ML for reconstruction at FCC-ee

Dolores Garcia

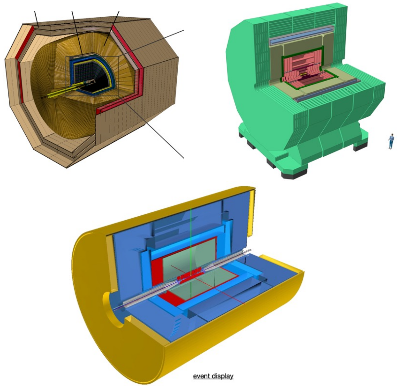
Swiss National Science Foundation

Motivation

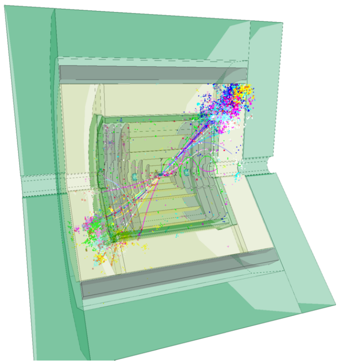
- FCC-ee has been recommended by the European Strategy as CERN's next flagship collider [1]
- FCC-ee will collide e+e- at Z, WW, ZH and $t\bar{t}$ thresholds with large luminosity
- Experiments aim to measure Higgs couplings, electroweak parameters, flavour observables with extemely high precision
- Achievable sensitivity on SM parameters is driven by the resolution with which visible final state particles and their invariant masses can be reconstructed

Precisions depends on:

High performance detectors

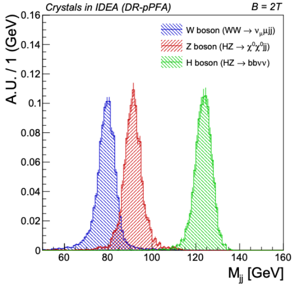


High quality reconstruction

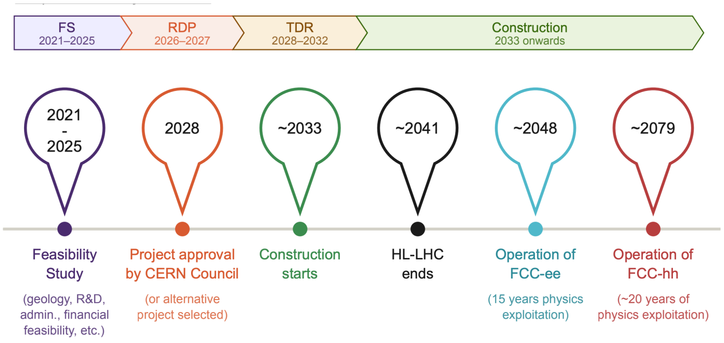


Best resolution translates into best physics sensitivity

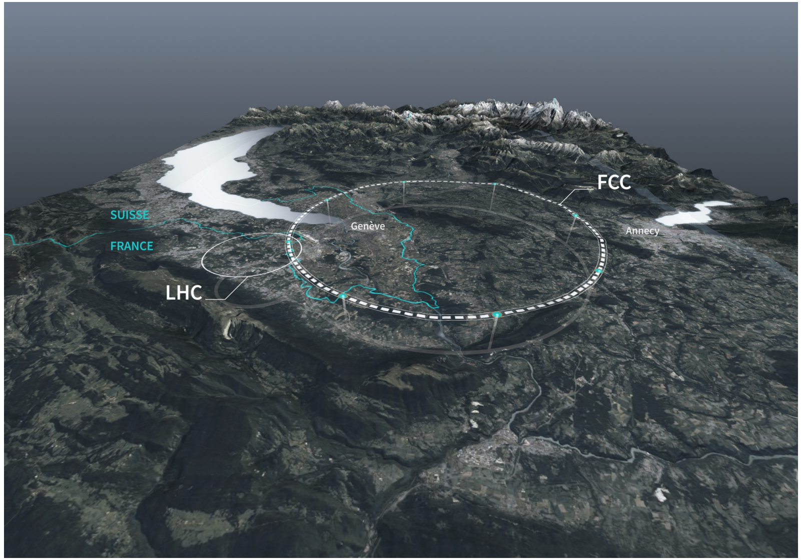
- Modern jet flavour tagging makes use of full jet per particle level data
- W/Z hadronic mass separation for H--> WW, ZZ
- Rare H→ $s\bar{s}$, H→ $c\bar{c}$ relative uncertainty scales with mass resolution



Entering the Reference Design Phase era



- ❑ Excellent test ground for development of reconstruction algorithms that could be translated to HL-LHC, complex detectors with many channels and low pile-up



- ❑ Goal: enable FCC-ee detector design optimization
- ❑ Solidify the physics-reach and physics case with robust full-simulation implementations
- ❑ Need general reconstruction algorithms to apply to multiple detector geometries and technologies

Adapted from F.Beirer@FCC Week 2026

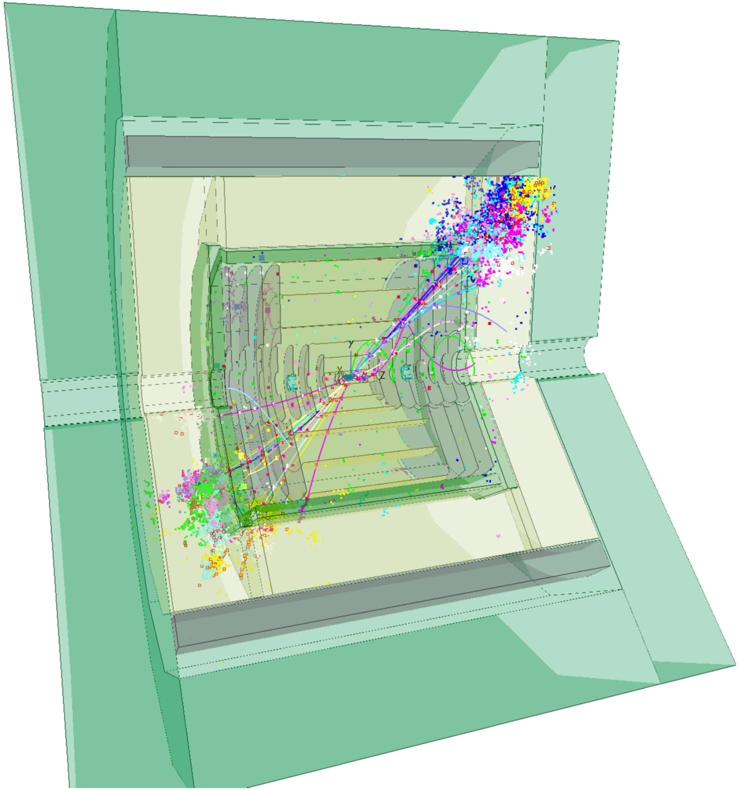
Summary of the talk

ML for reconstruction

Particle Flow

Tracking

Other works at FCC (will not be covered in this talk):
Tau reconstruction
Tagging
…

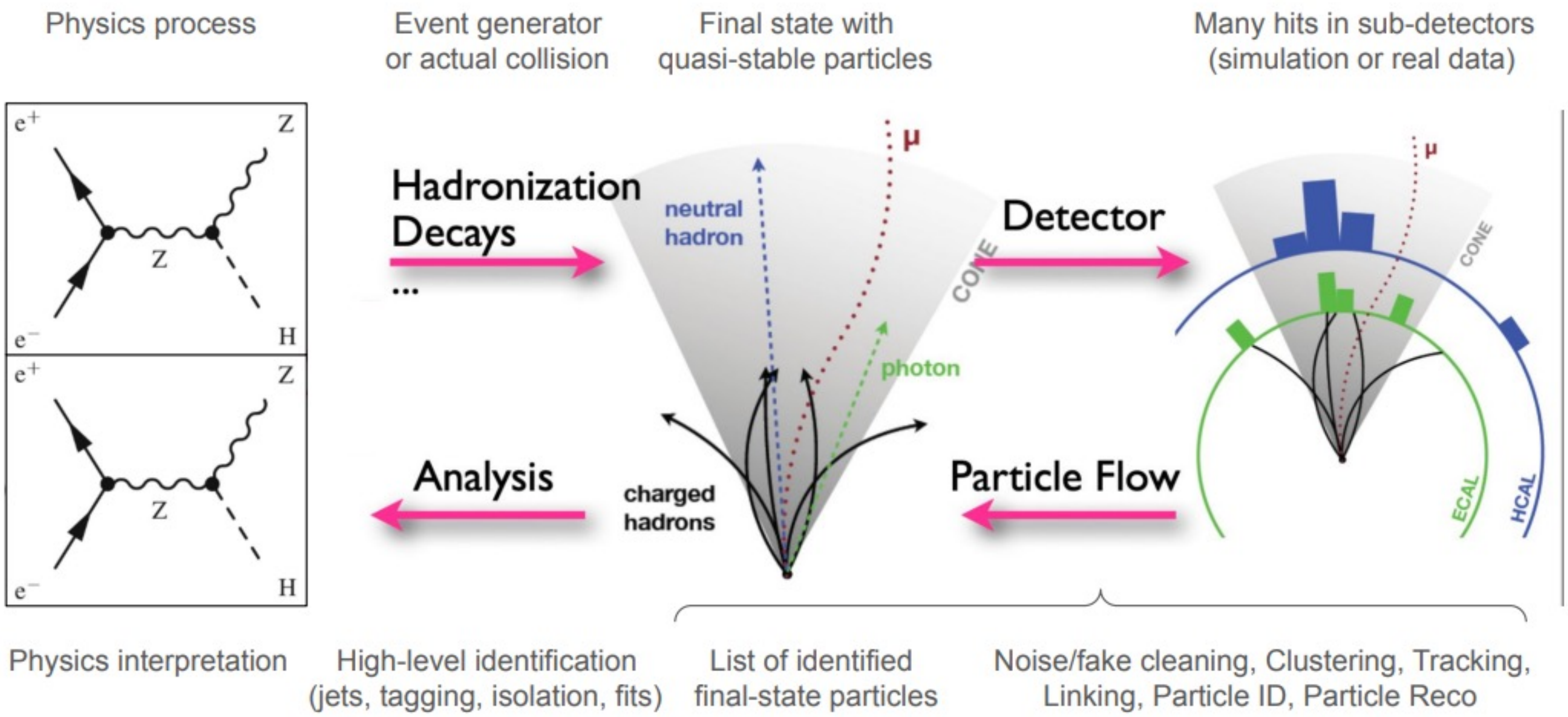


Event display in the CLD detector for a Z/γ∗→qq event 365 GeV. From Bacchetta et al 2019.

Particle Flow reconstruction

In the times of ML (and AI)

What is Particle Flow?



P.Janot PF@ FCC Workshop 2025

3

What is Particle Flow?

- **The goal:** Optimally combine measurements from all sub-detectors
  - Vertex, tracker, ECAL/HCAL, muon detector
  - From **hits** to **identified final state particles**

- For each ~**stable particle**, this combination aims at returning a **reco particle**
  - With the best **direction** and **energy** estimate
  - With the most likely **type** (photon, electron, muon, charged hadron, neutral hadron)
  - With its origin vertex (primary, secondary, etc.)

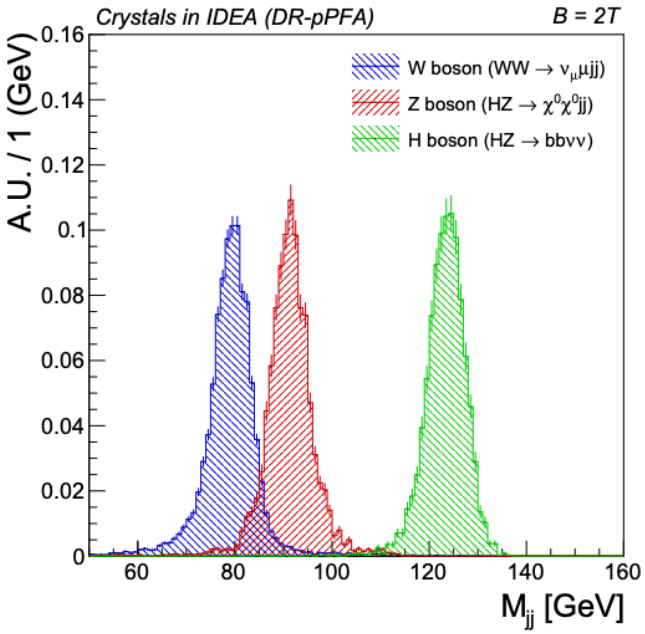
- The list of individual particles can then be used
  - To determine the **total**/missing **energy** and momentum; build jets; tag **jet flavour**; identify and reconstruct **taus** from their decay products; select leptons or photons; etc.
  - With the best possible energy/angular **resolution**, efficiency, purity, …
  - With **no** overlap or **double counting** - as with a generator-level analysis

P.Janot PF@ FCC Workshop 2025

Why Particle Flow reconstruction?

From raw data to physics observables

- Reconstructing individual particles from raw signals connects experimental measurements with physics observables
- Achievable sensitivity on key SM measurements is driven by the resolution with which the visible final state particles and their invariant masses can be reconstructed
- Rare higgs decays (H→cc, H → ss, relative uncertainty scale is sensitive to visible mass resolution)
- Hadronic final states large particle multiplicity and geometrical overlap

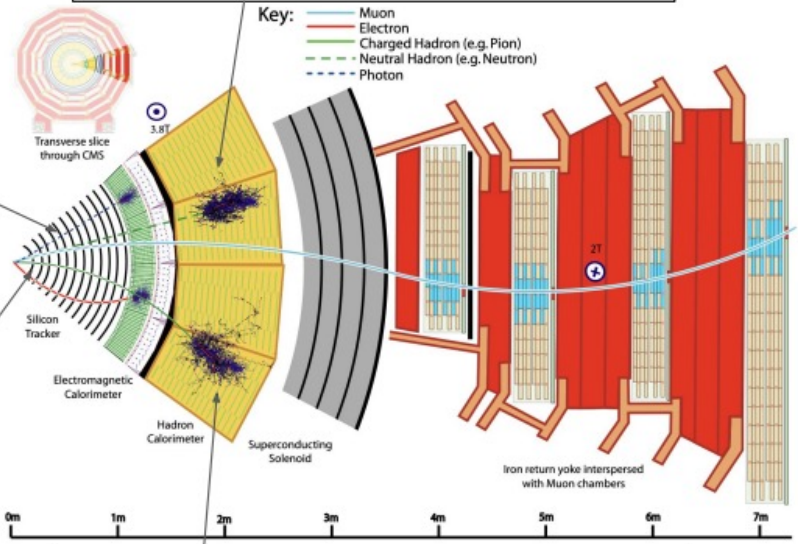


https://arxiv.org/pdf/2202.01474

The ‘game’

Neutral hadrons = one shower in HCAL and/or one shower in ECAL (or nothing)

Photons = one compact shower in ECAL



Muons = one track in the tracker, one track in the muon chambers plus MIP signals in ECAL and HCAL

Electrons = one compact shower in ECAL and one track

Charged hadrons = one track, one shower in HCAL and/or one shower in ECAL

Why is it hard?

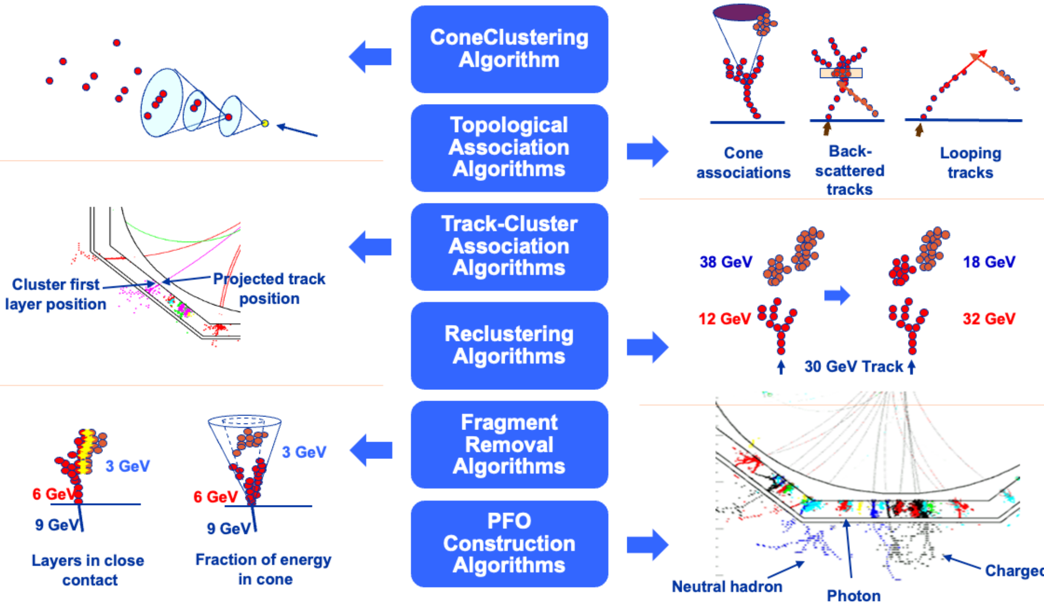
- Classically, multiple algorithms, many moving pieces:
  - Tracking (one per sub-detector), clustering (ECAL/HCAL), and linking tracks and clusters
- We want every particle in the event, but not fakes, hard to tune
- **Fake tracks**
- **Lonely tracks** (tracks that are real but left no clusters)
- **Displaced tracks** (less hits, tracking classically not tunned for that)
- Physics:
  - Photons conversion in the tracker (detached vertex + e+e-, maybe two tracks maybe none)
  - Electrons radiation (‘double counting’)
  - Hadrons interaction (detached vertex + more particles)
  - Neutral hadrons decay (e.g., K^0_S or Λ)
  - Many edge cases that are rare (hard to take all into account)

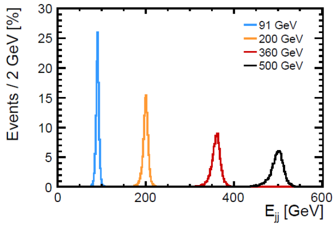
?

