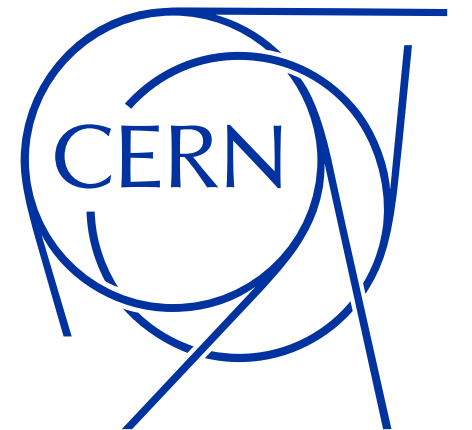


MLPF studies and ILD

Bohdan Dudar

bohdan.dudar@cern.ch

2026-09-14



A lot of progress in the MLPF community

The talks (I am aware of) over the last year

Title	Presenter	Event (links clickable)
AI for PF	Dolores Garcia	FCC Reco Workshop (upcoming 26 – 29 Oct)
k4PFHitML: Inference implementation of HitPF in key4hep	Lena Maria Herrmann	Fourth US FCC meeting
Time implementation for Particle Flow Reconstruction for the ILD Using Graph Neural Networks and Object Condensation	Arpad Schaeffer	Reconstruction Meeting for FCC
Particle Flow Reconstruction for the ILD Using Graph Neural Networks and Object Condensation	Cedric Trinh	ILD S&A Meeting
End-to-end event reconstruction for precision physics at future colliders	Dolores Garcia	ILD S&A Meeting
Machine Learning for event reconstruction at the FCCee	Dolores Garcia	ILD S&A Meeting
Towards Robust End-to-End Machine Learning Particle Flow	Katharina Schäuble	EP-IT Data Science Seminar
Machine Learning Based Particle Flow	Lena Maria Herrmann	ILD S&A Meeting
Particle Flow with Graph Neural Network	Tatsuki Murata	5th DRD Calo Collaboration Meeting
End-to-End Reconstruction with MaskFormers	Max Hart	ILD S&A Meeting
CMS ML-based Particle Flow reconstruction	Farouk Mokhtar	Reconstruction Meeting for FCC
Machine learning based particle flow	Dolores Garcia	EP-IT Data Science Seminar
Machine-Learning Particle Flow in CLICDet/FCC-ee CLD/CMS	Farouk Mokhtar	9th FCC Physics Workshop
Machine-Learning Particle Flow in FCC	Dolores Garcia	FCC-ee Workshop on Particle Flow Reconstruction
Machine-Learning Based Particle Flow in ILD	Taikan Suehara	FCC-ee Workshop on Particle Flow Reconstruction
End-to-end particle reconstruction for current and future colliders	Eilam Gross	ML4Jets2025

sorry if I missed someone!

