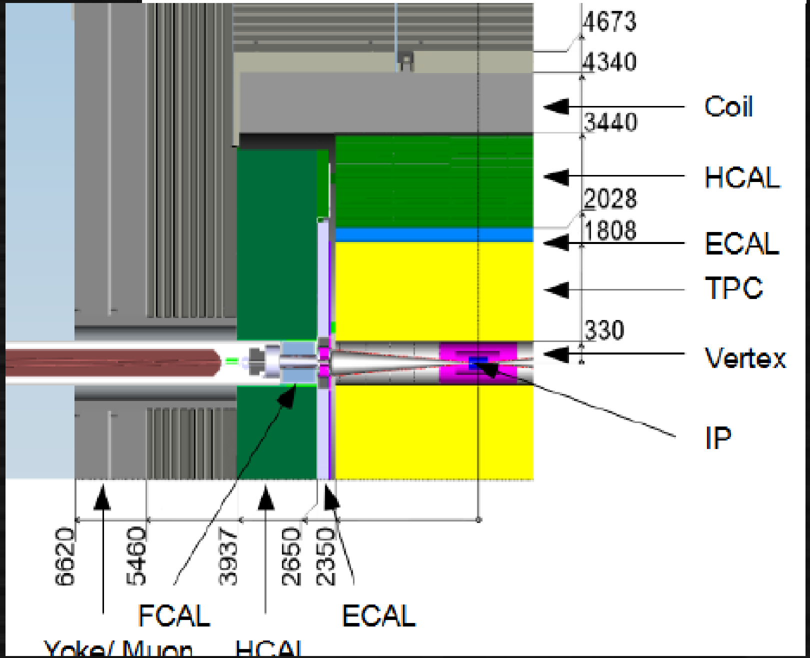
- Calorimeter concepts
  - – Particle flow with high granular calorimeters
    - Sensor options: silicon, scintillator, RPC, noble liquid
  - – Dual-readout calorimetry (scintillator + Cherenkov)
- Tracker options: drift chamber, TPC, silicon only
- Optional subdetectors: picosec timing, Cherenkov

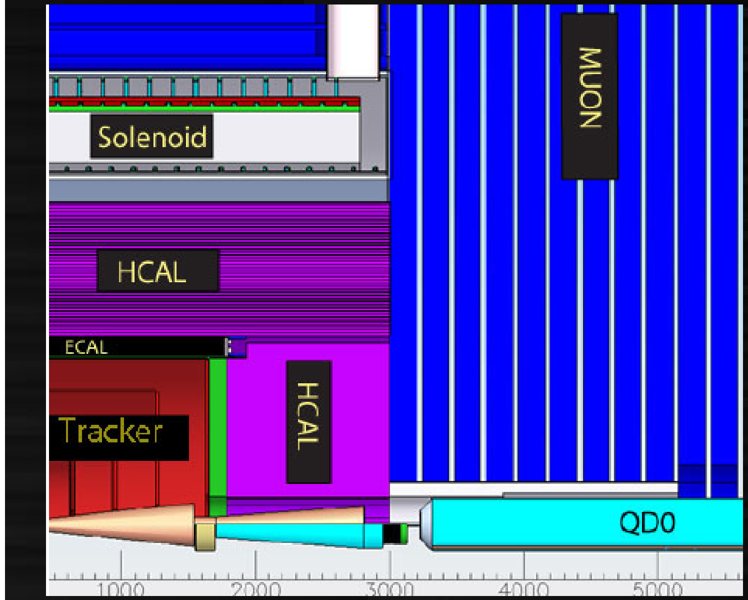


# Particle flow detector concepts









Two (similar) concept based on Particle Flow reconstruction

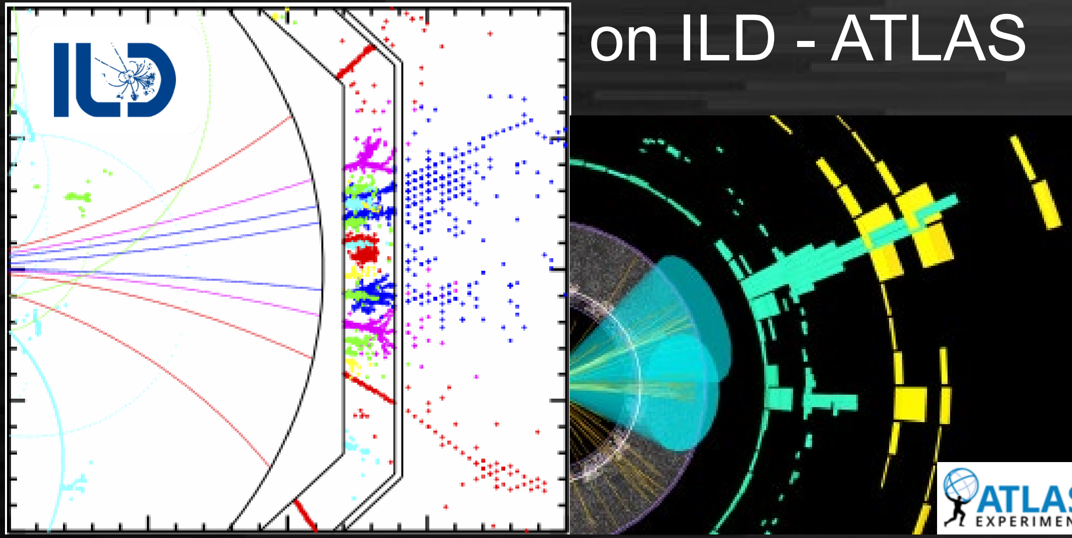
Already mature baseline design

- Monolithic silicon vertex
- Silicon tracker (inner/outer for ILD)
- Time projection chamber (only for ILD)
- Highly-granular ECAL/HCAL with several options
  - Silicon pads
  - Scintillator strips/tiles
  - Resistive plate chamber
  - Silicon pixels (MAPS)
- 3.5/5T solenoid outside HCAL

# Particle flow concept

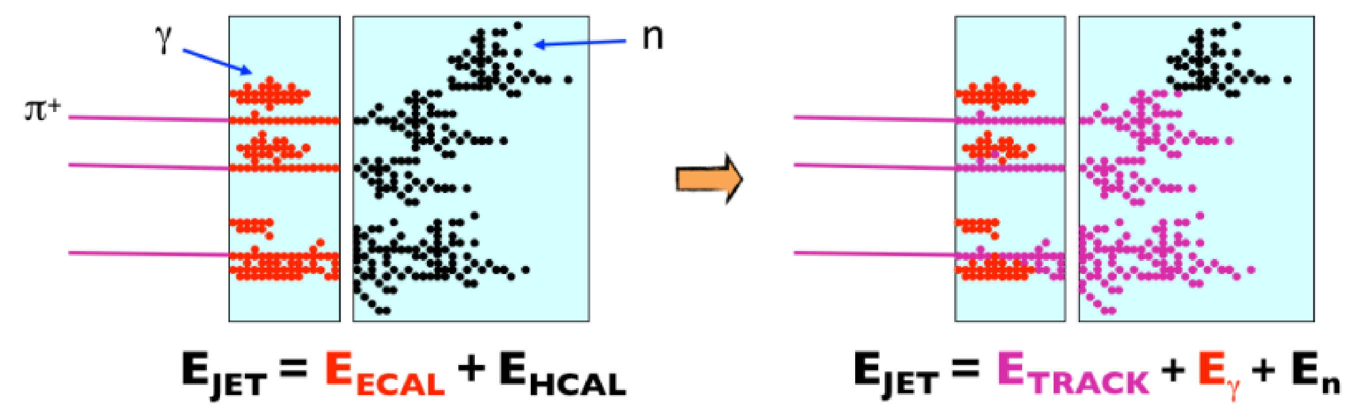
Separating particles inside jets to do track-cluster matching

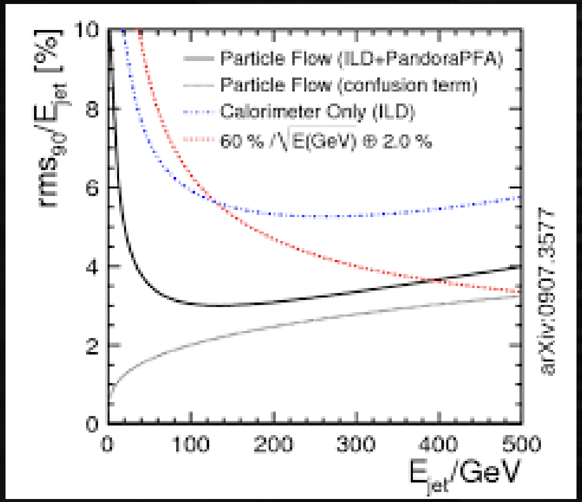
Requiring
- Highly-granular calorimeters
- Intelligent pattern recognition



Different granularity on ILD - ATLAS

Developed in ILC, first full application in CMS HGCAL at HL-LHC (partial use already in ATLAS/CMS)



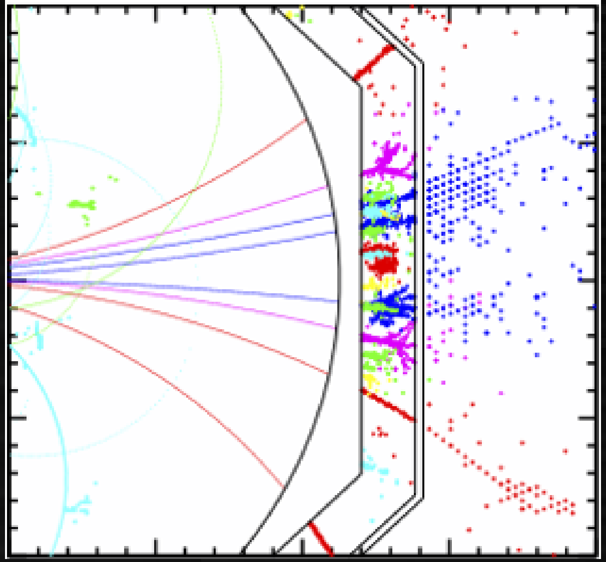


Possible to obtain jet energy resolution of

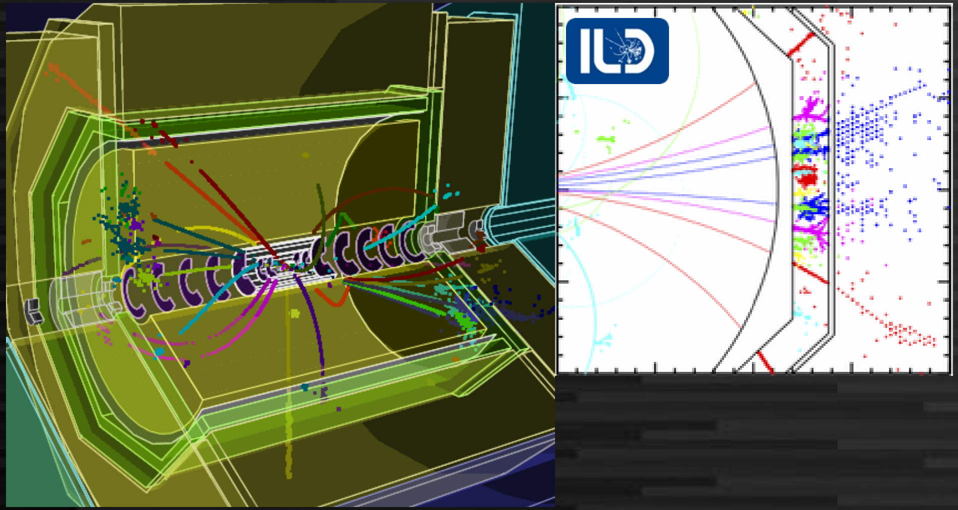
$$\frac{\delta E_{\mathrm{jet}}}{E_{\mathrm{jet}}} \cong \frac{30\%}{\sqrt{E_{\mathrm{jet}}[\mathrm{GeV}]}}$$

~2 times better than calo-only

# Particle flow with DNN

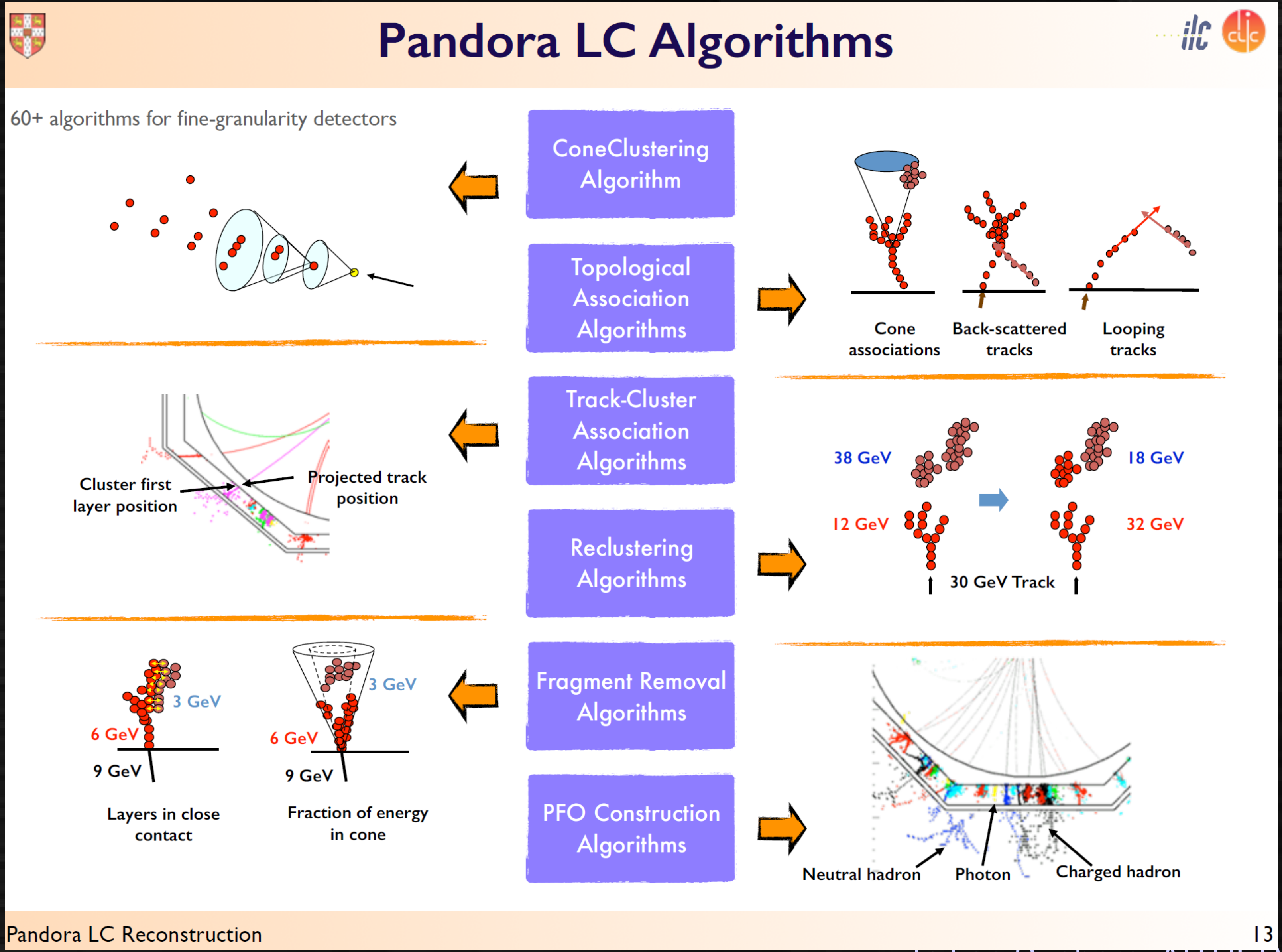


# Particle flow for Higgs factories



- High granular calorimetry
  - – 3D pixels for imaging EM/hadron showers at calorimeters
    - eg. $10^8$ channels for ILD ECAL
  - – Separation of particles inside jets → ~2x better energy resolution by separation of contribution from charged particles
    - Software algorithm essential (as well as hardware design)
- Particle Flow algorithm
  - – Essential algorithm for high granular calorimetry
  - – Complicated pattern recognition → good for DNN

# Pandora ParticleFlow algorithm



## Pandora LC Algorithms

60+ algorithms for fine-granularity detectors

ConeClustering Algorithm

Topological Association Algorithms

Cone associations

Back-scattered tracks

Looping tracks

Track-Cluster Association Algorithms

Cluster first layer position

Projected track position

Reclustering Algorithms

38 GeV

12 GeV

18 GeV

32 GeV

30 GeV Track

Fragment Removal Algorithms

3 GeV

6 GeV

9 GeV

Layers in close contact

3 GeV

6 GeV

9 GeV

Fraction of energy in cone

PFO Construction Algorithms

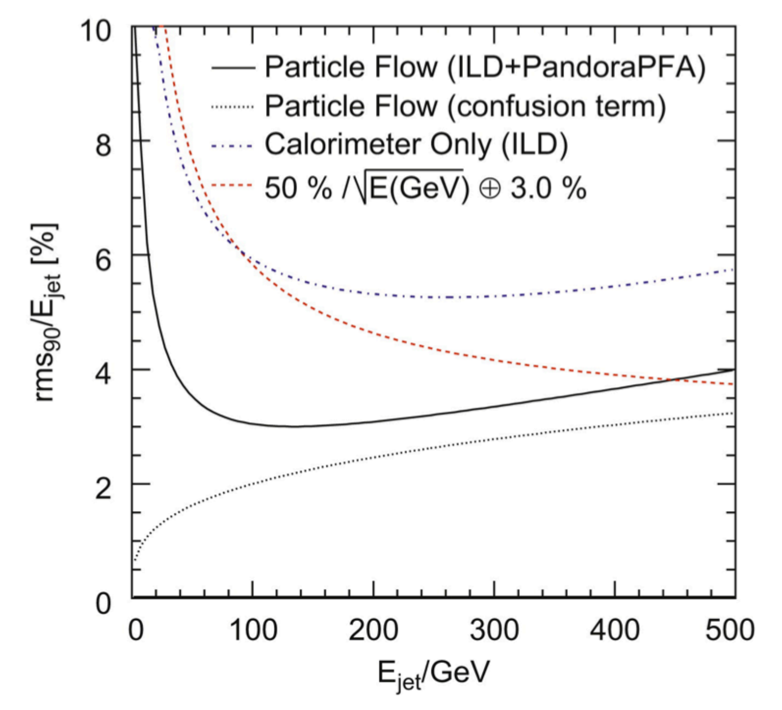
Neutral hadron

Photon

Charged hadron

Pandora LC Reconstruction

13



Widely used since 2008

Reasonably good performance up to ~50 GeV jets

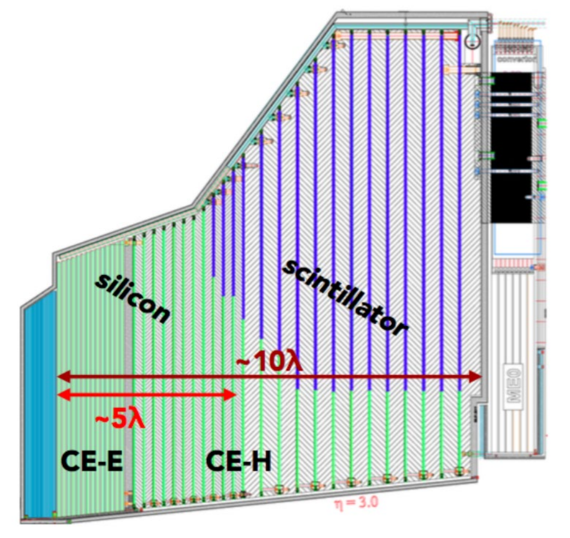
Confusion dominates at higher energies

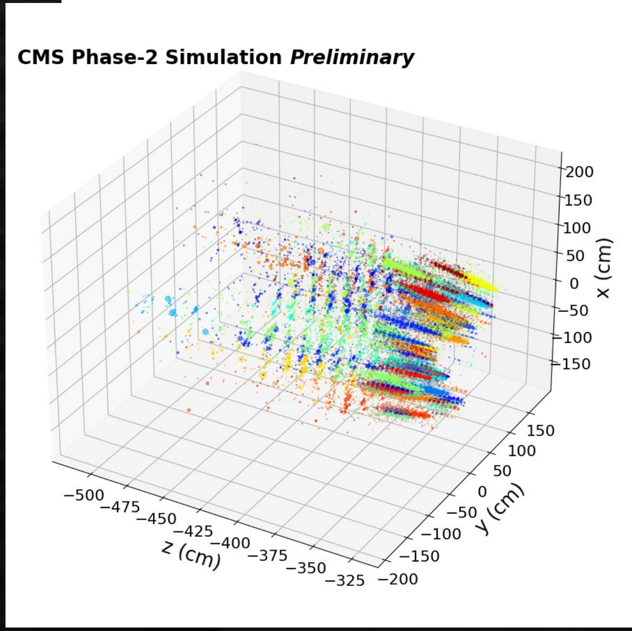
# Motivations for DNN particle flow

- Performance improvement
  - Confusion dominant at jet energy > 100 GeV
  - More efficient way to separate cluster from charged particles should be investigated
- Integrate other functions
  - Software compensation, particle ID etc. closely related to PFA
- Detector optimization
  - Comparison with different detector settings
    - PandoraPFA too much depends on internal parameters
  - Effect of timing information to be investigated
    - With different timing resolution (1 ns, 100 ps, 10 ps, …)