Talk by Mark Thomson at the FCC Particle Flow Workshop (Sept 2025)

How is it tackled?

- ❑ Classically, multiple algorithms:
  - ○ Tracking (one per sub-detector), clustering (ECAL/HCAL), and linking tracks and clusters
  - ○ ALEPH, DELPHI, CMS, ATLAS, CLIC,…
- ❑ Pandora SDK (baseline FCC)
  - ○ Series of distributed algorithms
  - ○ Iterative clustering
  - ○ Track cluster association and reclustering
  - ○ Neutral identification
  - ○ Relies on tuning, limiting flexibility during detector design phase (allows for geometry studies e.g transverse calo segmentation)

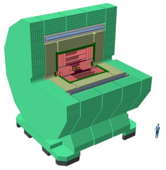




Talk by Mark Thomson at the FCC Particle Flow Workshop (Sept 2025)

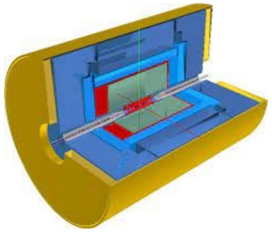
The need for an end-to-end approach: **Multiple detectors**

- Adaptable reconstruction is crucial for a systematic **design optimization**:
  - Parameter tuning is not straight forward; the detector will still change + different detectors
  - Costly conventional implementation

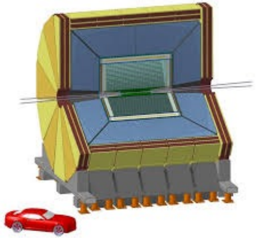


CLD
- Silicon pixel + silicon tracker
- ECAL Granular Silicon-Tungsten 5x5 mm^2 cells (15-17 %)
- Scintillator-steel, 44 layers, 30x30 mm^2
- Solenoid outside calo

IDEA
- Silicon pixel + drift chamber + silicon wrapper
- Dual-readout crystal
- Dual-readout fiber scintillating and cherenkov
- Solenoid between ECAL/HCAL



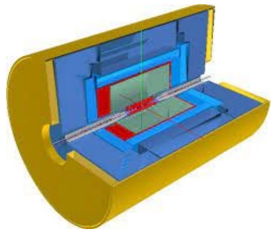
ALLEGRO
- Silicon pixel + drift chamber + silicon wrapper
- ECAL granular noble-liquid sampling LAr (8-10 %)
- Scintillating-tile HCAL
- Solenoid behind ECAL

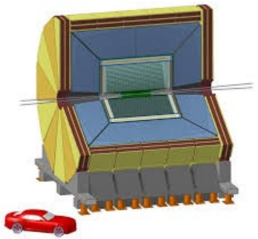


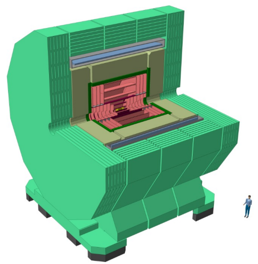
And others… ALFA, ILD

The need for an end-to-end approach: **Multiple detectors**

- ❑ The detector determines how the showers 'look':
  - ❑ What is the shape of an electron
    - o How wide (Molière radius)
    - o How long (longitudinal shower profiles)
  - ❑ How do hadrons shower in the HCAL, what is the shower geometry like, what are acceptable patterns?
  - ❑ What is the 'correct' threshold for a track-cluster separation to link them?





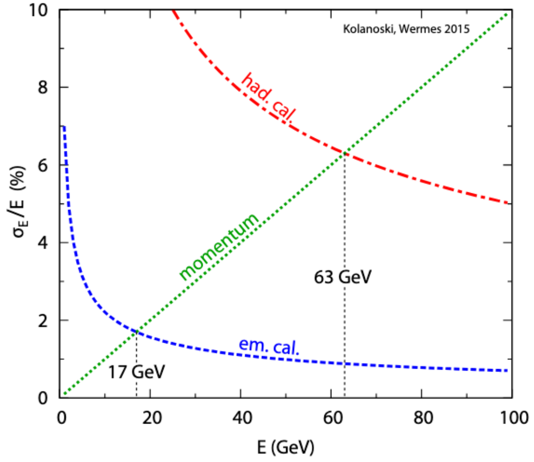


All these questions can be answered from data

The need for an end-to-end approach

Increased performance?

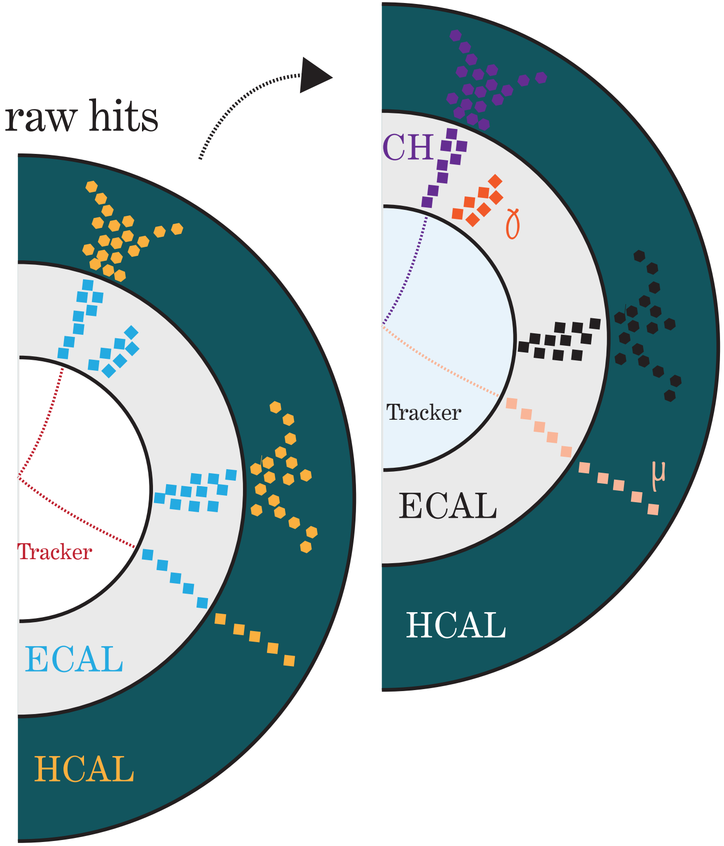
- ❑ Asymptotic resolution is 1-2% @100 GeV. Current performance 3-4% @100 GeV arXiv:0907.3577v1
- ❑ Key drivers of performance:
    - Tracking & calo cluster efficiency/purity
        - → Cluster merging: if two clusters are merged, with a track we lose resolution because we use the total calorimeter energy
        - → Shower splitting: double counting energy
    - Track cluster matching
    - Arbitration of track vs cluster energy



Average Jet Composition		Superior System
• 60% charged particles	→	tracker
• 30% photons	→	ECAL
• 10% neutral particles	→	HCAL

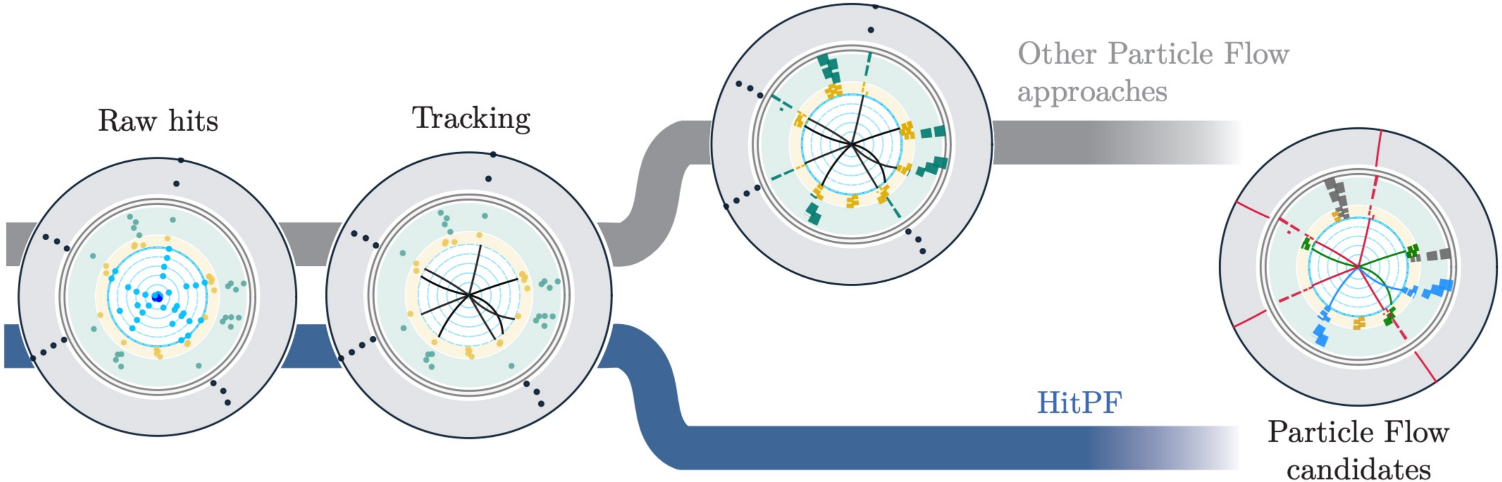
$$\sigma^2_{jet}(E_{vis}) = \sum_{i\,\in\,tr} \sigma^2_{tr}\left(E^{(i)}_{tr}\right) + \sum_{i\,\in\,\gamma} \sigma^2_{ECAL}\left(E^{(i)}_{ECAL}\right) + \sum_{i\,\in\,h^0} \sigma^2_{HCAL}\left(E^{(i)}_{h^0}\right)$$

How to set it as an ML problem



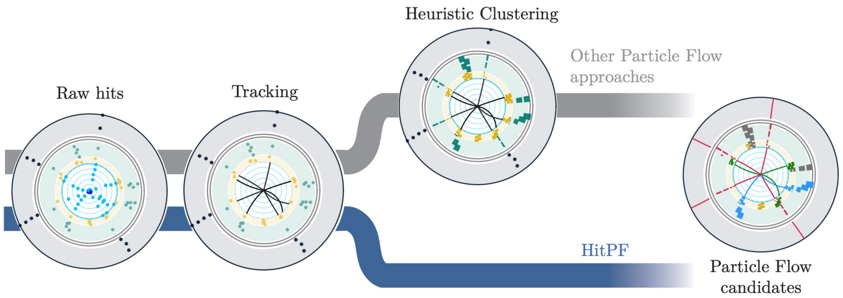
Hit-based approach to Particle Flow

Michele Selvaggi
Lena Herrmann
Gregor Krzmanc



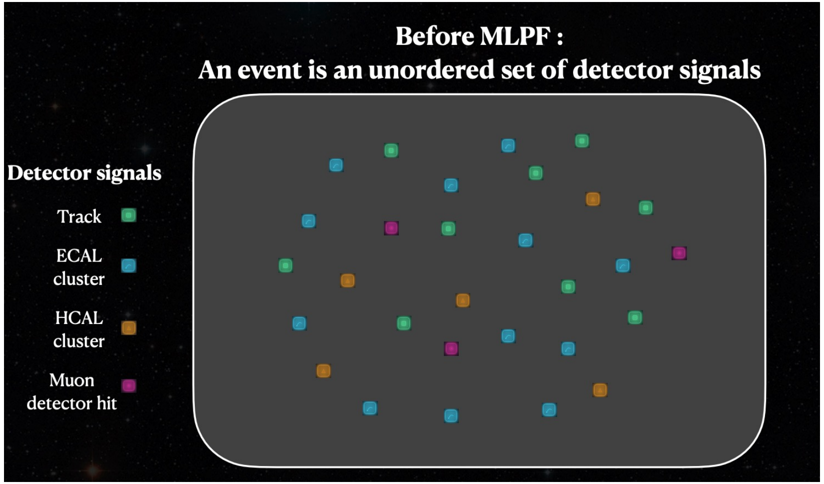
Comparison with other methods

Other PF approaches	HitPF
❑ Require heuristic clustering step + cluster track matching	❑ ML-based Clustering + track matching
❑ Complex parameter tuning for new detector	❑ Out of the box with multiple detector and hit geometries
❑ Requires geometry navigation	❑ Does not require the explicit detector geometry (navigation)

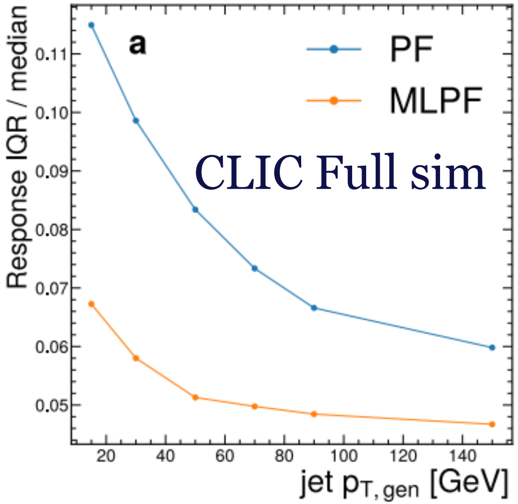


MLPF

- Cluster based approach:
  - Starts from pandora clusters and reconstructed tracks
  - In practice, very different problem
  - Improves the track-cluster linking and properties regression
  - 1 node in MLPF ~ 600 nodes in HitPF
  - Does not solve the geometry adaptation problem
  - New results for CMS