Five(-ish) independent working groups

Dolores Garcia et al

Jan Kiesler et al

Farouk Mokhtar et al

Max Hart et al

Taikan Suehara et al

Title	Presenter	Event (links clickable)
		FCC Reco Workshop (upcoming 26 – 29 Oct)
AI for PF	Dolores Garcia	Fourth US FCC meeting
k4PFHitML: Inference implementation of HitPF in key4hep	Lena Maria Herrmann	Reconstruction Meeting for FCC
Time implementation for Particle Flow Reconstruction for the ILD Using Graph Neural Networks and Object Condensation	Arpad Schaeffer	ILD S&A Meeting
Particle Flow Reconstruction for the ILD Using Graph Neural Networks and Object Condensation	Cedric Trinh	ILD S&A Meeting
End-to-end event reconstruction for precision physics at future colliders	Dolores Garcia	ILD S&A Meeting
Machine Learning for event reconstruction at the FCCee	Dolores Garcia	EP-IT Data Science Seminar
Towards Robust End-to-End Machine Learning Particle Flow	Katharina Schäuble	ILD S&A Meeting
Machine Learning Based Particle Flow	Lena Maria Herrmann	5th DRD Calo Collaboration Meeting
Particle Flow with Graph Neural Network	Tatsuki Murata	ILD S&A Meeting
End-to-End Reconstruction with MaskFormers	Max Hart	Reconstruction Meeting for FCC
CMS ML-based Particle Flow reconstruction	Farouk Mokhtar	EP-IT Data Science Seminar
Machine learning based particle flow	Dolores Garcia	9th FCC Physics Workshop
Machine-Learning Particle Flow in CLICDet/FCC-ee CLD/CMS	Farouk Mokhtar	FCC-ee Workshop on Particle Flow Reconstruction
Machine-Learning Particle Flow in FCC	Dolores Garcia	FCC-ee Workshop on Particle Flow Reconstruction
Machine-Learning Based Particle Flow in ILD	Taikan Suehara	FCC-ee Workshop on Particle Flow Reconstruction
End-to-end particle reconstruction for current and future colliders	Eilam Gross	ML4Jets2025

sorry if I missed someone!



Different Groups, Different Ideas, Different Approaches

	Pipeline	Architectures	Detectors
Dolores Garcia et al	hits+tracks → PFOs	GATr	CLD, ILD
Jan Kiesler et al	hits+tracks → PFOs	GravNet, local GATr	CMS HGCal, DICE, CLD, ILD
Farouk Mokhtar et al	clusters + tracks → PFOs	GNN, Transformer	CMS, CLICdet, CLD
Max Hart et al	hits+tracks → PFOs	MaskFormers	CLD
Taikan Suehara et al	hits+tracks → PFOs	GravNet, Transformer	ILD

**Everybody uses state-of-the-art GNN/transformer-based models.
A lot of work is ongoing to develop better NN architectures for MLPF.
A lot of unexplored area.**

Different Groups, Different Ideas, Different Approaches

	Pipeline	Architectures	Detectors
Dolores Garcia et al	hits+tracks → PFOs	GATr	CLD, ILD
Jan Kiesler et al	hits+tracks → PFOs	GravNet, local GATr	CMS HGCal, DICE, CLD, ILD
Farouk Mokhtar et al	clusters + tracks → PFOs	GNN, Transformer	CMS, CLICdet, CLD
Max Hart et al	hits+tracks → PFOs	MaskFormers	CLD
Taikan Suehara et al	hits+tracks → PFOs	GravNet, Transformer	ILD

1. Between the FCC-ee detectors, CLD is used the most.
Up until now, the only detector with a reliably working full reco chain.

Different Groups, Different Ideas, Different Approaches

	Pipeline	Architectures	Detectors
Dolores Garcia et al	hits+tracks → PFOs	GATr	CLD, ILD
Jan Kiesler et al	hits+tracks → PFOs	GravNet, local GATr	CMS HGCal, DICE, CLD, ILD
Farouk Mokhtar et al	clusters + tracks → PFOs	GNN, Transformer	CMS, CLiCdet, CLD
Max Hart et al	hits+tracks → PFOs	MaskFormers	CLD
Taikan Suehara et al	hits+tracks → PFOs	GravNet, Transformer	ILD



1. Between the FCC-ee detectors, CLD is used the most.
Up until now, the only detector with a reliably working full reco chain.
2. Some work on ILD has commenced (ILD@ILC so far)

*Having a stable “default” ILD@FCC detector model would be great for MLPF cross-detector studies while ALLEGRO, IDEA, ALFA are still finalising.
Would also make ILD more present.*

Different Groups, Different Ideas, Different Approaches

Dolores Garcia et al

Jan Kiesler et al

Farouk Mokhtar et al

Max Hart et al

Taikan Suehara et al

The biggest current problem: poor coordination

Not coordinates datasets:

- Different generator setups
- Different physics samples
- Different detector models
- Definition of the MC truth

Not coordinates metrics:

- IQR68/median, RMS90/mean90, STD/mean
- Momentum ranges
- efficiency/fake definitions
- and so on

Leads to incomparable results between groups

Starting work on standardizing MLPF datasets and metrics

- Goal:
- one input format that suits everybody
 - agreed set of metrics
 - (ideally) involve ML community



How well does your model reconstruct events at a future collider?

