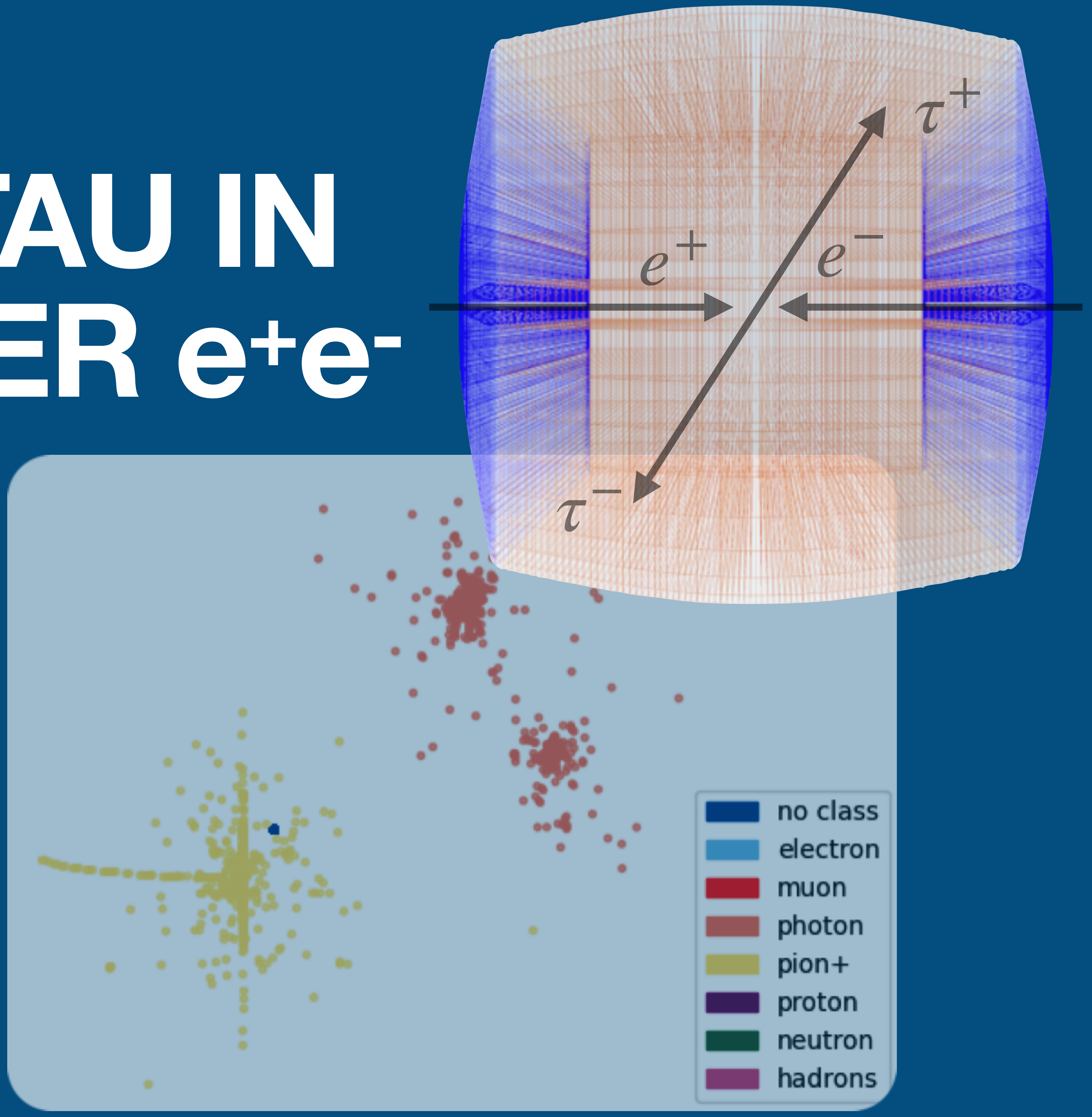
GRAPH NEURAL NETWORK PER L'IDENTIFICAZIONE DI LEPTONI TAU IN ESPERIMENTI AI FUTURI COLLIDER e^+e^-

S. Giagu (Sapienza UdRoma and INFN)



Riunione settimanale ML_INFN - 3 Luglio 2023

FUTURE AI RESEARCH

Istituto Nazionale di Fisica Nucleare

INTRODUCTION

- ongoing work aiming at studying and optimising the physics potential of future collider experiments (principally FCCee and CepC)
  - **case study:** τ-identification with the IDEA dual-readout calorimeter (DRC) concept
    - leverage modern machine learning methods based on differentiable deep neural networks
    - study performance using only standalone DRC information
    - helps also in optimising the detector and design of the readout electronics
  - **tasks studied:**
    - classification of τ-decays and separation from QCD jets based on Dynamic Graph Neural Networks (DGCNN)
    - development of a novel Bayesian-DGCNN for robust estimation of NN predictions
    - DGCNN-based object detection (identification of tau/jet constituents) for proto particle-flow algorithms
  - **evolutions & ononging work:**
    - improve object detection capabilities with hybrid architectures: GNN + encoder/decoder Transformer
    - model acceleration on FPGAs for real-time use (triggers/monitor of physics streams/…)
- **plan to use the results of these studies to prepare one or more tutorials on advanced topics (Bayesian-NN, Hybrid architectures, Model Compression, …) for ML_INFN/AI_INFN**

FCC-ee & -hh

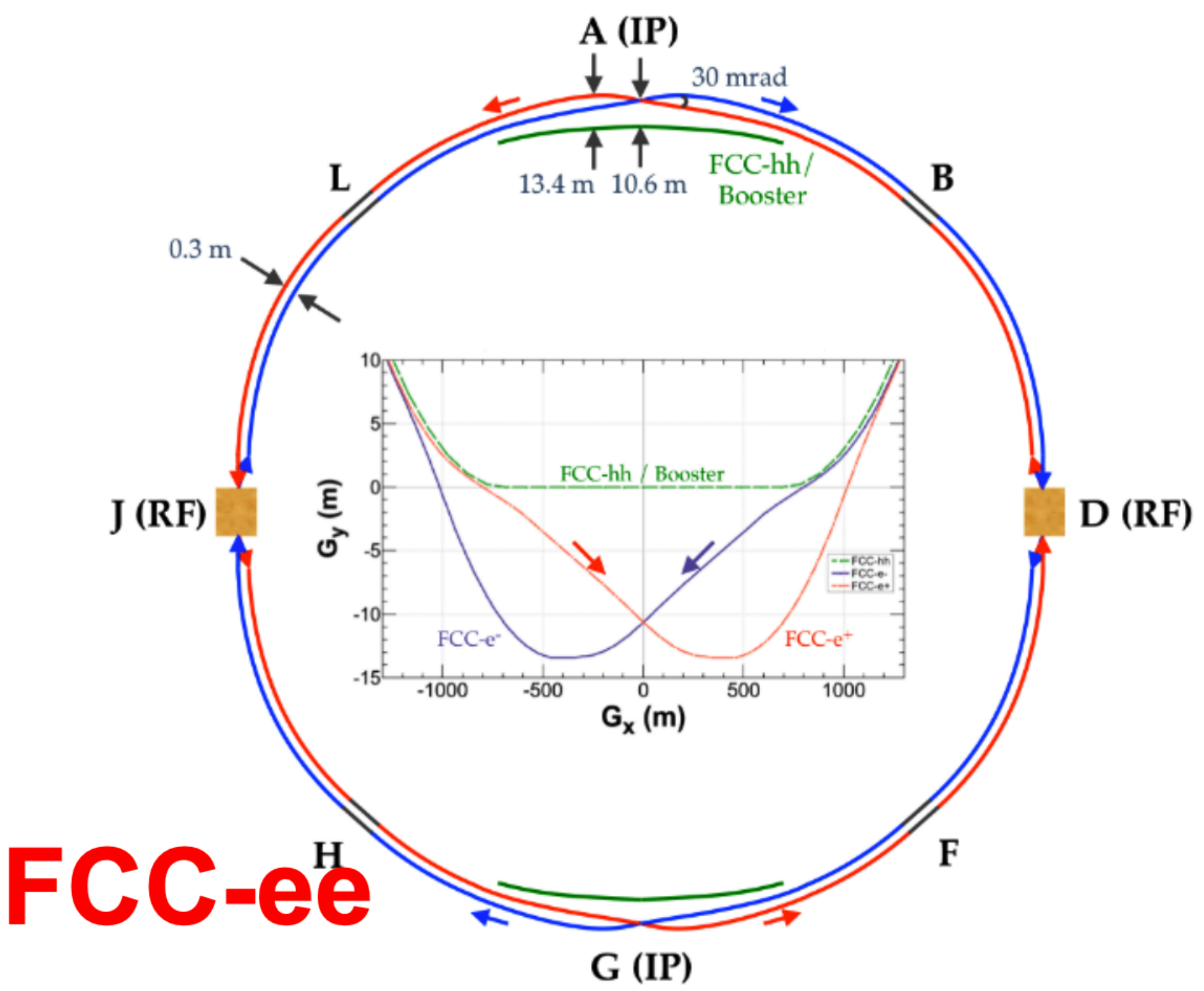
FCC CDRs: https://link.springer.com/article/10.1140/epjst/e2019-900045-4
https://link.springer.com/article/10.1140/epjst/e2019-900087-0

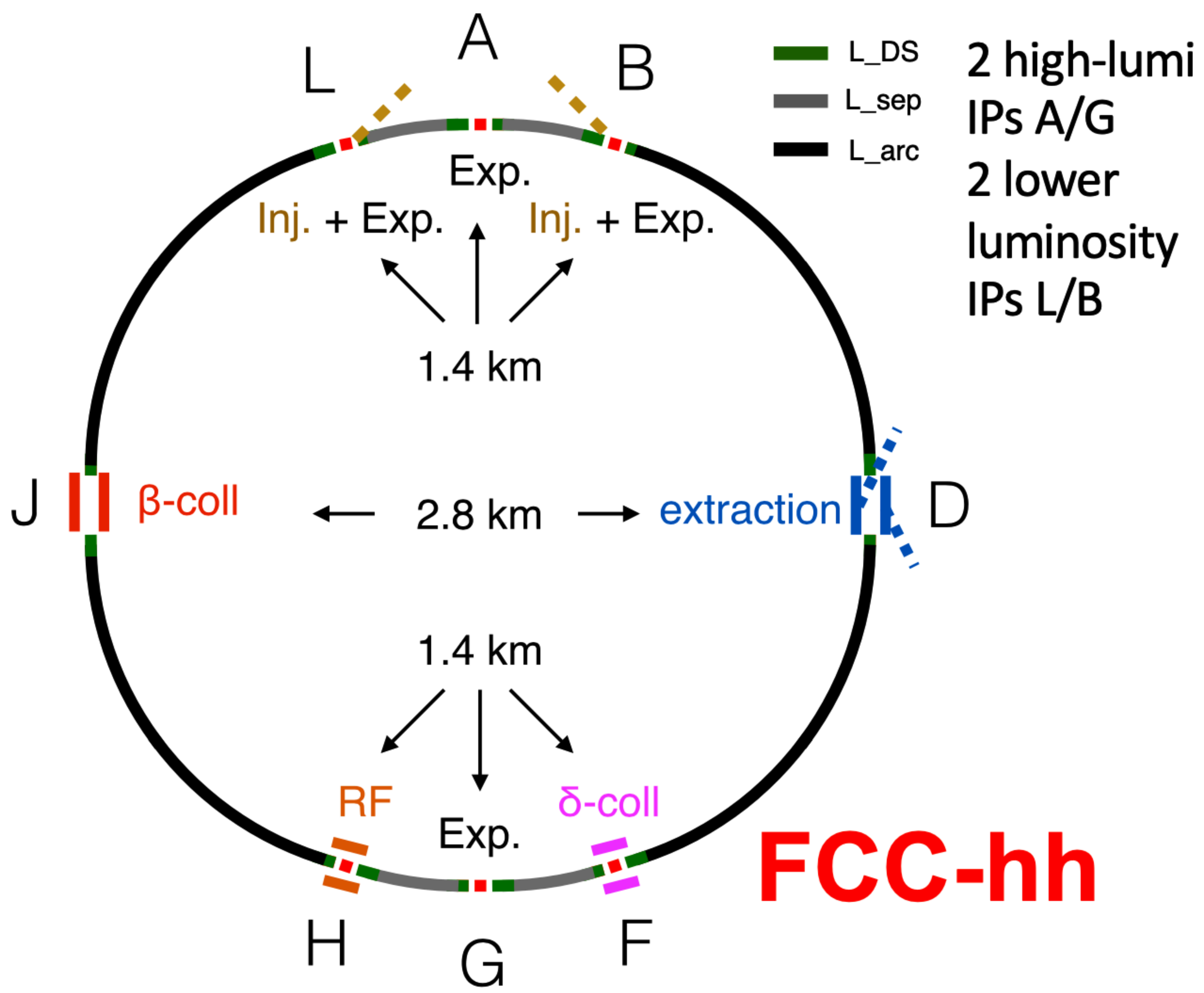
original study requested by ESPP in 2013, started in 2014 as main way to guarantee research continuity in HEP at CERN in the post HL-LHC era