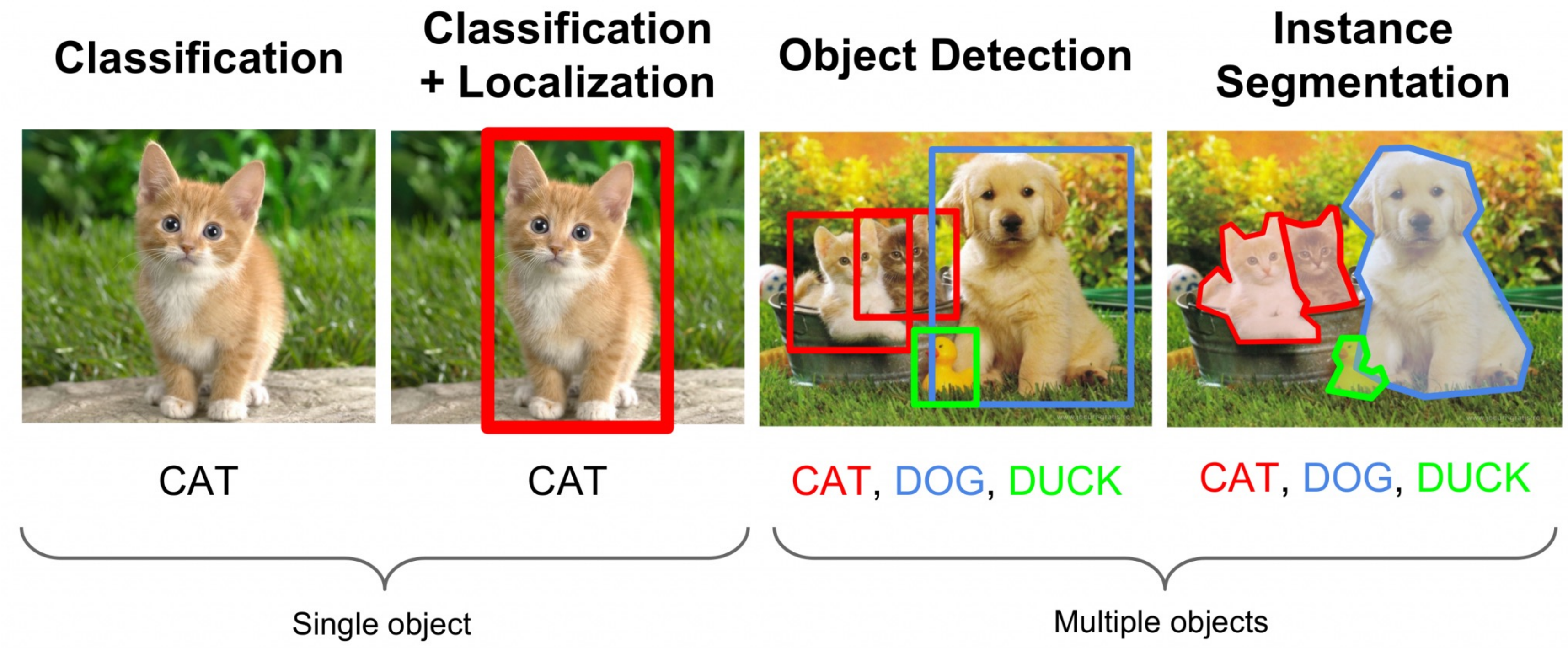


Farouk.M @ CERN Data Science Seminar

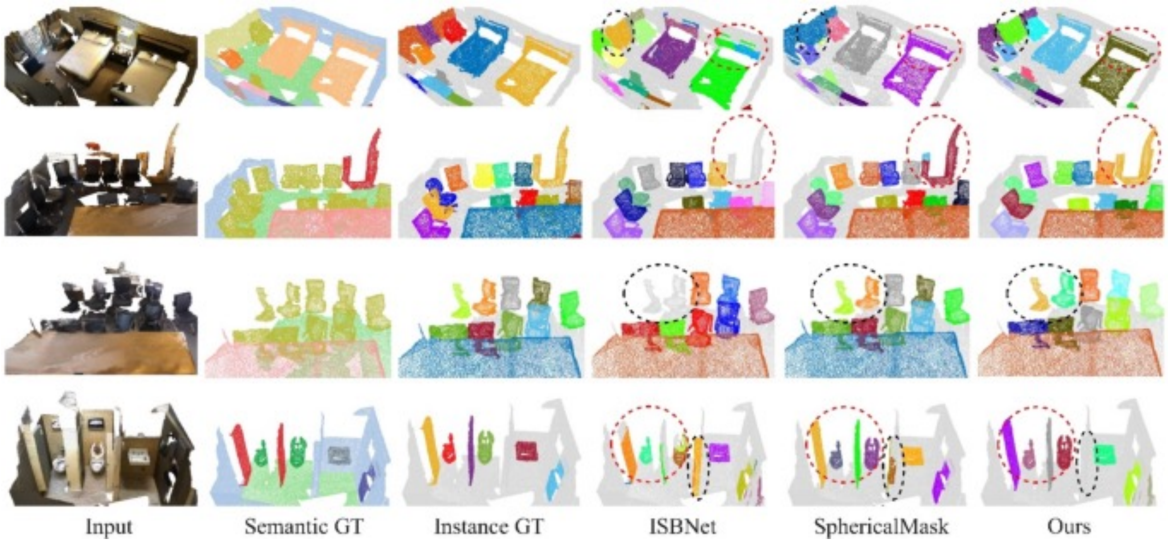


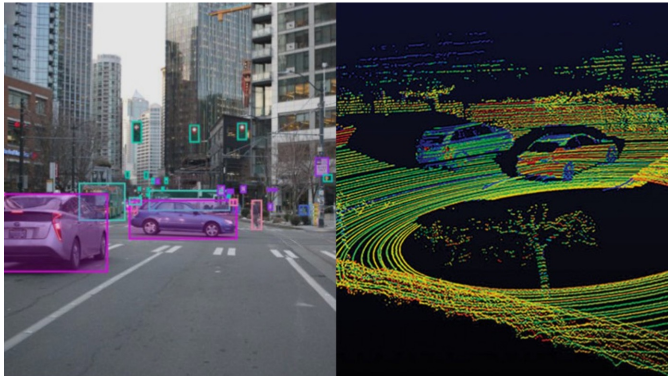
J.Pata et al 2024

Know problem in ML



Know problem in ML



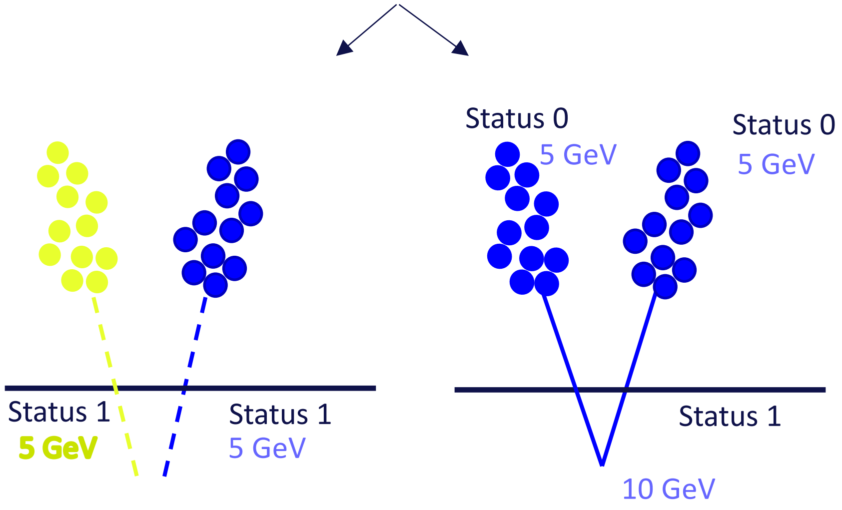


Supervision: Target definition

What is the best reconstruction target?

- Ideally, we would recover the stable particles that entered the detector (status 1) (previous studies)
- BUT this gives the algorithm conflicting training examples (why this is hard slide)
  - Conversions
  - Backscattering (clusters appear far in the detector)
- The **target** is defined as:
  - Resolvable detector objects
    - Ongoing work to add detector resolution
  - Defined in practice by the calorimeter boundary
  - This is what it supervisable
  - And means that we are solving the L1 problem

Same detector response



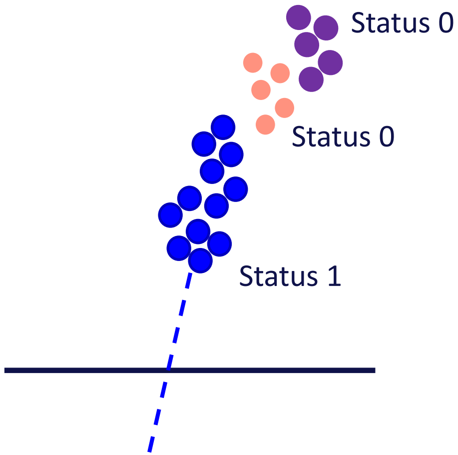
Different label?

Target definition

What is the best reconstruction target?

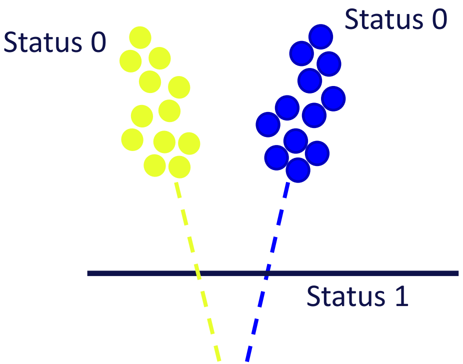
L0

- all Status 0 and status 1 particles eg. Hadronic showers resulting from nuclear interactions



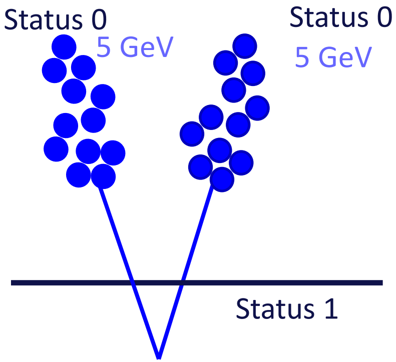
L1

- Calorimeter boundary abstraction
- Reconstruct two clusters for conversions, etc



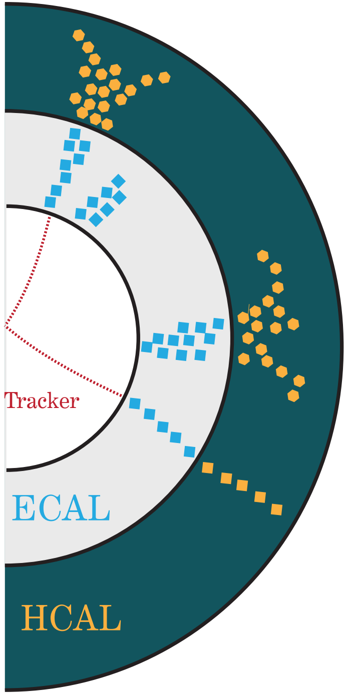
L2

- Status 1
- Complex topologies
- Brems, kinks, K^0_S



The workflow: Inputs

raw hits



Track 'hits:
Reference point in calorimeter (position)
p at vertex and calo
χ from track fit
Sub-detector ID

ECAL hit
Position (x,y,z)
E energy
Sub-detector ID

HCAL hit
Position (x,y,z)
E energy
Sub-detector ID

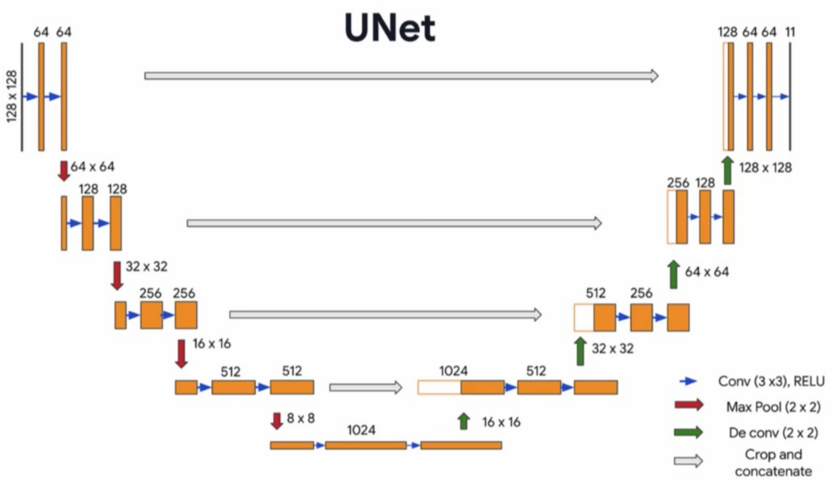
Muon hit
Position (x,y,z)
E energy
Sub-detector ID

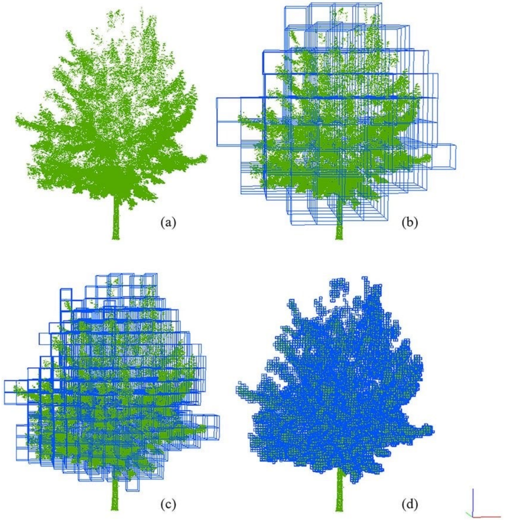
Other sub-detectors?
eg. ARC

pointcloud

The 3D ML architectures

UNet backbone + with voxelization





Inductive bias

Assumptions about the relationship between the input and output of a learning problem that are built into the model

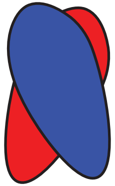
Can improve generalization

Can reduce required dataset size

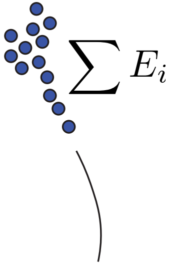
Adapting ML methods to PF

Inductive bias

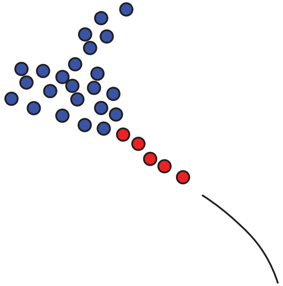
Self- intersections
No equivalent in 3D ML dataset



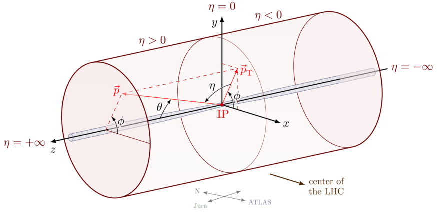
Multiple measurements in different subdetectors that ‘add up’ --> Optimal transport
No equivalent in 3D ML datasets (LiDAR scans of chairs do not add up to their mass)



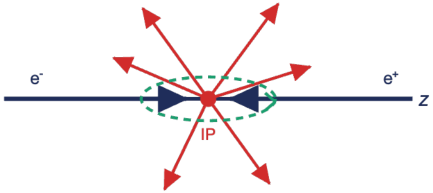
Need to preserve ‘fine’ structures in ECAL/tracker
Prevents voxelization and standard pulling operations



SO(2) + z mirror (endcaps)
~E(3) rotations, translations, and reflections



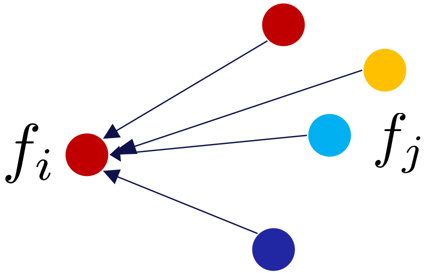
Beam direction
Preferred direction set by gravity



ML 3D data

Detector data

Model GNNs, transformers



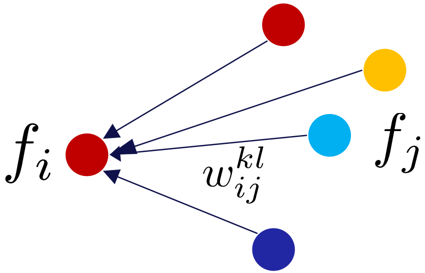
$$f_i' = \square_{j:(i,j)\in\mathcal{E}} h_{\Theta}(\mathbf{f}_i, \mathbf{f}_j).$$

Building block 1

$\square_{j:(i,j)\in\mathcal{E}}$ Aggregation operator

h_{Θ} Non-linear function with learnable parameters

Model GNNs, transformers



$$f_i' = \square_{j:(i,j)\in\mathcal{E}} h_\Theta(\mathbf{f}_i, \mathbf{f}_j).$$

Aggregation operator

$$\square_{j:(i,j)\in\mathcal{E}} \longrightarrow \quad f_i' = \sum_{j:(i,j)\in\mathcal{E}}$$

Non-linear function with learnable parameters

$$h_\Theta \longrightarrow \quad h_\Theta(\mathbf{f}_i^l, \mathbf{f}_j^l) = w_{ij}^{kl}\, \mathbf{V}^{kl} \mathbf{f}_j^l,$$

$$w_{ij}^{kl} = \mathrm{softmax}_j \left(\frac{(\mathbf{Q}_q^{kl} \mathbf{f}_i^l)^\top (\mathbf{K}_k \mathbf{f}_j^l)}{\sqrt{d}} \right)$$

Equivariant networks

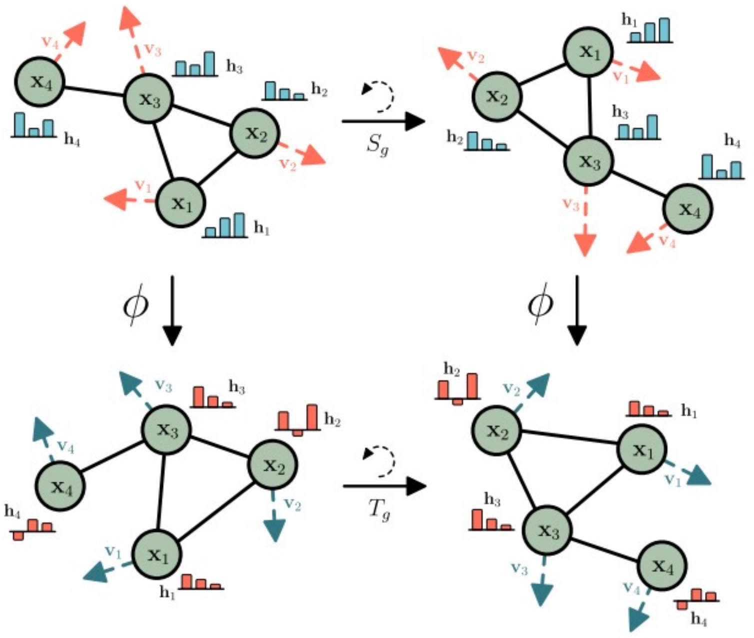


Figure from E(n) equivariant GNN

Building block 2

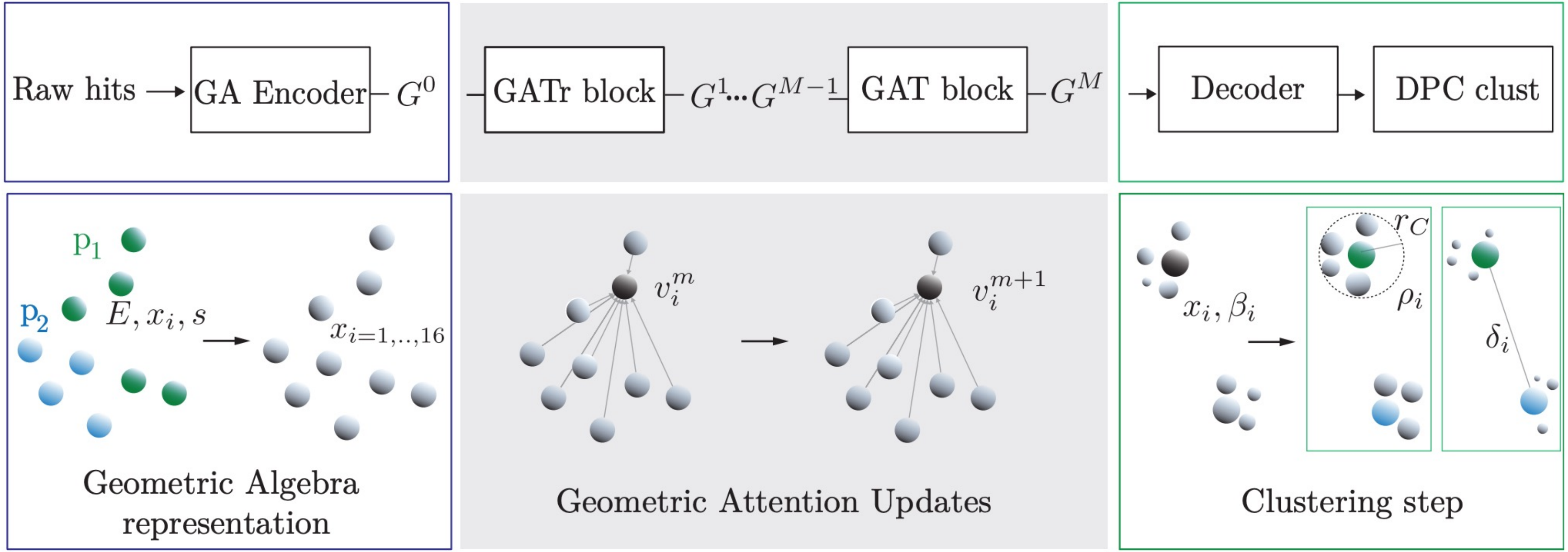
A symmetry group $\mathfrak{G}$
With a group representation

$$\rho : \mathfrak{G} \to GL(\mathcal{X}(\Omega))$$

Equivariant if

$$f(\rho(\mathfrak{g})x) = \rho(\mathfrak{g})f(x)$$

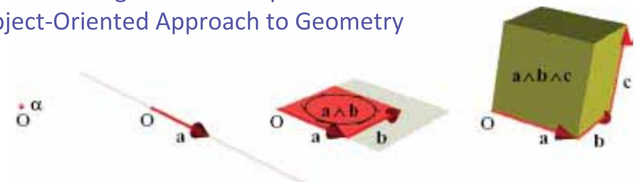
Geometric Algebra Transformer

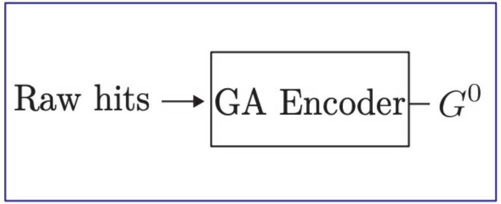


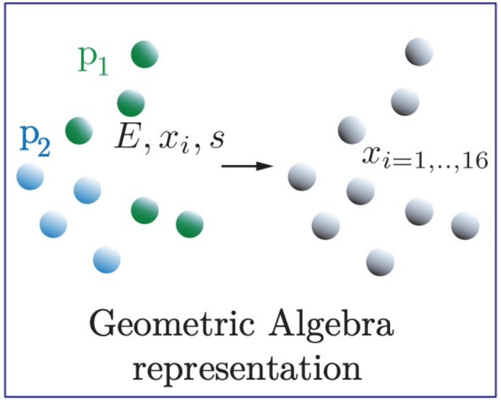
GATr [1]