ColliderBench tracks model results on simulated events from four detector concepts under study for post-LHC colliders. Every row is a reproducible train/test result on a fixed detector geometry and physics sample.

Submit model results

Contribute a dataset

16

4

4

16

result submissions

detector concepts

physics samples

distinct models

Leaderboard

Filter by detector concept and physics sample, or search a model name. Click a numeric column to sort.

Detector concept

Physics sample

Task

Model name

All detectors

All samples

Particle-flow reconstruction

e.g. EdgeConv, ParT...

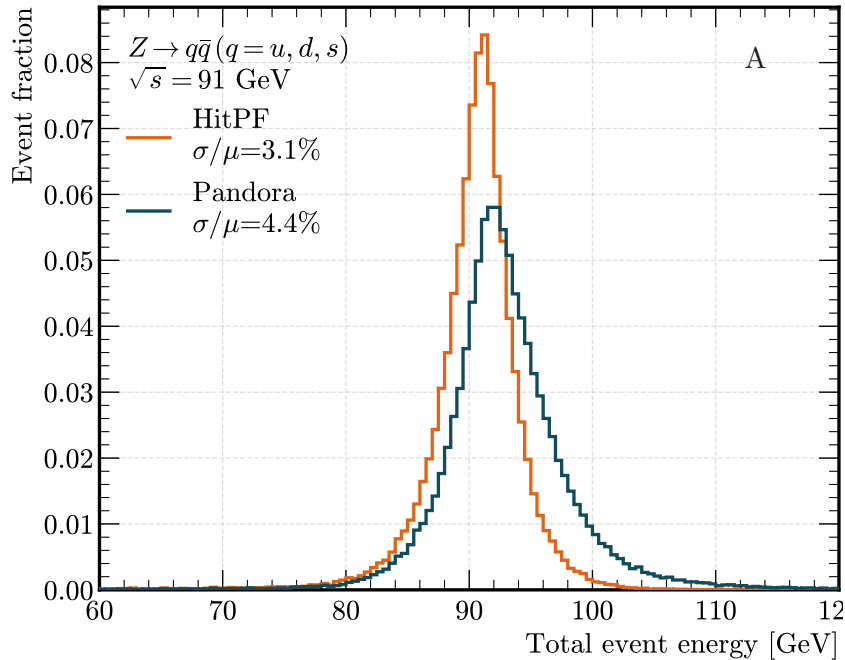
Reset filters

Model	Task	Detector concept	Physics sample	Train acc. (%)	Test acc. (%)	Test AUC	Energy res. (%)	Params (M)	Submitted
HGPflow-B	Particle-flow reconstruction	IDEA_o1_v02	ZH→VV gg @ 240 GeV	95	93.3	—	3.6	9.1	2026-08-20
GNN-PFlow v3	Particle-flow reconstruction	CLD_o1_v05	ZH→VV gg @ 240 GeV	94.2	92.6	—	4.1	7.4	2026-08-02
GNN-PFlow v2	Particle-flow reconstruction	CLD_o1_v05	ZH→qqqq @ 240 GeV	93.8	91.7	—	4.6	7.4	2026-05-14
MLPF-Net	Particle-flow reconstruction	ALLEGRO_v01	ZH→VV gg @ 240 GeV	92.7	90.9	—	5.2	5	2026-06-18
PointNet++-HEP	Particle-flow reconstruction	ILD_15_o1_v02	ZH→VV gg @ 240 GeV	91.5	89.8	—	6	2.6	2026-05-27
MLPF-Net-mini	Particle-flow reconstruction	IDEA_o1_v02	Z→uds @ 91 GeV	90.2	88.6	—	6.4	1.2	2026-03-22