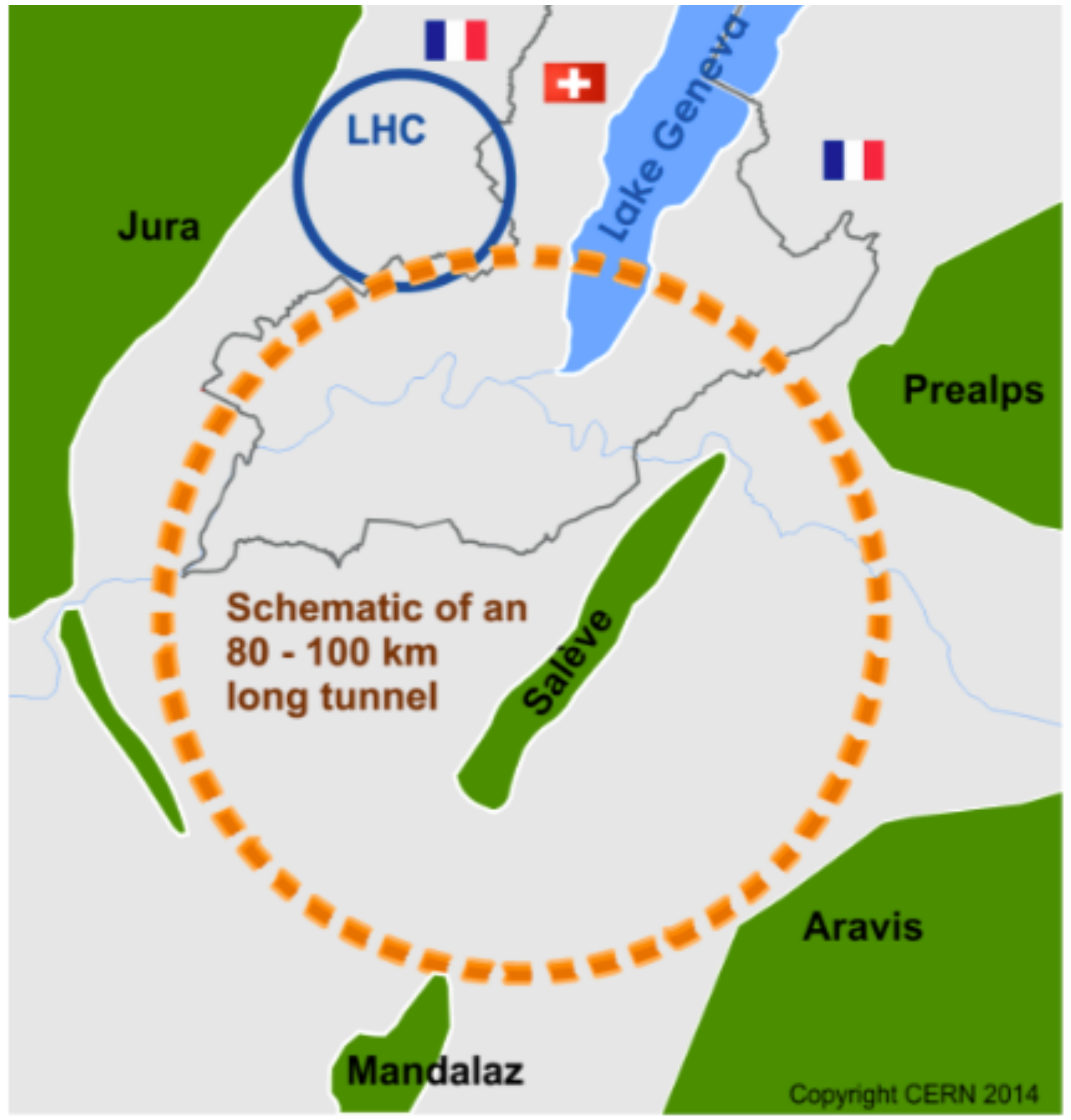
integrated project in two consecutive phases:

- stage 1: FCC-ee - ~90-400 GeV e^+e^- collider as Higgs, EW and top factory at the maximal achievable luminosity
- stage 2: FCC-hh - ~100 TeV hadron collider at the energy frontier + optional ions/eh machines

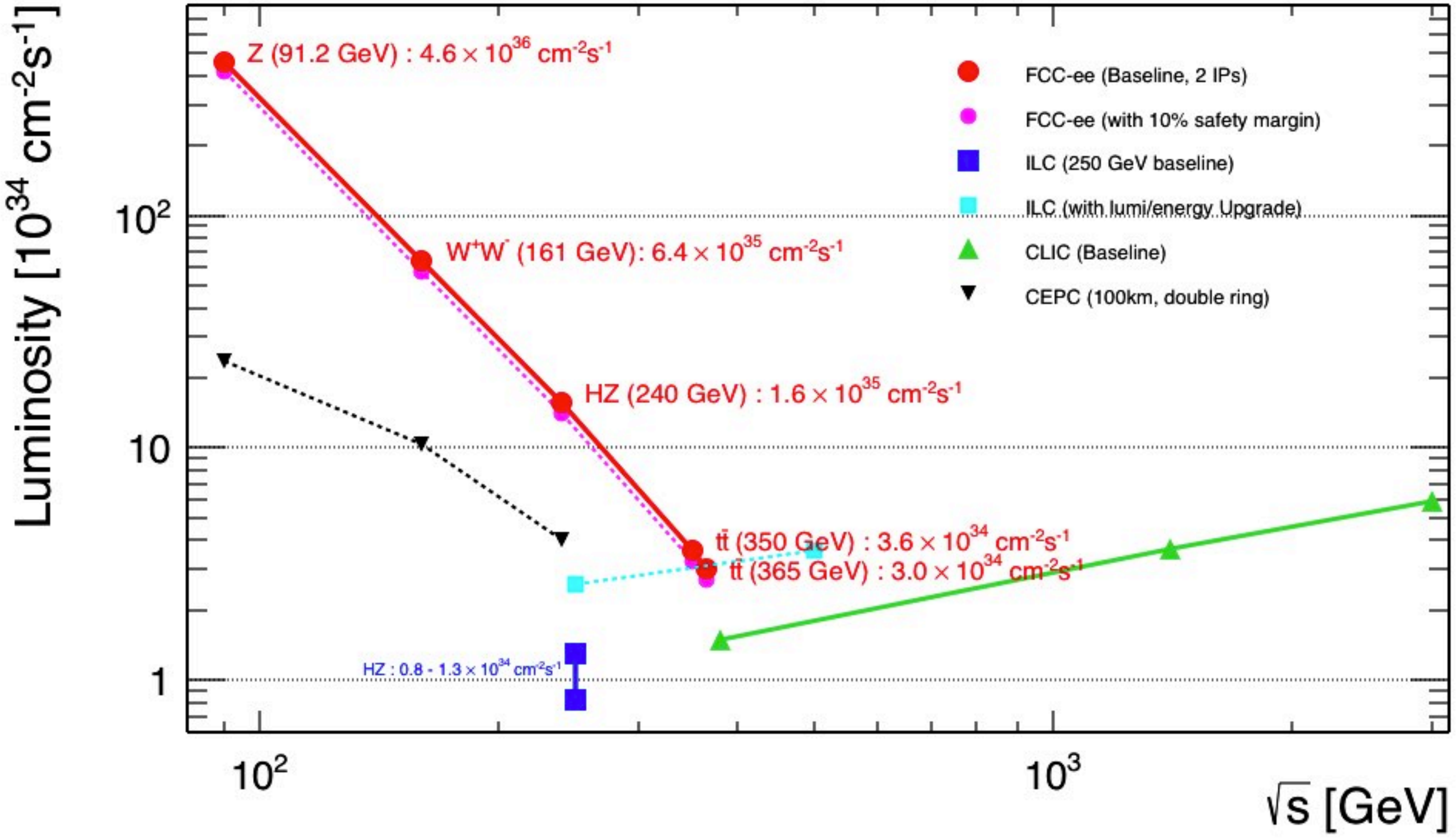
complementary physics goals & common infrastructures and civil engineering







FCC-ee LUMINOSITY



- 100 000 Z / second
  - 1 Z / second at LEP
- 10 000 W / hour
  - 20 000 W in 5 years at LEP
- 1 500 Higgs bosons / day
  - 10-20 times more than ILC
- 1 500 top quarks / day

$$\int Ldt \sim 1 - 40\,\mathrm{ab}^{-1}/\mathrm{year}$$

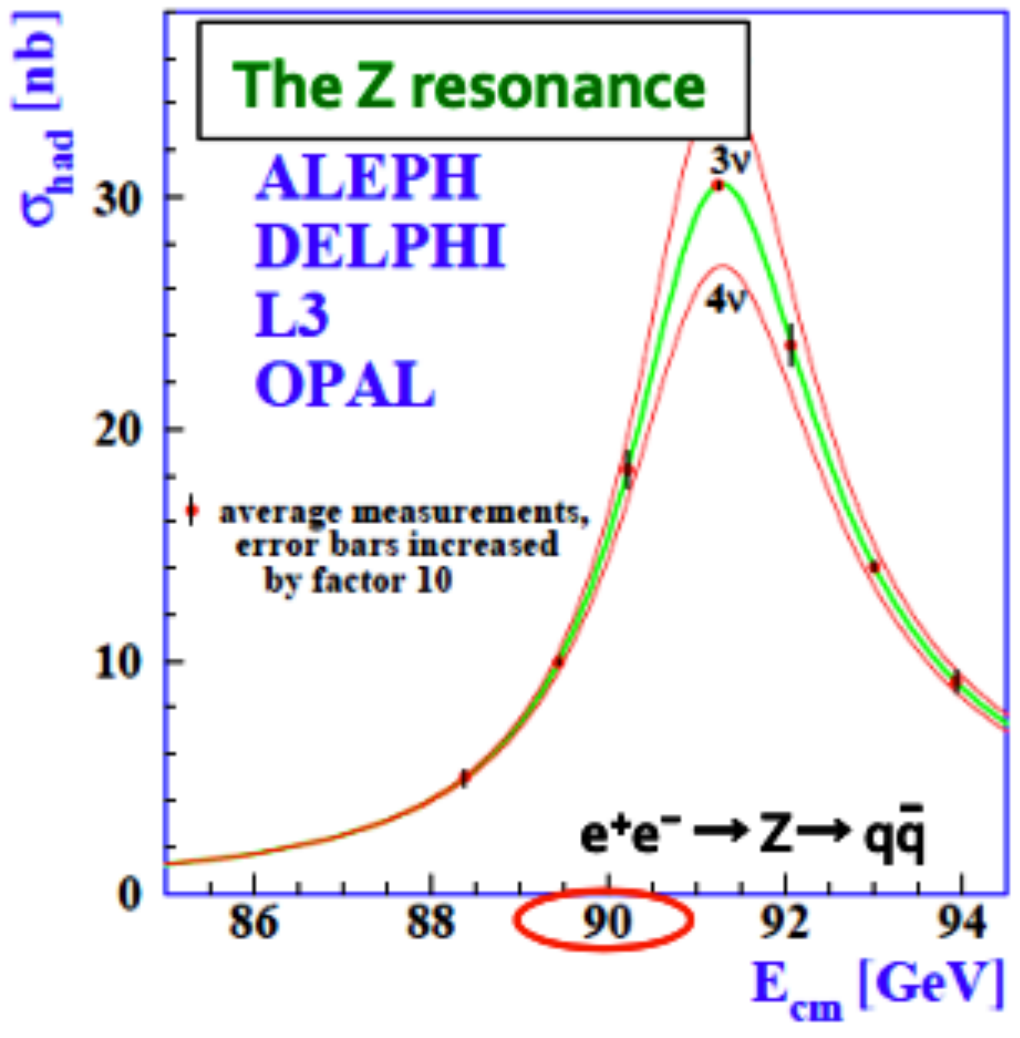
HZ   Z

luminosity x$10^{3\div5}$ LEP thanks to the use of techniques developed for B-factories