Multivectors

Geometric Algebra for Computer Science : An Object-Oriented Approach to Geometry



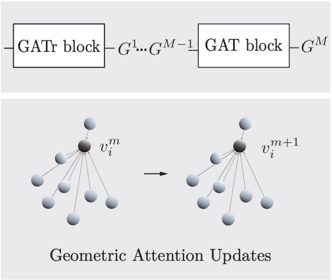


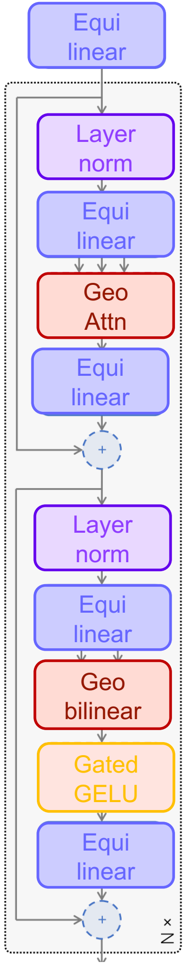


- Representation of multiple geometric objects of different grades
- Outer product a ∧ b (oriented span of a and b)
- Projective algebra (adds displacement from the origin)
- 3D object representation:
  - Planes are represented as vectors (normal + displacement)
  - Lines are intersection of two planes (bivectors)
  - Points are intersections of three planes (trivectors)
- Multivectors:
  - $x = x_s + x_1 e_1 + x_2 e_2 + x_3 e_3 + x_{12} e_1 e_2 + x_{13} e_1 e_3 + x_{23} e_2 e_3 + x_{123} e_1 e_2 e_3$, x_i real coefficients
  - Built with the geometric product (metric)

Object / operator	Scalar	Vector		Bivector		Trivector		PS
	1	e_0	e_i	e_{0i}	e_{ij}	e_{0ij}	e_{123}	e_{0123}
Scalar $\lambda \in \mathbb{R}$	λ	0	0	0	0	0	0	0
Plane w/ normal $n \in \mathbb{R}^3$, origin shift $d \in \mathbb{R}$	0	d	n	0	0	0	0	0
Line w/ direction $n \in \mathbb{R}^3$, orthogonal shift $s \in \mathbb{R}^3$	0	0	0	s	n	0	0	0
Point $p \in \mathbb{R}^3$	0	0	0	0	0	p	1	0
Pseudoscalar $\mu \in \mathbb{R}$	0	0	0	0	0	0	0	μ
Reflection through plane w/ normal $n \in \mathbb{R}^3$, origin shift $d \in \mathbb{R}$	0	d	n	0	0	0	0	0
Translation $t \in \mathbb{R}^3$	1	0	0	$\frac{1}{2}t$	0	0	0	0
Rotation expressed as quaternion $q \in \mathbb{R}^4$	q_0	0	0	0	q_i	0	0	0
Point reflection through $p \in \mathbb{R}^3$	0	0	0	0	0	p	1	0

Geometric Algebra Transformer





Equi linear

K-blade projections, no grade mixing

$$\phi(x) = \sum_{k=0}^{d+1} w_k \langle x \rangle_k + \sum_{k=0}^{d} v_k e_0 \langle x \rangle_k$$

Geo bilinear

Geometric product x,y --> xy= ⟨x,y⟩+x∧y --> grade mixing
Join product x,y --> (x* ∧ y*)*

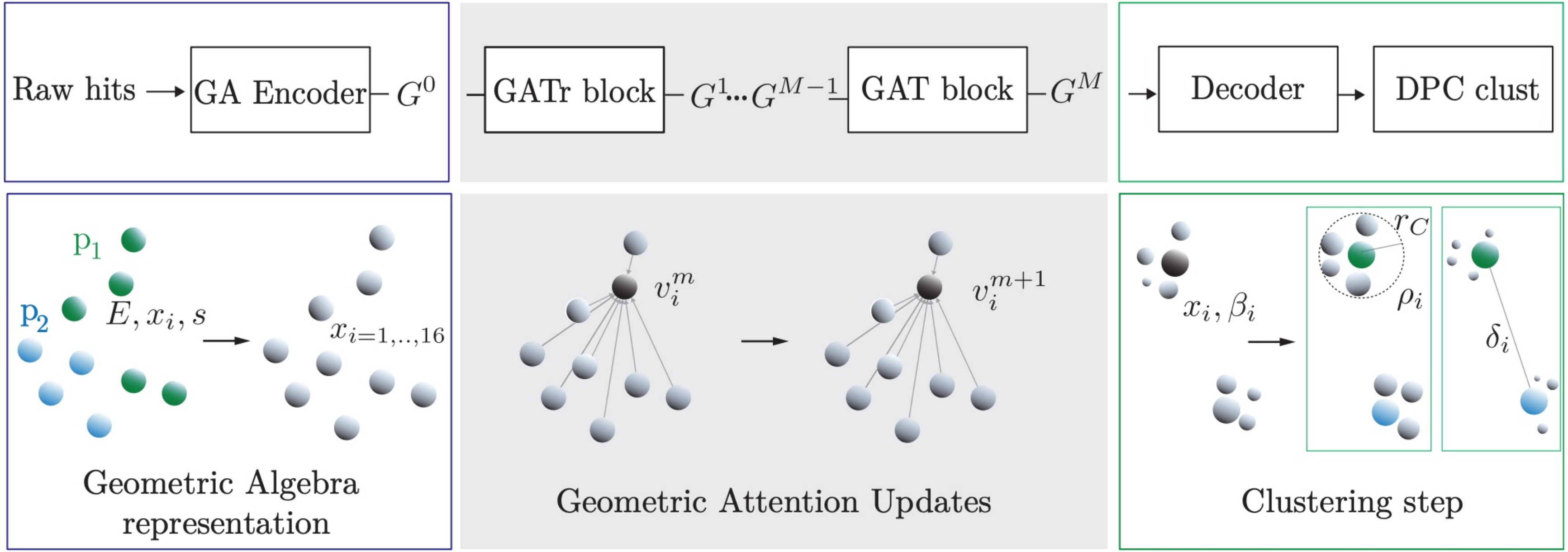
$$\mathrm{Geometric}(x, y; z) = \mathrm{Concatenate}_{\mathrm{channels}}\big(xy,\ \mathrm{EquiJoin}(x, y; z)\big)$$

Geo Attn

Attention mechanism as described previously but with inner product of the algebra

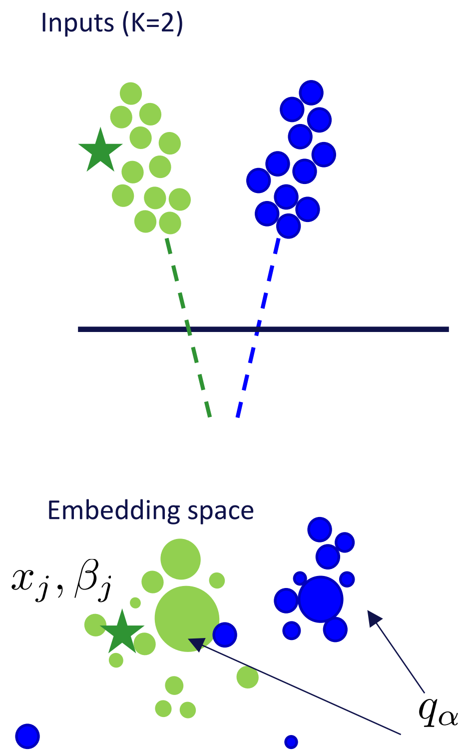
$$\mathrm{Attention}(q, k, v)_{i'c'} = \sum_i \mathrm{Softmax}_i \left(\frac{\sum_c \langle q_{i'c}, k_{ic} \rangle}{\sqrt{8 n_c}} \right) v_{ic'}$$

Geometric Algebra Transformer



GATr [1]

Loss: object condensation



$$\check{V}_k(x) = \|x - x_\alpha\|^2 q_{\alpha k},$$

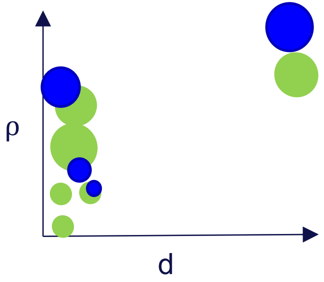
$$\hat{V}_k(x) = \max\left(0,\ \underline{1 - \|x - x_\alpha\|}\right) q_{\alpha k}$$

$$L_V = \frac{1}{N} \sum_{j=1}^{N} q_j \sum_{k=1}^{K} \left(M_{jk} \check{V}_k(x_j) + (1 - M_{jk}) \hat{V}_k(x_j) \right)$$

$$q_{\alpha k} = \max_i q_i M_{ik},$$

Object condensation @Jan K.

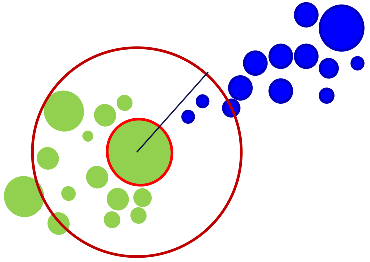
Clustering in embedding space



Two methods

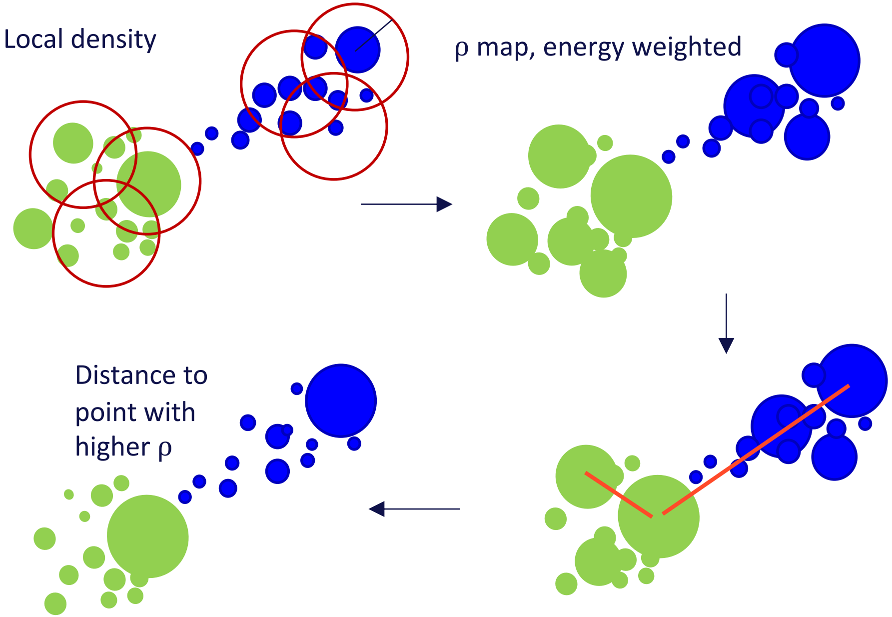
Beta based clustering [1]

x_j, β_j



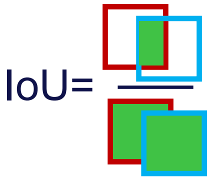
- Two parameters, β_{min} , distance threshold
- Depends on good training of β distribution

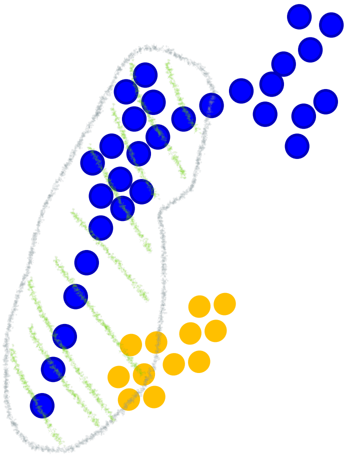
Density Peak clustering [2]



Matching

- We use a **Hungarian matching** to match a reco particle to a target particle with the intersection over union cost
- Builds a cost matrix
- One-to-one constraint
- Cost minimization



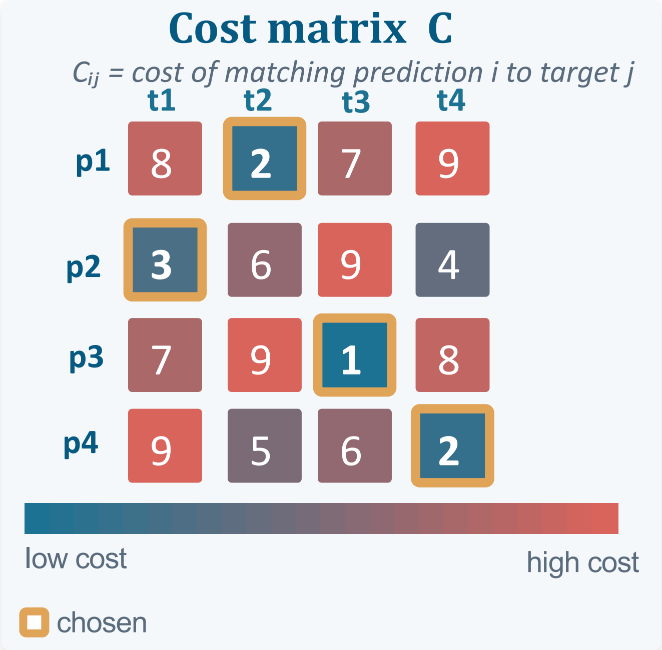


Matching

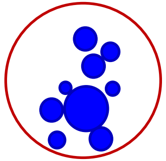
- We use a **Hungarian matching** to match a reco particle to a target particle with the intersection over union cost
- Builds a cost matrix
- One-to-one constraint
- Cost minimization

$$\min_{X\in\{0,1\}^{N\times N}} \sum_{i=1}^{N}\sum_{j=1}^{N} C_{ij}\, X_{ij} \quad \text{s.t.}$$

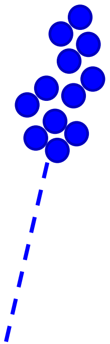
$$\sum_{j} X_{ij} = 1 \ \ \forall i, \qquad \sum_{i} X_{ij} = 1 \ \ \forall j.$$



Regression & Classification



Clustering
proposed PF

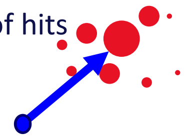


x_i, E_i, χ, ID

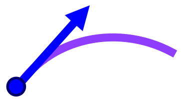
Per PF candidate property determination

$\vec{p}/|p|$

Neutral
Direction of energy-weighted average of hits



Charged
Direction from track state at IP



Charged
Energy measurement based on tracking

Energy

CLUSTER EMBEDDINGS (GATr)
HIGH-LEVEL FEATURES

$E_{cluster}, \frac{E_{ECAL}}{E_{cluster}}, \frac{E_{HCAL}}{E_{cluster}}$, dispersion