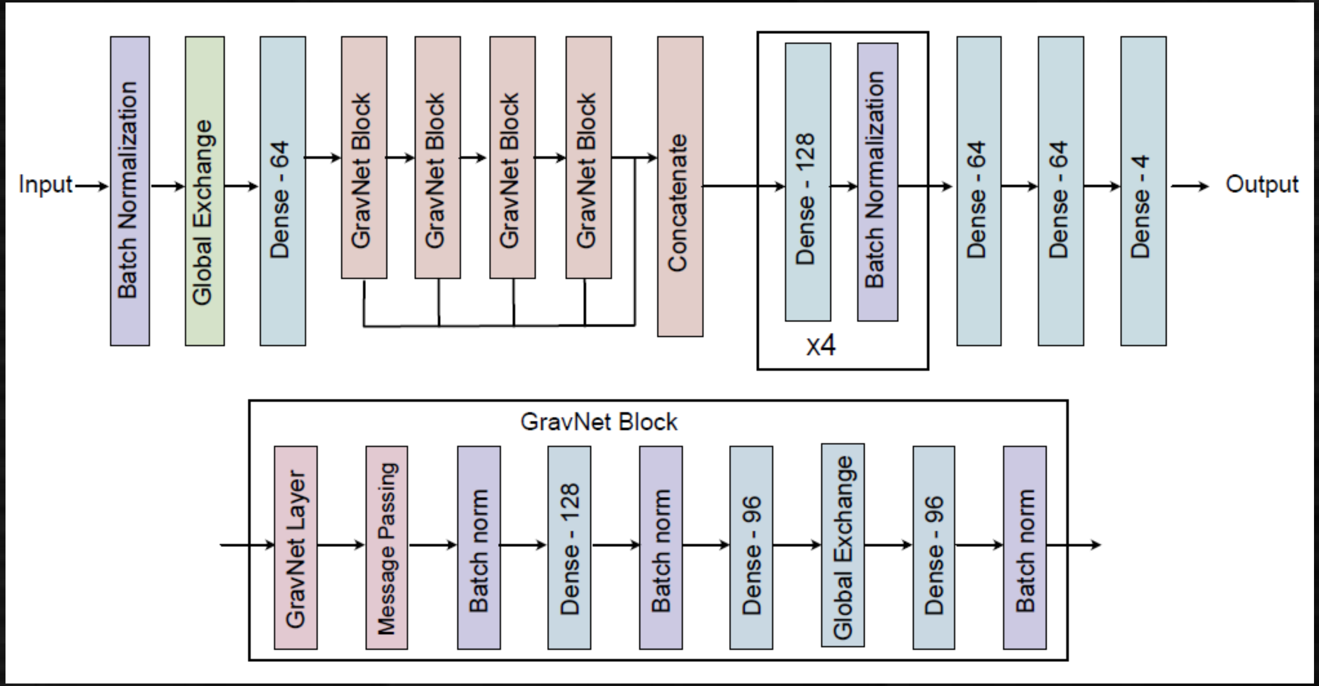
# GravNet for CMS HGCAL

- CMS HGCAL
  - – High granular forward calorimeter for HL-LHC upgrade at CMS
  - – Similar to ILD calorimeter (silicon pixel + scintillator)
    - Inspired by CALICE development
- Reconstruction at HGCAL
  - – Pileup/noise to be separated by software
  - – Numerous particles from ~200 pileups
    - Difficult to handle: software algorithm critical
    - DNN reconstruction being investigated
      - – Reasonable performance obtained up to ~50 pileups?





# The network



Rather complicated network with ~30 hidden layers

“Object condensation” loss function is applied (shown in next page)

Input/output obtained for each hit at calorimeter

Input: Features at each hit (position, energy deposit, timing)

Output: “condensation coefficient” $\beta$, position at virtual coordinate (2-dim)
optional output of features such as energy, PID (not used now)

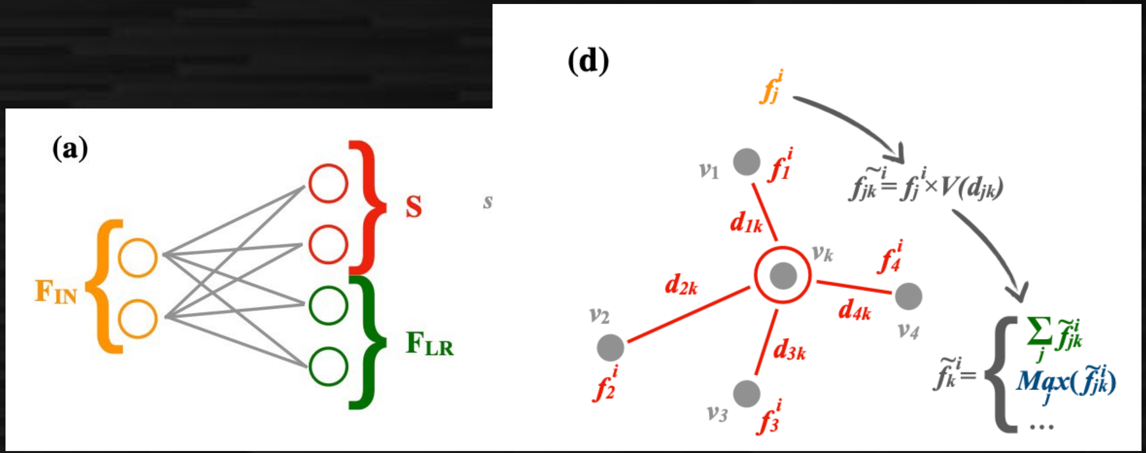
Dense (fully-connected layer) inside each hit, GravNet connects hits

# GravNet and Object Condensation

## GravNet

arXiv:1902.07987

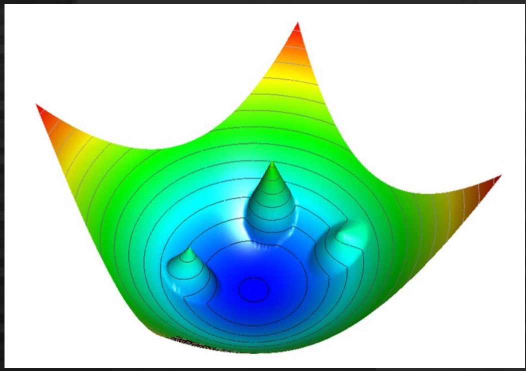
- The virtual coordinate (S) is derived from input variables with simple MLP
- Convolution using “distance” at S (bigger convolution with nearer hits)
- Repeat 2 times and concatenate the output with simple MLP



## Object Condensation (loss function)

$$L = L_p + s_C(L_\beta + L_V)$$

arXiv:2002.03605



- Condensation point: The hit with largest β at each (MC) cluster
- $L_V$: Attractive potential to the condensation point of the same cluster and repulsive potential to the condensation point of different clusters
- $L_\beta$: Pulling up β of the condensation point
- $L_p$: Regression to output features (energy etc.) → currently not used

# What we implemented: track-cluster matching

- PFA is essentially a problem "to subtract hits from tracks"
- HGCAL algorithm does not utilize track information
  - Only calorimeter clustering exists
- Putting tracks as "virtual hits"
  - Located at entry point of calorimeter
  - Having "track" flag (1=track, 0=hit)
  - Energy deposit = 0
- Modification on object condensation to forcibly treat tracks as condensation points

$$L = L_p + s_C(L_\beta + L_V)$$

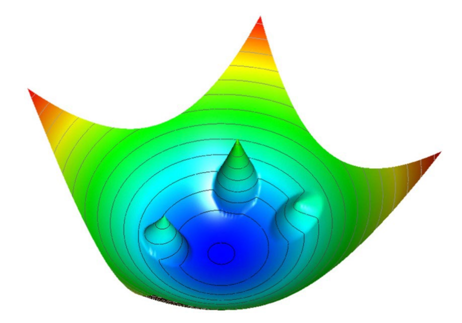
$L_V$: attractive/repulsive potential to condensation points / tracks

$L_\beta$: Pulling up β of the condensation points / tracks

Tracks are prioritized over other condensation points

Current number of parameters: ~420K

# Object condensation and our implementation



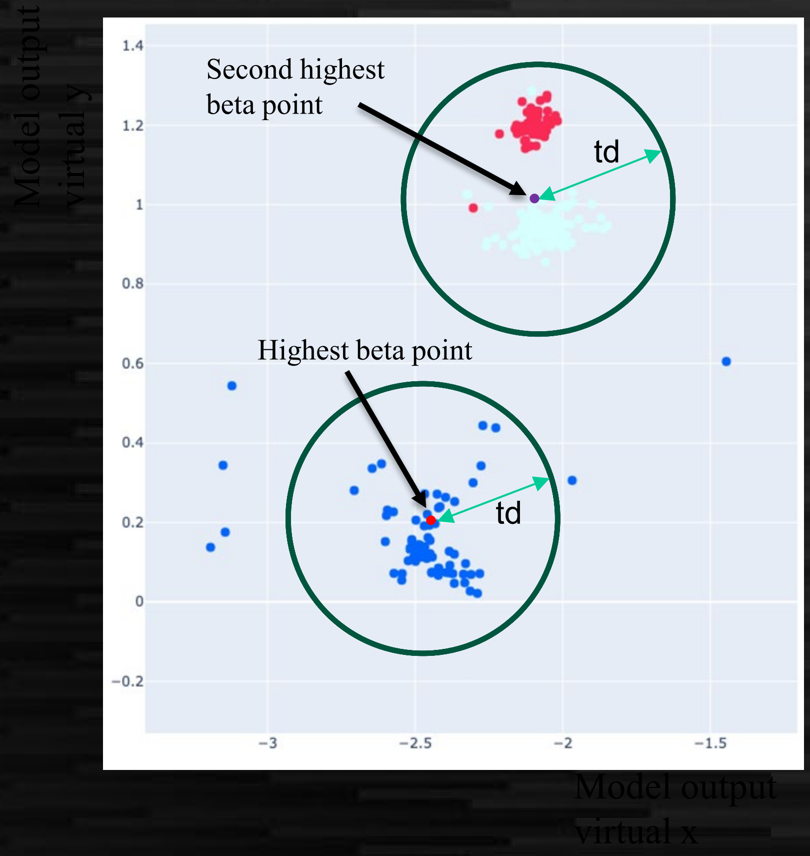
Object condensation loss function (the function to minimize)

$$L = L_p + s_C(L_\beta + L_V)$$

- Condensation point: The hit with largest β at each (MC) cluster
→ For each MC cluster having a track,
the track is forcibly the condensation point regardless of β
- $L_V$: Attractive potential to the condensation point of the same cluster and repulsive potential to the condensation point of different clusters (no modification)
- $L_\beta$: Pulling up β of the condensation point (up to 1)
(no modification, but β of tracks become spontaneously close to 1)
- $L_p$: Regression to output features (energy etc.) → currently not used

# Clustering algorithm

- Output of the network is position and $\beta$ of each hit → need clustering
- Hits that are within a certain distance (td) from the highest $\beta$ point assume as a cluster
- Continues clustering until all hits are clustered or $\beta$ of remaining hits are below threshold (tbeta)
- td/tbeta are tunable parameters

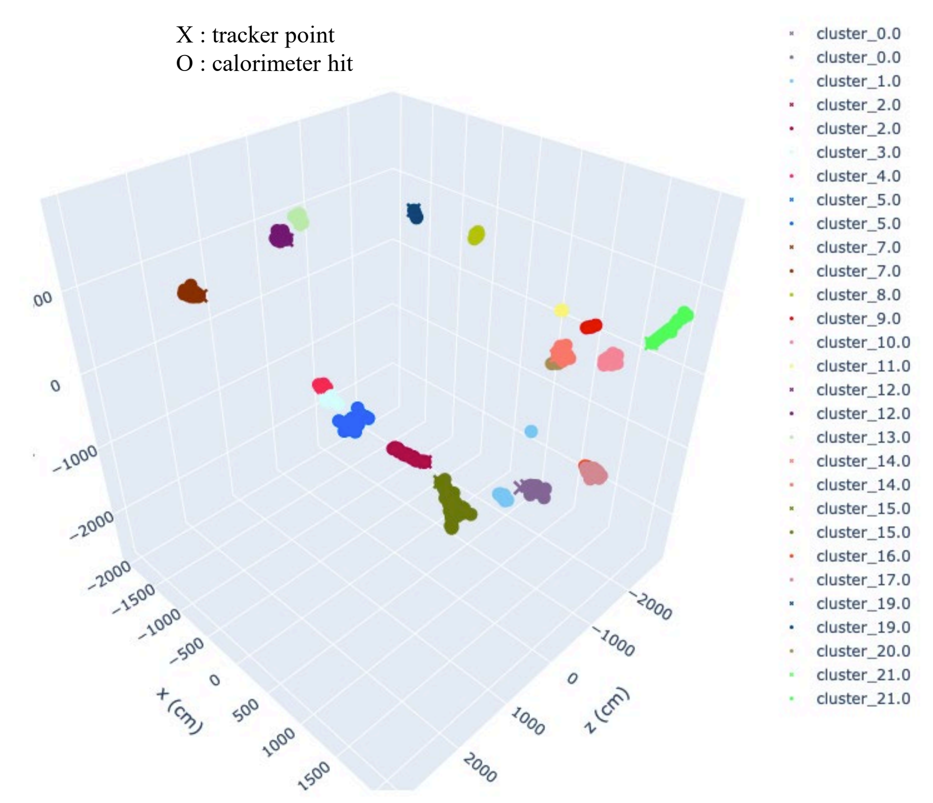


# Our samples for performance evaluation

- ILD full simulation with SiW-ECAL and AHCAL
  - ECAL: 5 x 5 mm$^2$, 30 layers, HCAL: 30 x 30 mm$^2$, 48 layers
  - Taus overlayed with random direction
    - 100k events, 10 GeV x 10 taus / event → 1 million taus
  - qq (q=u, d, s) sample at 91 GeV
    - ~75k events
    - Official sample for PFA calibration (other energies available)
  - Converted to awkward array stored in HDF5 format
    - A few 10 GB each

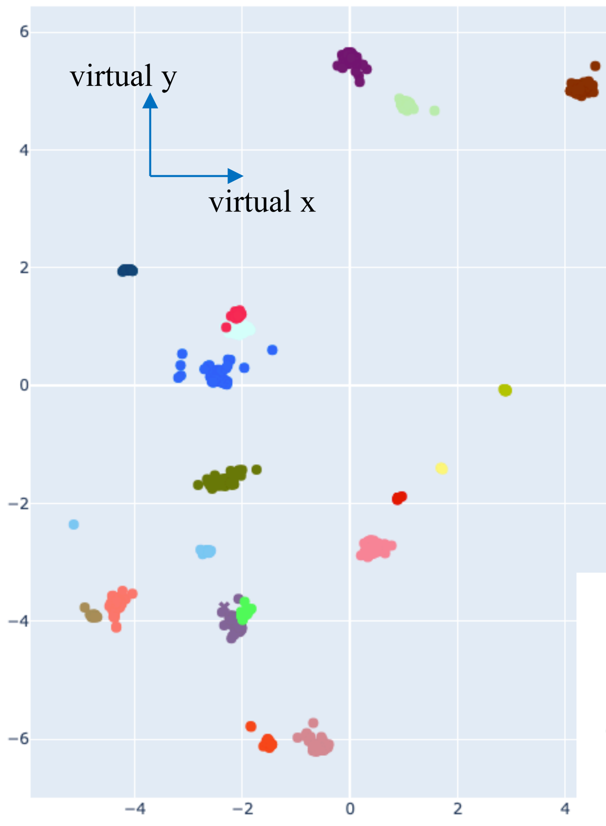
Taus: good mixture of hadrons, leptons and photons
with some isolation
Good for training

# Event display



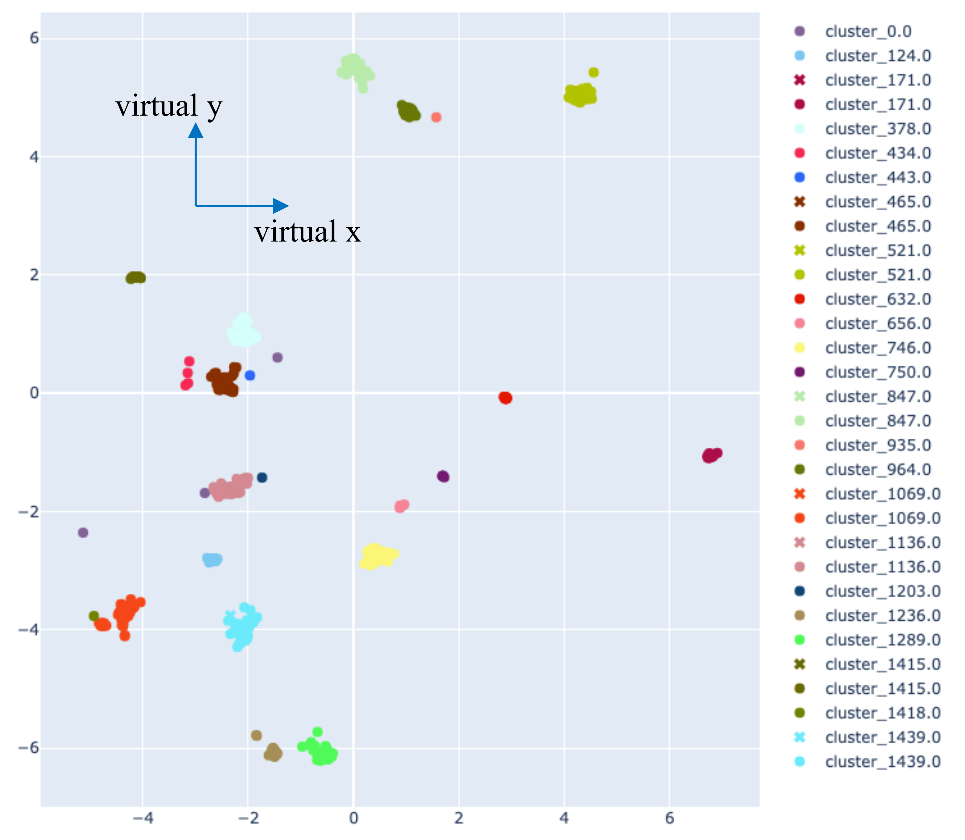
Input features
Real coordinate in detector