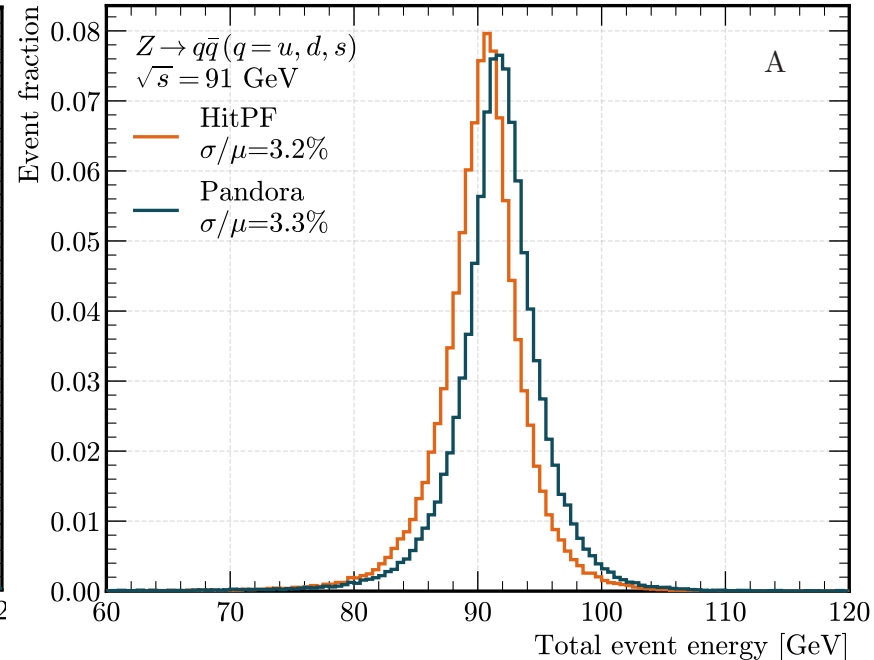
It is just a vibecoded example of the idea. The website doesn't exist...