# of hits

# of tracks, $p_{track}, \chi^2_{track}$

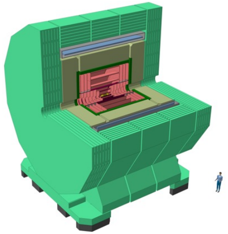
# of muon hits, E_{muon}

MLP Energy — neutral energy

PID

MLP PID — CH, e, μ, NH, γ

Key4hep

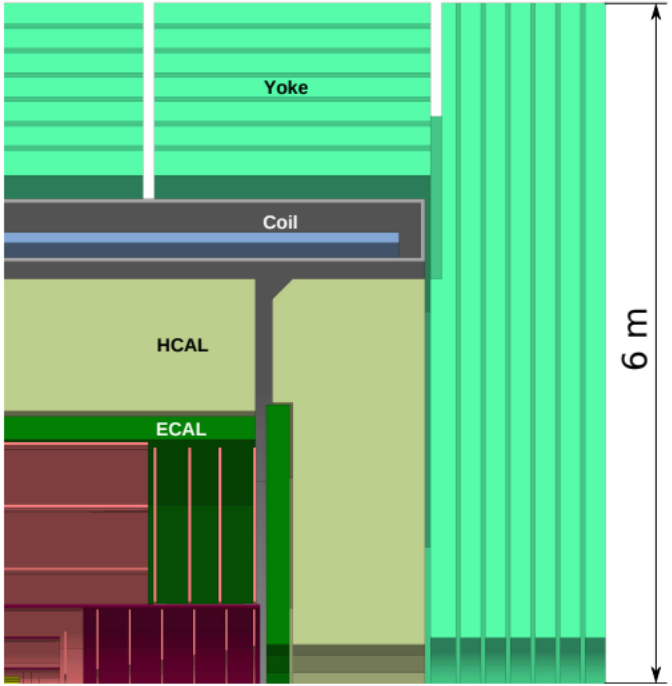


CLD a classic detector

Full simulation

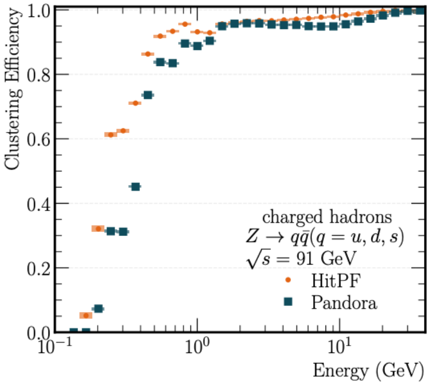
- ❑ Implemented in DD4hep
- ❑ Full simulation in Key4hep
- ❑ Comparison againts Pandora tunned for CLD (CLD_o2_v05)

- ❑ $Z \rightarrow q\overline{q}, q = (u, d, s)$ 91 GeV
- ❑ Jet like events:
  - ❑ Close particles
  - ❑ High particle multiplicity

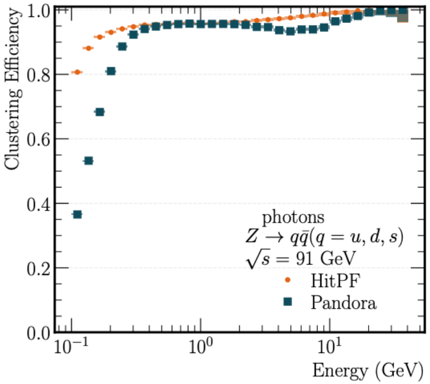


Clustering performance

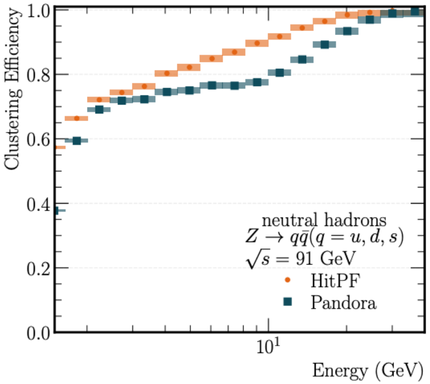
Charged Hadron



Photons



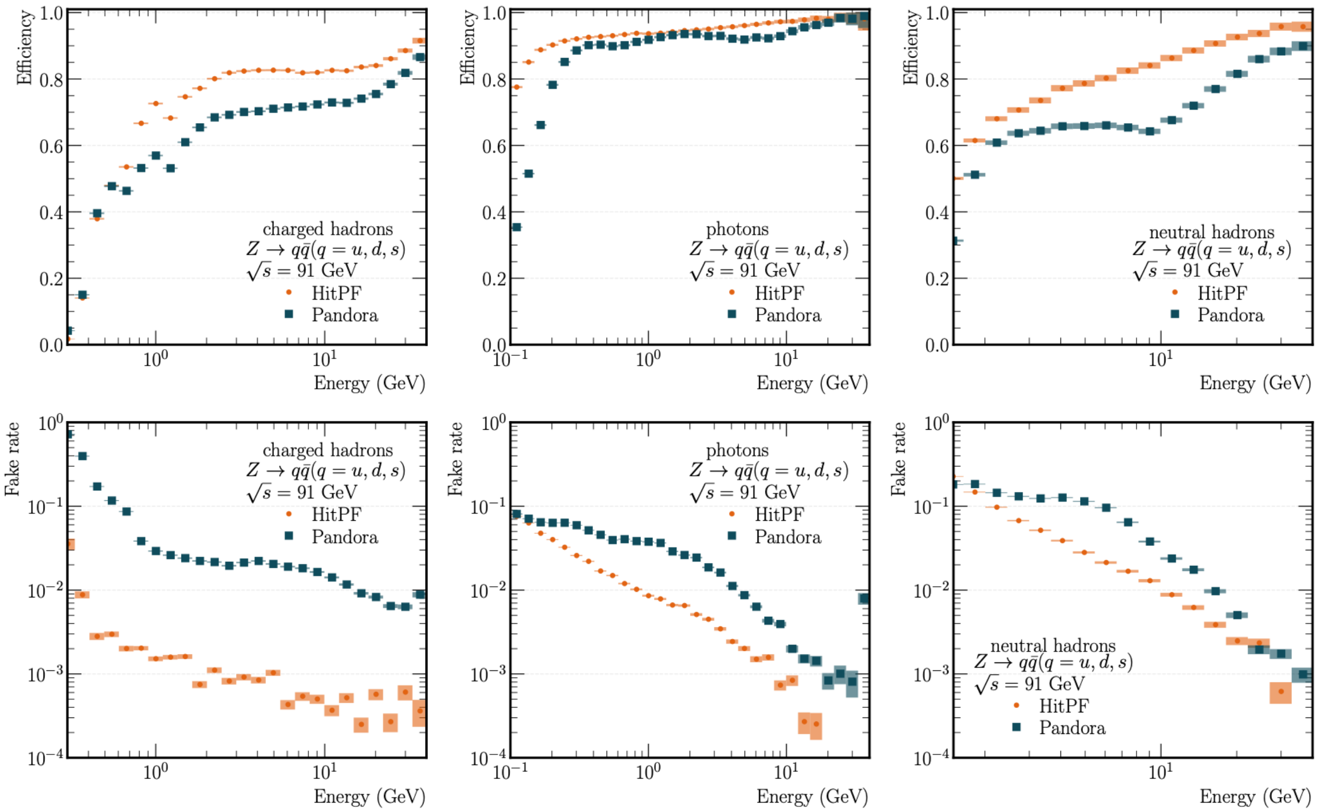
Neutral Hadrons



- It is efficient if for a given target particle it finds a reco particle (no PID), purely a measure of clustering performance
- vs the energy of the deposited hits (reco)
- Higher efficiency for all momentum range, substantial at energies below 1 GeV

PID: Efficiency and fake rate

Charged Hadron | Photons | Neutral Hadrons



CH:

- Up to 20% higher eff.
- Fake rate $\mathcal{O}(2)$ smaller

→Eff. limited by track reco.

Photons:

- ~5% higher eff. at high E
- Improvement by up to factor 2 for low E
- Fake rate up to ~ $\mathcal{O}(1)$

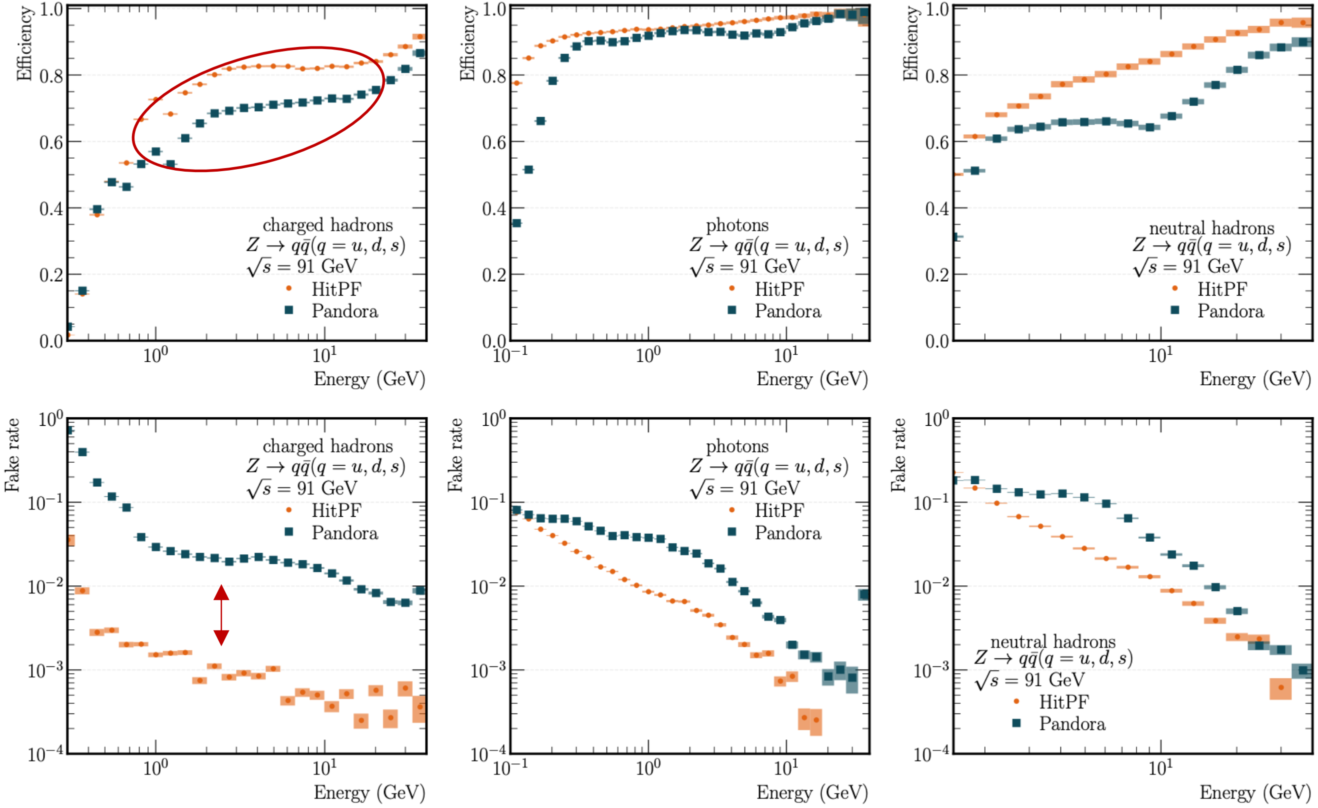
NH:

- Up to 20% higher eff. with Fake rate reduced by almost factor 5 in this region

→ Better clustering (no fake clusters from splitting)

PID: Efficiency and fake rate

Charged Hadron | Photons | Neutral Hadrons



CH:

- **Up to 20% higher eff.**
- **Fake rate $\mathcal{O}(2)$ smaller**

→ Eff. limited by track reco.

Photons:

- ~5% higher eff. at high E
- Improvement by up to factor 2 for low E
- Fake rate up to ~ $\mathcal{O}(1)$

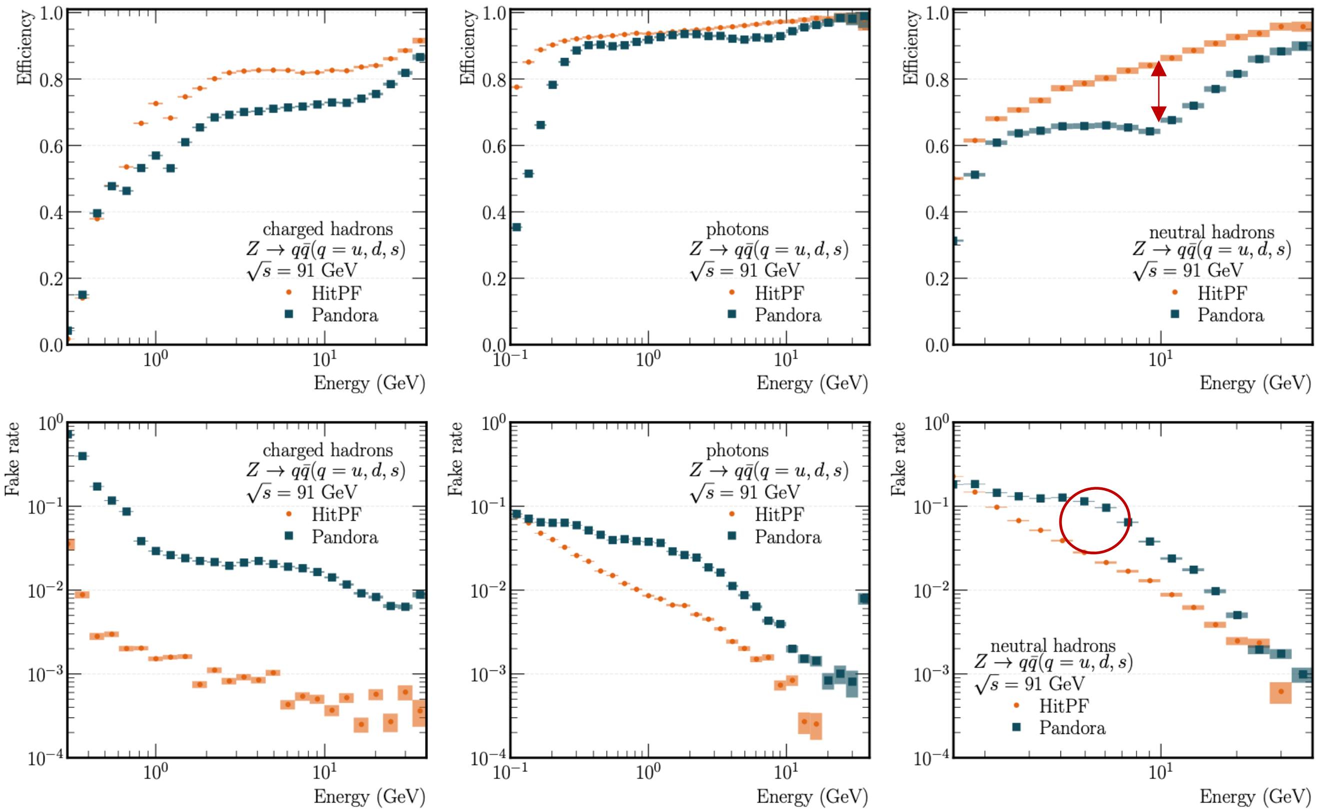
NH:

- Up to 20% higher eff. with Fake rate reduced by almost factor 5 in this region

→ Better clustering (no fake clusters from splitting)

PID: Efficiency and fake rate

Charged Hadron | Photons | Neutral Hadrons



CH:

- Up to 20% higher eff.
- Fake rate $\mathcal{O}(2)$ smaller

→Eff. limited by track reco.

Photons:

- ~5% higher eff. at high E
- Improvement by up to factor 2 for low E
- Fake rate up to ~ $\mathcal{O}(1)$

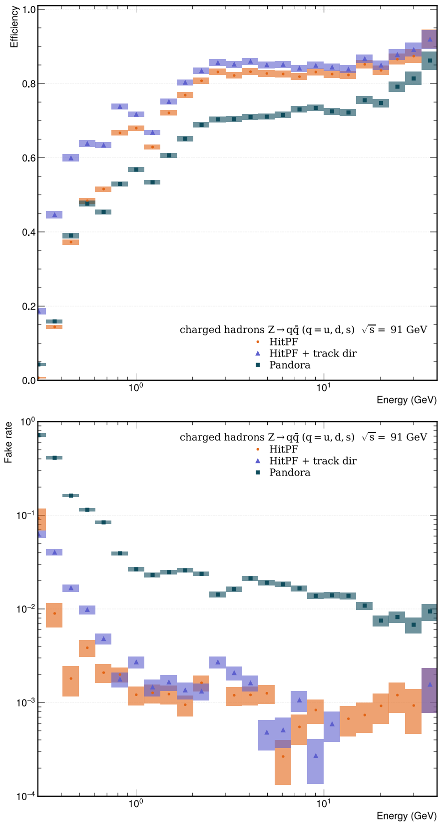
NH:

- **Up to 20% higher eff. with Fake rate reduced by almost factor 5 in this region**

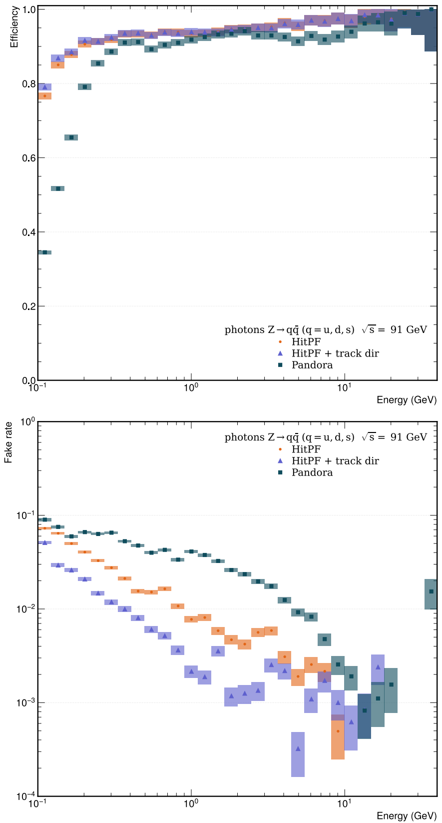
→ Better clustering (no fake clusters from splitting)