Colored by true clusters



Output features
Virtual coordinate

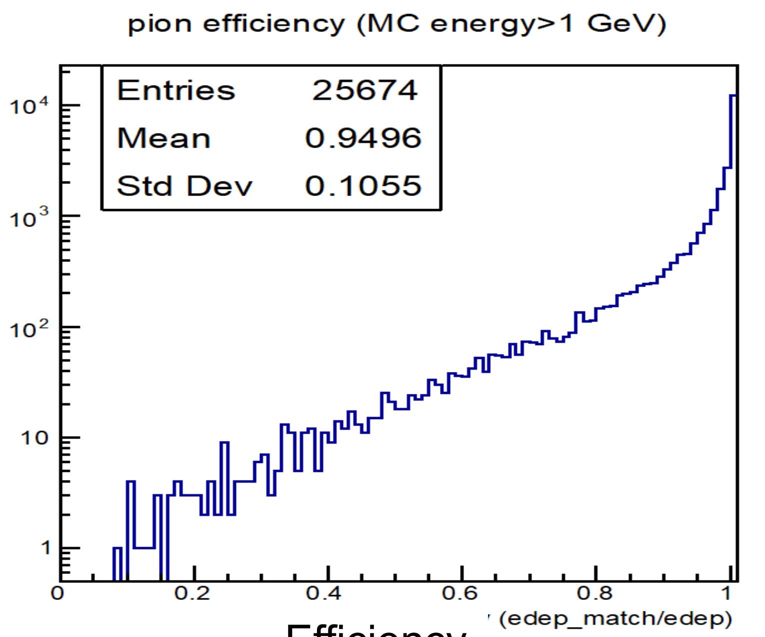
Colored by
true clusters



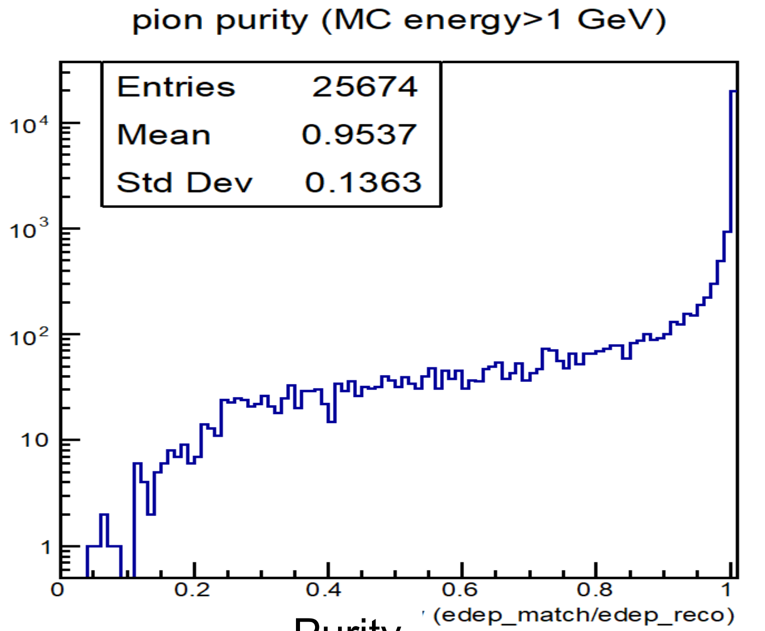
Colored by
reconstructed clusters

# Quantitative evaluation

- Make 1-by-1 connection of MC and reconstructed cluster
  - Reconstructed cluster with highest fraction of hits from the MC is taken
  - Multiple reconstructed cluster may connect to one MC cluster
- Quantitative comparison with PandoraPFA
  - Compared “efficiency” and “purity” of particle flow
    - Efficiency : (reconstructed cluster energy that matches the MC cluster) / (MC cluster energy)
    - Purity : (reconstructed cluster energy that matches the MC cluster) / (reconstructed cluster energy )

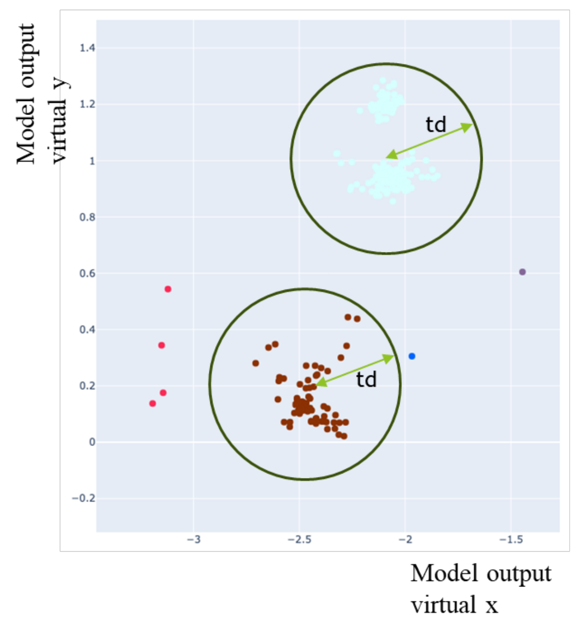


Efficiency



Purity

# Optimization of performance

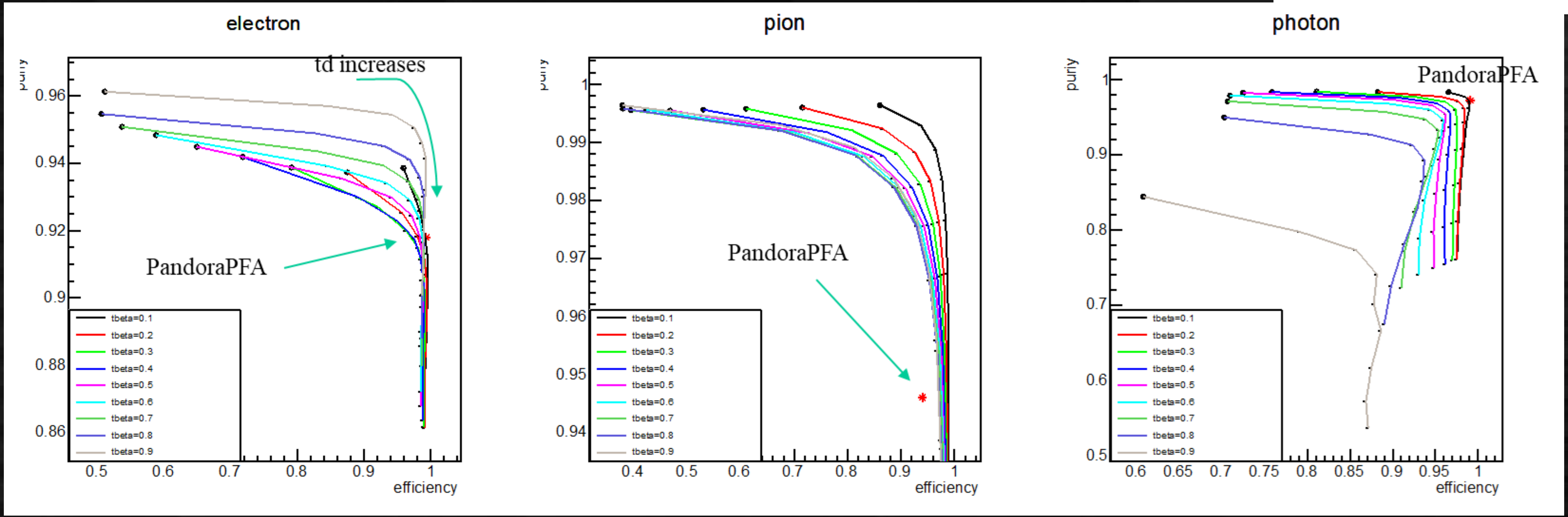


## Output dimension of the coordinate

- The initial work done with output coordinate dimension D = 2 (for visibility)
- Tried D=3,4,8,16 → D=4 selected

## Clustering parameters (td, tbeta)

- td: radius which hits are treated as coming from the same cluster
- tbeta: threshold of beta to form clusters



Scanning result: tbeta=0.1, td=0.3/0.4 selected

# Results on efficiency and purity

| **Algorithm train/test** | **Electron eff.** | **Pion eff.** | **Photon eff.** | **Electron pur.** | **Pion pur.** | **Photon pur.** |
|---|---|---|---|---|---|---|
| GravNet<br>10 taus/10 taus | 99.1% | 96.5% | 99.0% | 91.8% | 98.9% | 97.1% |
| PandoraPFA<br>10 taus | 99.3% | 94.0% | 99.1% | 91.8% | 94.6% | 97.2% |
| GravNet<br>jets/jets | 94.5% | 93.1% | 95.2% | 77.4% | 93.2% | 92.4% |
| PandoraPFA<br>jets | 80.2% | 90.4% | 79.0% | 75.0% | 90.6% | 77.7% |
| PandoraPFA<br>jets<br>(ILCSoft truth) | 96.7% | 95.5% | 96.4% | 97.1% | 90.4% | 97.7% |

At least in our measure, performance of GravNet-based algorithm exceeds PandoraPFA

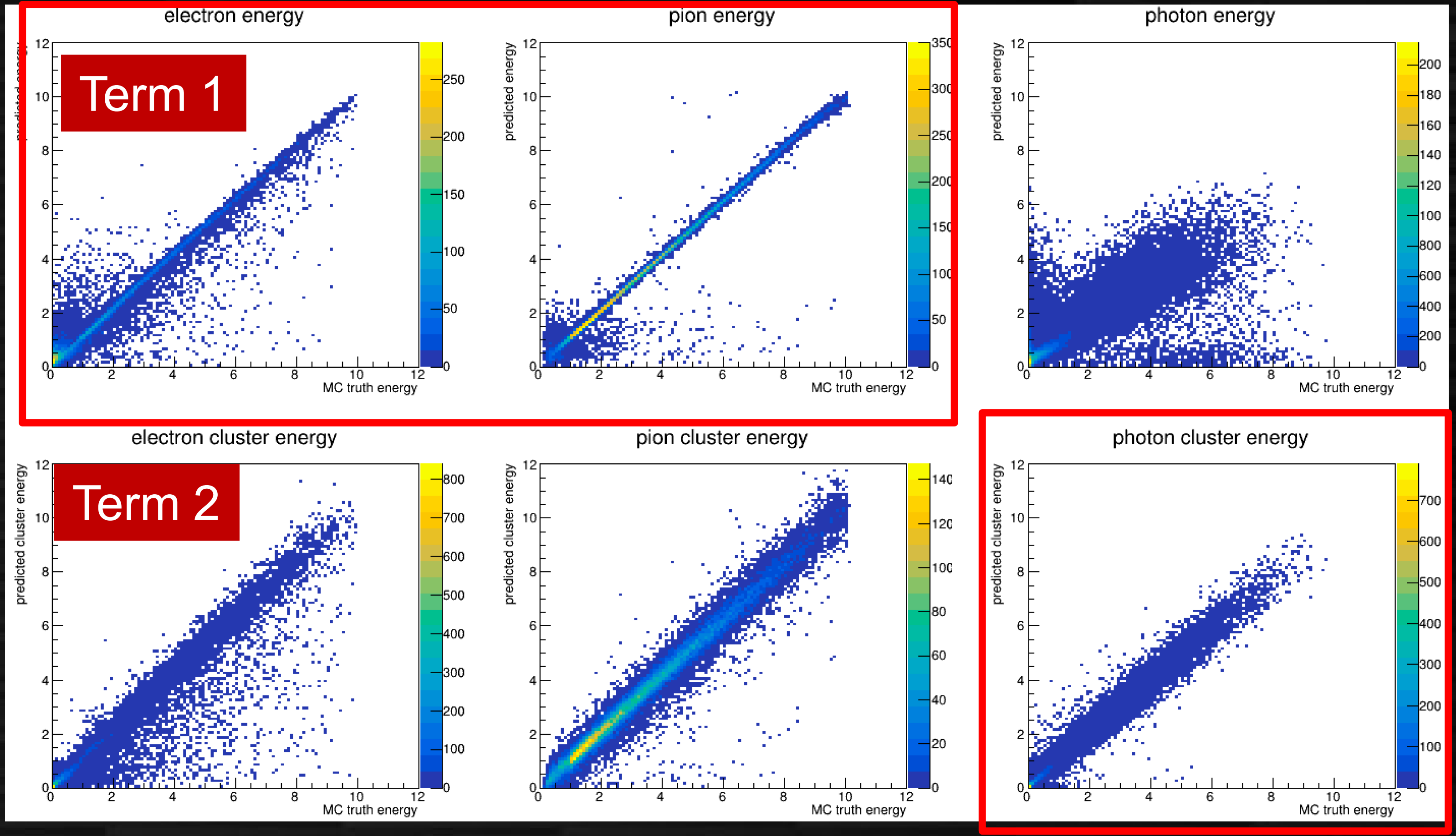
→ Promising as full PFA (but energy regression to be done)

Definition of MC truth clusters needs to be tuned (see ILCSoft truth)

# Energy regression: ongoing work

Add $E_{tr}$ and $E_{hit}$ to the output of the network (for each hit)
Add terms (1, 2) to object condensation loss

Cluster energy (MC vs reco) at 10 taus event



Two additional loss term

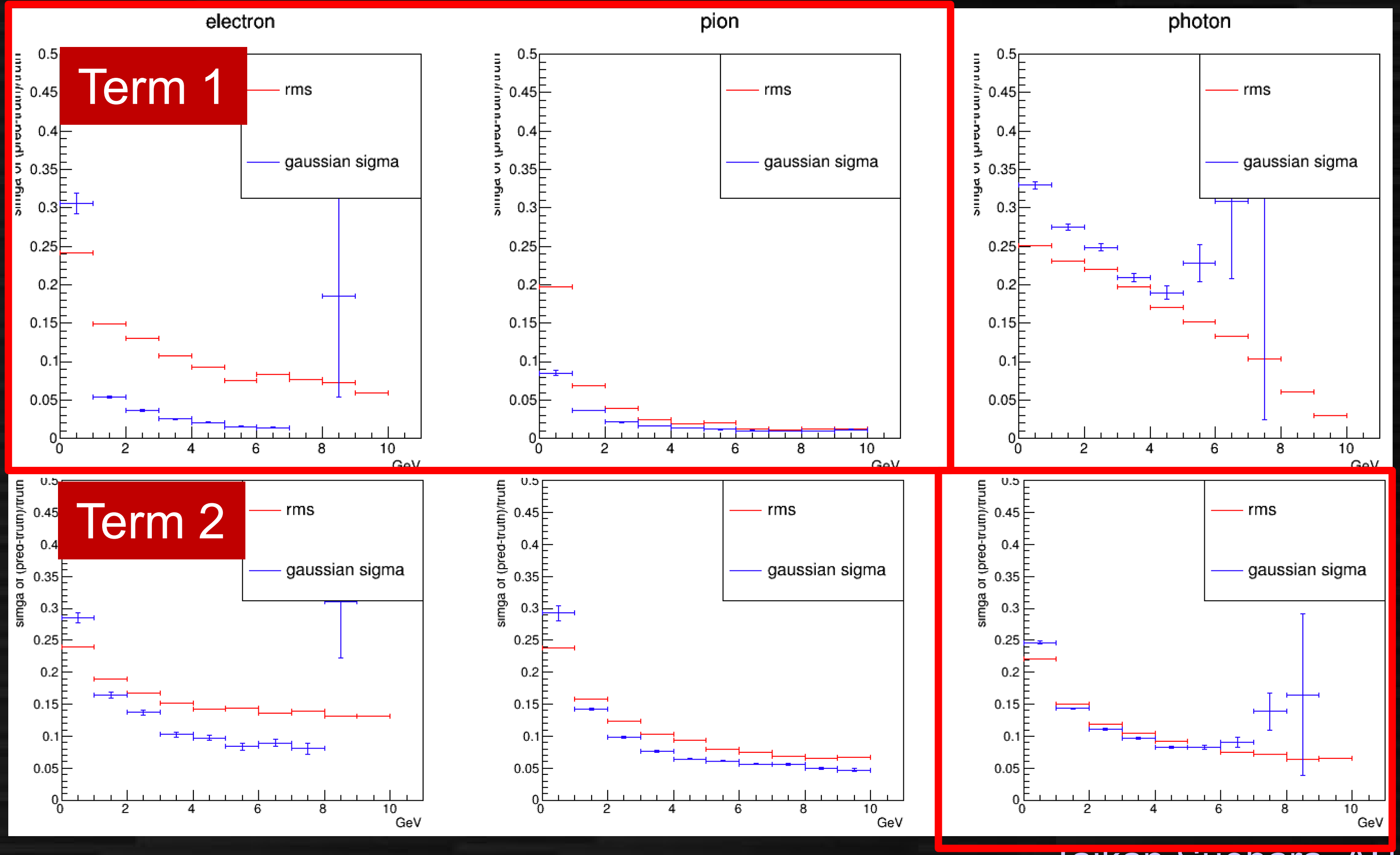
1. $E_{tr}$ at condensation points to be regressed to MC cluster energy
2. Sum of $E_{hit}$ of all energies to be regressed to MC cluster energy
3. Use $E_{tr}$ for charged clusters and use sum of $E_{hit}$ for neutral clusters

# Energy regression: ongoing work (energy reso.)

Add “energy” to the output of the network (for each hit)

Add a term to object condensation

Cluster energy (MC vs reco) at 10 taus event



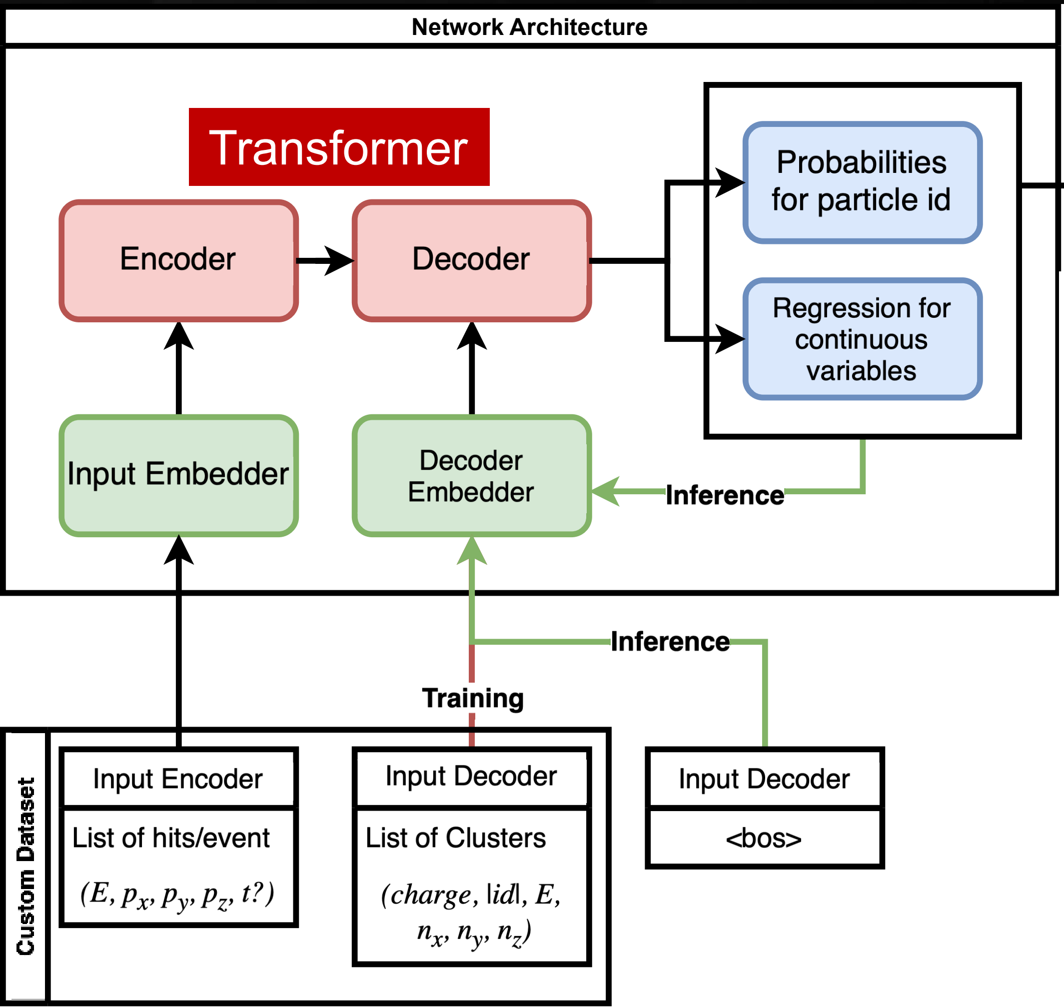
Two additional loss term

1. $E_{tr}$ at condensation points to be regressed to MC cluster energy
2. Sum of $E_{hit}$ of all energies to be regressed to MC cluster energy
3. Use $E_{tr}$ for charged clusters and use sum of $E_{hit}$ for neutral clusters

Tendency is OK but slightly worse than expected resolution of 15-18% / sqrt(E [GeV]) for photons

# More NLP-like model: transformer

Submitted to E1 in FY2025-26



**Network Architecture**

Transformer

Encoder → Decoder

Input Embedder

Decoder Embedder

Probabilities for particle id

Regression for continuous variables

**Inference**

**Training**

**Custom Dataset**

Input Encoder

List of hits/event

$(E, p_x, p_y, p_z, t?)$

Input Decoder

List of Clusters

$(charge, |id|, E, n_x, n_y, n_z)$

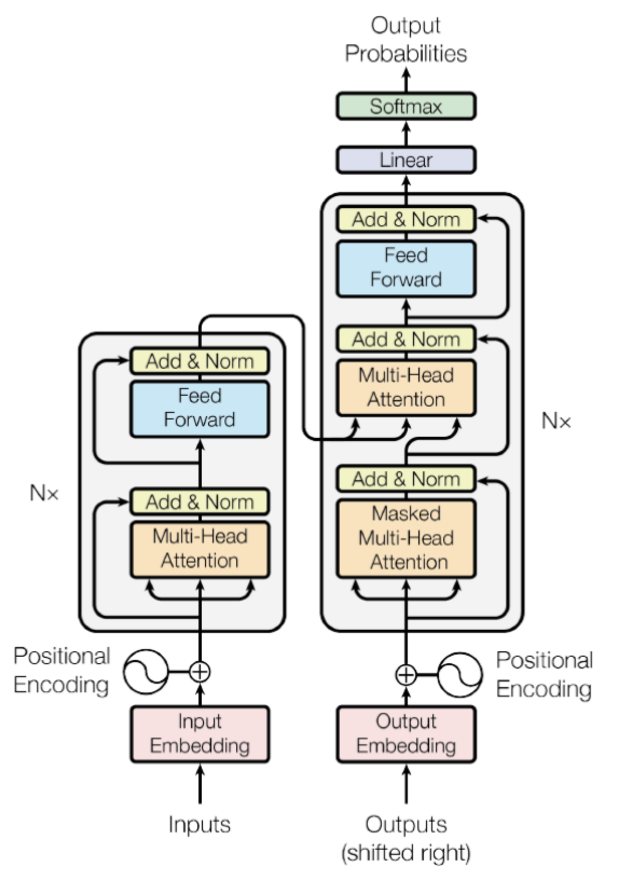
Input Decoder

<bos>

Output

List of Clusters, Classed by charge and ordered by energy

Planned structure for PFA



Transformer

Transformer: training relation among elements (hits in PFA) with (multi-head) self-attention mechanism (used in GPT etc.)