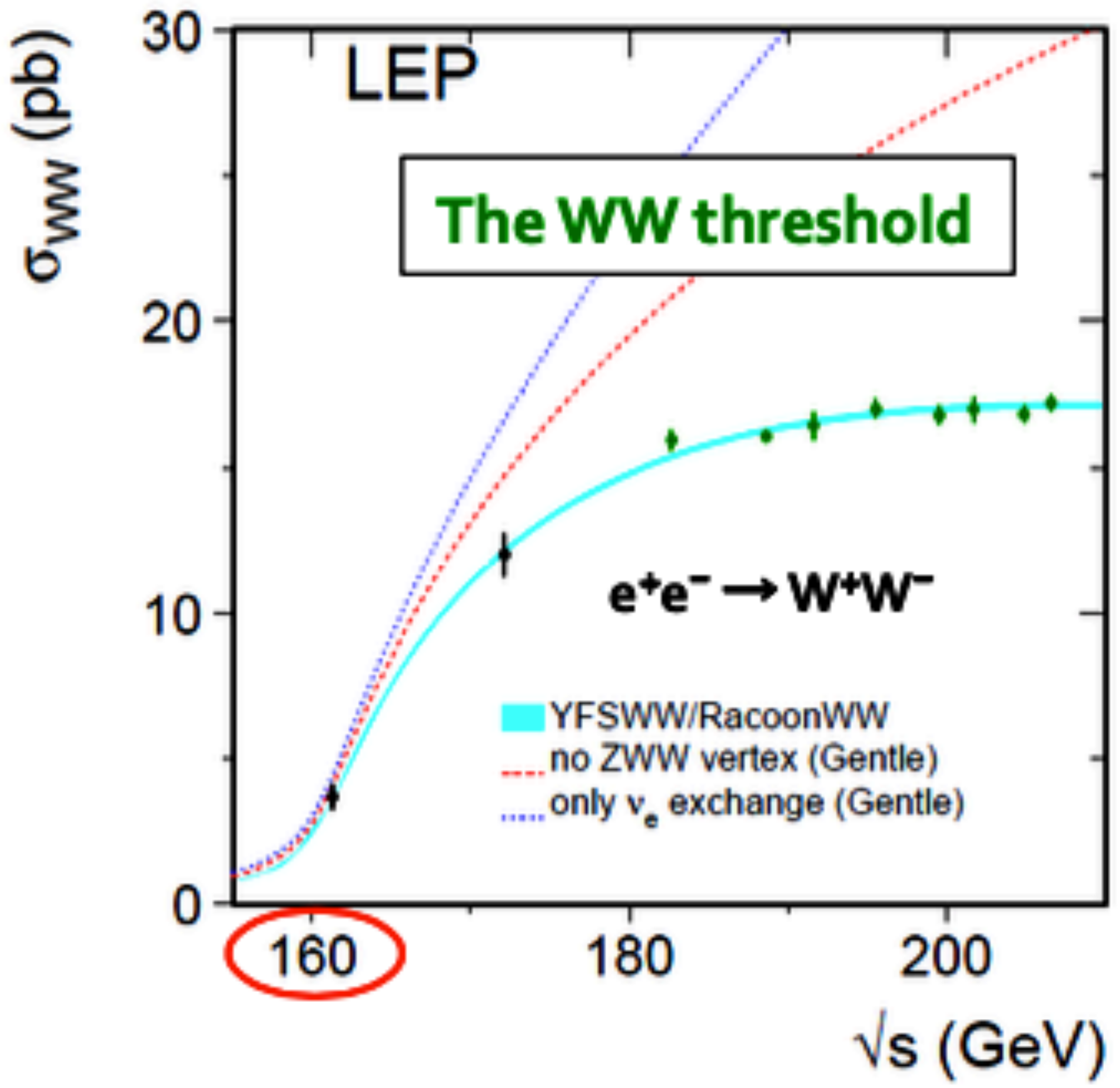
- independent rings for e^+ and e^-: more bunches, higher currents w/o parasitic collisions
- crab waist and asymmetric IP, and continuous injection
- parameters optimised to keep same totale power for synchrotron radiation at all CM energies (100 MW)
  - total consumption with 50% of the klystrons active is 200 MW (compare with LHC: 210 MW and HL-LHC: 260 MW)

FCC-ee CM ENERGY

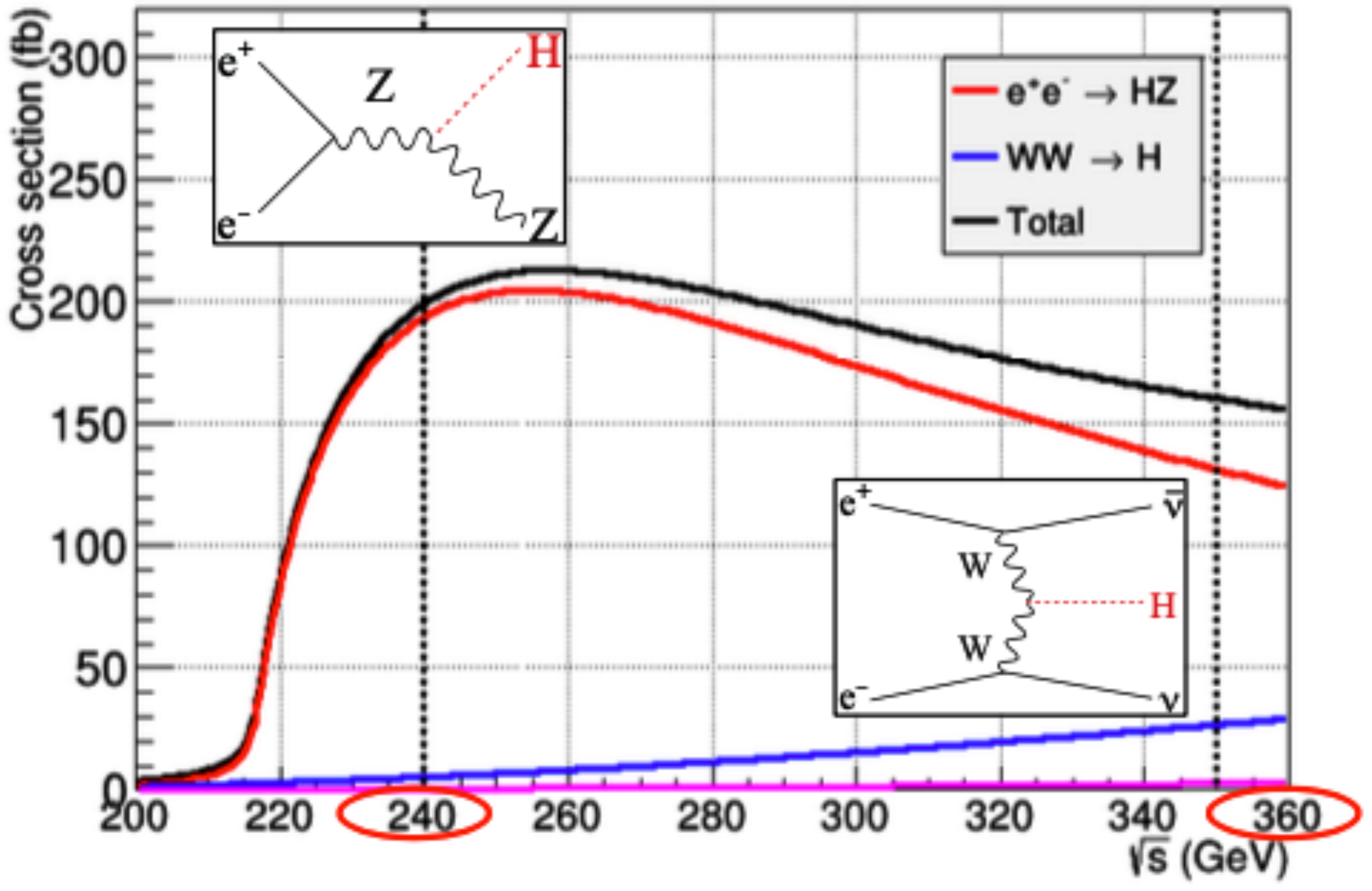
15 years physics: 4 (Z) + 2 (WW) + 3 (H) + 1 LS + 5 (tt) not necessarily in this order …