PID: Efficiency and fake rate

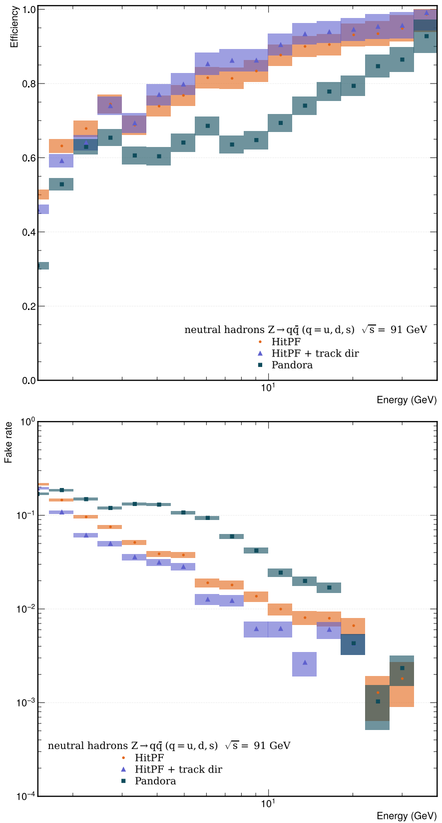
Charged Hadron



Photons



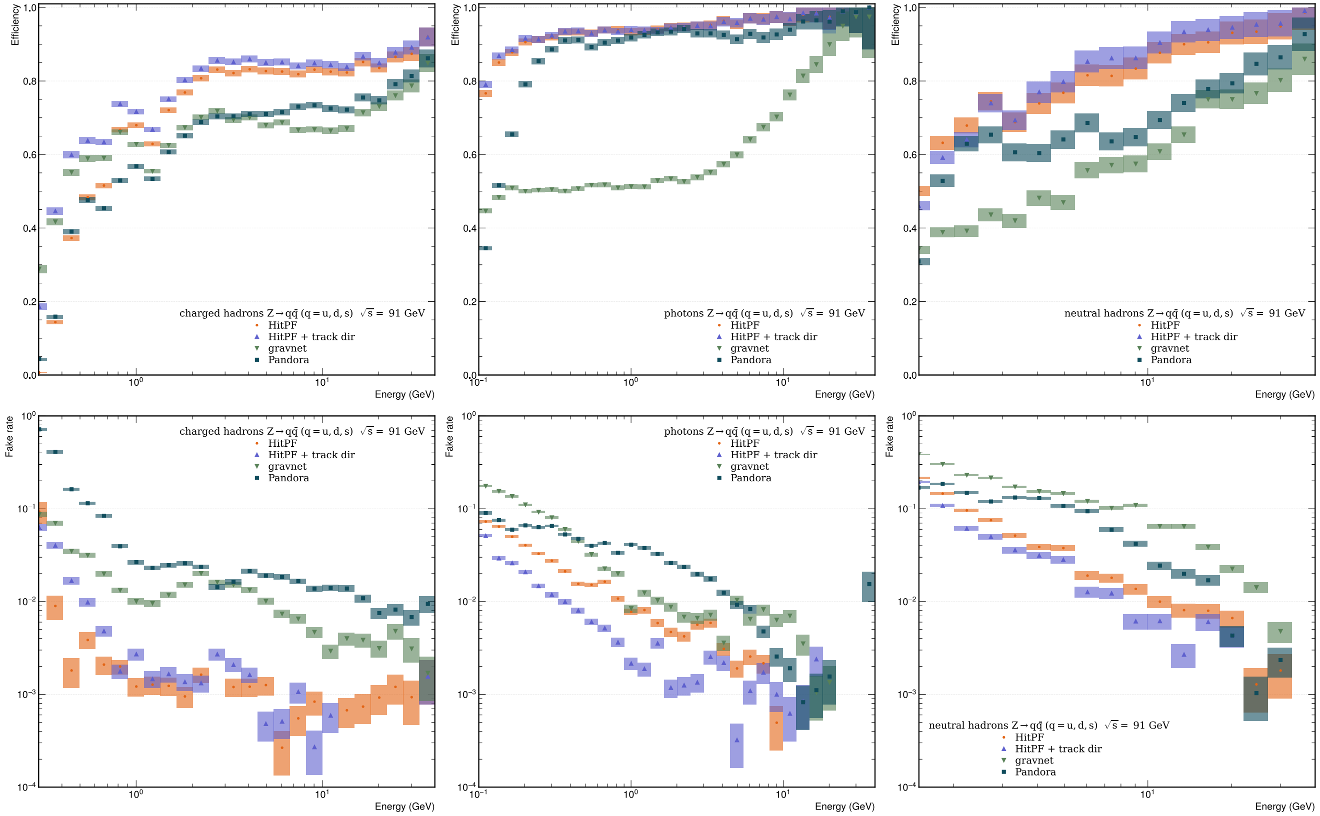
Neutral Hadrons



- ❑ Track direction at calo included as a vector in the GA.
- ❑ Improved efficiency and lower fake rate

PID: Efficiency and fake rate

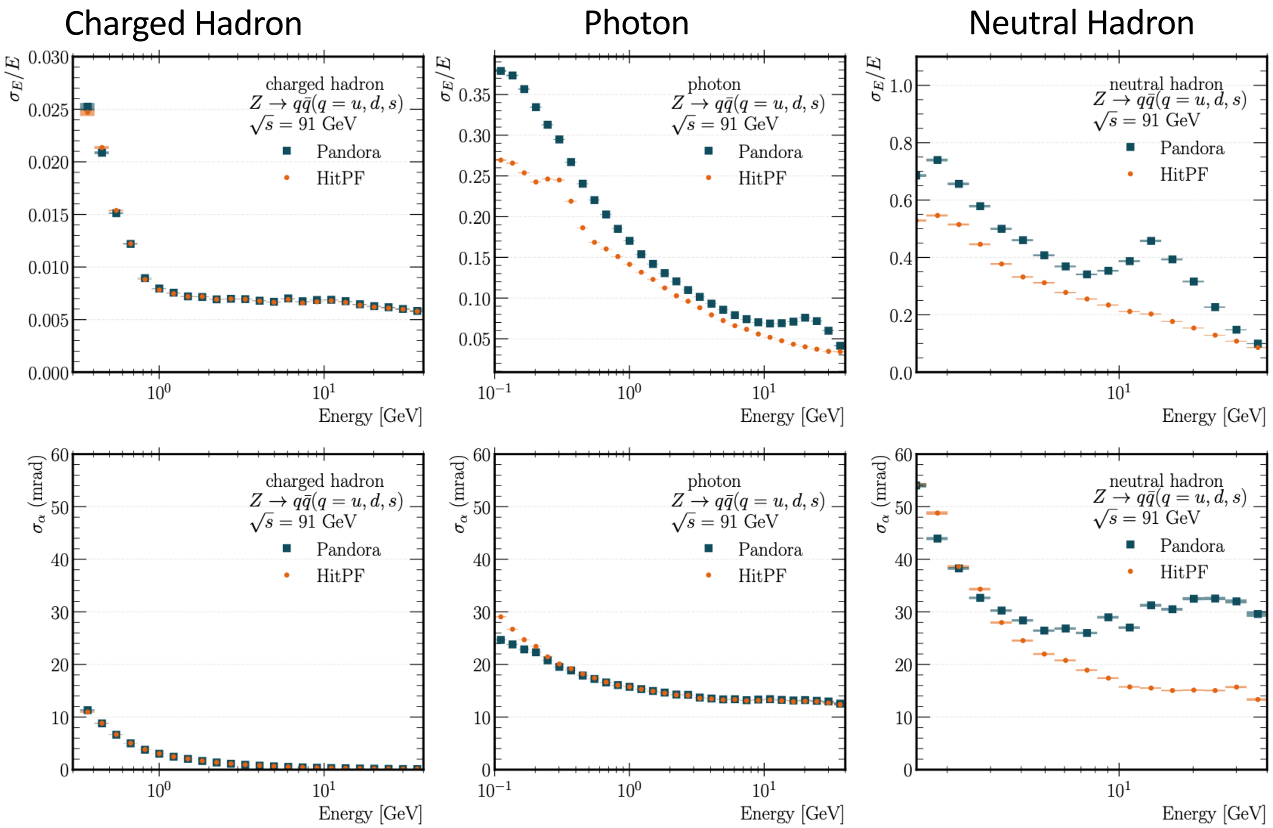
Charged Hadron | Photons | Neutral Hadrons



- ❑ Track direction at calo included as a vector in the GA.
- ❑ Improved efficiency and lower fake rate
- ❑ More ‘classic’ GNN architecture (Gravnet) does not achieve the same performance

Energy and angular resolution

Energy Resolution: half-width of 16th-84th percentile normalized to median
Angular resolution: 68th percentile of the opening angle between rec and true

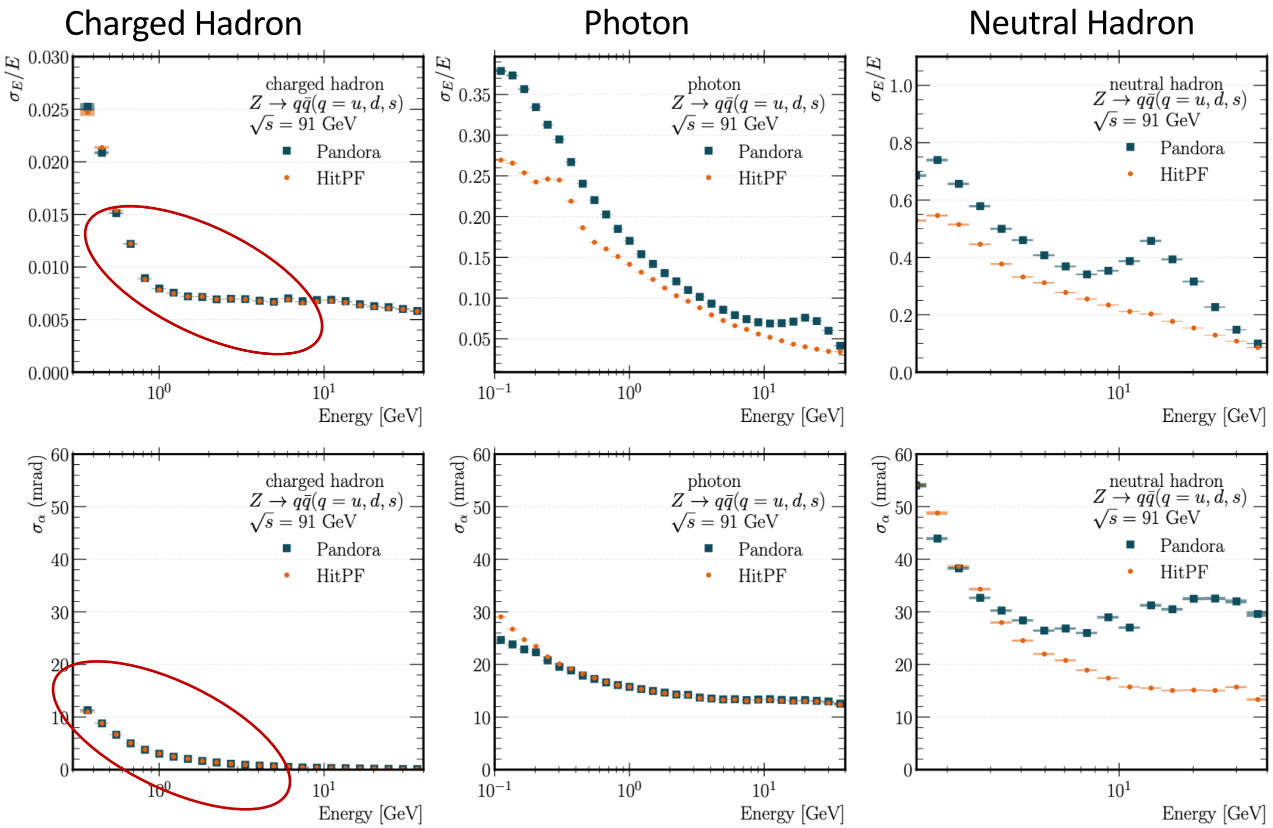


CH:
- Comparable by construction
- solely tracker information used

NH:
- Largest gain in energy/ angular resolution
- Less shower merging, better hit capture

Energy and angular resolution

Energy Resolution: half-width of 16th-84th percentile normalized to median
Angular resolution: 68th percentile of the opening angle between rec and true



CH:

- **Comparable by construction**
- **solely tracker information used**

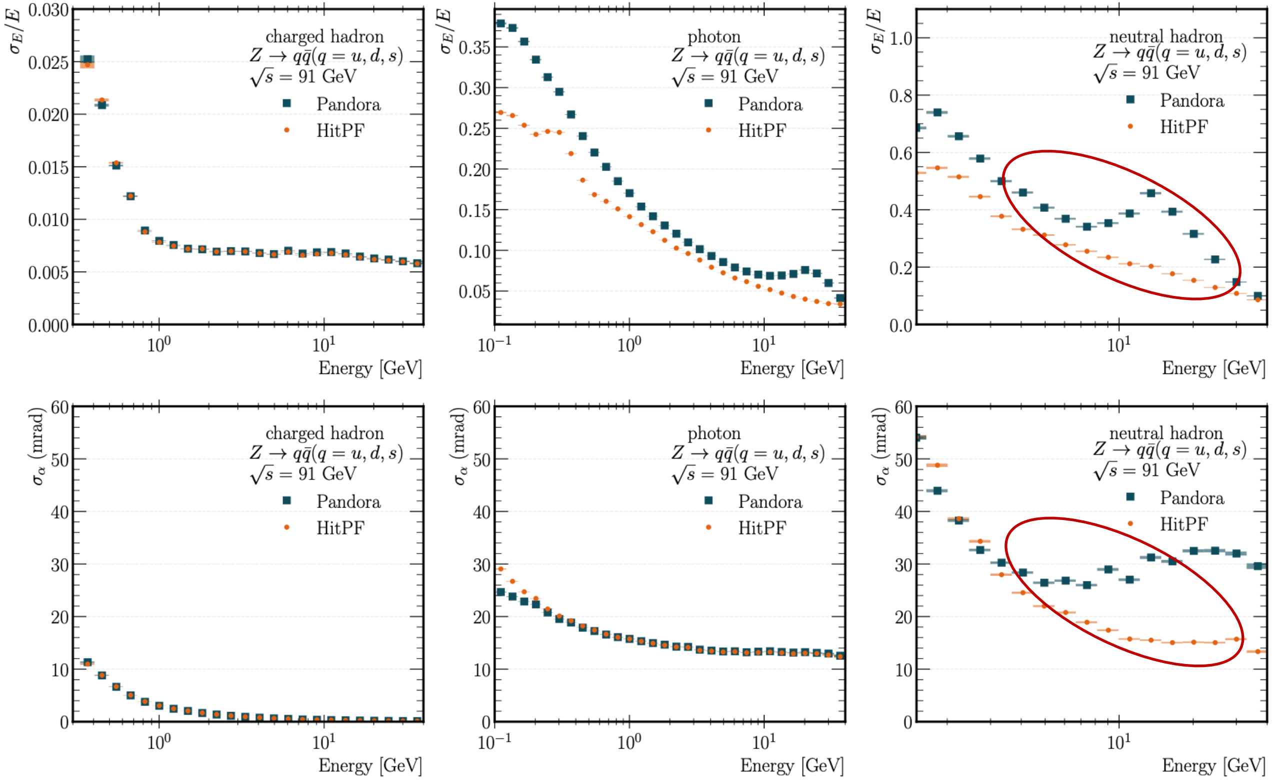
NH:

- Largest gain in energy/ angular resolution
- Less shower merging, better hit capture

Energy and angular resolution

Energy Resolution: half-width of 16th-84th percentile normalized to median
Angular resolution: 68th percentile of the opening angle between rec and true

Charged Hadron | Photon | Neutral Hadron



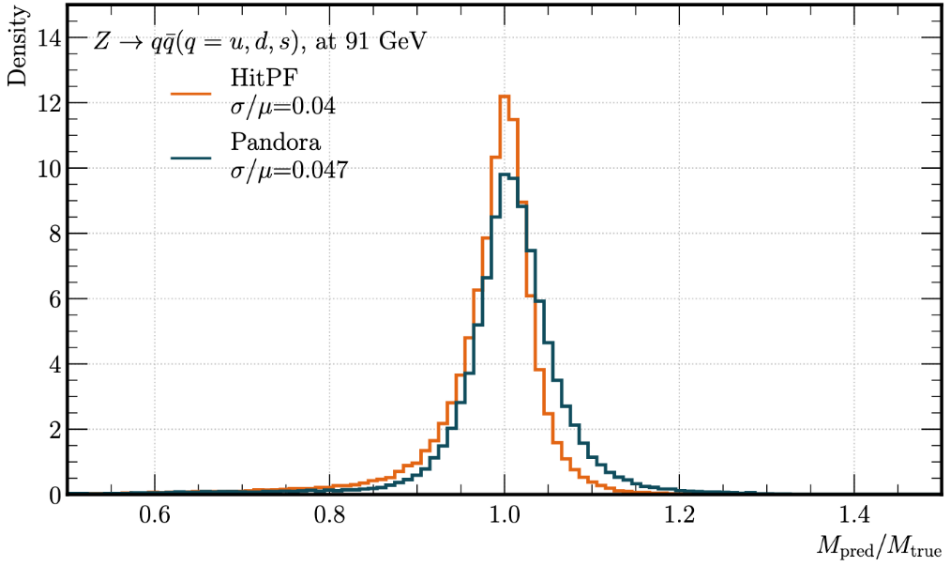
CH:
- Comparable by construction
- solely tracker information used