Encoder: accumulate info of all hits/tracks by transformer

Decoder:

Input cluster info one by one

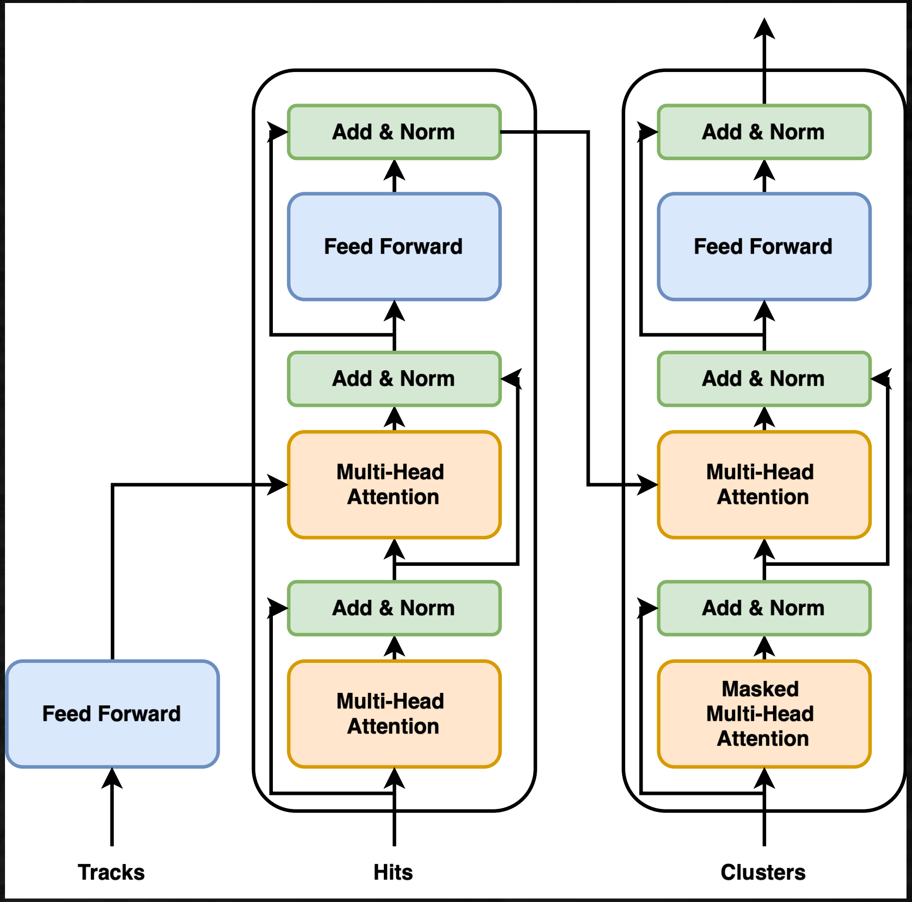
Output info of next cluster

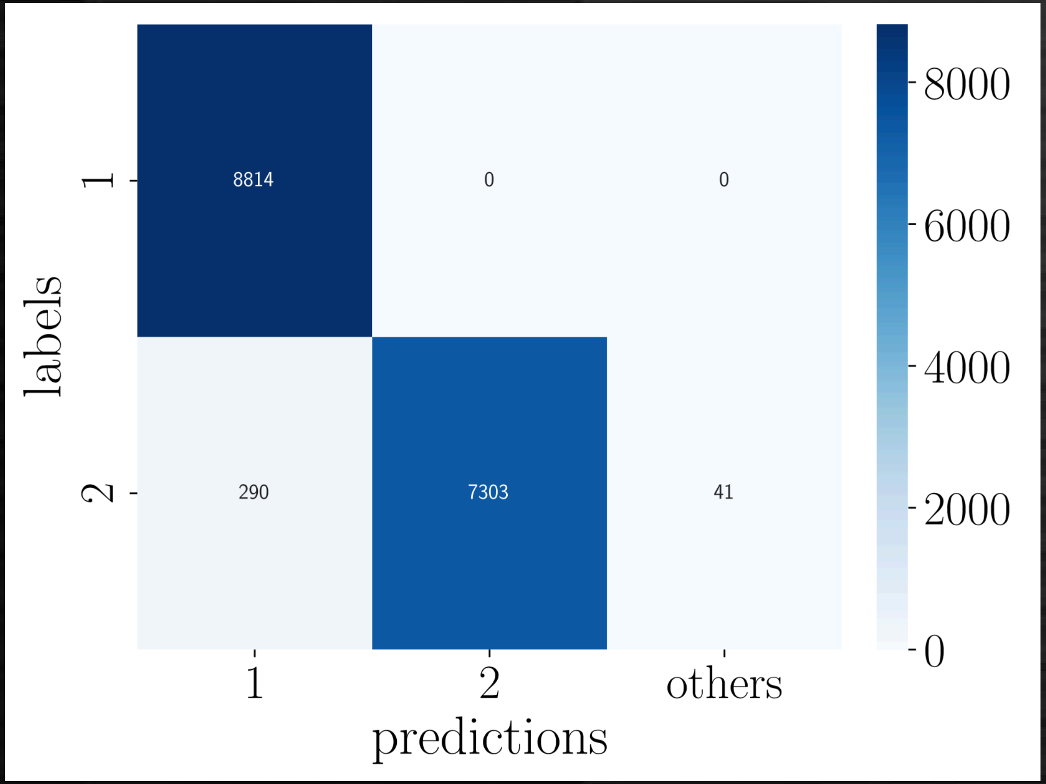
(training) MC truth clusters

(inference) just provide <bos> to derive first cluster, using output as next input until <eos> obtained

(Inspired by translation NN)

# Transformer-based PFA: some quick view





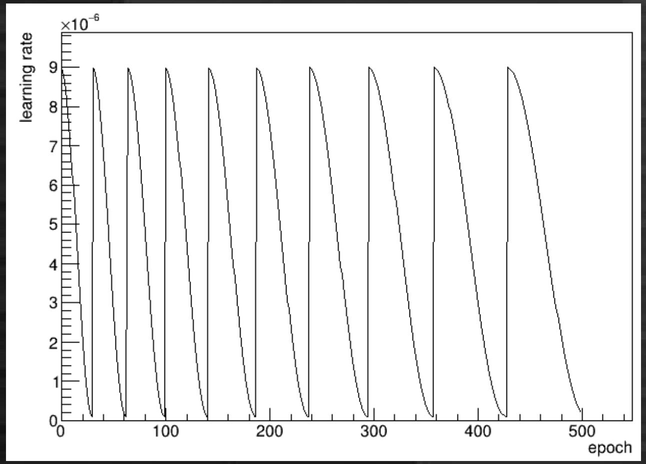
Separation of single and double photons
- random opening angle – not too bad
but worse than GNN-based study now

Proposal from collaborator: should investigate independent training of encoder part by e.g. masking some particles in each event (as often done in NLP)

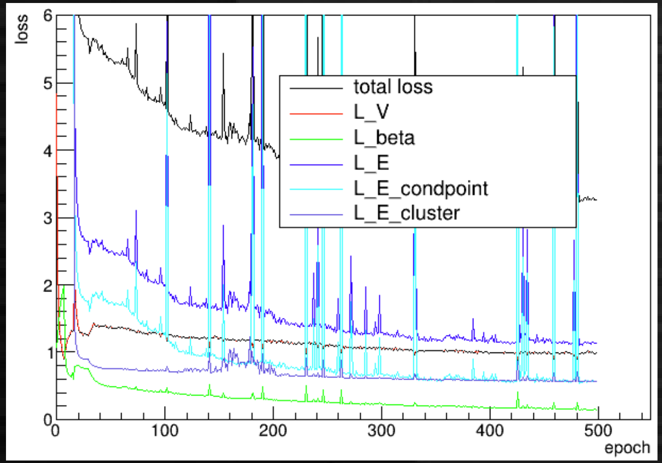
# Prospects

- GNN-based – energy regression
  - – Identifying spikes on validation loss
  - – Hyper-parameter tuning (learning rate etc.)
  - – Check jet energy resolution
- Transformer-based PFA
  - – Try to replace GravNet with Transformer
  - – Training with “masking hits”
    - Firstly with the NLP-like method
    - Also consider to include object condensation (e.g. encoder training with masking, decoder training with object condensation)

Learning rate (decay + restart)



Problem on validation loss



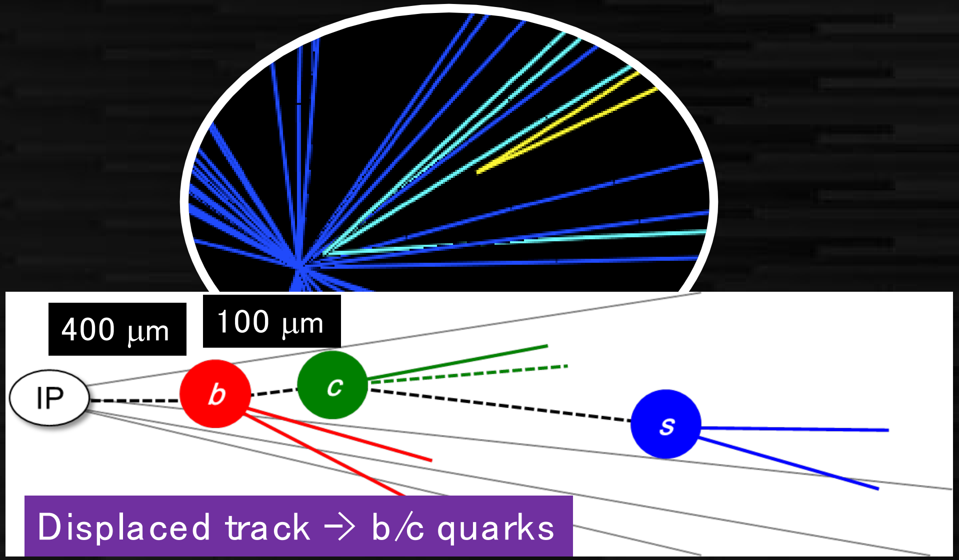
# Prospects (2)

- Try to differentiate detector models
  - Larger/smaller pixel sizes, magnetic field, etc.
  - MAPS calorimeter (e.g. 50 μm pixel instead of 5 mm) or strip calorimeter (e.g. 5 x 45 mm interleaved)
  - Timing calorimeter (with pico-sec timing capability)
  - Dual-readout (hadron) calorimeter
    - Cherenkov layers in addition to scintillator layers (to partially separate hadronic activities)
  - Liquid-Argon calorimeter (proposal from FCCee)
  - Crystal/glass calorimeters (proposal from CEPC)

# Particle flow: summary and plans

- GNN-based particle flow has possibility to replace PandoraPFA
  - Performance seems significantly exceeded at least in our measure
  - Definition of MC-truth to be investigated in more detail
- Regression of cluster energy being investigated
  - Reasonable single-particle energy starting to be obtained
  - Jet energy resolution would be compared with PandoraPFA
- Possible improvements on network
  - More hyperparameter tuning
  - Transformer-based one and/or training with “masking hits”
- Applications
  - Incorporation of new detector technologies (timing, pixel, dual-readout,…)
  - Application to physics analysis

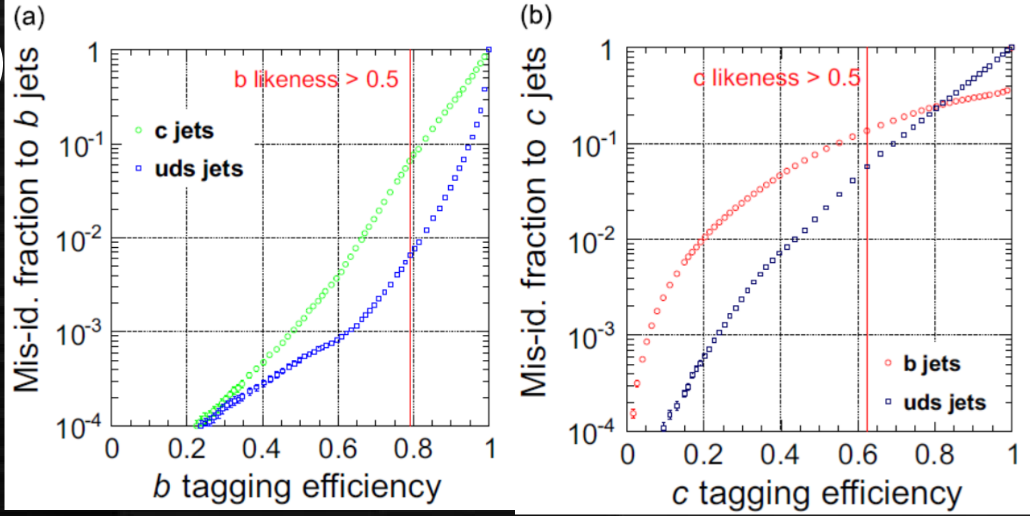
# Flavor tagging with Particle Transformer (ParT)

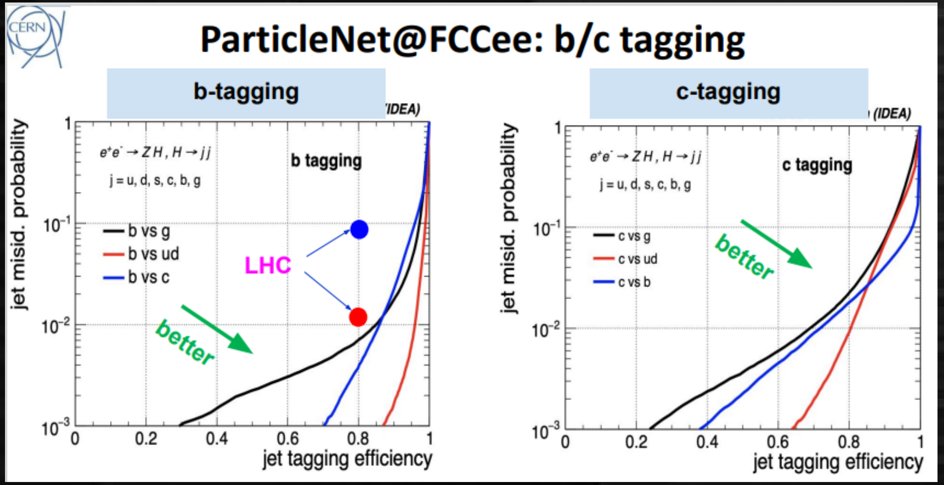


# Flavor tagging for Higgs factories

- Jet flavor tagging is essentially important for Higgs studies (including self coupling)
- LCFIPlus (published 2013) was long used for flavor tagging
  - All physics performance in ILD/SiD/CLIC are based on LCFIPlus
- FCCee reported >10x better rejection using ParticleNet (GNN) in 2022
  - Delphes is used for simulation
- We studied DNN-based flavor tag with ILD full simulation to confirm it
  - Using latest algorithm: Particle Transformer (ParT)

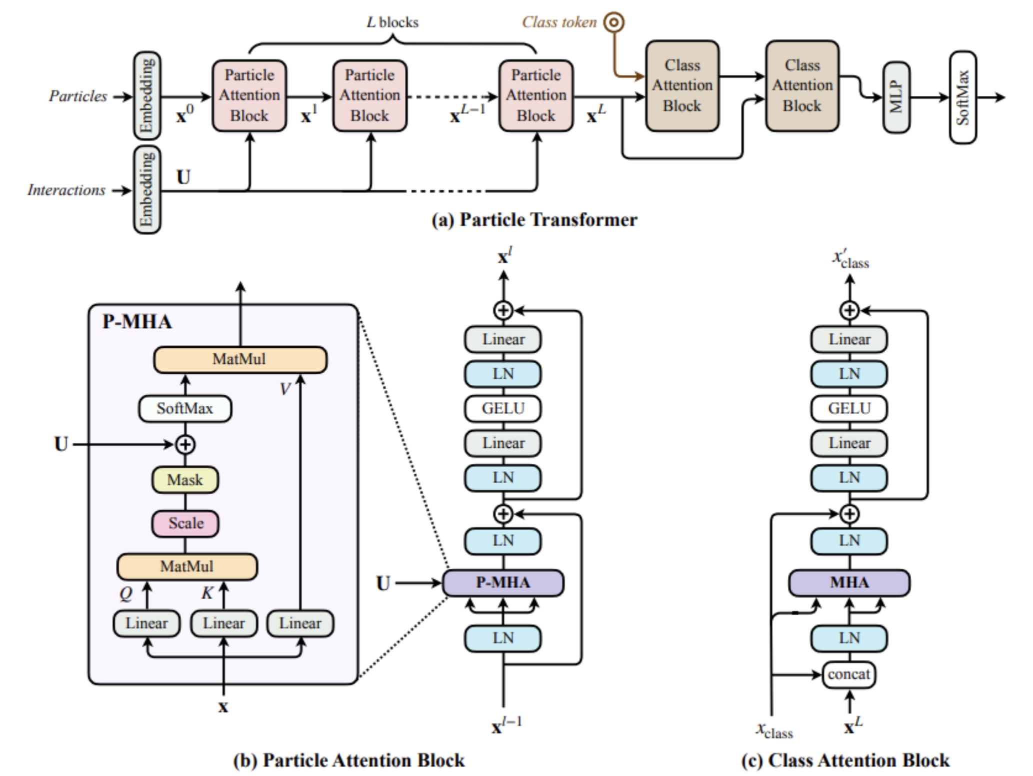
LCFIPlus performance plots

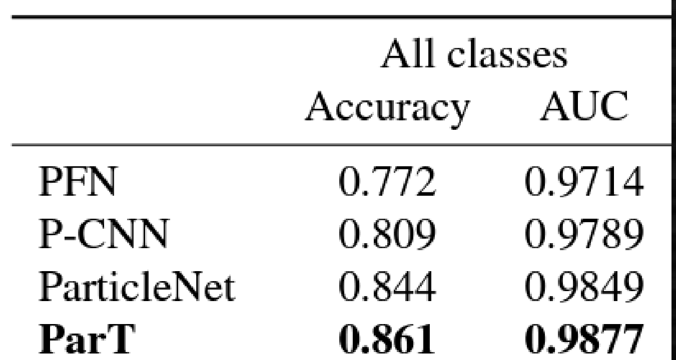




# Particle Transformer (ParT)

- Transformer: self-attention-based algorithm intensively used for NLP (e.g. chatGPT)
  - Weak biasing: possible to train big samples efficiently (with more learnable weights) but demanding big training sample for high performance
- ParT is a new Transformer-based architecture for Jet tagging, published in 2022.
  - Pair-wise variable (angle, mass etc.) is added to plain Transformer encoder to boost attention
- Surpasses the performance of ParticleNet
  - ParticleNet only looks “neighbor” particles while Transformer uses attention to learn where to look





| | All classes | |
|---|---|---|
| | Accuracy | AUC |
| PFN | 0.772 | 0.9714 |
| P-CNN | 0.809 | 0.9789 |
| ParticleNet | 0.844 | 0.9849 |
| **ParT** | **0.861** | **0.9877** |

Performance with JetClass event classification (100M sample)

# Data Samples and Input Variables

## Data samples

- ILD full simulation

1. e+ e- → qq (at 91 GeV) (used in LCFIPlus study)

q = b,c,uds
j=b,c,u,d,s,g

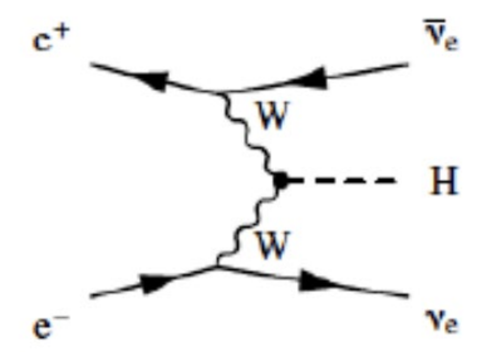
2. e+ e- → ννH → ννjj (at 250 GeV) (2020 production)

1M jets (500k events) for each flavor

- FCCee fast simulation (Delphes with IDEA detector):

e+ e- → ννH → ννjj (at 240 GeV)