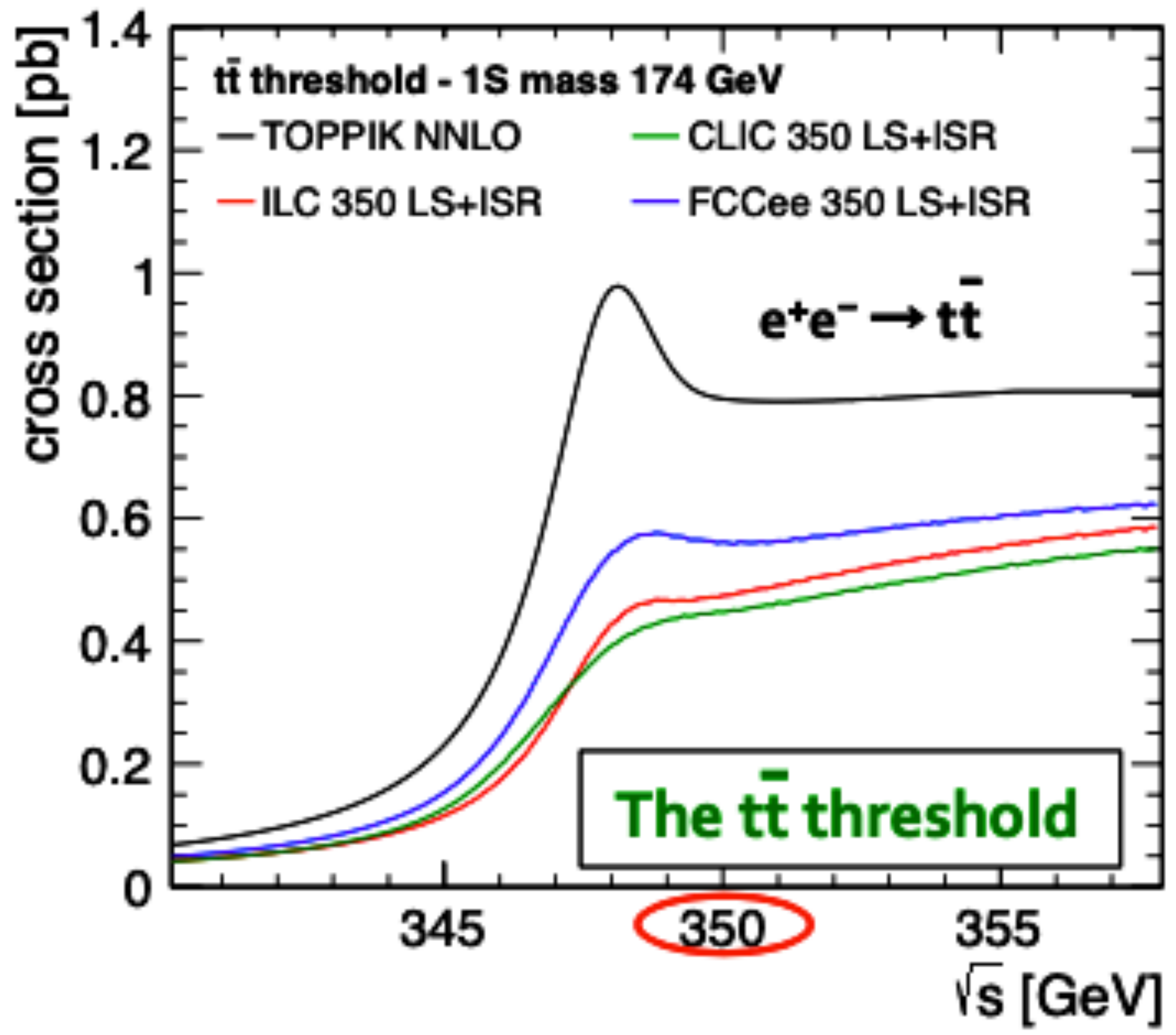
Z



WW

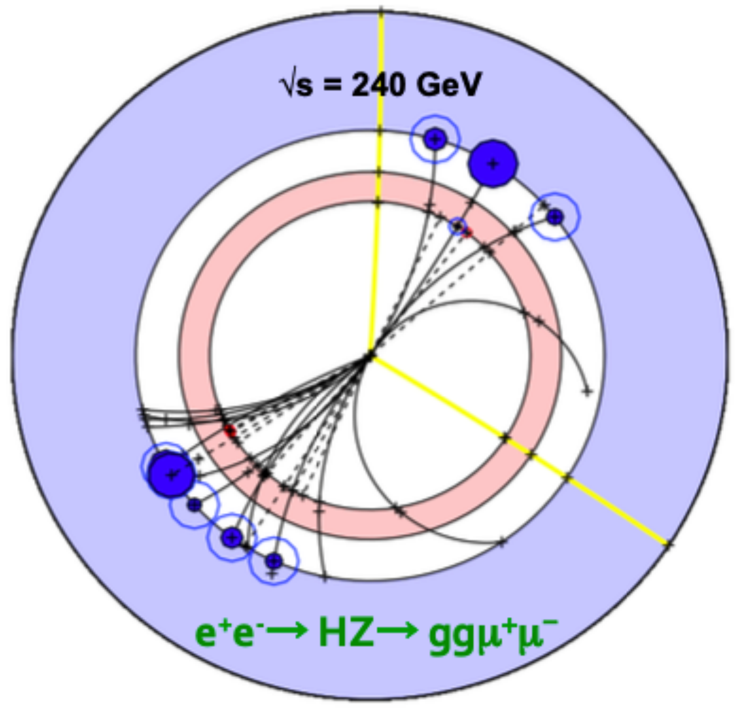


ZH e VBF-H



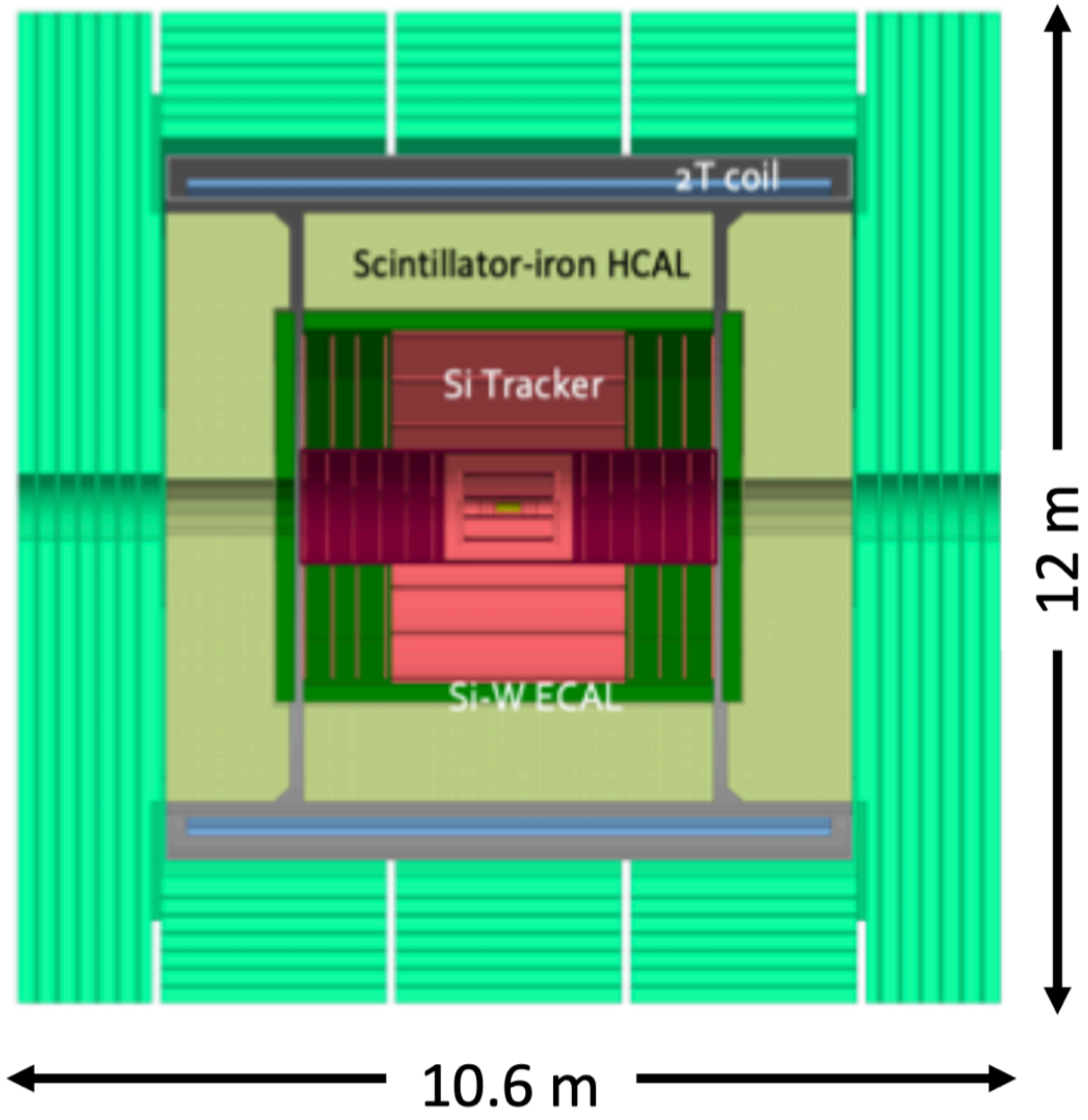
$t\bar{t}$

- physics at the Z pole allows study of light fermions (τ and b - factory)
- clean environment and substantial yields open the possibility to study the properties of gluons in higgs decays: $e^+e^- \rightarrow HZ \rightarrow gg\mu^+\mu^-$



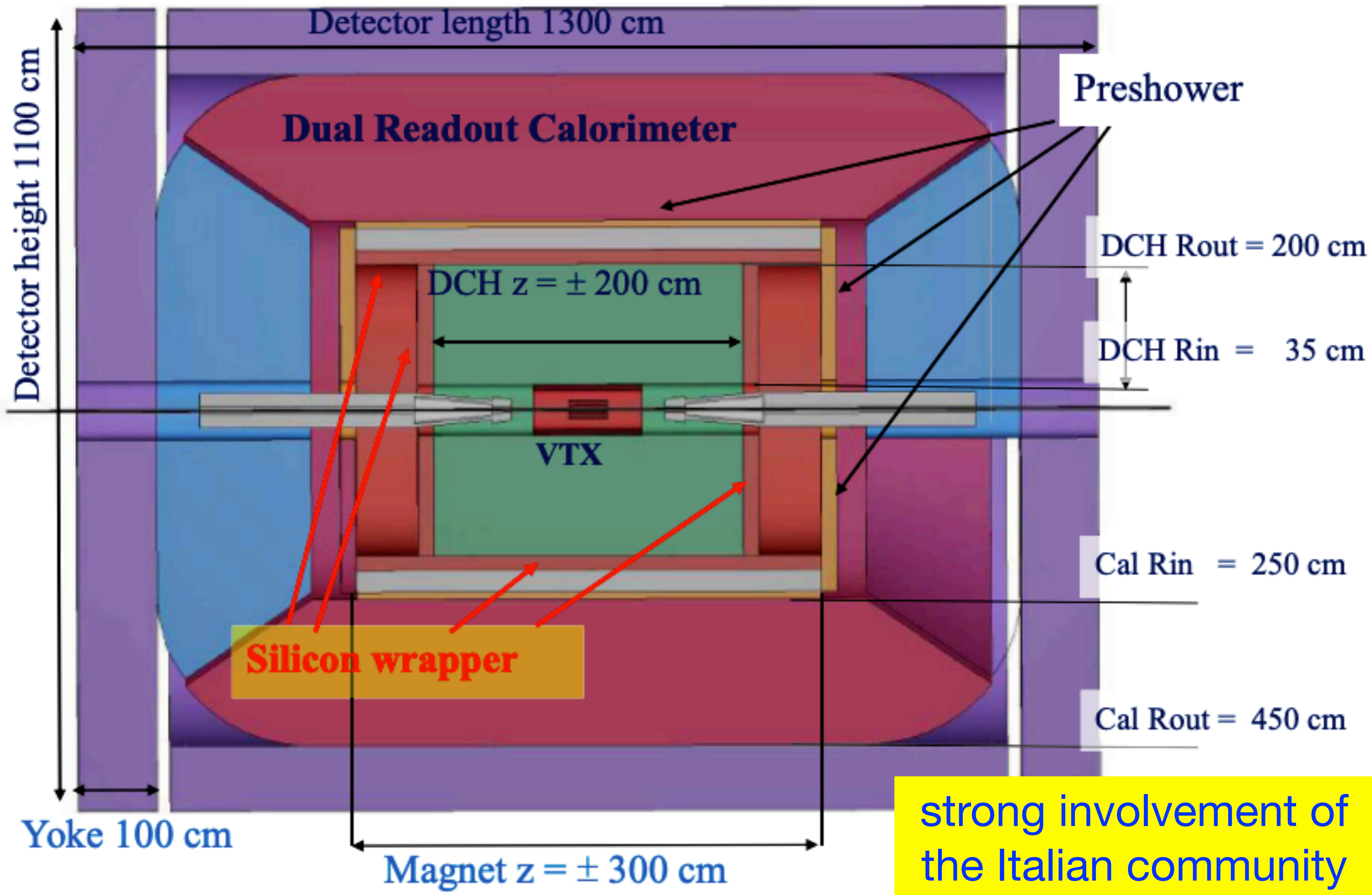
FCC-ee CONCEPTUAL REFERENCE DETECTORS

Clic-Like Detector: adapted from CLIC design



- 2T B-field (CMS-style)
- Silicon ID (pixel + tracker)
- 3D imaging Silicon-tungsten ECAL
- Scintillator + FE HCAL
- MS: steel yoke instrumented with RPCs

International Detector for Electron-positron Accelerators: specific design for FCC-ee / CepC



strong involvement of the Italian community

- 2T SC solenoid 2T ultra-thin and transparent before calorimeters
- Silicon vertex detector + short-drift, ultra-light wire chamber
- Silicon wrapper pre-shower/timing counter
- Single **Dual-readout calorimeter for EM&HAD calorimetry** + optional crystal DR EM
- MS: thin iron yoke equipped with RPCs

DRC PRINCIPLE

$\langle f_{em} \rangle$ fluctuations largely determine energy resolution for jets