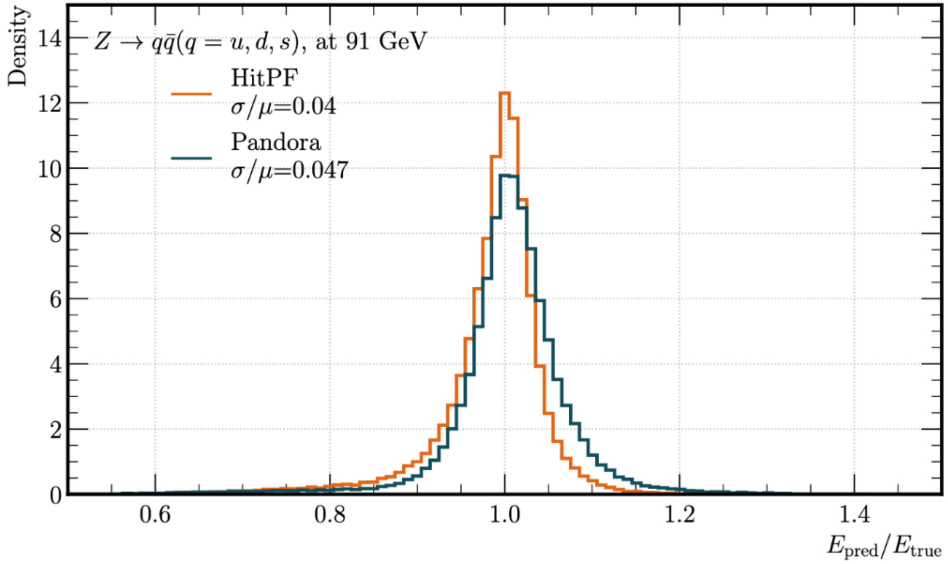
NH:
- **Largest gain in energy/ angular resolution**
- **Less shower merging, better hit capture**

Visible mass and energy

- ❑ Good complementary metric to per-particle metrics
- ❑ Improved mass resolution due to:
  - ❑ Higher efficiency in low energy neutrals
  - ❑ Improved energy resolution
  - ❑ Better track cluster-linking

22% relative improvement of HitPF compared to Pandora

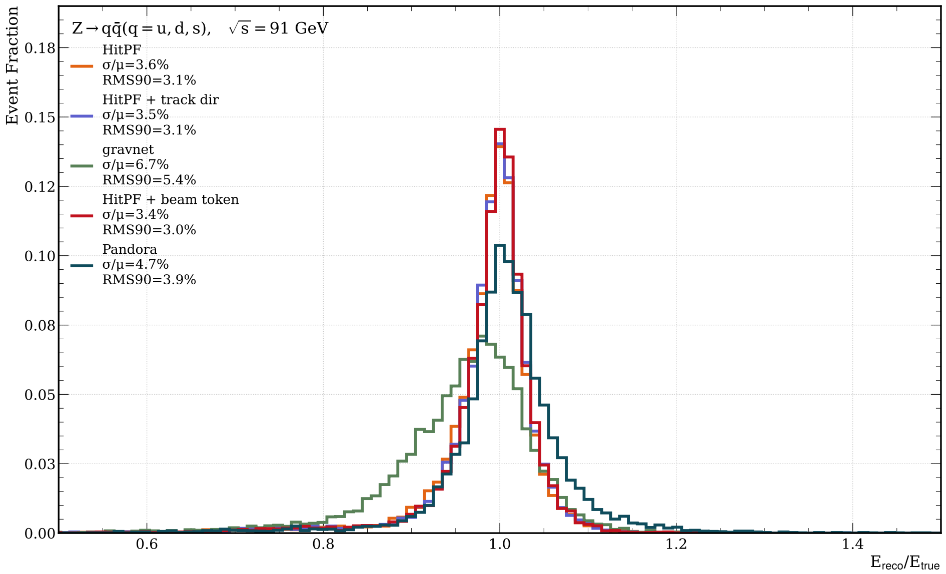


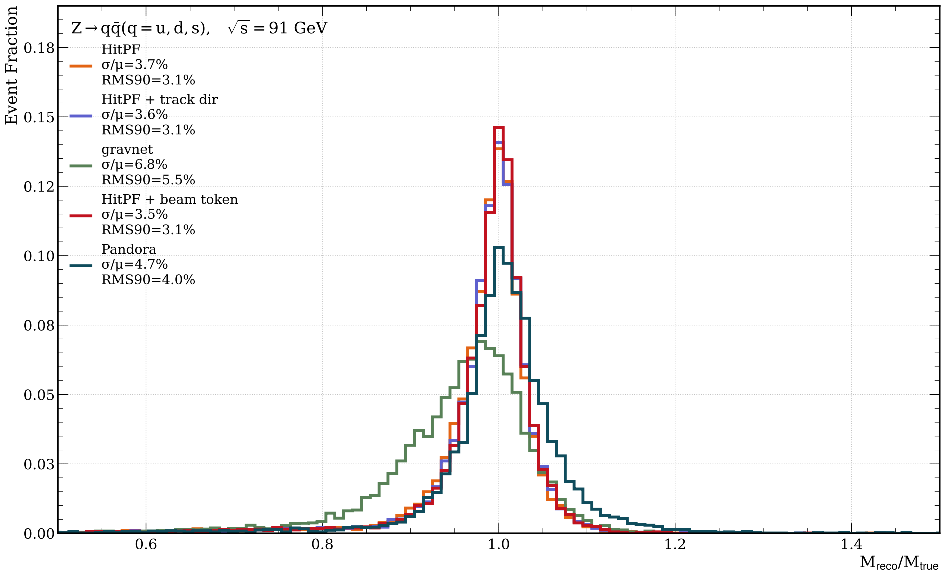


Visible mass and energy

- Good complementary metric to per-particle metrics
- Improved mass resolution due to:
    - Higher efficiency in low energy neutrals
    - Improved energy resolution
    - Better track cluster-linking

26% relative improvement of HitPF compared to Pandora

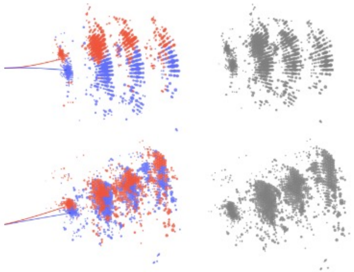


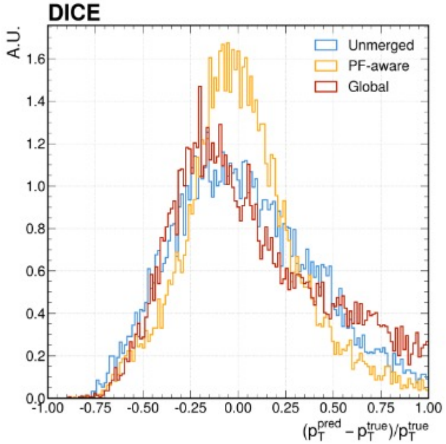


The impact of the truth defintion

Detector aware-target

- ❑ Define a shower merging algorithm based on a resolvability score and a connection score
  - Resolvability: what fraction of shower is not shared
  - Connection: show strongly shower B overlaps with A
- ❑ Graph with sim showers as nodes and possible absorptions as directed edges
- ❑ Algorithm to determine merging steps



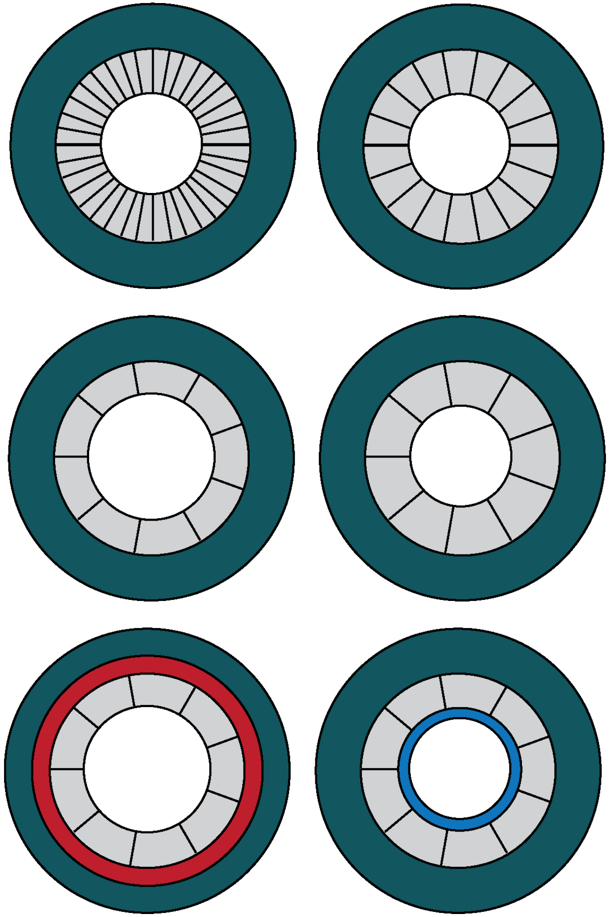


Transverse momentum response in tau jet sample for models trained with different target definition

@ Alessandro Brusamolino, Katherina Schauble, Jan Kieseler

Karlsruher Institut für Technologie

What does this give access to?



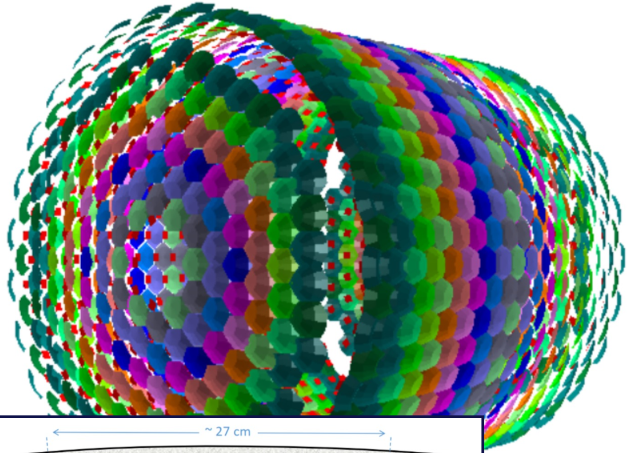
Full simulation physics studies

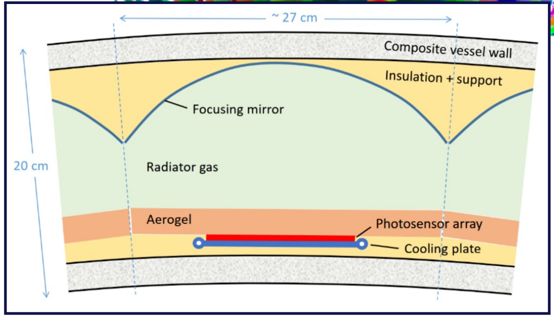
- ❑ Impact of beam background
- ❑ Effect of timing information in clustering
- ❑ Impact of low momentum tracks reco on PF
- ❑ Impact of acceptance
- ❑ Tau reconstruction
- ❑ Physics validation
  - Retrain jet taggers with HitPF input
  - Deploy for physics analysis in CLD

Impact of the ARC's material

Vincent Richers

- ARC adds charged-hadron PID capability to CLD, but its integration places material upstream of the ECAL and compresses the tracker
- Compact RICH-based PID concept for charged-hadron separation (π/K/p) via Cherenkov radiation.
- Uses gas and aerogel radiators with spherical focusing mirrors and SiPM-based photon detection.
- Tracker is compressed to accommodate ARC: outer radius reduced from 2145 to 1900 mm. Studies have assesed impact on track resolution but not effect on PF

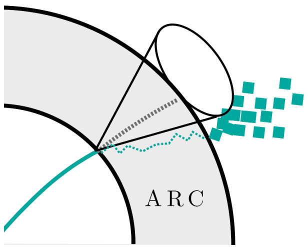


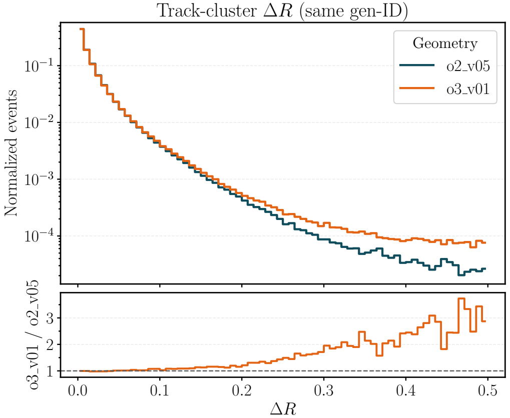


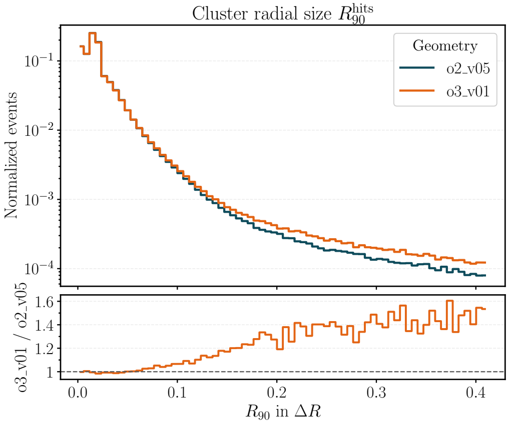
The ARC

Impact of the ARC's material

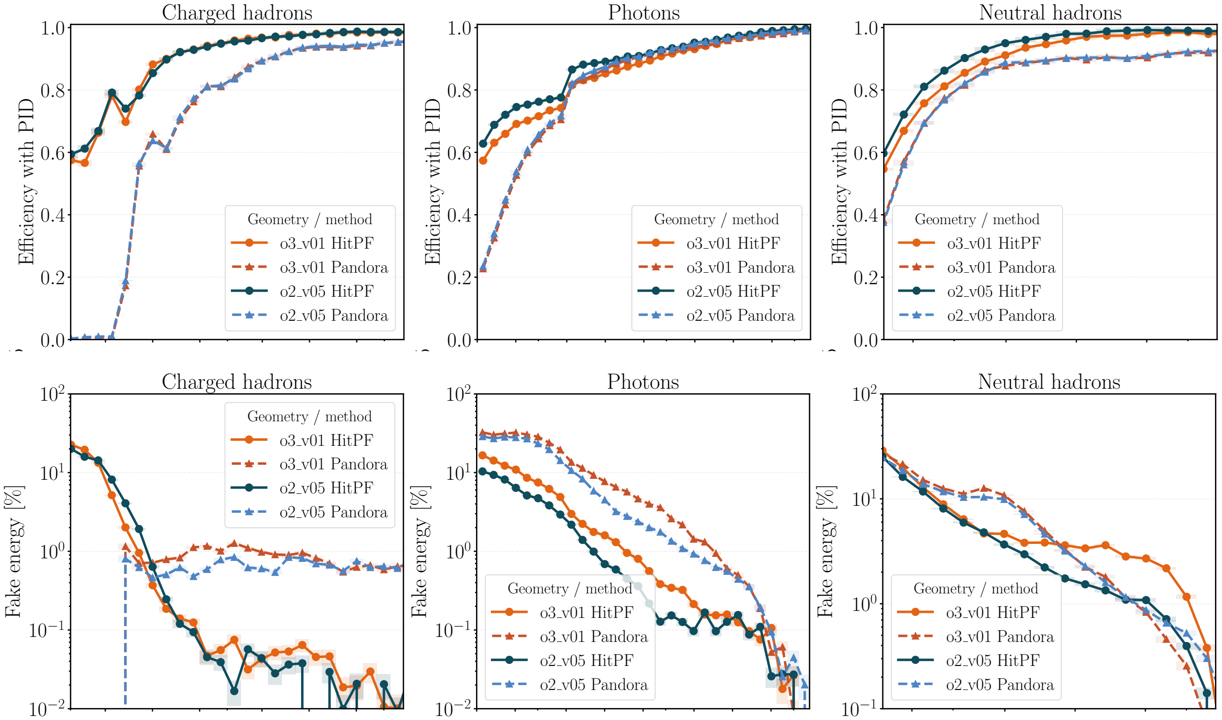
- Tracker is shrunk resulting in impact on momentum resolution [1] (track lever arm) (~15%) not necessarily a big impact on PF in itself
- But tracking-cluster linking might be impacted (multiple scattering)
- Tracking efficiency is also impacted
- R90 is impacted by earlier showering







Impact of the ARC's material



- ❑ Charged-particle reconstruction remains stable once tracking effects are controlled.
- ❑ The visible ARC impact is concentrated in photons and neutral hadrons.
- ❑ Photons and neutral hadrons show reduced PID efficiency, increased fakes, and broader response

TRACKING

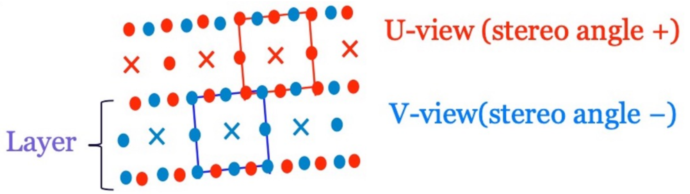
Gaseous trackers

L.Centeno Gas trackers

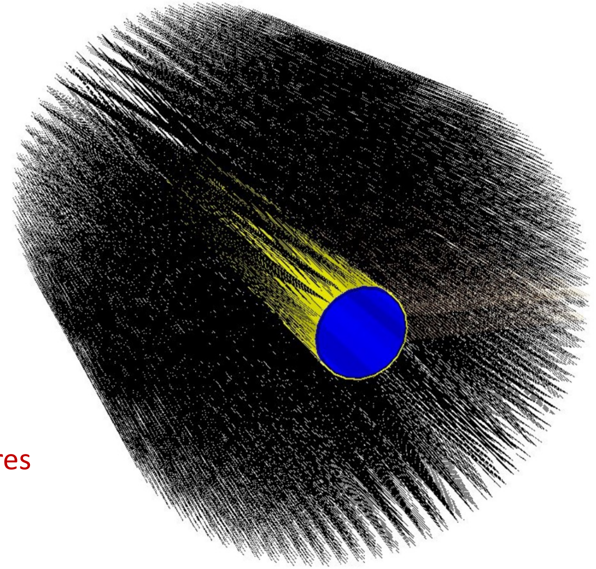
Why Gaseous Trackers?