10M jets (5M events) each flavor



80% for training
5% for validation
15% for test

## Input variables

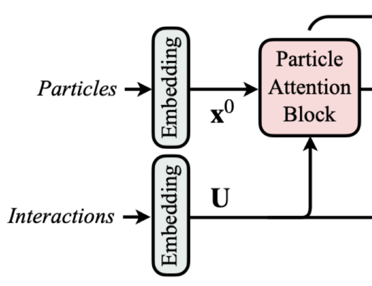
Particles: for every track/neutral

- Impact parameters (6)
  - 2D/3D, from primary vertex
- Jet distance (2)
  - Displacement from jet axis
- Covariant matrix (15)
- Kinematics (4)
  - Energy fraction, angles, charge
- Particle ID (6)
  - Probability (or binary selection) of e, μ, hadron, gamma, neutral hadron

Interactions: for every particle pair

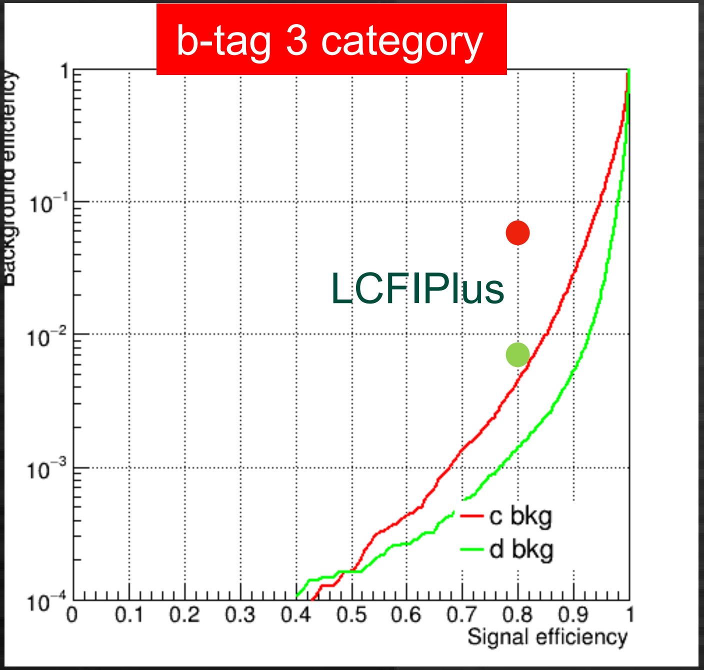
- $\delta R^2$, $k_t$, Z, mass

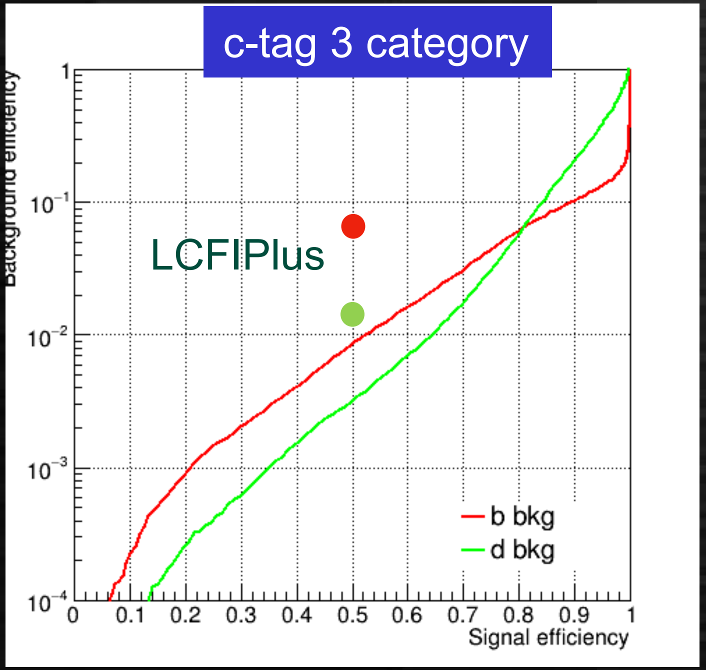
Input of ParT



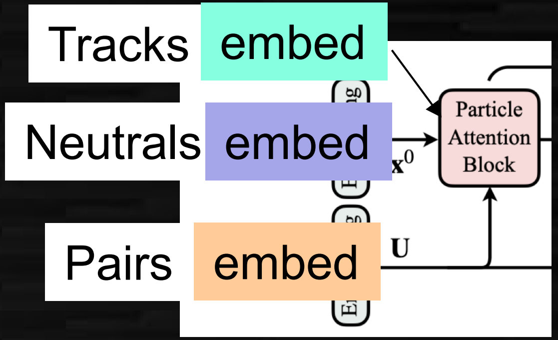
# Improvements wrt. LCFIPlus

- Factor (3-9) improvement at ParT from LCFIPlus without any tuning
- Another factor (max 3) improvement by tuning
  - Optimizing input variables
  - Separate embedding for tracks/neutrals





| | b-tag 80% eff. | | c-tag 50% eff. | |
|---|---|---|---|---|
| background | c jets | uds jets | b jets | uds jets |
| +LCFIPlus (BDT) | 6.3% | 0.79% | 7.4% | 1.2% |
| *ParT (initial) | 1.3% | 0.25% | 1.0% | 0.43% |
| **ParT (improved) | 0.48% | 0.14% | 0.86% | 0.34% |



+LCFIPlus (BDT) 250 GeV nnqq
*ParT (initial) 91 GeV qq, default settings
**ParT (improved) 250 GeV nnqq, b/c/d separation

# Comparison with FCCee results

Conditions

- FCCee data provided by M. Selvaggi (as ROOT files including input variables)
- Processed with our script (using weaver (by H. Qu) and based on provided configuration)

Results

- FCCee 1M gives ~2x better
  - Comparable with reduced inputs
- FCCee 4/6/8M gives much better
  - Sample size dependence needs to be investigated with ILD (maybe difficult with full simulation)
  - (JetClass has 100M events)

1M: 800k jets for training
4/6/8M: 4/6/8M jets for training

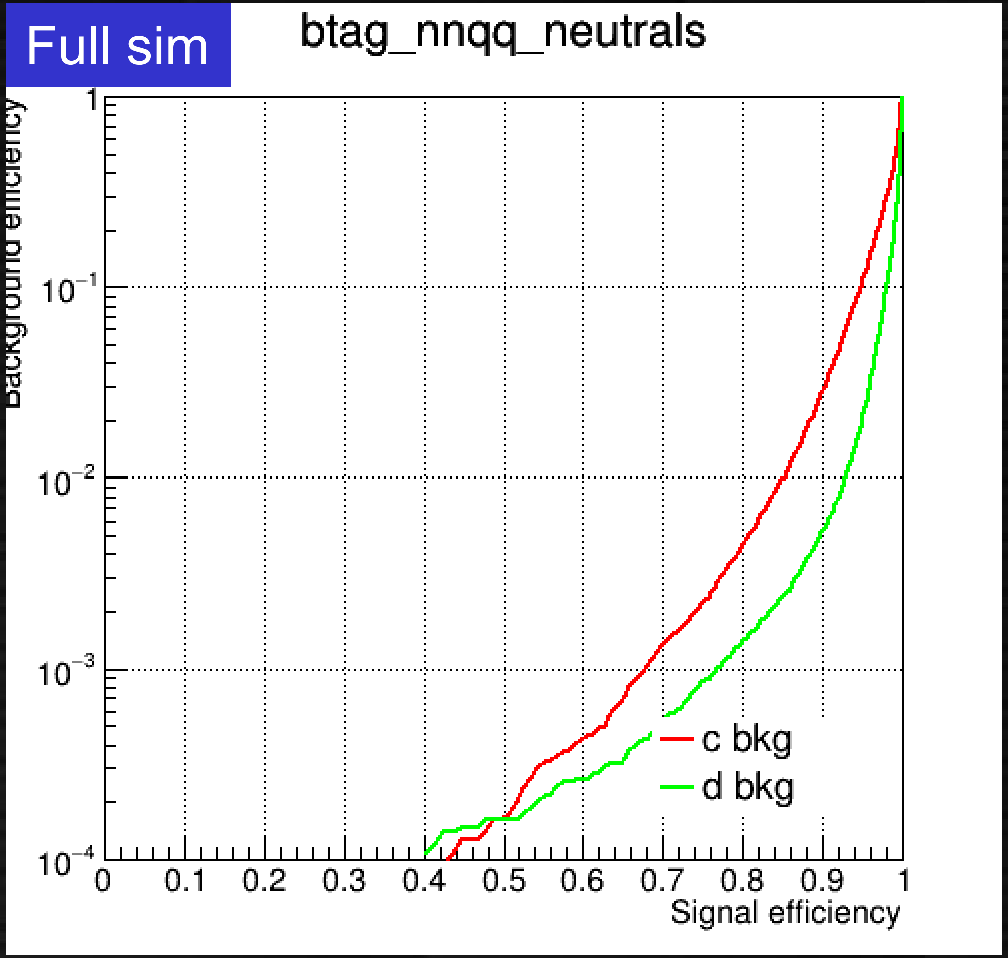
| Sample / sample size | b-tag 80% eff. | | c-tag 50% eff. | |
|---|---|---|---|---|
| background | c jets | uds jets | b jets | uds jets |
| ILD full-sim 1M (optimized) | 0.48% | 0.14% | 0.86% | 0.34% |
| FCCee Delphes 1M (reduced) | 0.47% | 0.12% | 0.64% | 0.10% |
| FCCee Delphes 1M (full) | 0.21% | 0.054% | 0.36% | 0.059% |
| FCCee Delphes 4M | 0.045% | 0.025% | 0.20% | 0.033% |
| FCCee Delphes 6M | 0.014% | 0.010% | 0.13% | 0.022% |
| FCCee Delphes 8M | 0.007% | 0.006% | 0.076% | 0.021% |

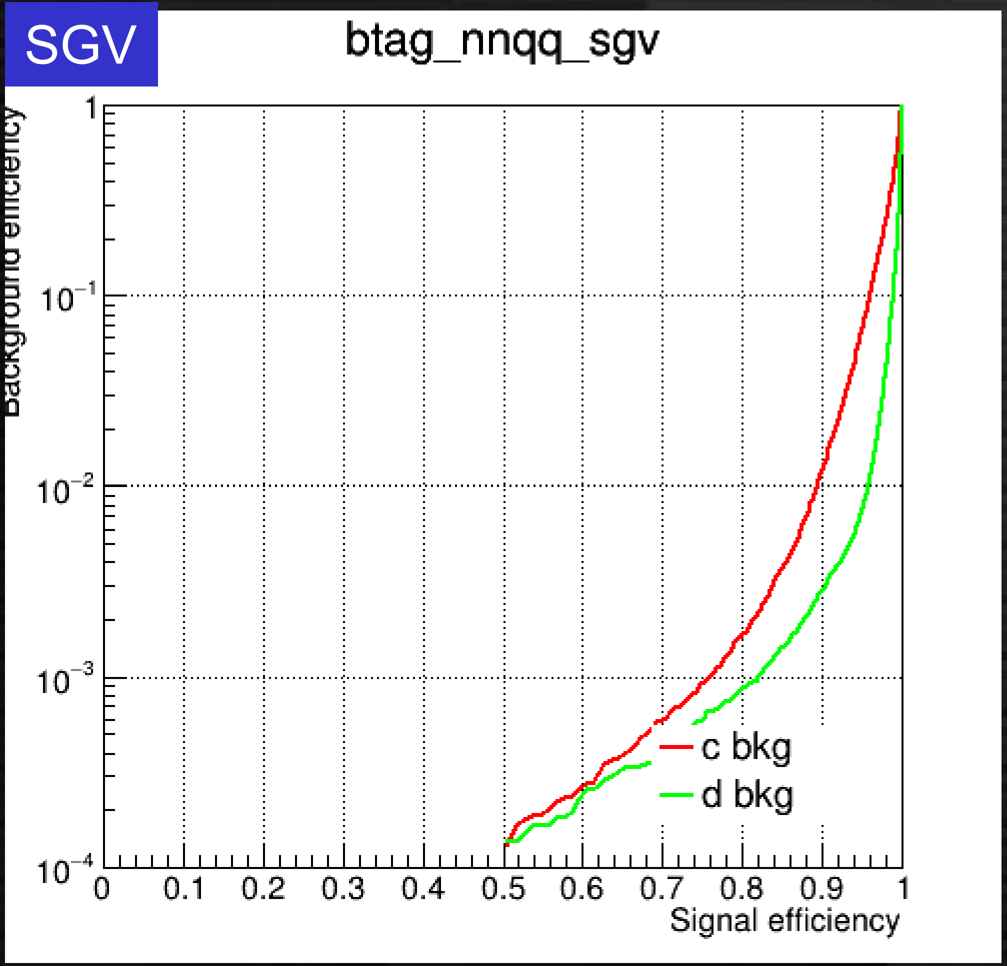
We see mild consistency between ILD and FCC!

FCCee configurations:

- Simulation: Delphes (IDEA geometry)
- Input: Kinematic/Impact parameter/Track error /Particle ID (including TOF and dn/dx) (not with reduced)
- Slight difference with ILD variables (e.g. interaction)

# Comparison with ILC fastsim (SGV) vs full sim





SGV gives better performance (the sample size is same)

Maybe related to difference of neutral particles (confusion may not be included)

Comparison on scaling raw to be checked

# Phys-based comparison on Higgs self-coupling

Higgs self coupling: ZHH at 500/550 GeV

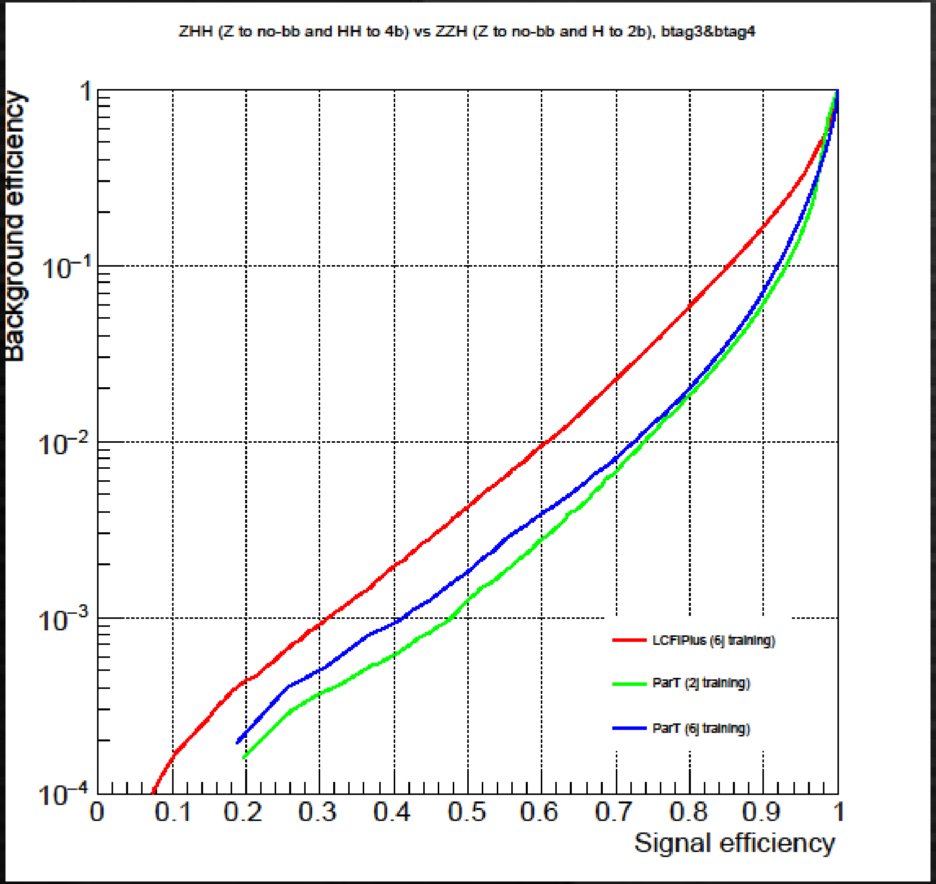
- Very small signal cross section: ~0.2 fb
- Interfering diagrams: sensitivity further reduced
- Background: ZZH / tt (>100 fb) / ZH
  - Strong suppression of tt by b-tag essential

Simple benchmark here

- ZHH (Z to non-bb, HH to 4b) as signal
- ZZH (2Z to 0b, H to 2b) as background
- 3rd and 4th largest b-tag are used to calculate efficiency (2D-scan)
- Comparison between
  - Old BDT-based flavor tagging (LCFIPlus)
  - ParT trained with nnqq 250 GeV
  - ParT trained with 6b/6c/6q at 500 GeV
- ParT with nnqq training better by ~50% at background acceptance 0.1%

Signal: qqHH: qq→0b, HH→4b

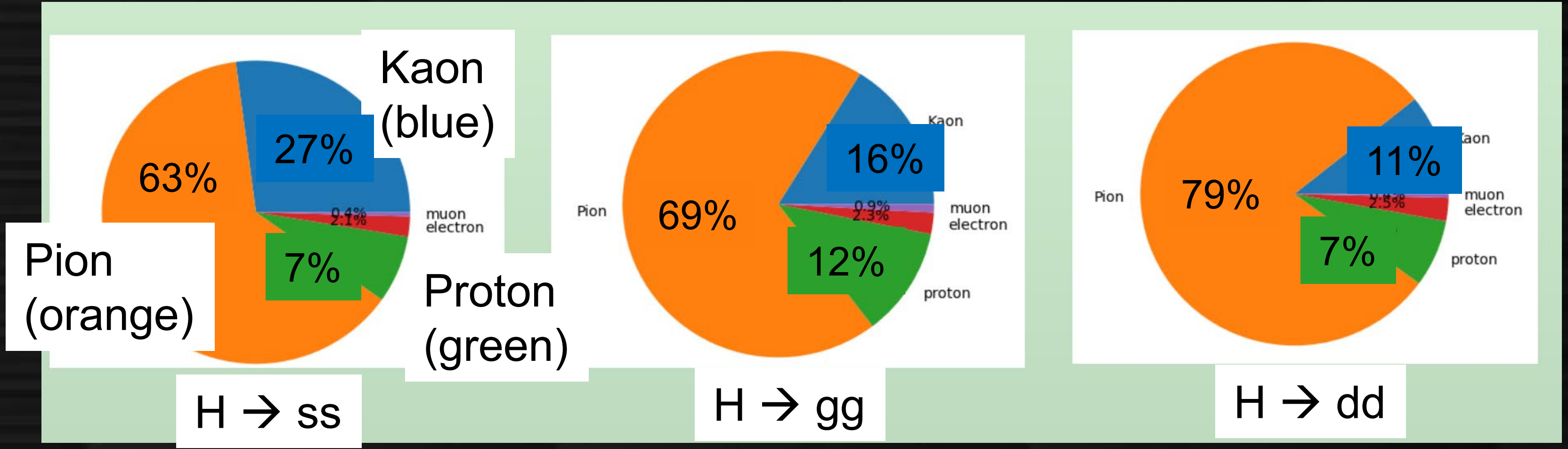
Bkg: qqqqH: qqqq→0b, H→2b



Improvement from 27% → 18% on self coupling resolution expected

# Strange tagging

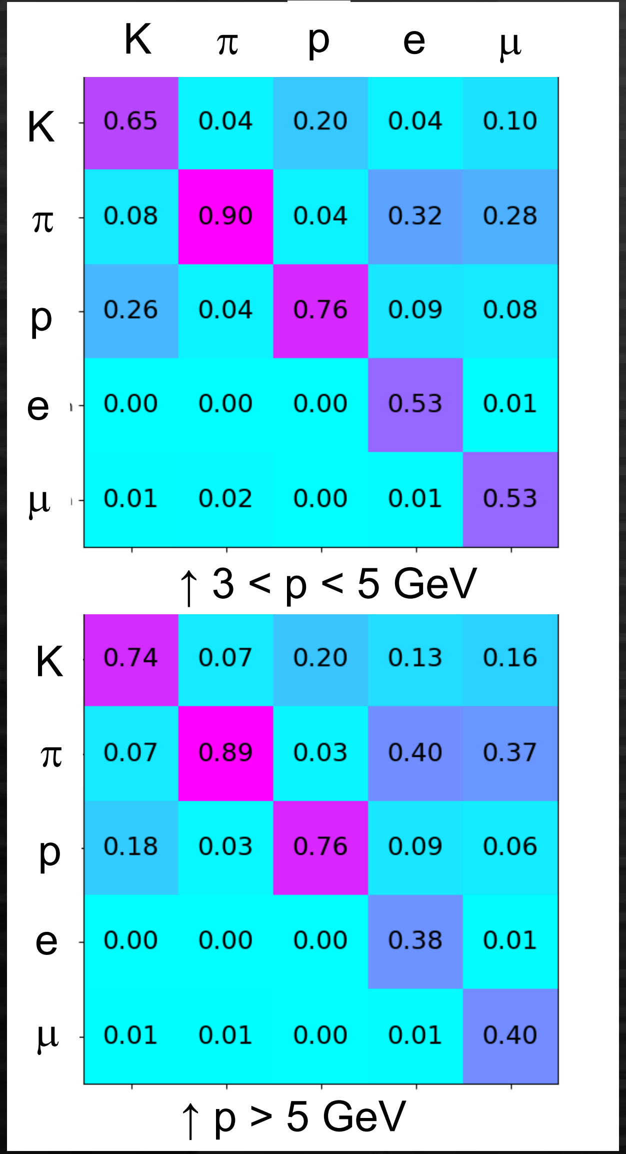
- High-momentum kaon in jet is a clue to strange jets
  - Contamination from g→ss give relatively low momentum
- dE/dx is essential for Particle ID in ILD
  - As well as ToF, but only effective in low energy tracks (which are less important in strange tagging)
- Using newly-developed comprehensive PID
  - Giving much better separation than previous PID