MLPF JER is similar or better than Pandora

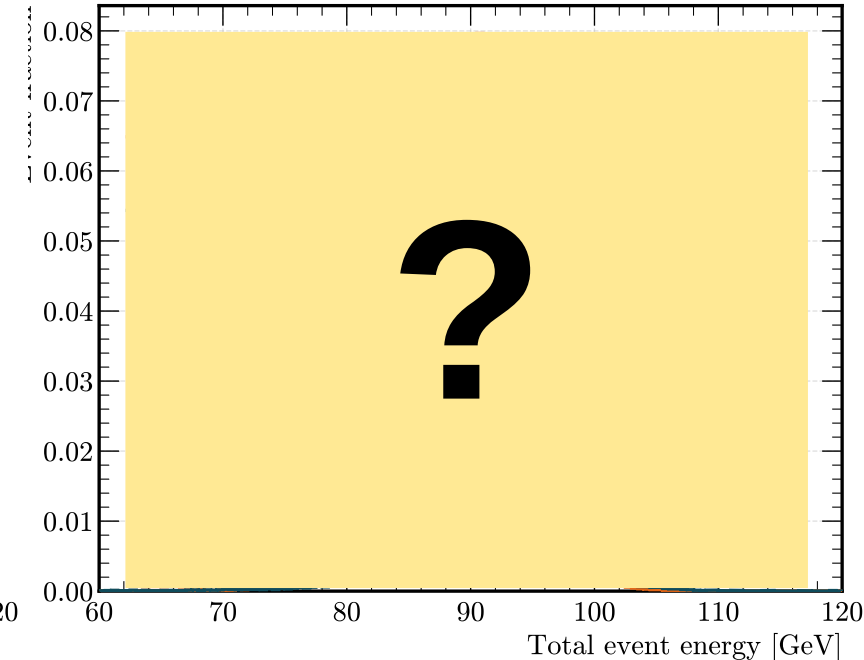
CLD



ILD@ILC



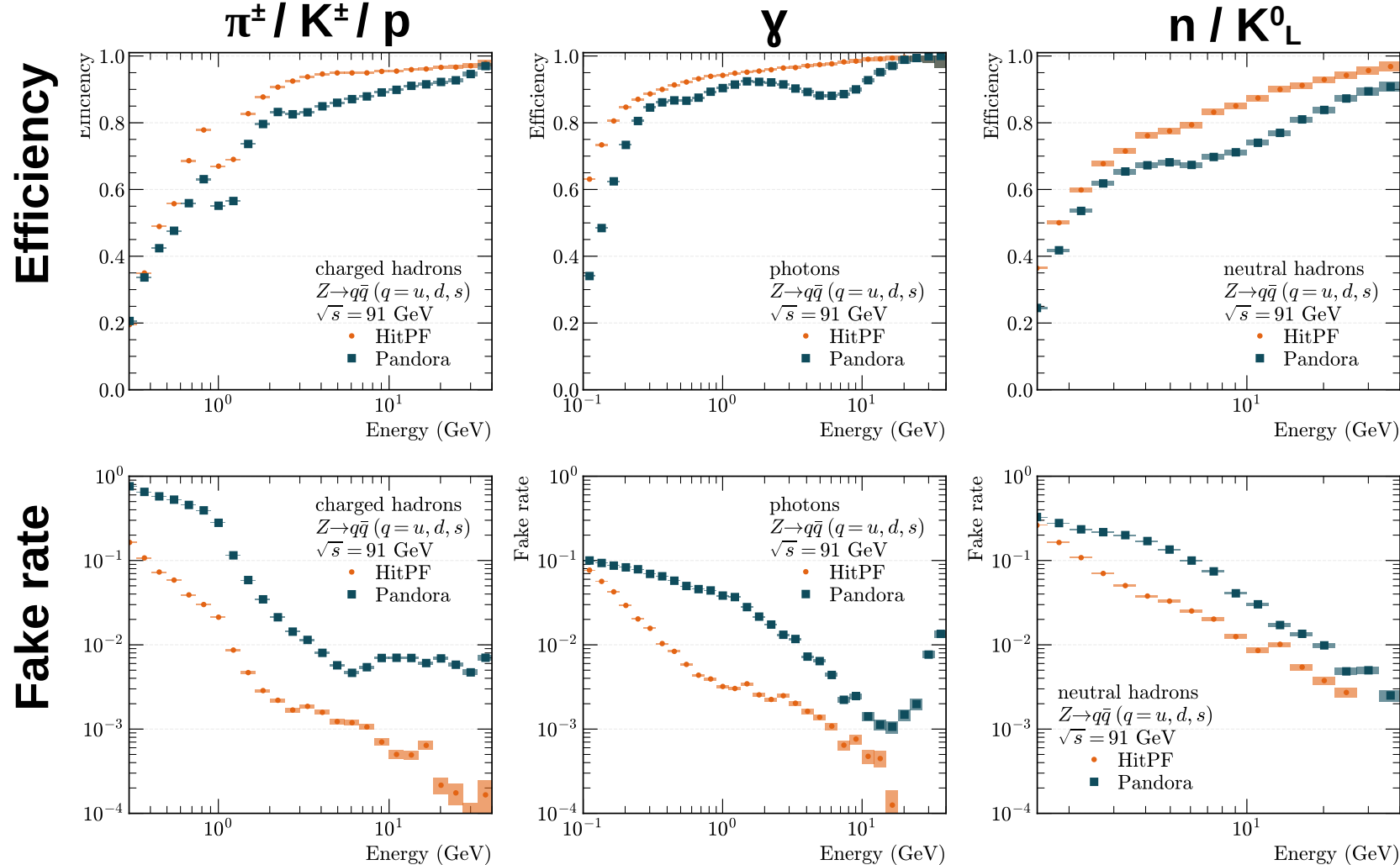
ILD@FCC



credit: Dolores Garcia et al.
a paper is in the process.