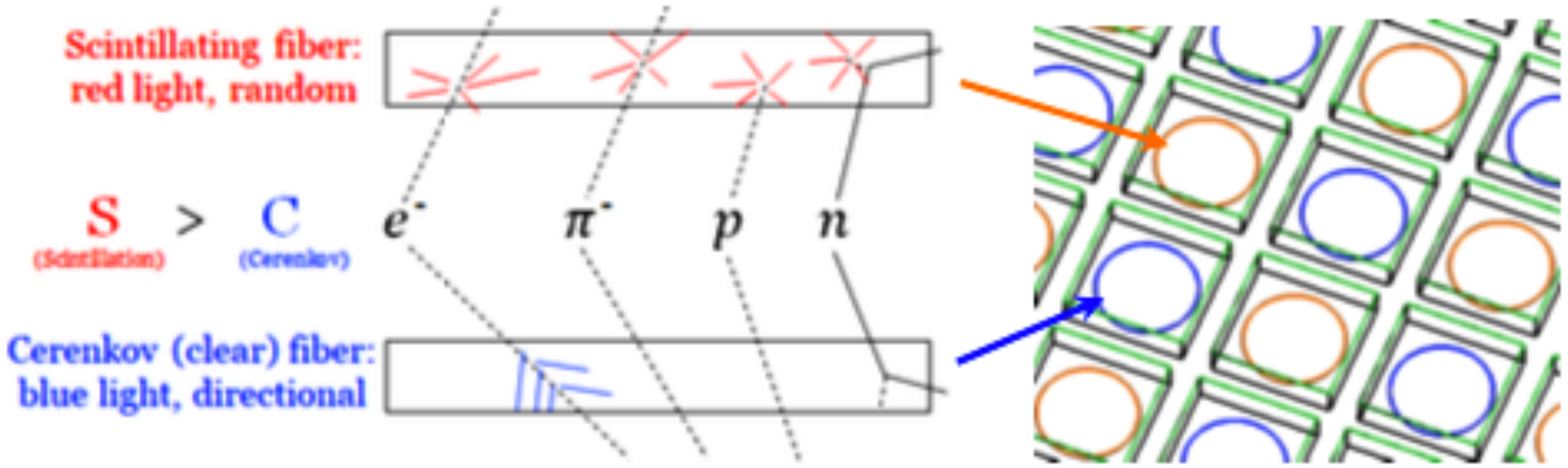
$$S = E[f_{em} + (h/e)_s(1 - f_{em})]$$

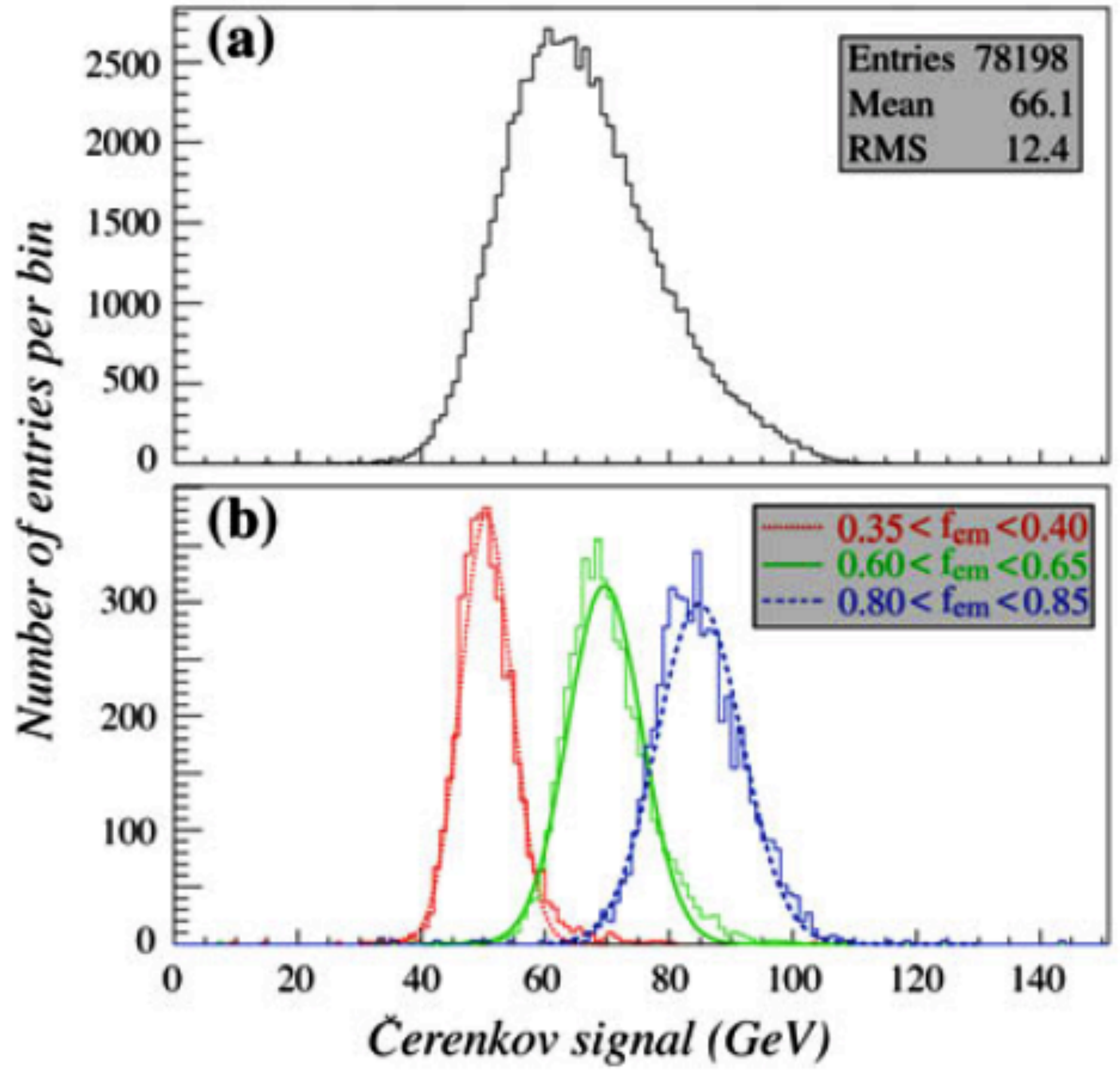
$$C = E[f_{em} + (h/e)_c(1 - f_{em})]$$

two equations in two unknowns: f_{em} and E

$$E = \frac{S - \chi C}{1 - \chi} \qquad \frac{1 - (h/e)_S}{1 - (h/e)_C} = \chi$$

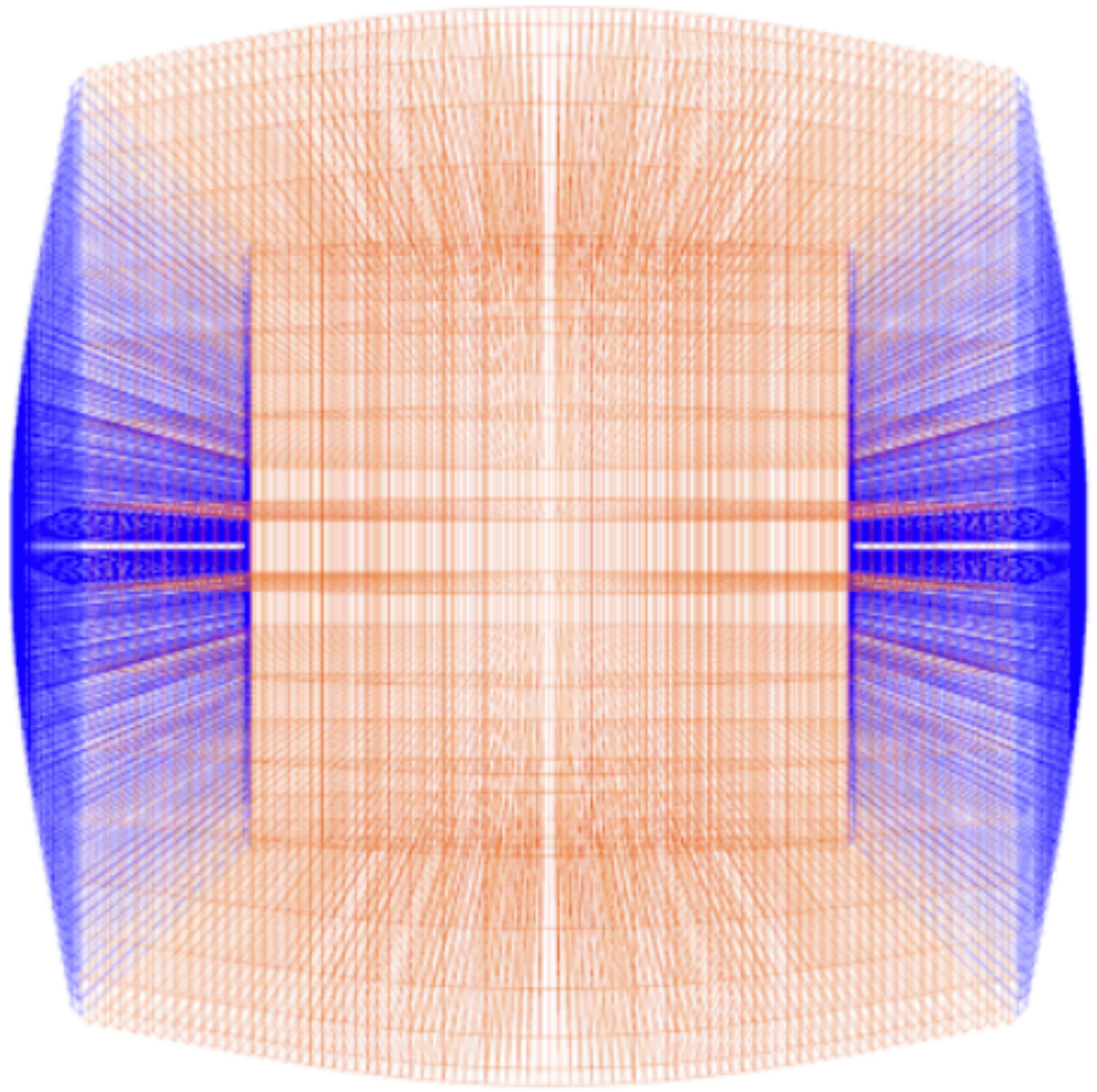
correct shower energy event by event for non-compensation by measuring the EM fraction in hadronic shower by sampling with two readouts of different e/h response: Cherenkov (C) mostly sensitive to the em shower component, Scintillation (S) sensitive to all