Fractions of tracks having > 5 GeV

More Kaons in ss
More protons in gg

Fraction of true particles



True particle

CPID prediction

| | K | π | p | e | μ |
|---|---|---|---|---|---|
| K | 0.65 | 0.04 | 0.20 | 0.04 | 0.10 |
| π | 0.08 | 0.90 | 0.04 | 0.32 | 0.28 |
| p | 0.26 | 0.04 | 0.76 | 0.09 | 0.08 |
| e | 0.00 | 0.00 | 0.00 | 0.53 | 0.01 |
| μ | 0.01 | 0.02 | 0.00 | 0.01 | 0.53 |

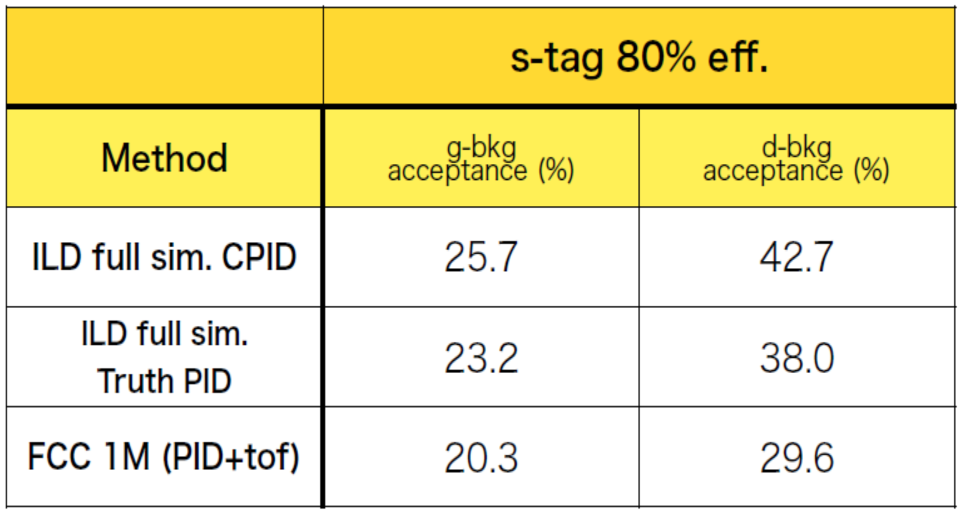
↑ 3 < p < 5 GeV

| | K | π | p | e | μ |
|---|---|---|---|---|---|
| K | 0.74 | 0.07 | 0.20 | 0.13 | 0.16 |
| π | 0.07 | 0.89 | 0.03 | 0.40 | 0.37 |
| p | 0.18 | 0.03 | 0.76 | 0.09 | 0.06 |
| e | 0.00 | 0.00 | 0.00 | 0.38 | 0.01 |
| μ | 0.01 | 0.01 | 0.00 | 0.01 | 0.40 |

↑ p > 5 GeV

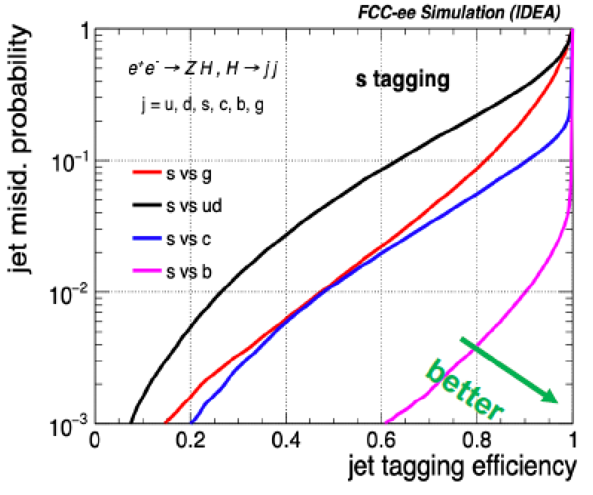
# Strange tagging: initial results

- First results obtained with CPID
  - No significant improvements from old PID: investigating
  - Compared with truth PID: some difference
  - FCC (1M) better than ILD Truth PID
    - Reason needs to be investigated
      (maybe non-perfect assignment of truth PID)
- Still needs study

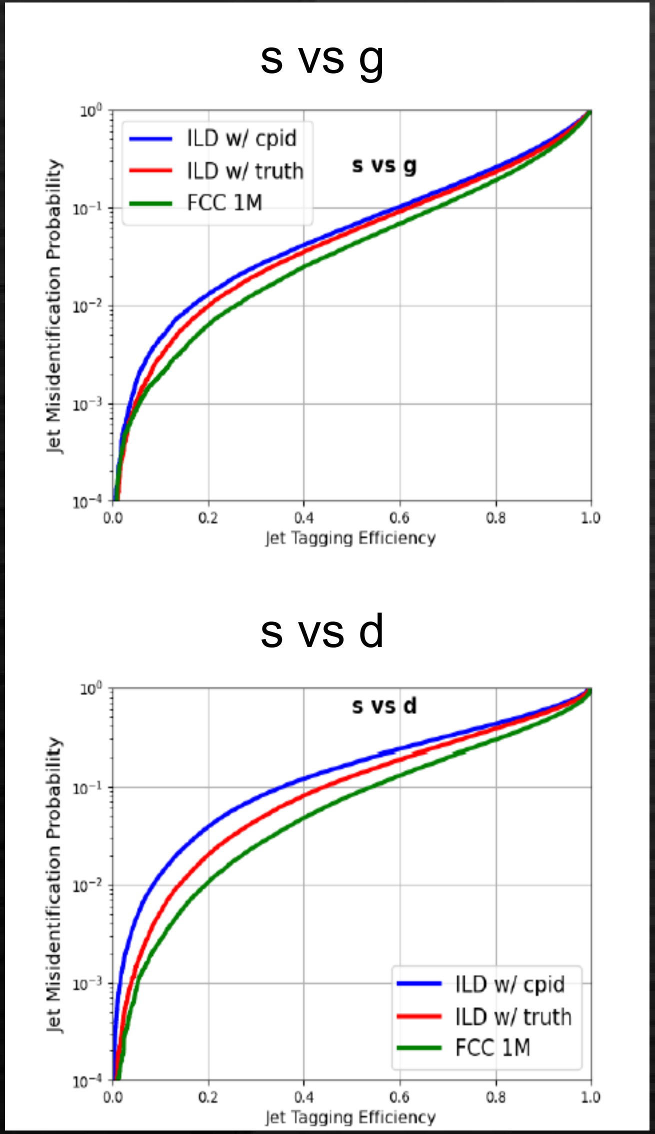


|  | s-tag 80% eff. |  |
|---|---|---|
| Method | g-bkg acceptance (%) | d-bkg acceptance (%) |
| ILD full sim. CPID | 25.7 | 42.7 |
| ILD full sim. Truth PID | 23.2 | 38.0 |
| FCC 1M (PID+tof) | 20.3 | 29.6 |

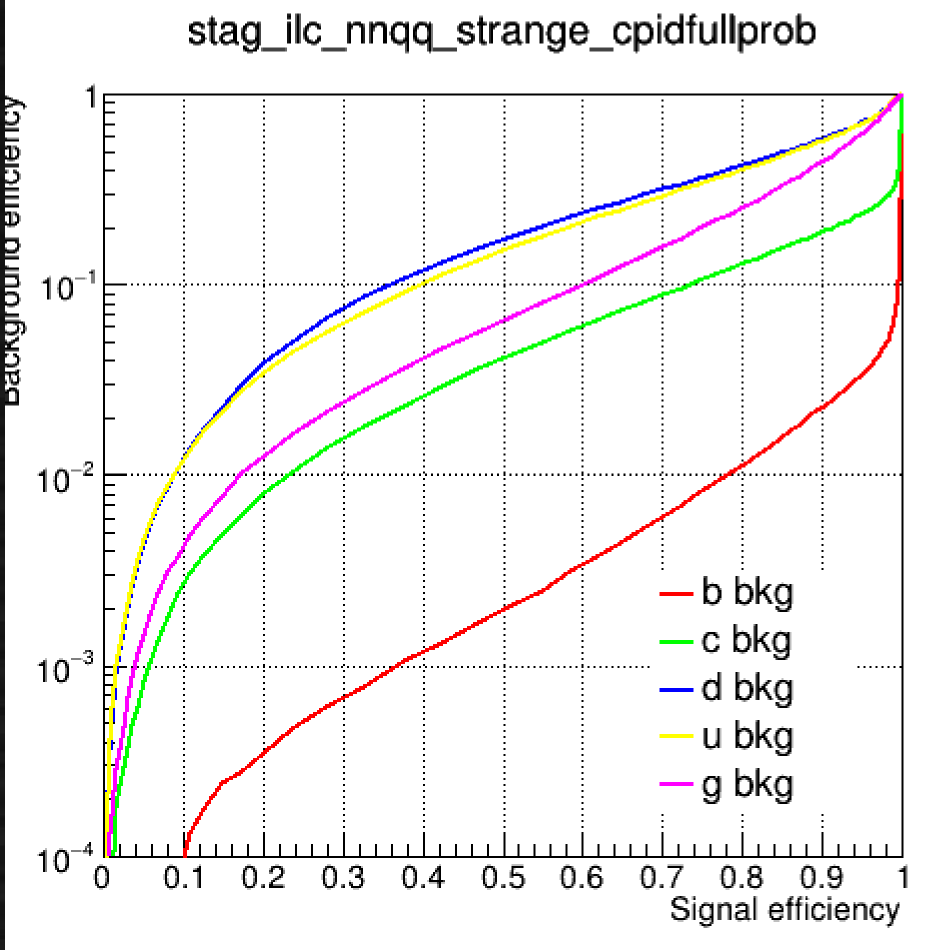
Strange tagging performance



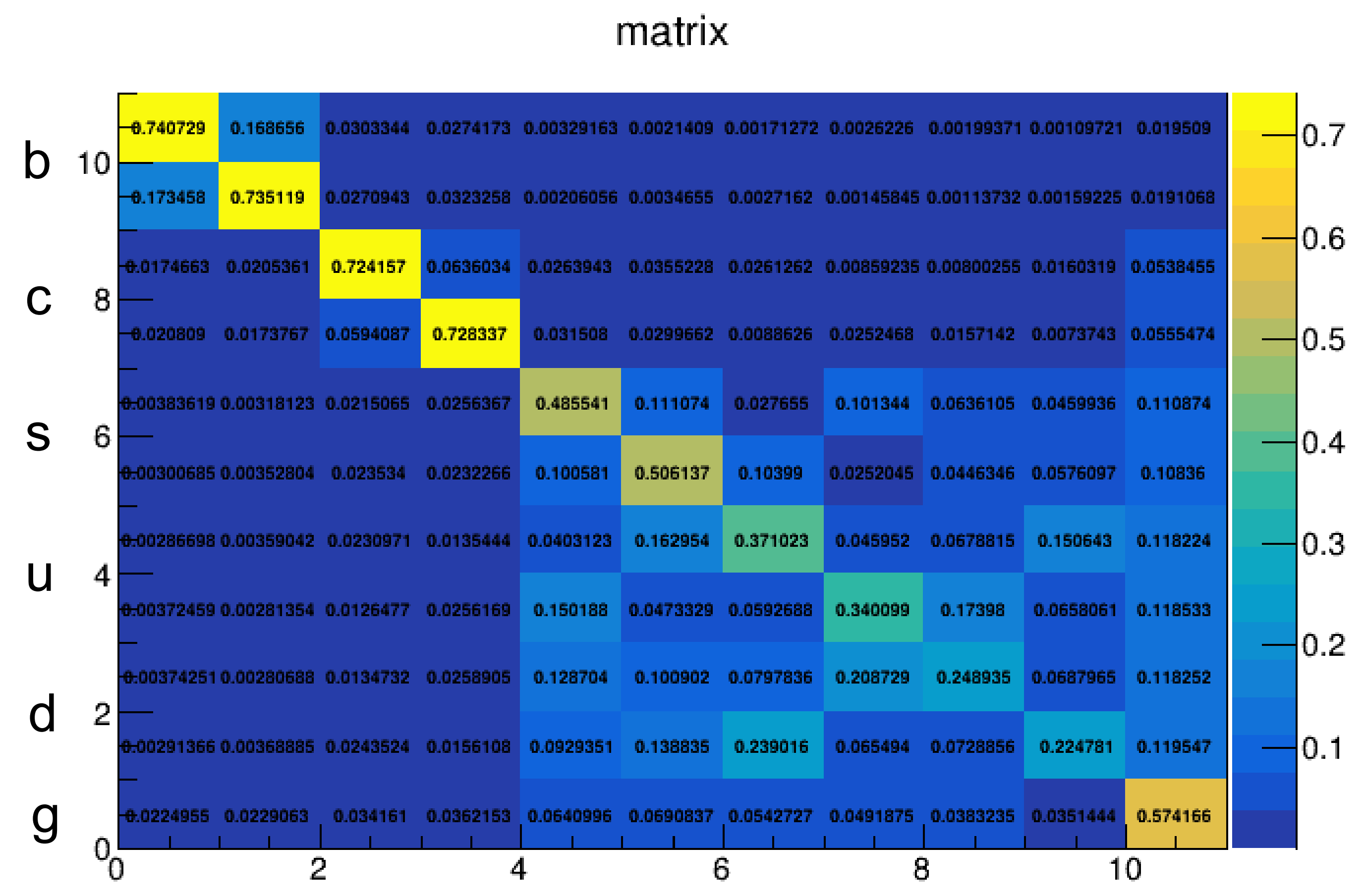
FCCee plot (in their study)



# 11-category q/qbar tag (nnqq sample)



CPID (K/π/p probability)
with 100 ps TOF (x 10 hits) + dE/dx



| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| b | 0.740729 | 0.168656 | 0.0303344 | 0.0274173 | 0.00329163 | 0.0021409 | 0.00171272 | 0.0026226 | 0.00199371 | 0.00109721 | 0.019509 |
| | 0.173458 | 0.735119 | 0.0270943 | 0.0323258 | 0.00206056 | 0.0034655 | 0.0027162 | 0.00145845 | 0.00113732 | 0.00159225 | 0.0191068 |
| c | 0.0174663 | 0.0205361 | 0.724157 | 0.0636034 | 0.0263943 | 0.0355228 | 0.0261262 | 0.00859235 | 0.00800255 | 0.0160319 | 0.0538455 |
| | 0.020809 | 0.0173767 | 0.0594087 | 0.728337 | 0.031508 | 0.0299662 | 0.0088626 | 0.0252468 | 0.0157142 | 0.0073743 | 0.0555474 |
| s | 0.00383619 | 0.00318123 | 0.0215065 | 0.0256367 | 0.485541 | 0.111074 | 0.027655 | 0.101344 | 0.0636105 | 0.0459936 | 0.110874 |
| | 0.00300685 | 0.00352804 | 0.023534 | 0.0232266 | 0.100581 | 0.506137 | 0.10399 | 0.0252045 | 0.0446346 | 0.0576097 | 0.10836 |
| u | 0.00286698 | 0.00359042 | 0.0230971 | 0.0135444 | 0.0403123 | 0.162954 | 0.371023 | 0.045952 | 0.0678815 | 0.150643 | 0.118224 |
| | 0.00372459 | 0.00281354 | 0.0126477 | 0.0256169 | 0.150188 | 0.0473329 | 0.0592688 | 0.340099 | 0.17398 | 0.0658061 | 0.118533 |
| d | 0.00374251 | 0.00280688 | 0.0134732 | 0.0258905 | 0.128704 | 0.100902 | 0.0797836 | 0.208729 | 0.248935 | 0.0687965 | 0.118252 |
| | 0.00291366 | 0.00368885 | 0.0243524 | 0.0156108 | 0.0929351 | 0.138835 | 0.239016 | 0.065494 | 0.0728856 | 0.224781 | 0.119547 |
| g | 0.0224955 | 0.0229063 | 0.034161 | 0.0362153 | 0.0640996 | 0.0690837 | 0.0542727 | 0.0491875 | 0.0383235 | 0.0351444 | 0.574166 |

Vertical: truth jet PDG, horizontal: predicted jet PDG
PDG with highest score taken

Application to H → ss ongoing

# Prospects on flavor tagging

- Checking “scaling law” with full simulation
  - Generation/production ongoing
- Optimize input variables / network
  - Hyperparameter tuning, learning rate etc.
  - “vertex information” to be included
- Direct application to event categorization
  - Reduce uncertainty on jet clustering (for multi-jet final states)
  - Can combine flavor tagging results? Treated as fine tuning?
- Applying or producing foundation models

# Flavor tagging: summary and plans

- Significantly better performance of flavor tagging with ParT
  - Scaling law to be checked
  - Application to physics analyses ongoing
- Strange tagging / quark-antiquark tagging
  - Performance still needs to be understood more
  - Dependence on PID performance to be investigated
    - Coming with various detector configurations
- More study on event categorization / foundation model foreseen

# Overall summary

- Future Higgs factories are gaining momentum
  - Still unclear whether we can get or not
- Flavor tagging with ParT significantly better than LCFIPlus
  - Applying to key studies (self coupling etc.)
  - Strange tagging / quark-antiquark tagging available
- Particle flow with GNN gives competitive performance
  - Energy regression under study
  - Hope to replace PandoraPFA in ~a year
  - To be used also for detector design

# Backup

# Results on efficiency and purity (another view)

| **Algorithm train/test** | **Electron eff.** | **Pion eff.** | **Photon eff.** | **Electron pur.** | **Pion pur.** | **Photon pur.** |
|---|---|---|---|---|---|---|
| GravNet<br>10 taus/10 taus | 98.8% | 99.6% | 99.1% | 92.6% | 99.3% | 97.7% |
| PandoraPFA<br>10 taus | 99.3% | 94.0% | 99.1% | 91.8% | 94.6% | 97.2% |
| GravNet<br>jets/jets | 94.6% | 93.1% | 95.2% | 77.4% | 93.1% | 92.4% |
| PandoraPFA<br>jets | 80.2% | 90.4% | 79.0% | 75.0% | 90.6% | 77.7% |
| PandoraPFA<br>jets<br>(ILCSoft truth) | 96.7% | 95.5% | 96.4% | 97.1% | 90.4% | 97.7% |

At least in our measure, performance of GravNet-based algorithm exceeds PandoraPFA

→ Promising as full PFA (but energy regression to be done)

Definition of MC truth clusters needs to be tuned (see ILCSoft truth)

# Software for Particle Transformer

- Public in github, with instruction provided
  - https://github.com/jet-universe/particle_transformer
- Input: ROOT files for training (80%), validation (5%), test (15%)
  - Input variables can be provided via steering file (XML)
    - Input for each particle (tracks, neutral clusters)
    - Input for “interaction” → currently momentum only
    - Input for “coordinate” → theta/phi plan wrt. jet axis
- Output: ROOT files including evaluation results (likeness) for test events
  - To be analyzed with ROOT or so
- We implemented a processor (inside LCFIPlus) to produce ROOT files for input as much as compatible to FCCee variables
  - Except for PID values, which are not fully implemented
- Easy for testing, but not direct to be used for physics analyses