**MLPF yields a similar or better JER than PandoraPFA.
Cross-detector comparison is important.**

MLPF per-particle metrics are better than Pandora



Consistently better reco
of the individual particles:

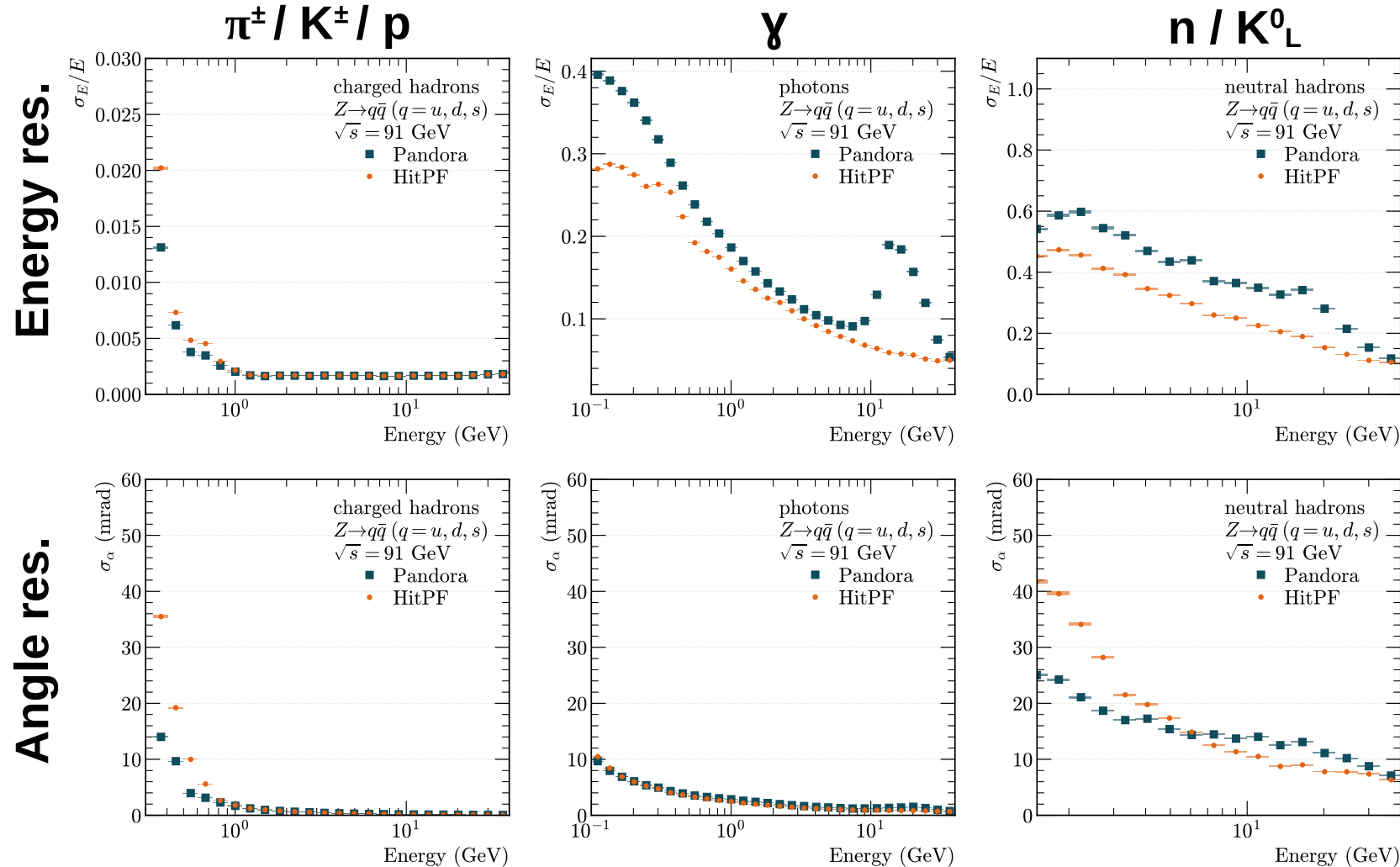
- efficiency
- fake rate
- energy resolution
- angular resolution

TODO:
the impact on physics?

credit: Dolores Garcia et al.
a paper is in the process.