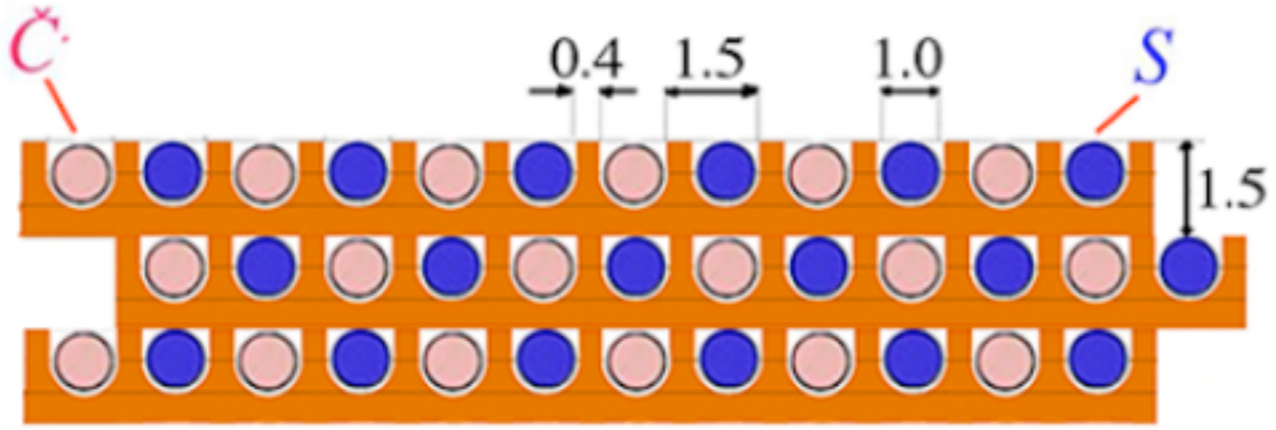


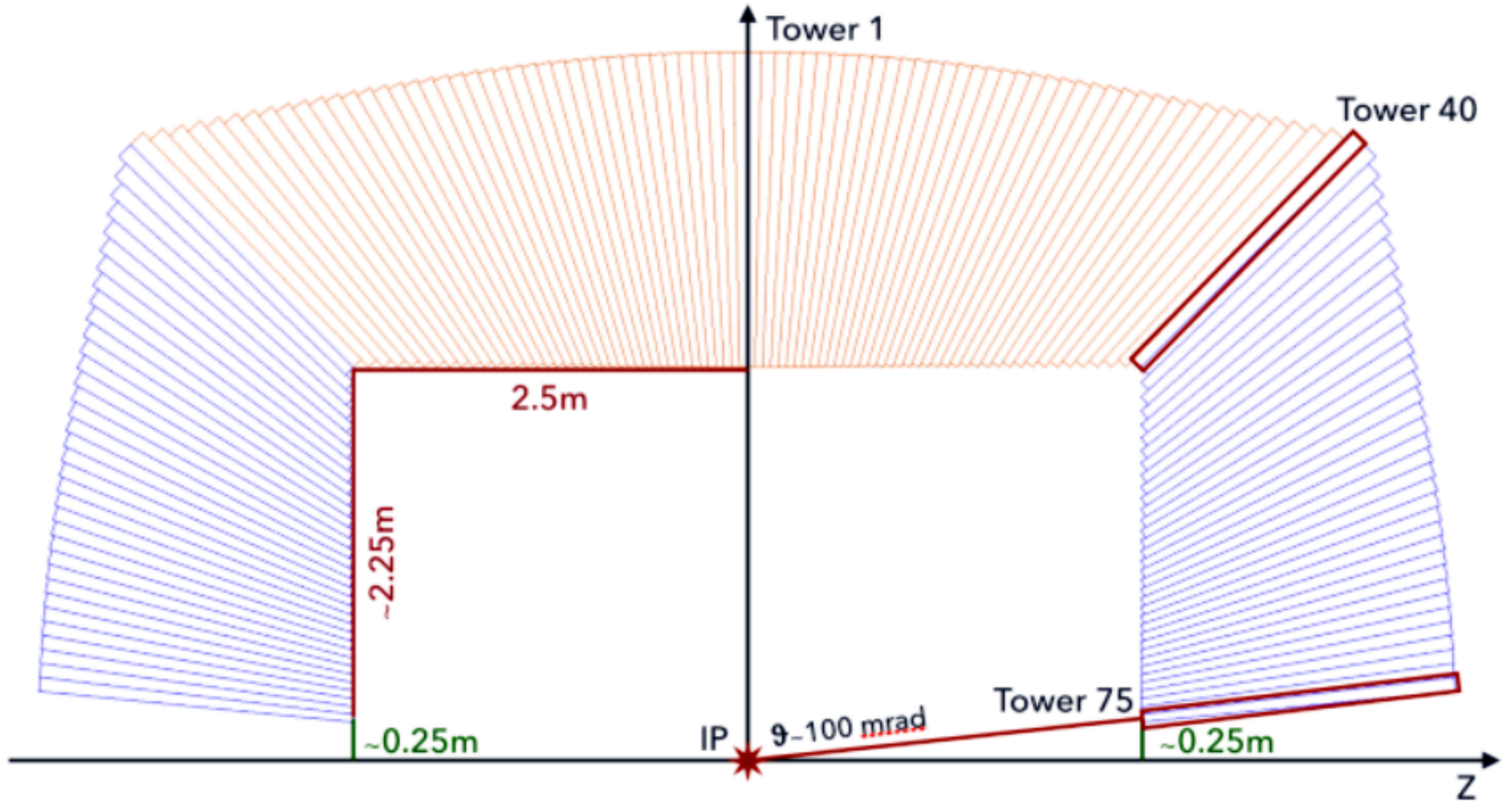


different patterns of S vs C light from different particles, combined with the fine segmentation provided by the fibres can be leveraged also for powerful particle identification …

IDEA DRC FULL SIMULATION





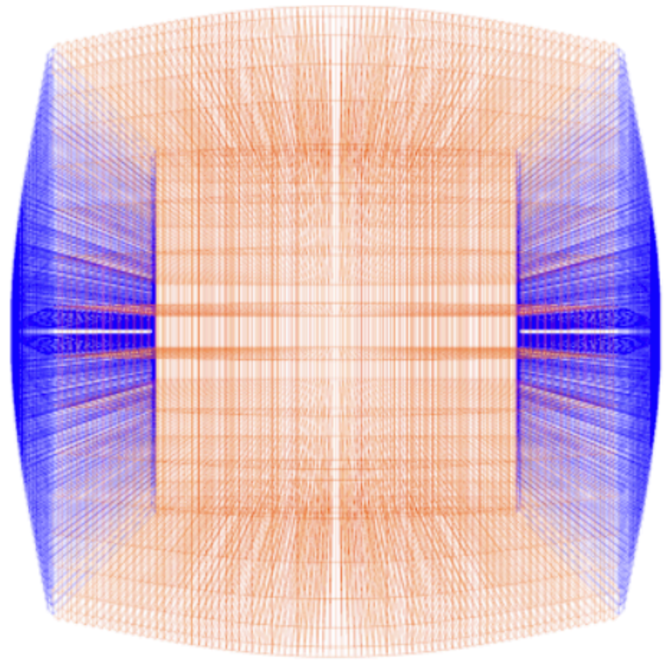


- full G4 simulation of the calorimeter geometry:
  - includes B field and solenoid material in front of the calorimeter
  - fiber-sampling calorimeter: Cu absorber, 1mm fibres, 1.5mm pitch
  - read out of each single fibre via SiPM
  - 130 M channels, excellent granularity and lateral shape sensitivity:

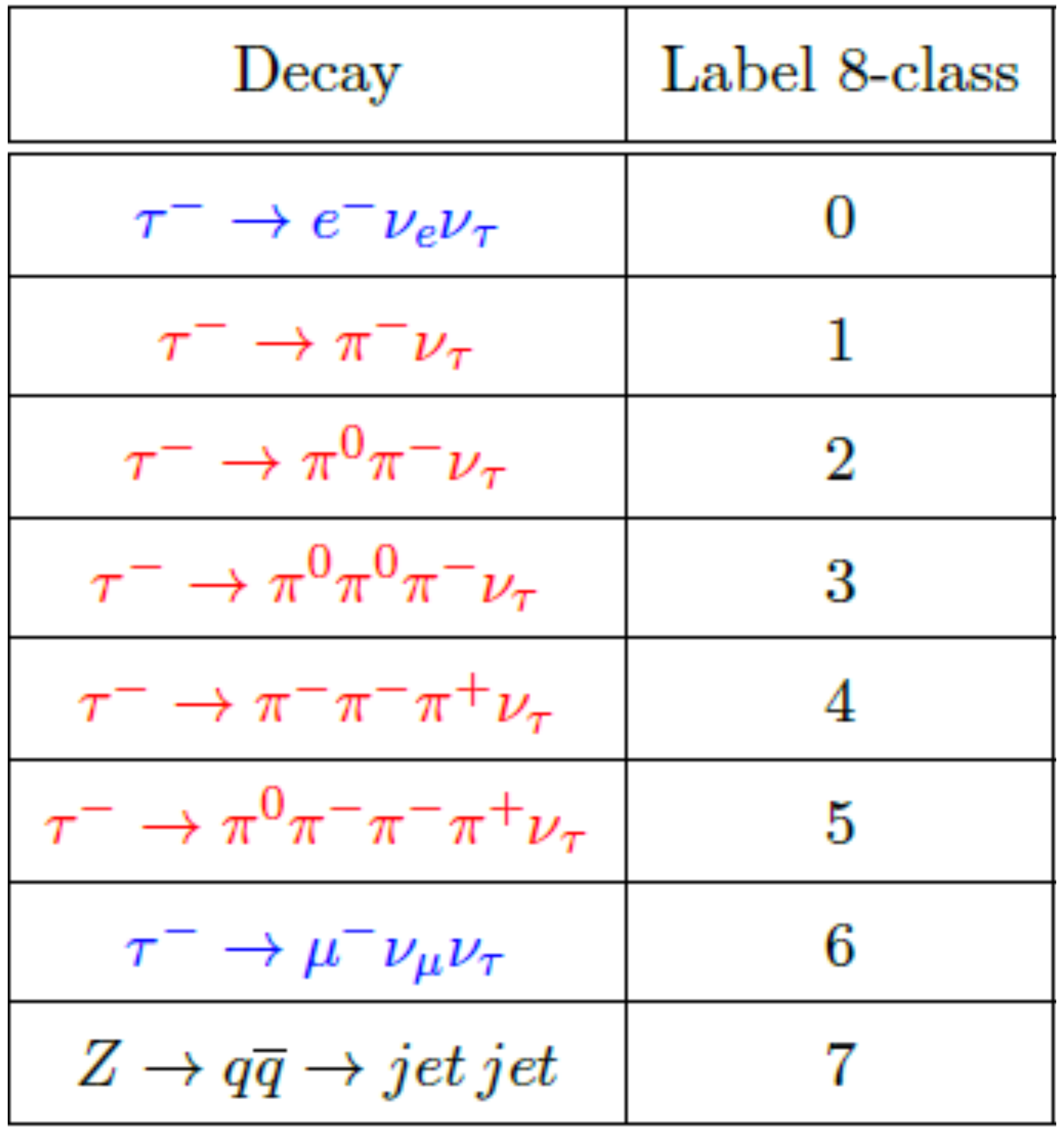
$$\Delta\theta, \Delta\phi = \sim 0.035^{\circ}$$

- parametrised simulation of SiPM readout and signal processing
  - dark counts, crosstalk, afterpulses, saturation, noise, ...

DATASETS



- Pythia8 e^+e^- → Z → ττ and qq at Z pole
- 5000 events for each decay mode



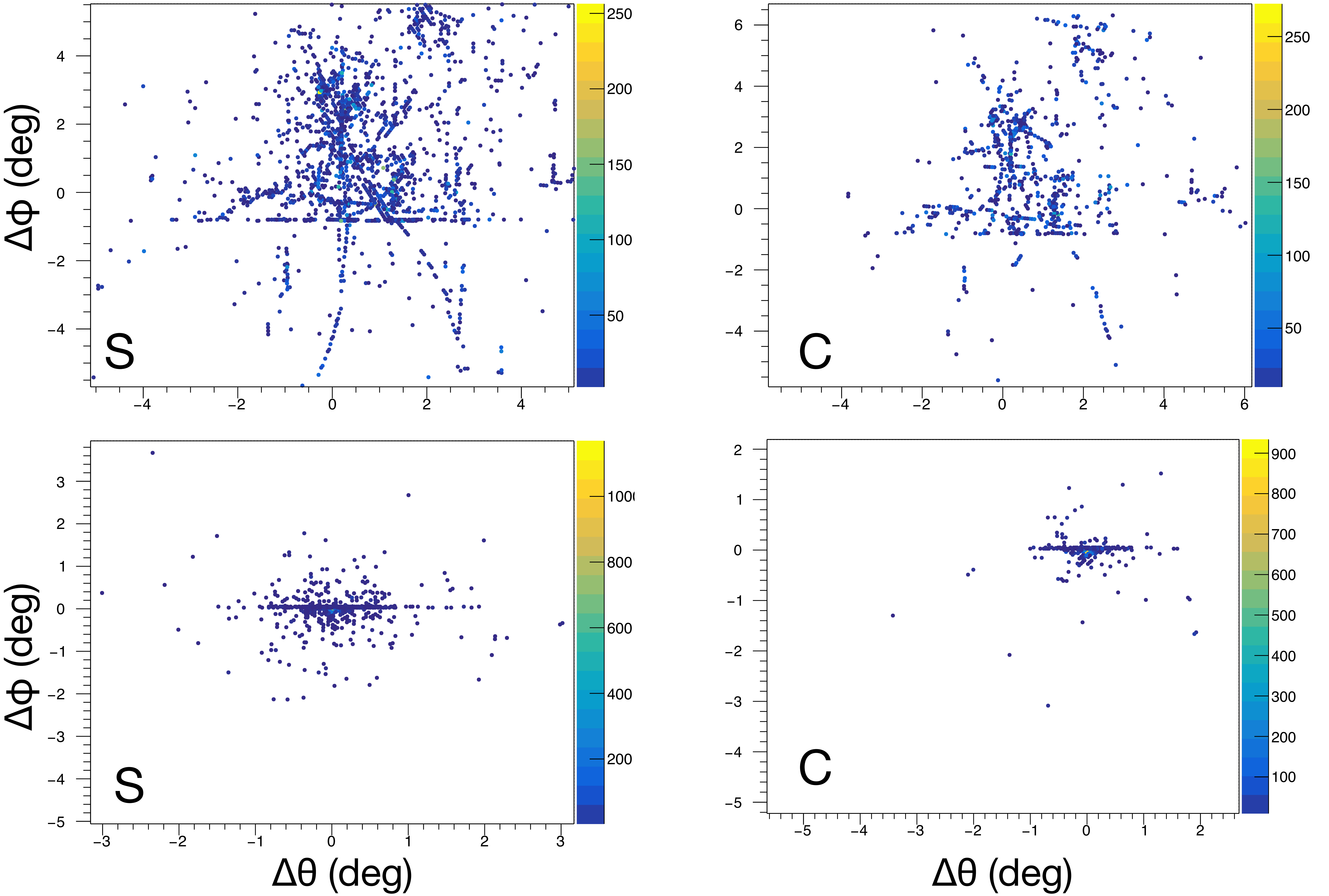
Decay	Label 8-class
$\tau^- \to e^- \nu_e \nu_\tau$	0
$\tau^- \to \pi^- \nu_\tau$	1
$\tau^- \to \pi^0 \pi^- \nu_\tau$	2
$\tau^- \to \pi^0 \pi^0 \pi^- \nu_\tau$	3
$\tau^- \to \pi^- \pi^- \pi^+ \nu_\tau$	4
$\tau^- \to \pi^0 \pi^- \pi^- \pi^+ \nu_\tau$	5
$\tau^- \to \mu^- \nu_\mu \nu_\tau$	6
$Z \to q\bar{q} \to jet\, jet$	7