# Training parameters - epochs

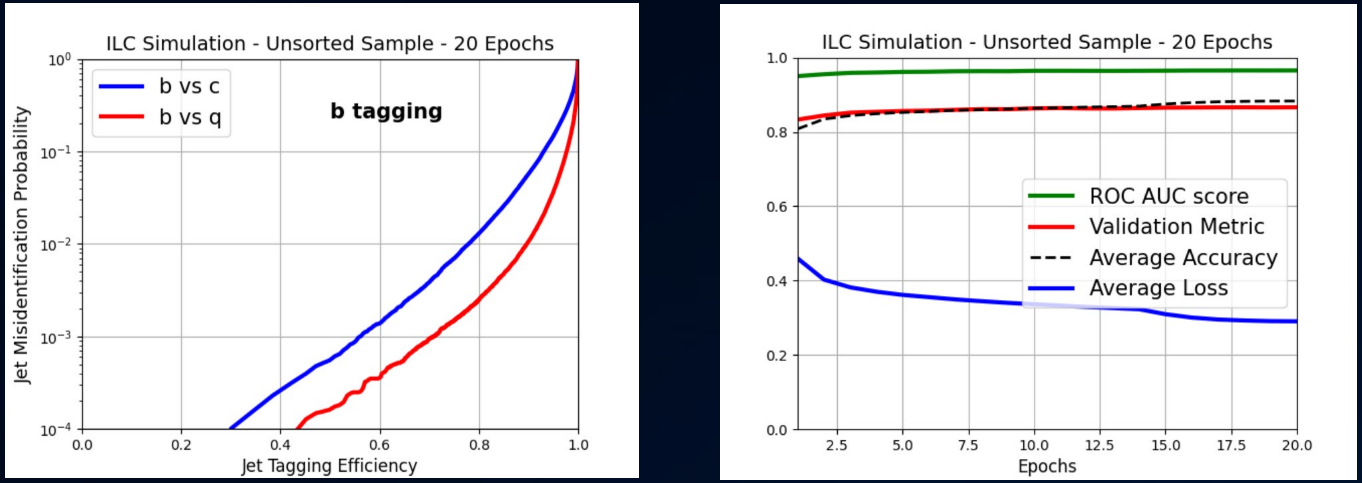
- Run on NVIDIA TITAN RTX (memory: 24 GB)
  - 20 Epochs: 3 hours
  - 200 Epochs: 30 hours

- No significant improvement in tagging efficiency

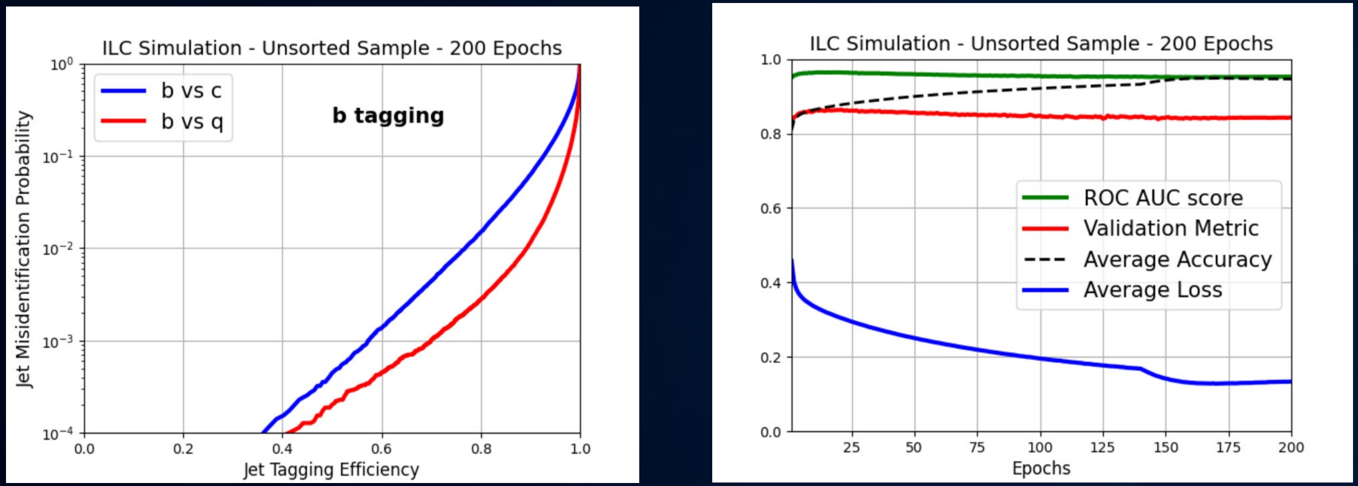
- Both ROC AUC score and Validation Metric reaches a maximum around 20 epochs.

- Overtraining after 20 epochs.

- Hence 20 epochs of training is selected to avoid overtraining.



20 epochs (ILD qq 91 GeV)



200 epochs (ILD qq 91 GeV)

# Input Variables - Features

*Naming follows FCCee scheme – may not express exact meaning

- Impact Parameter (6):

pfcand_dxy
pfcand_dz
pfcand_btagSip2dVal
pfcand_btagSip2dSig
pfcand_btagSip3dVal
pfcand_btagSip3dSig

*d0/z0 and 2D/3D impact parameters, 0 for neutrals

- Jet Distance (2):

pfcand_btagJetDistVal
pfcand_btagJetDistSig

*Displacement of tracks from line passing IP with direction of jet
0 for neutrals

- Particle ID (6):

pfcand_isMu
pfcand_isEl
pfcand_isChargedHad
pfcand_isGamma
pfcand_isNeutralHad
pfcand_type

* Not including strange-tagging related variables (TOF, dE/dx etc.)
* Simple PID for ILD, not optimal

- Kinematic (4):

pfcand_erel_log
pfcand_thetarel
pfcand_phirel
pfcand_charge

*Fraction of the particle energy wrt. jet energy (log is taken)

- Track Errors (15):

pfcand_dptdpt
pfcand_detadeta
pfcand_dphidphi
pfcand_dxydxy
pfcand_dzdz
pfcand_dxydz
pfcand_dphidxy
pfcand_dlambdadz
pfcand_dxyc
pfcand_dxyctgtheta
pfcand_phic
pfcand_phidz
pfcand_phictgtheta
pfcand_cdz
pfcand_cctgtheta

*each element of covariant matrix
0 for neutrals

# Input Variables - Interactions

- FCC data uses p (scalar momentum) as interaction:

  - pfcand_p

- ILD data contains $p_x$, $p_y$, $p_z$ (vector momentum) as interaction:

  - pfcand_px
  - pfcand_py
  - pfcand_pz

- But it's possible to transfer ILD's interaction to FCC's form for fair comparison:

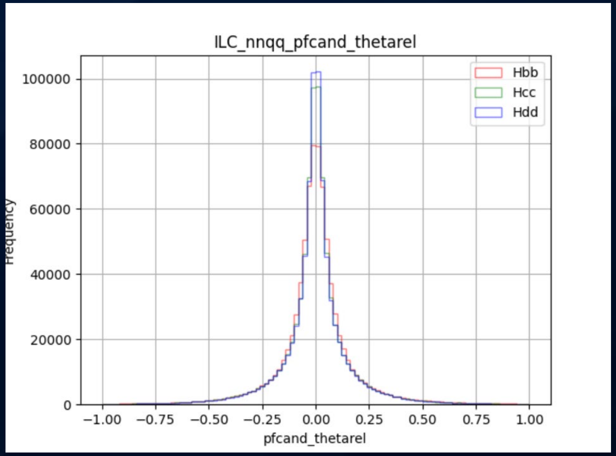
$$p = \sqrt{{p_x}^2 + {p_y}^2 + {p_z}^2}$$

# Use px, py, pz instead of p (Interaction)

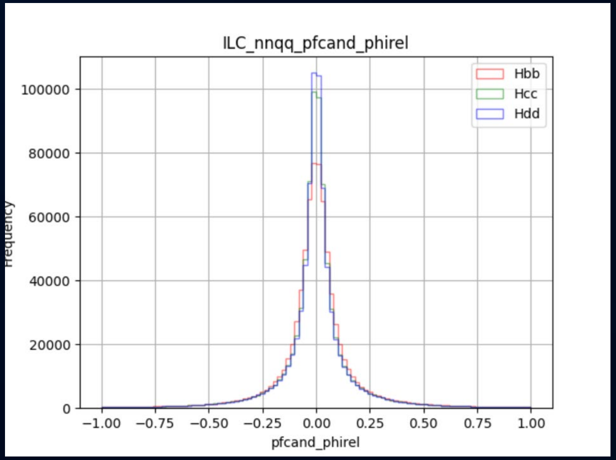
| | | | | c-bkg acceptance @ b-tag 80% eff. | | b-bkg acceptance @ c-tag 50% eff. | |
|---|---|---|---|---|---|---|---|
| **Particle ID** | **Impact Parameters** | **Jet Distance** | **Track Errors** | **p** | **$p_x$ $p_y$ $p_z$** | **p** | **$p_x$ $p_y$ $p_z$** |
| ✗ | ● | ● | ● | **0.62%** | **0.49%** | **1.14%** | **1.01%** |
| ✗ | ●<br>+log(abs) | ●<br>+log(abs) | ●<br>+log(abs) | **0.54%** | **0.52%** | **1.06%** | **1.00%** |
| ✗ | ●<br>+log(abs) | ● | ● | **0.47%** | **0.50%** | **1.03%** | **0.97%** |

- ILD (ννqq 250 GeV) data shows that application of px, py, pz has better performance than p.
- However, application of log(abs) of the parameters becomes less significant.
- Can be because that application of px, py, pz changes the way log(abs) interacts with other parameters.
- Other potential treatments can be investigated.

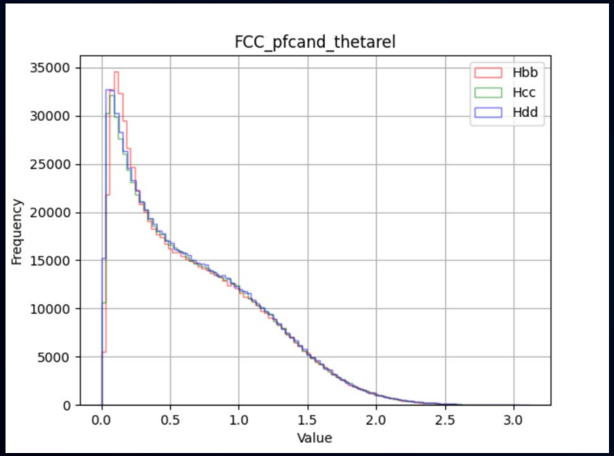
# ILD vs. FCC – theta/phi distribution



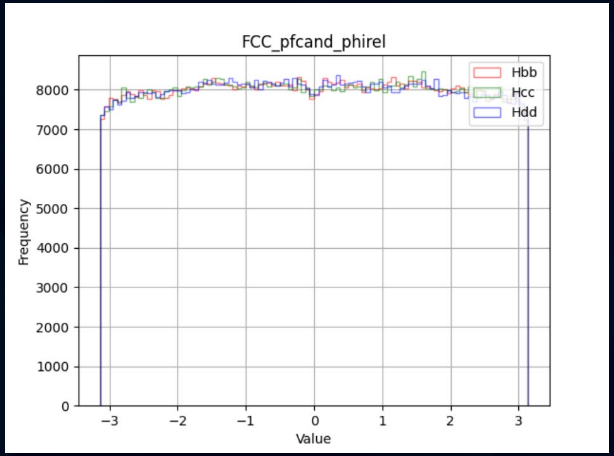
ILD theta



ILD phi



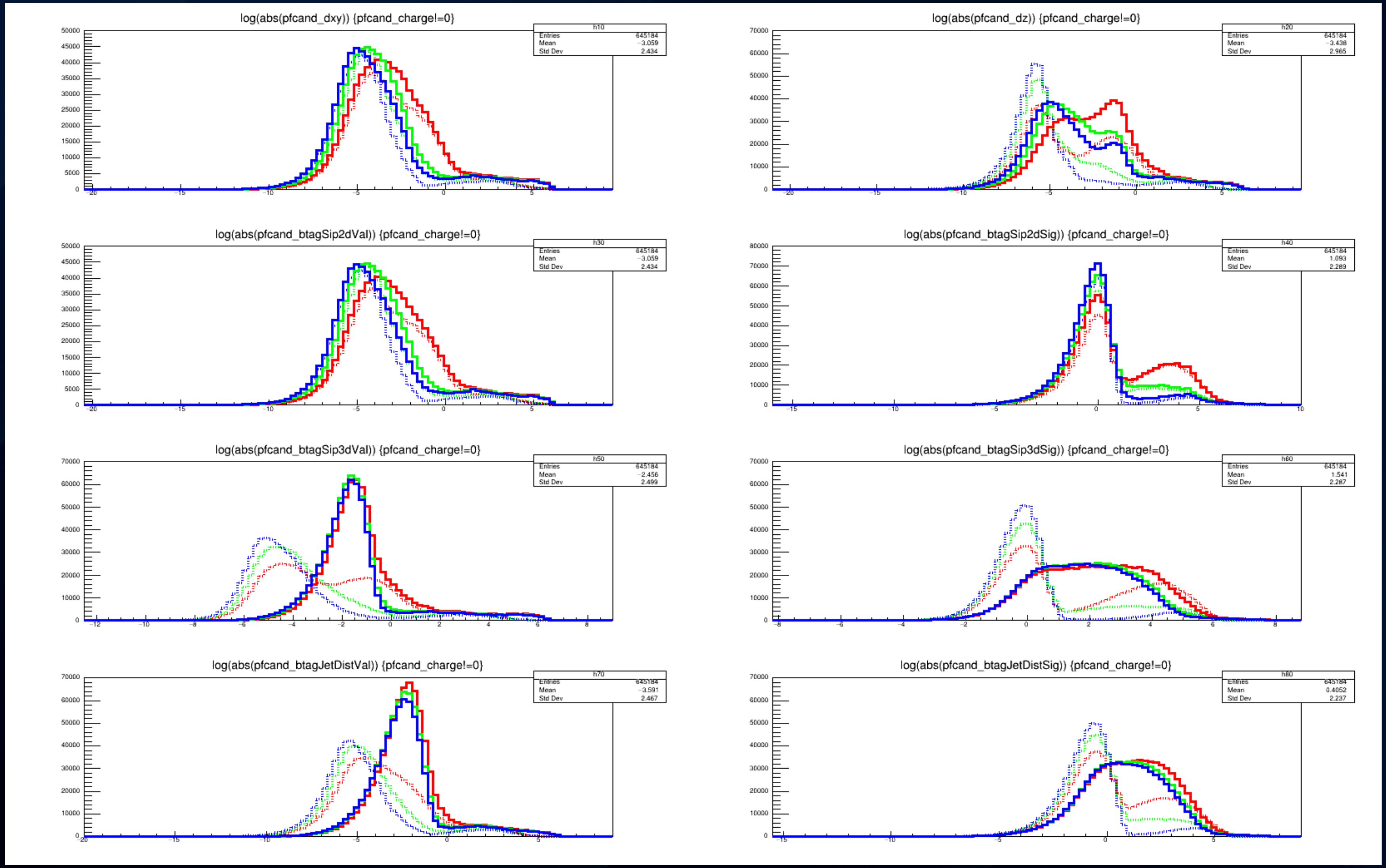
FCC theta



FCC phi

- ILD theta/phi are calculated from the difference between particle and jet theta/phi in the frame of the detector.
- FCC theta/phi are obtained from relative trace of the particle compared to the jet.
- This can cause some differences in the interaction of other parameters in the model.

# Difference in impact parameters



Dotted – FCCee
Solid – ILD

Red – nnbb
Green – nncc
Blue – nndd

Significant difference on dz seen
- beam spot smearing?

# Fine tuning

Two objectives

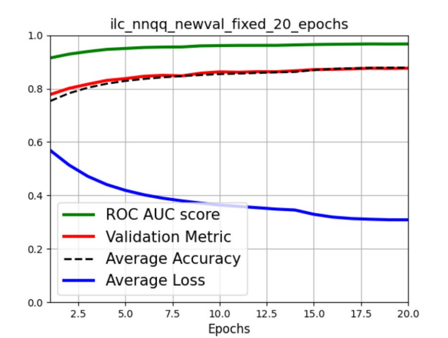
- Pretrained with fast sim and fine-tune with full sim
- Pretrained with large central production and fine-tune with dedicated physics samples in each analysis

| | | | | | | | c-bkg acceptance @ b-tag 80% eff. | | b-bkg acceptance @ c-tag 50% eff. | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Particle ID** | **Impact Parameters** | **Jet Distance** | **Track Errors** | **Fine-Tuning Sample** | **Training Sample** | **Similar theta/phi ?** | **No Fine-Tuning** | **With Fine-Tuning** | **No Fine-Tuning** | **With Fine-Tuning** |
| ✗ | ● | ● | ● | FCC 240 GeV (8M) | ILD 250 GeV (800k) | ✗ | 0.62% | 1.37% | 1.14% | 1.95% |
| ✗ | ● | ● | ● | FCC 240 GeV (8M) | ILD 250 GeV (800k) | ● | 1.77% | 1.32% | 2.22% | 2.01% |
| ● | ● | ● | ● | ILD 250 GeV (800k) | ILD 91 GeV (80k) | ● | 4.49% | 0.97% | 3.79% | 1.53% |

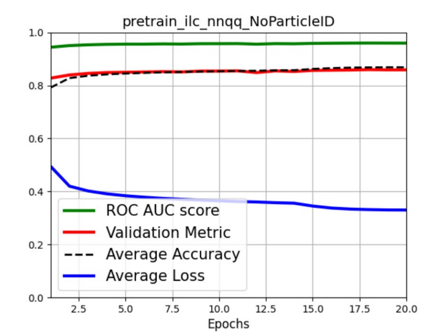
- Use result of 8M FCC data to train ILD 800k data
- Improves performance only when setups are similar
- Training of same setup (pretrain ILD 91 GeV data with ILD 250 GeV data) gives best performance
- Further investigation should be conducted on how to maximise the outcome for fine-tuning between different data sets

# Fine tuning – Training curves

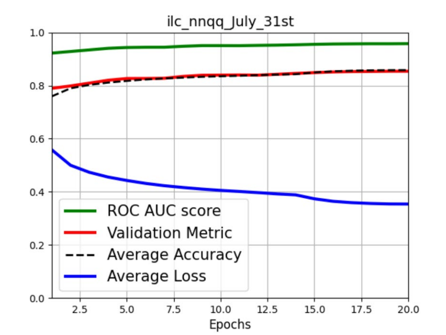
(1)



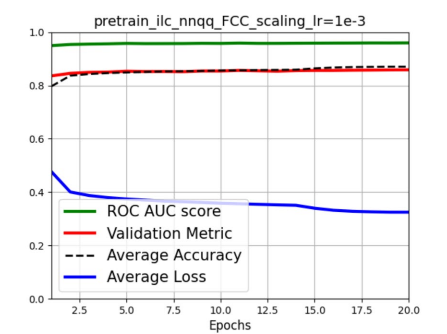
(2)



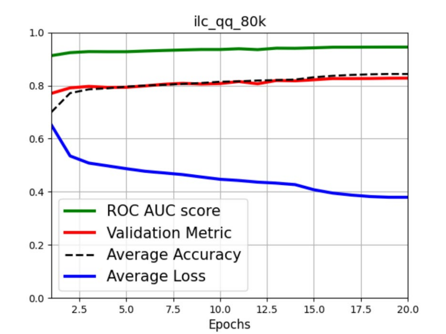
(3)



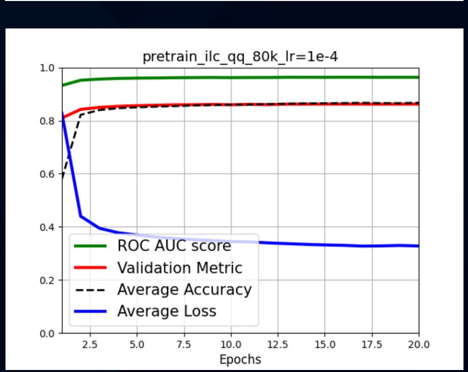
(4)



(5)