ILD@ILC

MLPF per-particle metrics are better than Pandora



Consistently better reco of the individual particles:

- efficiency
- fake rate
- energy resolution
- angular resolution

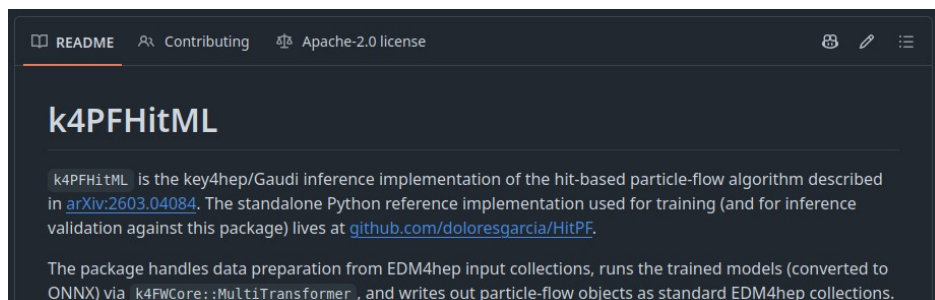
TODO:
the impact on physics?

credit: Dolores Garcia et al.
a paper is in the process.

ILD@ILC

Inference is available in early access (CLD)

<https://github.com/key4hep/k4PFHitML>



Generalizability

What about HitPF inference implementations other than CLD ?

Input collections called in PerformMLPF.py → can be changed here

- Hit type is fixed to 4 → detector with different sub-detector structure needs to take this into account
- Requires conversion of ONNX model
- DPC clustering parameters need to be configured (parameters act on the embedding space)

credit: Lena Maria Herrmann et al.

MLPF inference for ILD will probably be the next!
Test it and give us the feedback once it is ready!

The current status of MLPF from Dolores et al.

- $Z \rightarrow qq$ or ll @ 91 GeV for CLD + inference
- $Z \rightarrow qq$ or ll @ 91 GeV for ILD

The next steps

- More validation
- ZH @ 240 GeV
- $t\bar{t}$ @ 365 GeV
- ALLEGRO, ALFA, IDEA (w/wo ARC, w/wo dual readout, w/wo timing, ...)
- Translate the results into the physics studies
- Need to check: different generator setups, ISR/FSR effects

The MC production is coming (Autumn 2026)

Submit requests

- What? relevant physics samples for ILD@FCC studies
(allows cross-detector MLPF and physics studies)
- When? end of October
(one can ask for samples afterwards, but the software key4hep stack will be frozen)
- Where? <https://fcc-drs.web.cern.ch/>

Contribute to the development of the production process with iLCDirac

- FCC has started to use iLCDirac for dataset production
- Goal: a system suited for all FCC detector concepts
- Including: migration from Dirac → DiracX
- FCC community is very interested in the experience and expertise, which the ILD community has with iLCDirac and their other production tools. The ILD community is very welcome to join the efforts!

For questions, please contact: Juraj Smiesko

Summary

- 1) Including the ILD@FCC for the MLPF studies would be great.
Let us know, which ILD@FCC detector model we should consider the “default” and when.
- 2) We are starting an attempt to standardize the input/metrics for the MLPF studies.
Idea is to simplify the comparison process of the results between different MLPF groups.
Any feedback will be very welcome.
- 3) From the current studies, the MLPF performance looks already usable.
We are constantly working on improving it.
Checking the impact on the physics studies will become possible very soon.
Let us know if you are interested, so we try to make certain inferences available quicker!
- 4) MC production is coming.
Request ILD datasets of interest. (at least a copy of all major processes CLD has)
Contribute with the expertise in the production tools.