- Information available for each fibre:
  - geometrical quantities: Δθ, Δϕ wrt the tau/jet cluster center
  - energetic quantities: # of photo-electrons in fibres and energy (scintillation and Cherenkov)
  - SiPM information (1 SiPM per fibre): Integral and Peak of the SiPM output, Time of Arrival, Time over Threshold, Time of Peak
- Ground truth labels:
  - fiber type (scintillating or cherenkov)
  - decay type label

EXAMPLES OF EVENTS WITH FULL GRANULARITY

Z→qq Jets

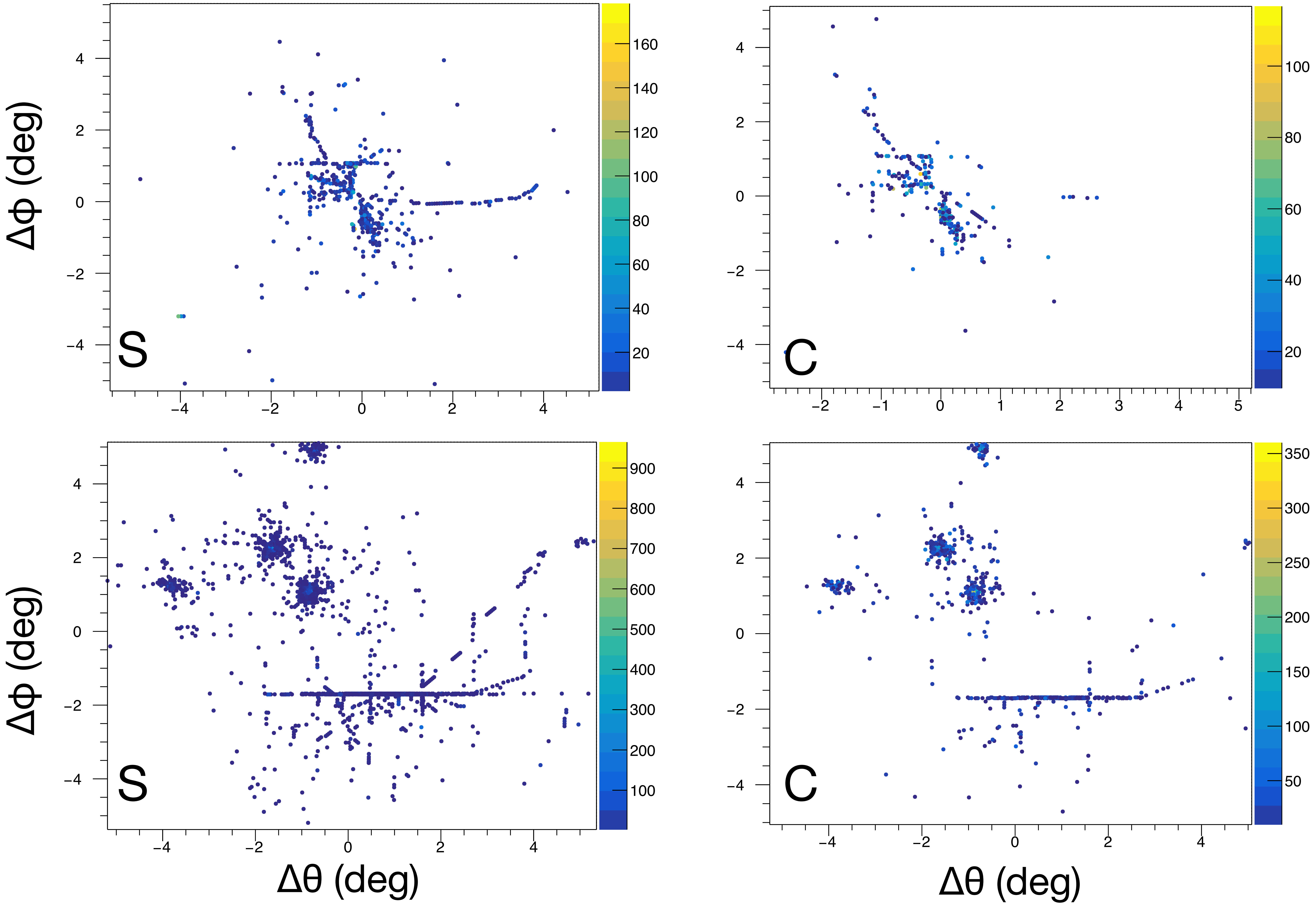
τ →eνν



EXAMPLES OF EVENTS WITH FULL GRANULARITY

$\tau \rightarrow \pi \nu$

$\tau \rightarrow \pi \pi^0 \pi^0 \nu$



MAIN ISSUES IN TRAINING A DL MODEL TO IDENTIFY TAUS IN DRC

- sparsity of data representation: fired fibres are 5-10% of the total ← makes use of CNN architectures inefficient and hard to train
  - **solution:** use point-cloud/graph representations

- ability to extract confidence measures on the prediction of the ANN models ← modern modern ANNs are known to be not well calibrated (e.g. softmax outputs vs true class probabilities)
  - **solution:** calibrate the ANN output, for example by using dropout to adjust the output, or by using conformal predictions or bayesian weights to directly quantify the model uncertainty

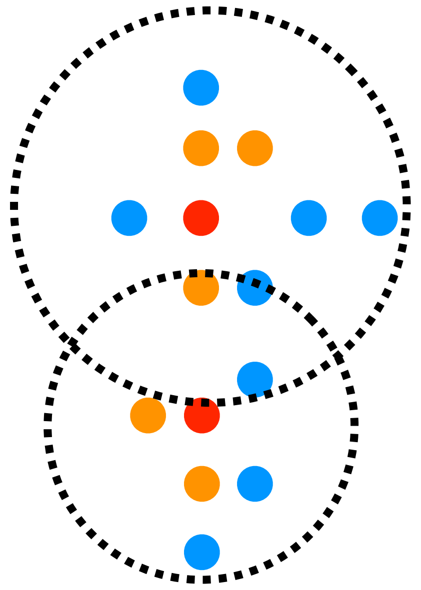
DATA REPRESENTATION

- Image-based: treating the energy deposition on each fiber as the pixel intensity creates an image of the event in fixed-shape mesh
  - natural representation for Convolutional Neural Networks
  - unclear how to incorporate additional information from the fibers
  - very sparse and inefficient representation: jets/tau decays have O(10) to O(100) particles → more than 90% of the pixels are blank

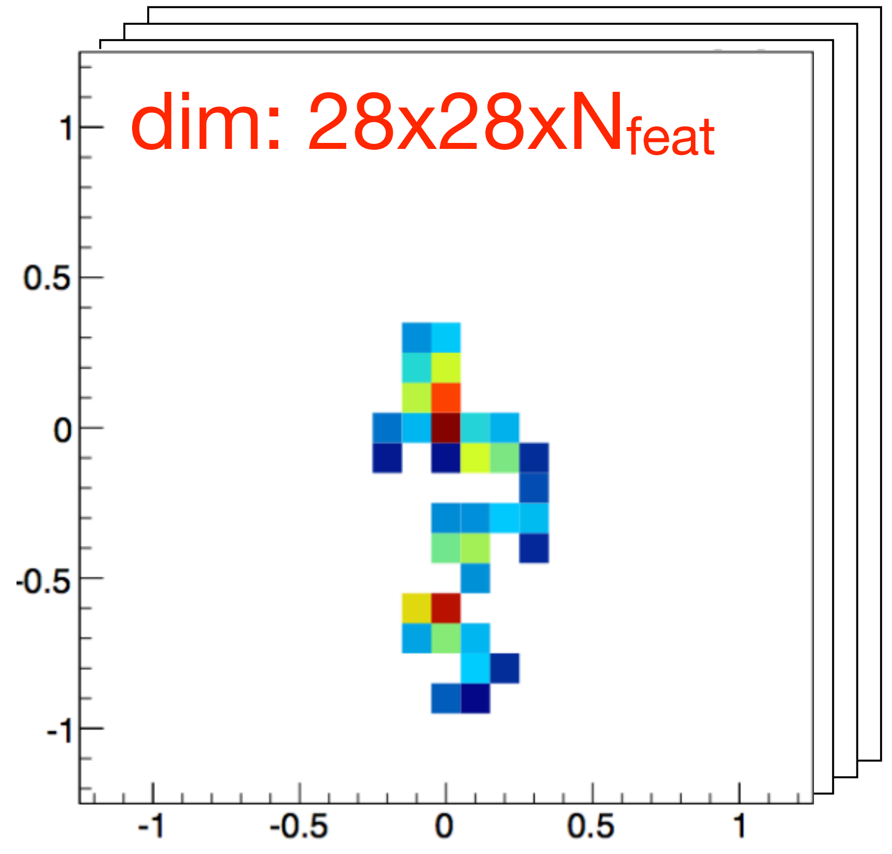
- Point cloud-based: unordered sets of entities distributed irregularly in space, analogous to the point cloud representation of 3D shapes
  - clouds allow rich internal structures
  - straightforward to incorporate additional information of the fibers (fibre type, energy, time information, …)
  - the architecture of the neural network has to be carefully designed to fully exploit the potential of this representation → Graph Neural Networks (also RNN, Transformers, …)

Example

boosted W→qq’



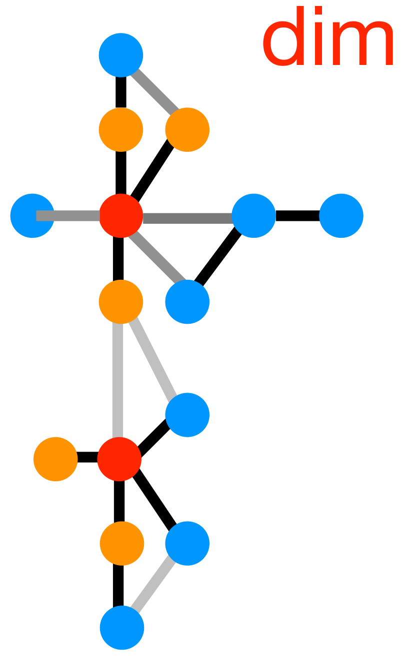
fixed-mesh image rap.