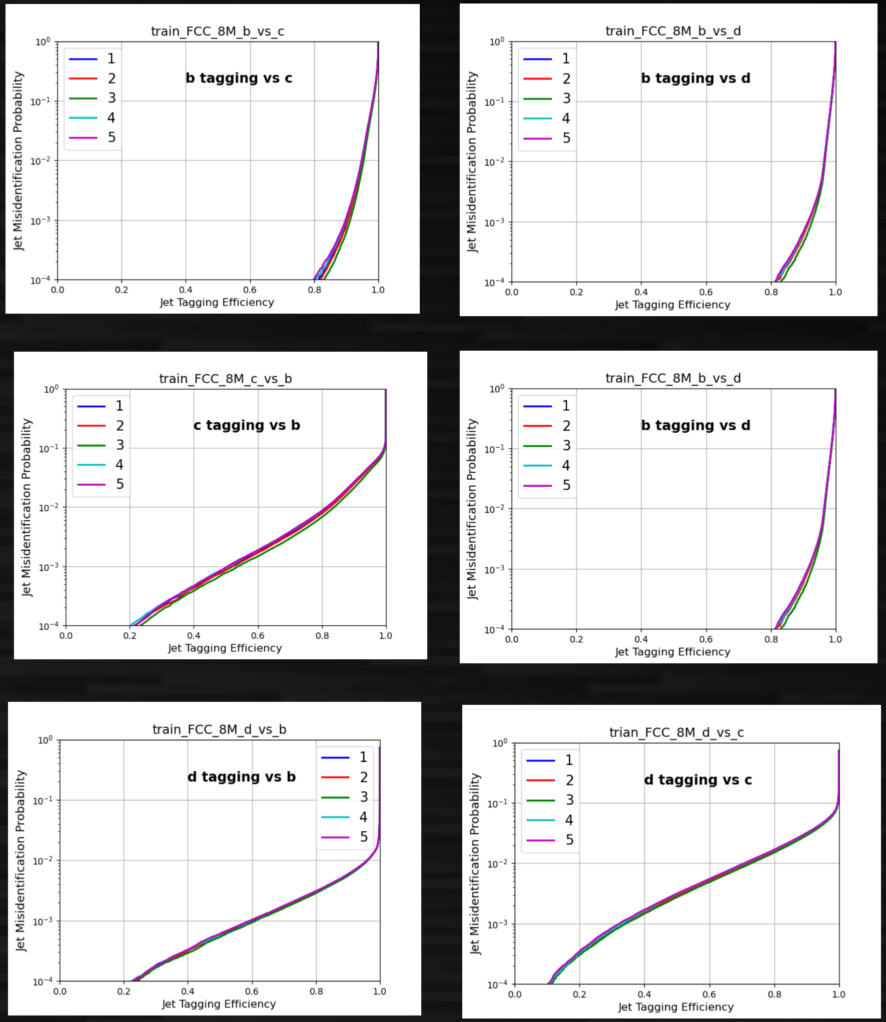
(6)



| | | | | | | | Plot Indices | |
|---|---|---|---|---|---|---|---|---|
| Particle ID | Impact Parameters | Jet Distance | Track Errors | Fine-Tuning Sample | Training Sample | Similar theta/ phi? | No Fine-Tuning | With Fine-Tuning |
| ✗ | ● | ● | ● | FCC 240 GeV (8M) | ILD 250 GeV (800k) | ✗ | (1) | (2) |
| ✗ | ● | ● | ● | FCC 240 GeV (8M) | ILD 250 GeV (800k) | ● | (3) | (4) |
| ● | ● | ● | ● | ILD 250 GeV (800k) | ILD 91 GeV (80k) | ● | (5) | (6) |

- With fine-tuning, the training is obviously accelerated for the initial epochs (even for those with worse eventual performance)
- This is particularly obvious between plots (5) & (6) – similar simulation setup data

# Multiple Training Runs

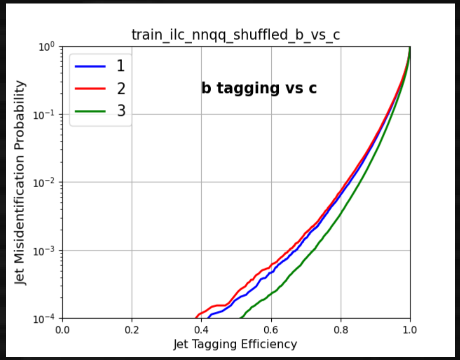


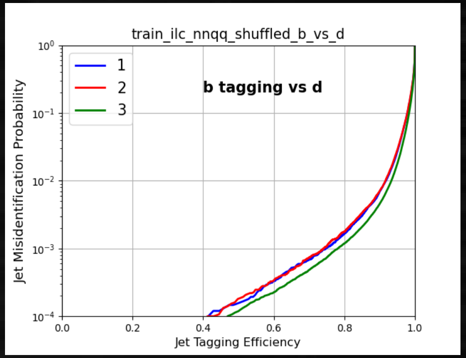
5 times training of FCC_8M data

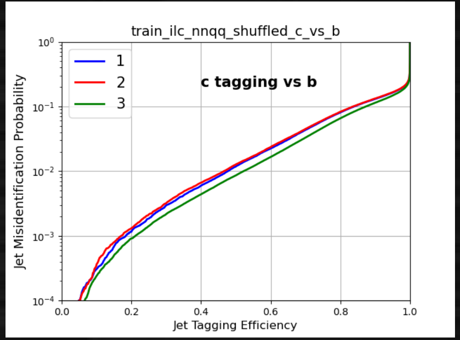
- Multiple training runs don't give significant impacts on results.
- The smaller data size is, the bigger impacts on results multiple runs give.
- The results of no Particle ID trainings varies more than those of with Particle ID.

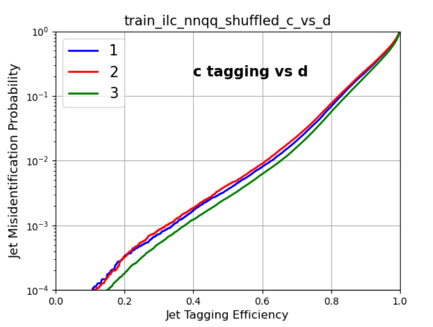
| data | Particle ID | b vs c 0.8 Score | variation |
|---|---|---|---|
| FCC 4M | ○ | 4.82e-4 | 0.43e-4 |
| FCC 8M | ○ | 8.14e-5 | 1.58e-5 |
| FCC 4M | × | 1.69e-3 | 0.14e-3 |
| FCC 8M | × | 7.04e-4 | 3.49e-4 |

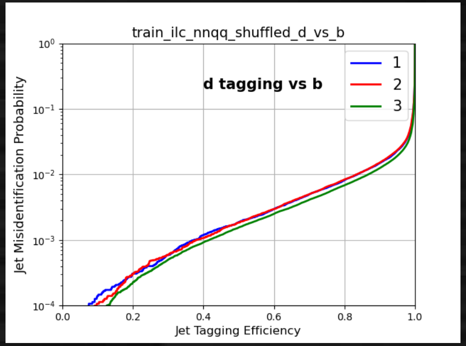
# Data Shuffled

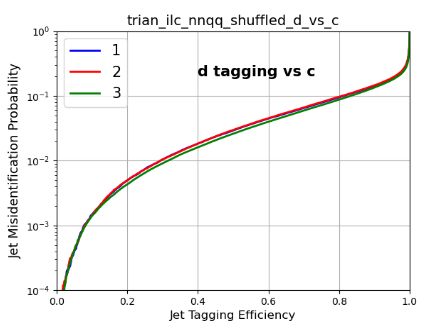












- ILC nnqq dataset
  - 80% training, 5% validation, 15% test
- Shuffled the order of train/test/val making root files
  - Pattern 1: train/val/test
  - Pattern 2: val/train/test
  - Pattern 3: train/test/val

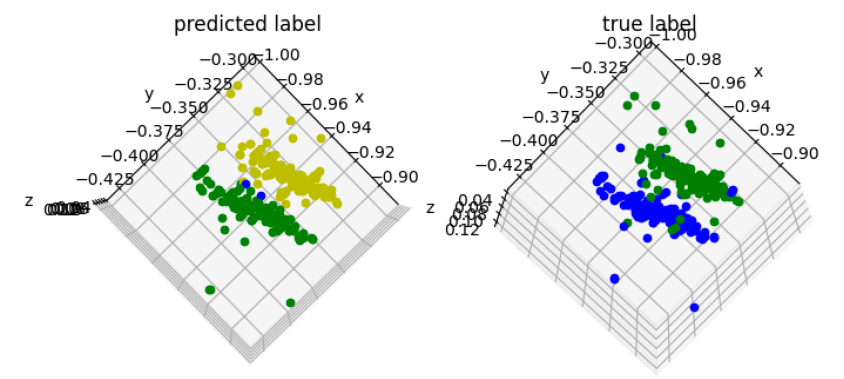
e

| data | b vs c 0.8 score |
|---|---|
| Shuffle pattern 1 | 0.00647 |
| Shuffle pattern 2 | 0.00734 |
| Shuffle pattern 3 | 0.00338 |

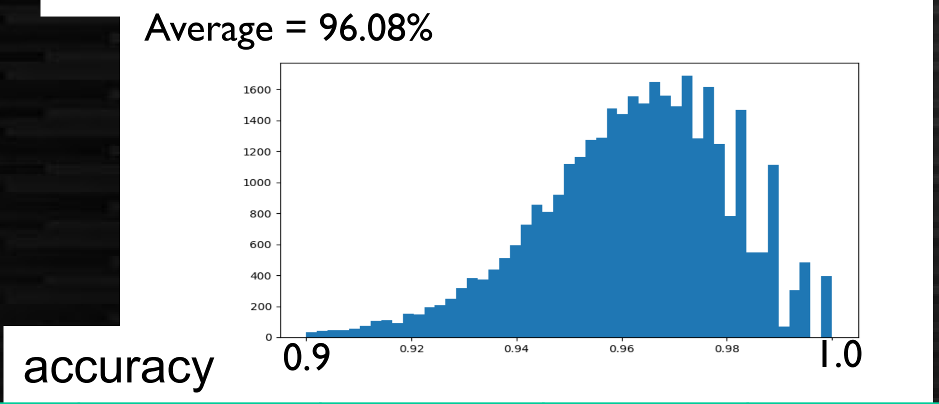
# Importing to ILD full simulation

- Prepare features from ILD full simulation
  - With recent versions (> v02-02)
- Input features: (x, y, z, edep)
- True cluster info
  from MCParticle and LCRelation
- Produced events
  - Two photons
    (5/10 GeV, fixed opening angles)
  - (n x ) taus (5/10 GeV)
- Evaluation
  - Fraction of hits associated to the correct cluster (accuracy)

Example of a two-photon event
(5 GeV, 30 mrad)



Average = 96.08%

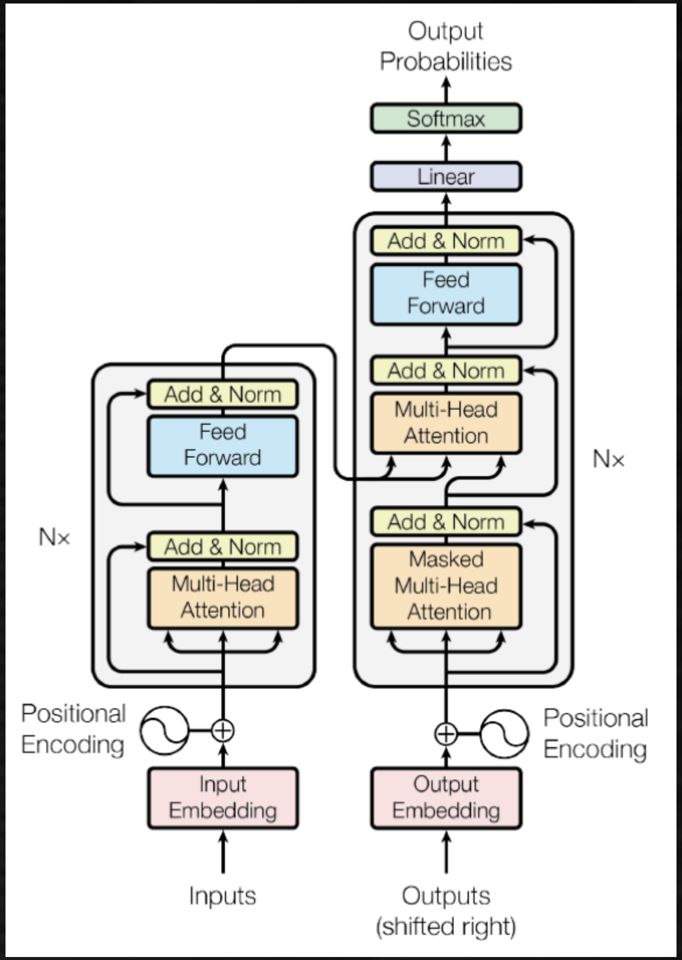


Reasonable
performance seen

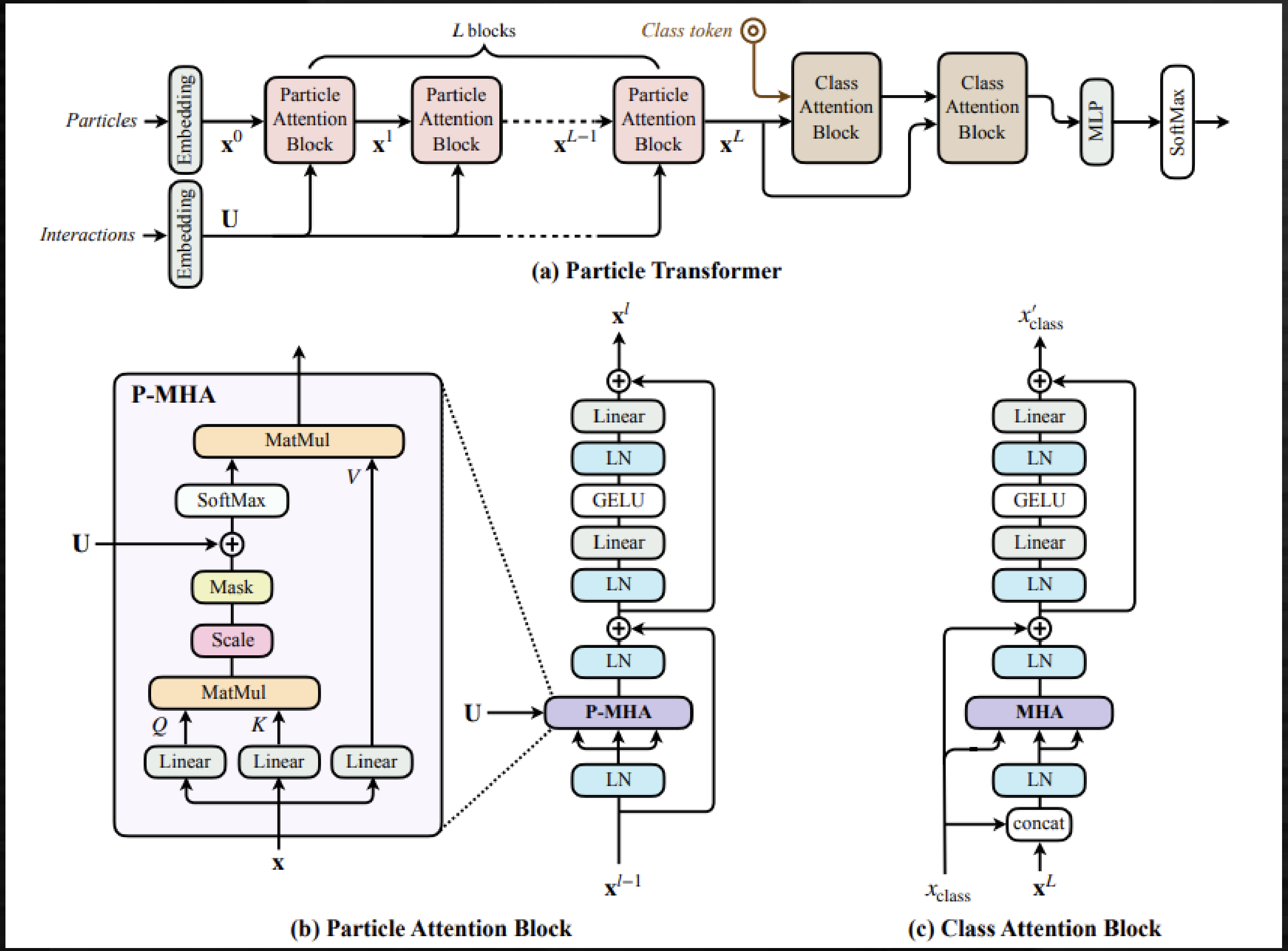
| **Angle[mrad]** | **30** | **60** | **90** | **120** | **150** |
|---|---|---|---|---|---|
| Accuracy[%] | 96.08 | 98.64 | 99.30 | 99.68 | 99.56 |

For details, refer eg. https://indico.slac.stanford.edu/event/7467/contributions/5948/attachments/2887/8032/230517-lcws2023-hlreco-suehara.pdf

# Comparison between regular Transformer and Particle Transformer



Regular Transformer



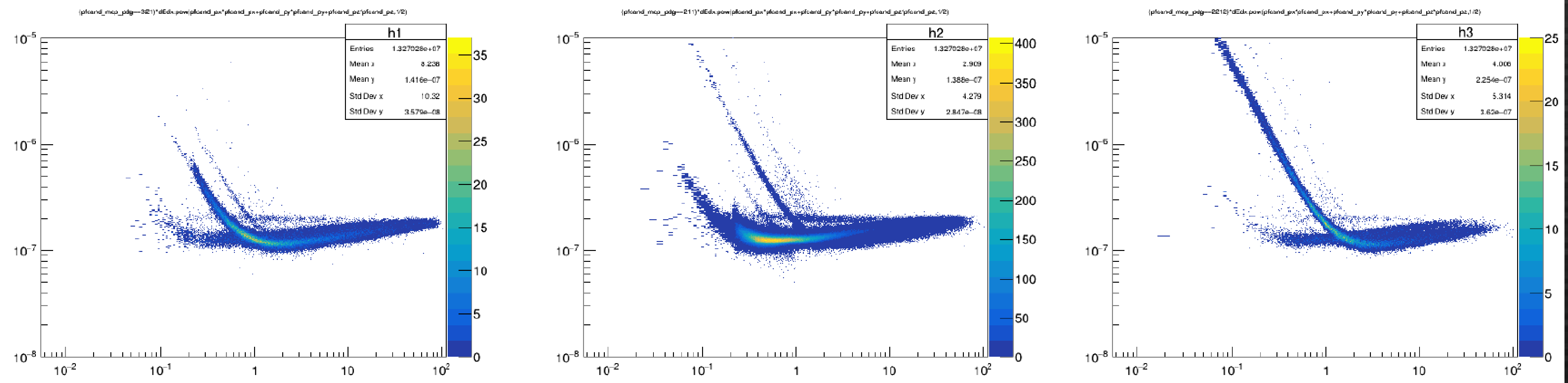
Particle Transformer

Note:
MHA – MultiHeadAttention
P-MHA – Augmented version of MHA by Particle Transformer that involves Interactions Embeddings instead of Positional Embeddings

# Progress in strange tag

| | s vs c | s vs g | s vs u |
|---|---|---|---|
| 0.8 efficiency | 0.138 | 0.288 | 0.466 |

Current performance with ParT
(under investigation yet)



Kaon Pion Proton

dE/dx inside strange jets (separated by MC PID)

# Inference within LCFIPlus

- Training done in python/weaver framework
  - New LCFIPlus algorithm (MLMakeNtuple) to create input ROOT files
  - ROOT files used for training ParT
    - nnqq 250 GeV, ~1M jets / each flavor
    - MC/jet matching inside LCFIPlus (only for q/qbar training)
      - Color-singlet tagging by RecoMCTruthLink, q/g identified based on angle
        - If multiple jets assigned to the same q/g, jet with highest energy taken
  - Training with GPU (~a half day for 20 epochs with Tesla V100)
- Weights (checkpoint) converted to onnx
  - Using onnx 1.15.0, onnxruntime 1.17.1 (to be compatible with key4hep)
- Inference with CPU in LCFIPlus framework
  - New processor MLInferenceWeaver with onnx files (uploaded in LCFIPlusConfig)
- Currently on private repository (pulling to official repository being processed)
  - LCFIPlus github with ParT, https://github.com/suehara/LCFIPlus/tree/onnx
  - LCFIPlusConfig with weight/steering files, https://github.com/suehara/LCFIPlusConfig

# ILC: International Development Team

| | |
|---|---|
| *Americas Liaison* | Andrew Lankford (UC Irvine) |
| *Working Group 3 Chair* | Jenny List (DESY) |
| *Working Group 2 Chair* | Shinichiro Michizono (KEK) |
| *Executive Board Chair and Working Group 1 Chair* | Tatsuya Nakada (EPFL) |
| *KEK Liaison* | Yasuhiro Okada (KEK) |
| *Europe Liaison* | Steinar Stapnes (CERN) |
| *Asia-Pacific Liaison* | Geoffrey Taylor (U. Melbourne) |

See LCWS2023: https://indico.slac.stanford.edu/event/7467/

WG3 physics group hosts series of physics meetings
https://agenda.linearcollider.org/category/266/
(Next: July 13th)
Mailing list subscription:
https://agenda.linearcollider.org/event/9154/

Established in 2020: aiming for ILC pre-lab
Pre-lab proposal in 2021
https://arxiv.org/abs/2106.00602
→ MEXT expert panel (2021)

- Not mature enough for proceeding to pre-lab
  - Mainly in international situation
- Accelerator technology should be developed in preparation for next step

→ Two steps towards pre-lab

- International Technology Network (ITN)
  - Collaboration framework with US/Europe
  - Doing time-critical works of pre-lab
  - Japanese part is funded by MEXT
- International Expert Panel
  - Among researchers connected to FA
  - Discussing how to proceed "global" projects