Low material in tracking system
- ❑ Minimise contribution of multiple scattering (MS) to momentum resolution
- ❑ Favourable for Particle Flow

Continuous tracking volume
- ❑ Large number of hits beneficial for tracking combinatorics
- ❑ Enhanced sensitivity to displaced vertices
- ❑ PID capabilities via dN/dx (cluster counting)



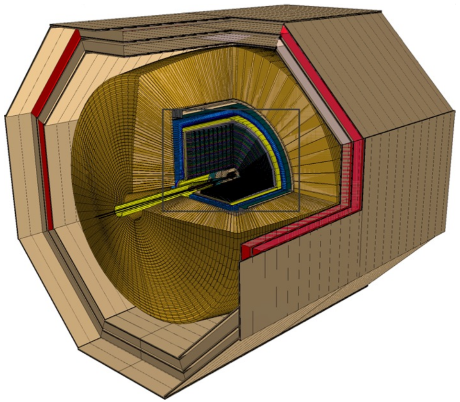
O(300k wires)
56 448 sense wires

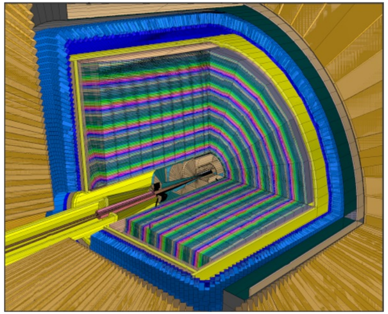


Gaseous Trackers in FCCee

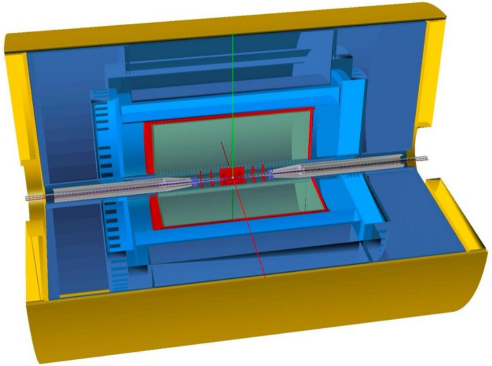
L.Centeno Gas trackers

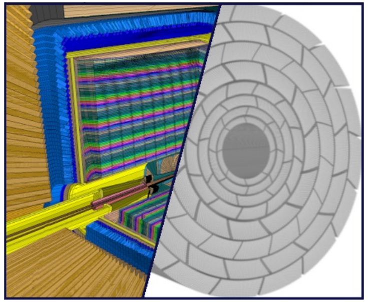
IDEA



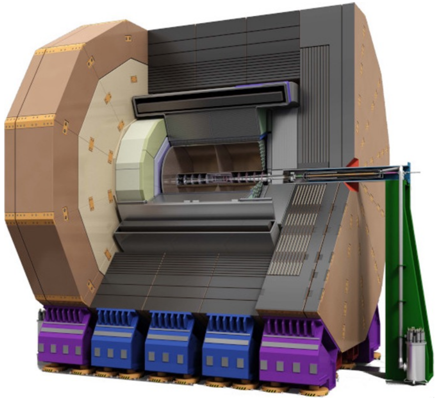


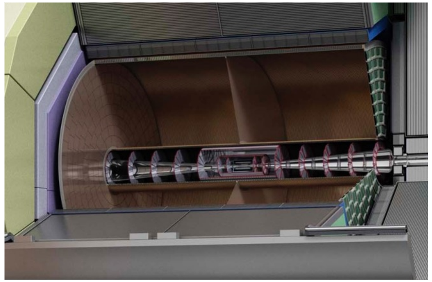
ALLEGRO





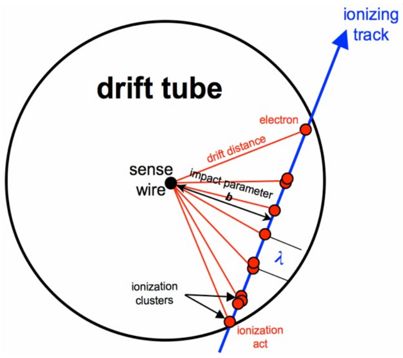
ILD

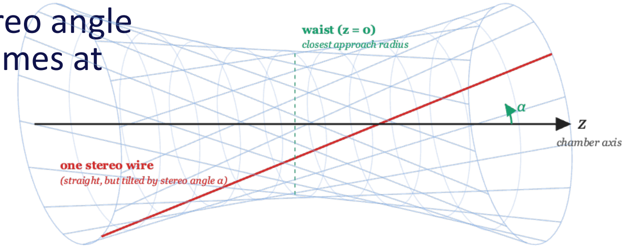




Drift Chamber

- Drift chambers measure a charged particle's path by timing how long ionization electrons take to drift through a gas to sense wires:
  - The particle ionizes the gas along its track, the liberated electrons drift in an applied electric field toward thin anode wires, and the measured drift time (× known drift velocity) gives **the distance of closest approach to each wire**
  - The **z coordinate** obtained from stereo angle and the difference in signal arrival times at the two ends of the wire
  - Hits are 'isochrones'





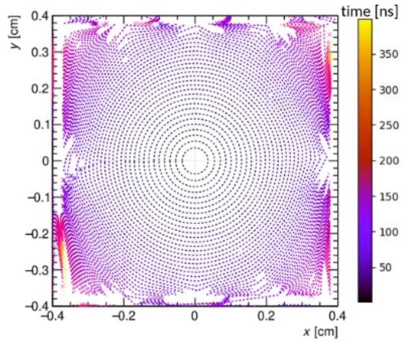
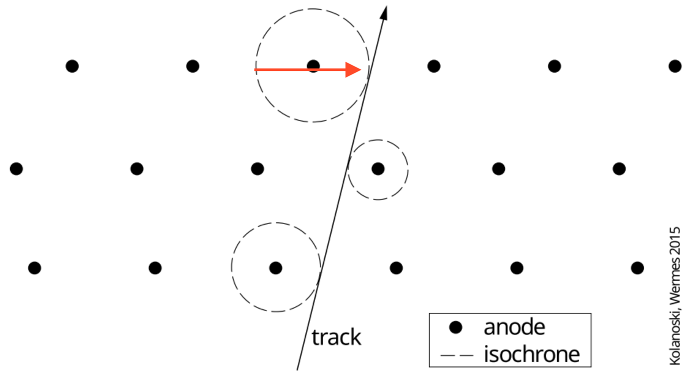
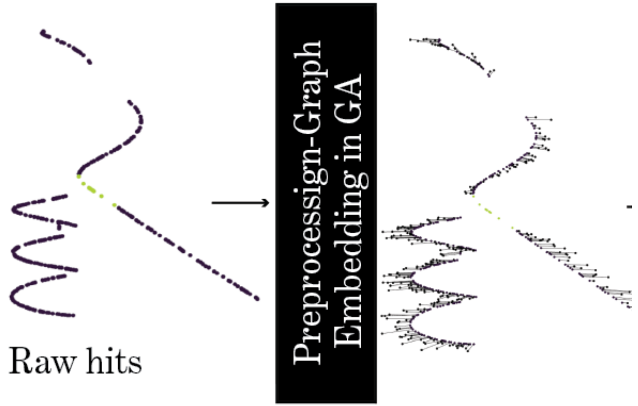


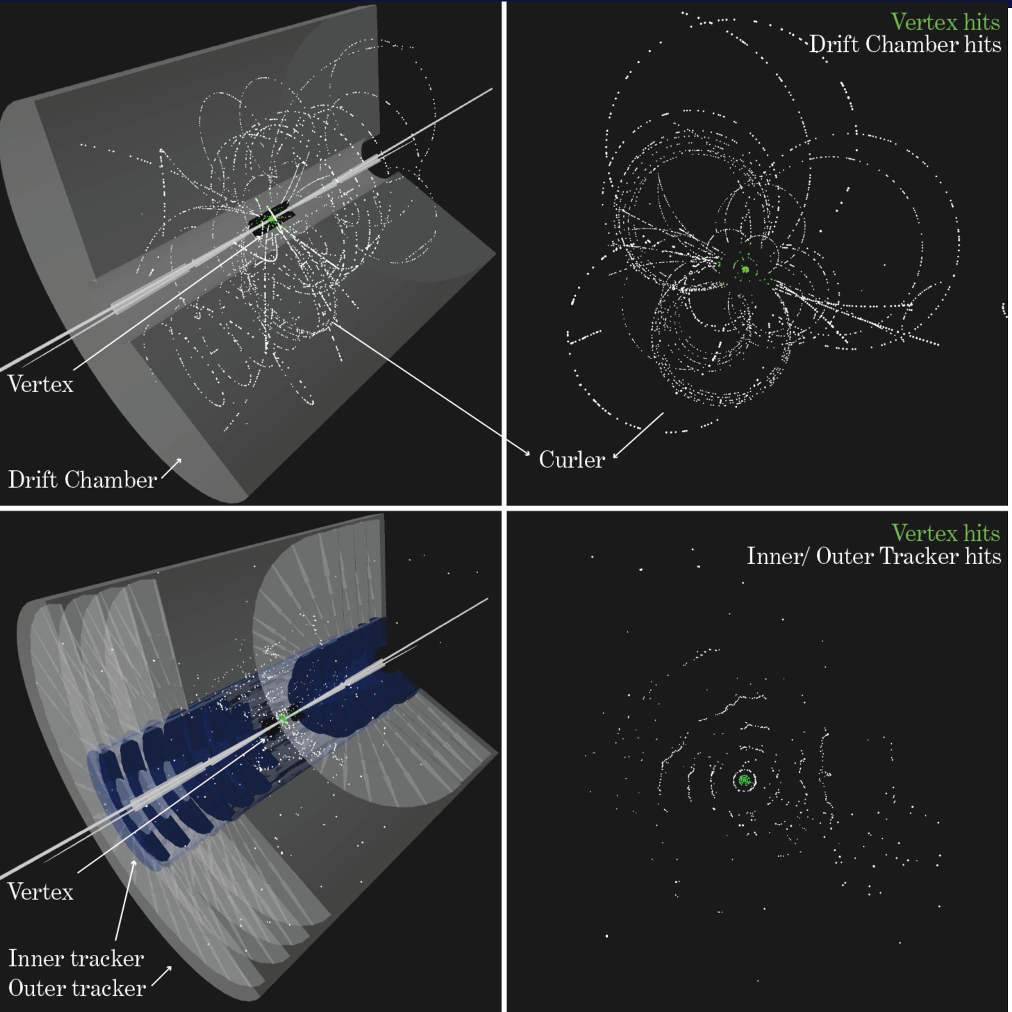
Figure 15 An example of the lines connecting points with the same drift time towards the sense wire ("isochrones").

The workflow: Input representation

- The geometric algebra representation allows for the left right hit ambiguity in pattern recognition
- Encoded as a vector (choose direction L to R)
- A new conformal algebra representation allows to input the circle (the ’natural’ representation)
- Ongoing work Marko Cechovic et al. FCC week 2026 Poster.







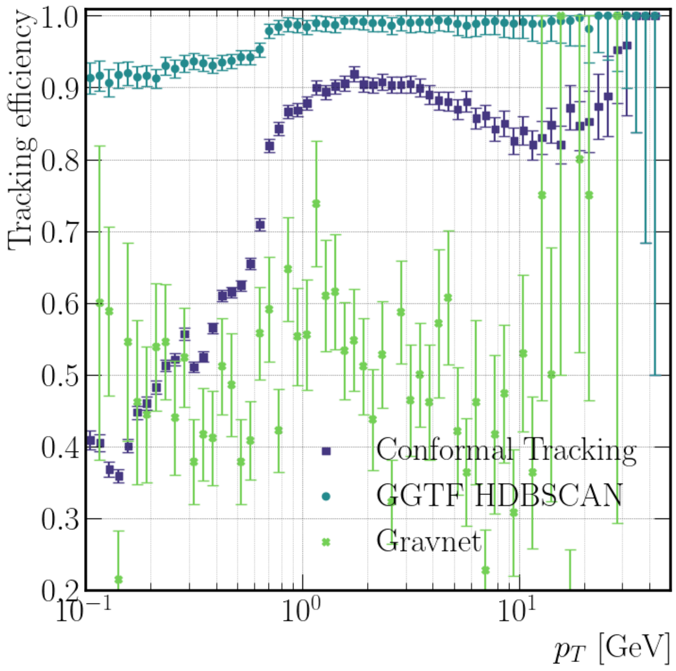
The workflow: Inputs

- Vertex detector hits
  Position
  Type

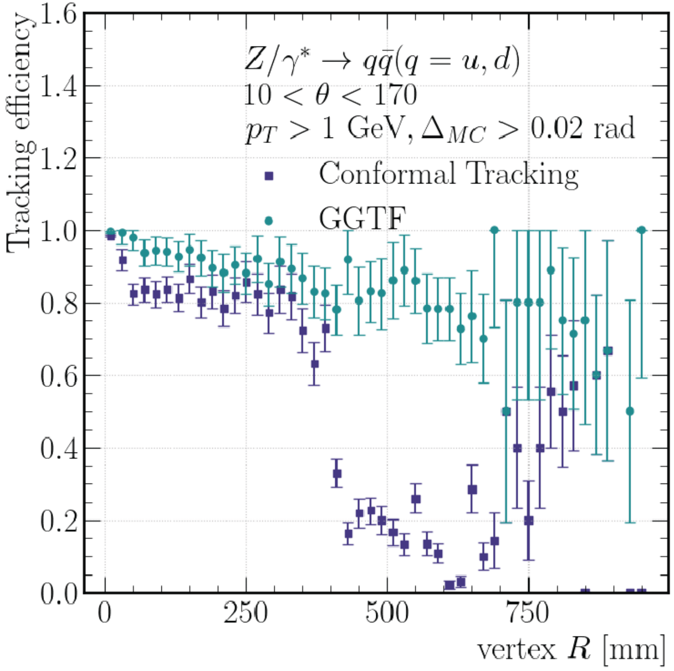
- Drift chamber hits
  Position (left right)
  Type

Geometric Graph Neural Network based track finding

Comparison to other methods



Comparison to GNN without geometric bias

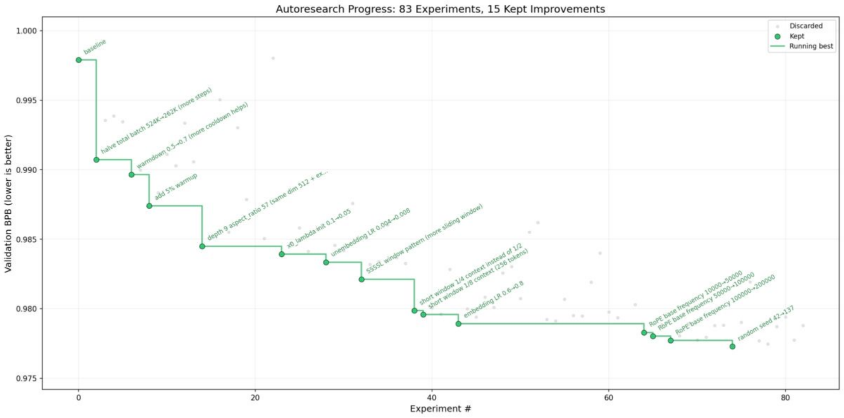


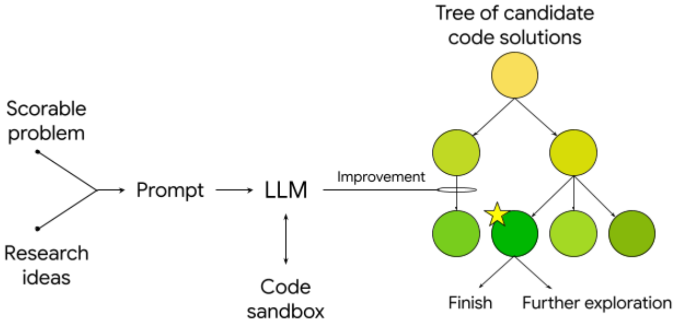
Improved efficiency for displaced tracks

The role of AI

AI is being used to further improve scientific codebases

- autoresearch (Andrej Karpathy):
  - Agents modify architecture, optimizer, hyperparameters to improve a loss function
  - Fast, fixed-budget trial-and-error loop experiments result in improved loss
- LLMs + Tree search for scientific software (Google, Nature):
  - Automatically write and iteratively improve scientific code that maximizes a quality metric
  - Applied to multiple fields (Covid prediction to geospatial segmentation) obtains improved results over baselines
  - A large share of the top-performing results came from recombination





Summary

- Developments in ML allow for improved data representation for hit-based particle flow
- Improved performance in PF compared to classical approaches across all metrics: efficiency in complex hadronic topologies, neutral hadron efficiency, fakes and visible mass resolution (improvements by 26%)
- New geometric models could futher improve performance, but development is needed as the properties of HEP are not the same as classic 3D datasets
- Similar methods apply to tracking and have been shown to improve over baseline approaches
- Can AI help? Ongoing work, stay tuned.

Thank you to the key4hep team and all FCC collaborators (Michele Selvaggi, Lena Herrmann, Gregor Krzmanc, Andrea de Vita, Brieuc Francois, Vincent Richers, Maria Cepeda, Olmo Arquero) Jan Kieseler, Alessandro Brusamolino, Katherina Schauble @ KIT & Justin Solomon @MIT

Swiss National Science Foundation

Massachusetts Institute of Technology

Thank you
for your attention.