dim: 28x28xN_{feat}

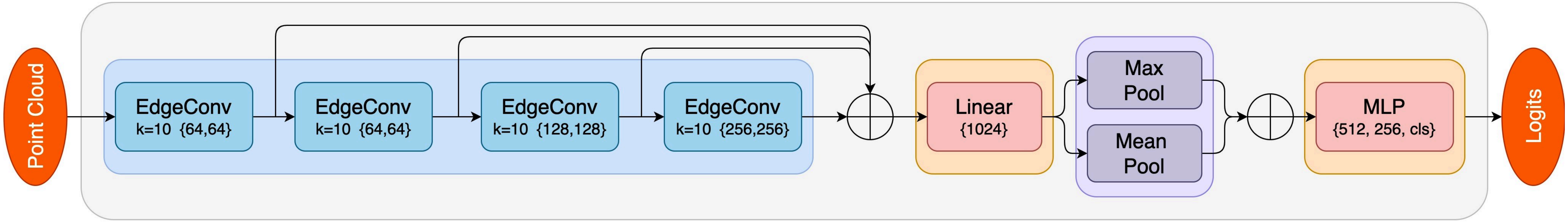
VS

graph representation



dim: 15xN_{feat}

DGCNN ARCHITECTURE



Input:
x = {θ, ϕ, geometrical features,
SiPM features, …}

local feature extractor

feature aggregator

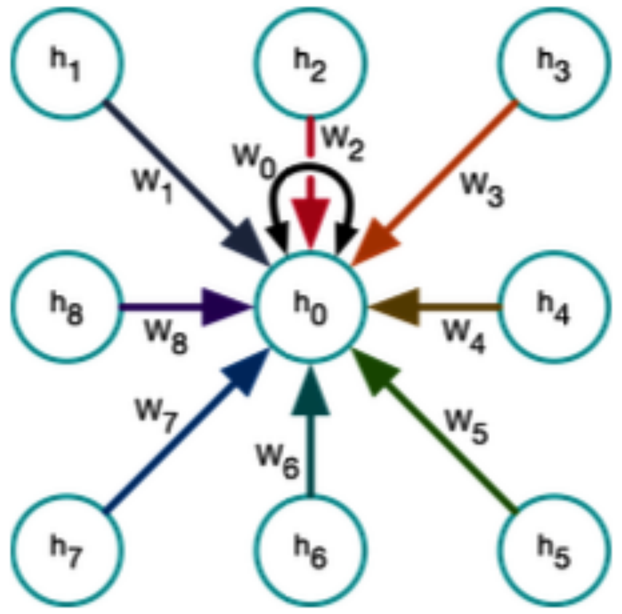
MLP global classifier

- simple and flexible architecture optimised for point cloud inputs able to learn both local (trough the edge convolution) and global (through skip connections & feature aggregator) structures on the input data
- simplify inclusion of additional features and SiPM information
- # of input fibres fixed and treated as model hyper parameter, discarding those with lowest signals or adding zero valued vectors and masking in case of events with lower active fibres
- hyper-parameters chosen using a validation set

EDGE CONVOLUTION

Regular convolution operators cannot be applied on point clouds:

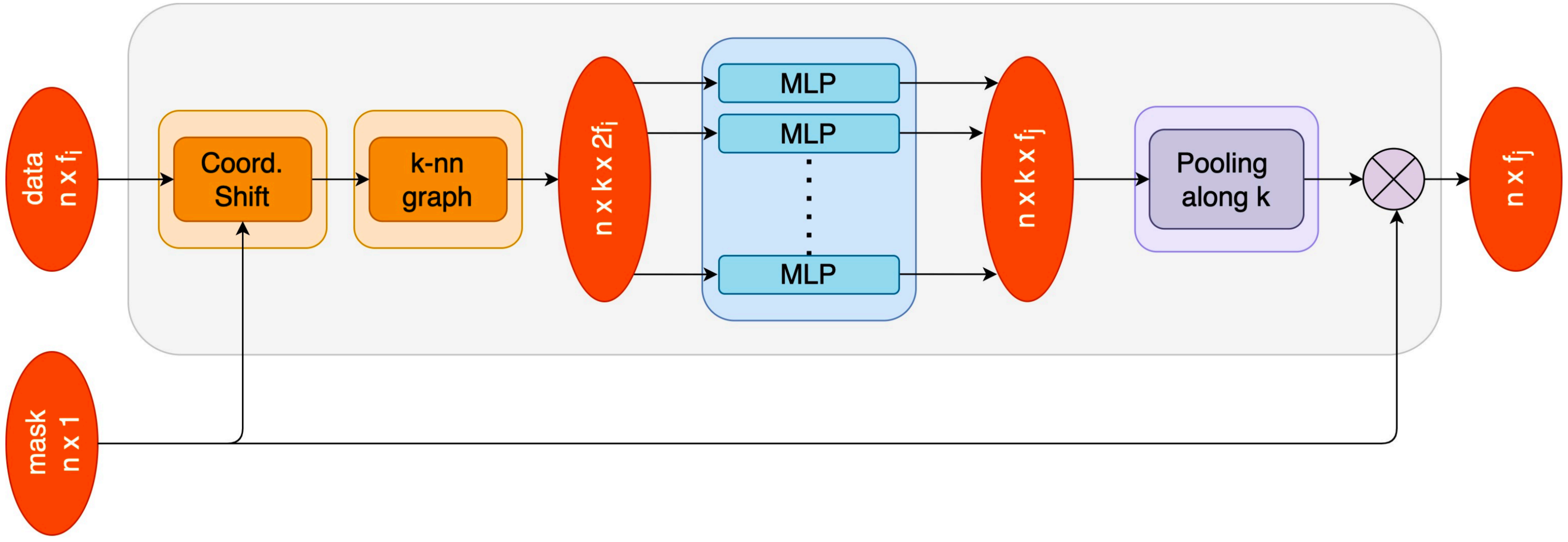
- points distribution is usually irregular (unlike uniform grids of the pixels in an image)
- they're not invariant under permutation of the points



CNN-conv

$$h'_o = \sum_{i=0}^{8} h_i w_i$$

Viable solution: **EDGE convolution**: point cloud represented as a graph with **Vertices (**the points themselves) and **Edges** (connections between each point to its k nearest neighbouring points): results in a **regular distribution** for each point, for which is possible to define convolution operations

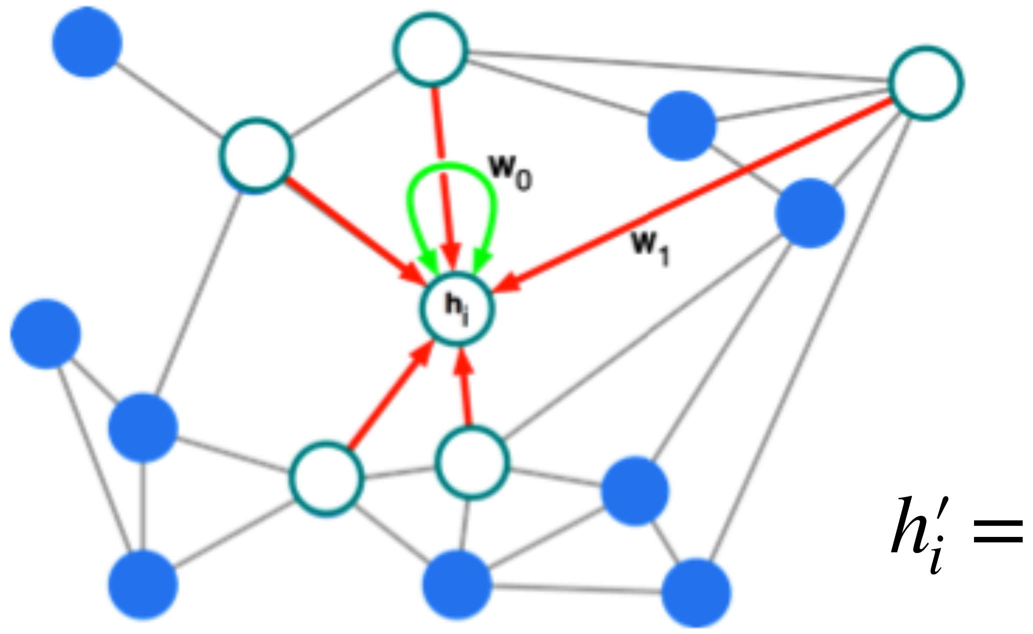


Compute k-nn directed graph

Extract edge features with a shared linear layer

Aggregate edge features by max-pooling

ex. k=6



EDGE-conv

$$\sum_{j \in N_k(h_i)} \ ; \quad \max_{j \in N_k(h_i)} \ ; \quad \cdots$$

$$h'_i = \square_{j \in N_k(h_i)} f_w(h_i, h_j)$$

$$f_w(h_i, h_j) = \text{ReLU}(\mathbf{w}(h_j - h_i) + w_0 h_i)$$

τ DECAY IDENTIFICATION WITH DGCNN

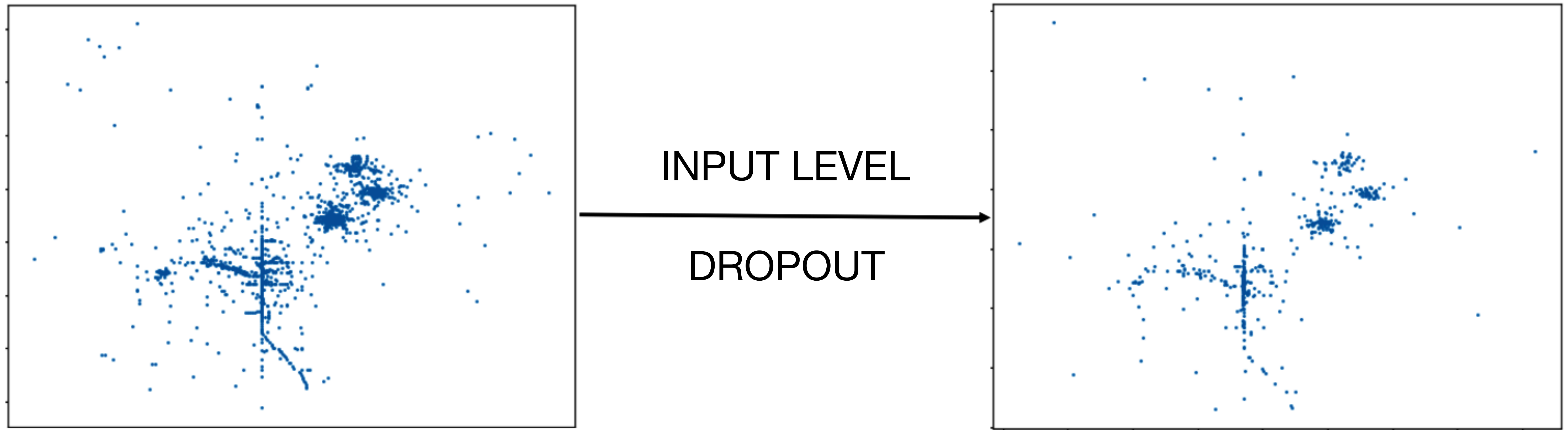
- Classification task:
  - 8-classes: 7 tau decays + QCD jets
  - training/validation/test sets: 22k/6k/7k events (balanced among classes)
- Data-preprocessing:
  - simple geometrical clustering, no specific selection or fiducial volume applied
  - saved fibres signal around each cluster ($\sqrt{\Delta\theta^2+\Delta\phi^2}$) < 1)
- DGCNN inputs:
  - jet/tau representation: 2D point-cloud of fibres coordinates
  - fiber type (S, C), #photo-electrons, SiPM's: Integral and Peak of the SiPM output, ToT, ToA, ToP (in different combinations)

S.Giagu, M. di Filippo, L.Torresi, Tau leptons identification with Graph Neural Networks at future electron-positron colliders, Front. Phys., Volume 10 - 2022

DATA AUGMENTATION VIA DROPOUT

- overfitting and memorisation for the DGNN model controlled using two different implementations of the dropout regularisation:
    - (traditional) in the network: some of the parameters of the last MLP block are randomly zeroed during the training phase
    - (as data augmentation/perturbation regularisation) at input level: some of the fired fibers are switched off before to be exposed to the network

- much better generalisation obtained leveraging both methods
- dropout levels optimised on validation set



INPUT LEVEL

DROPOUT

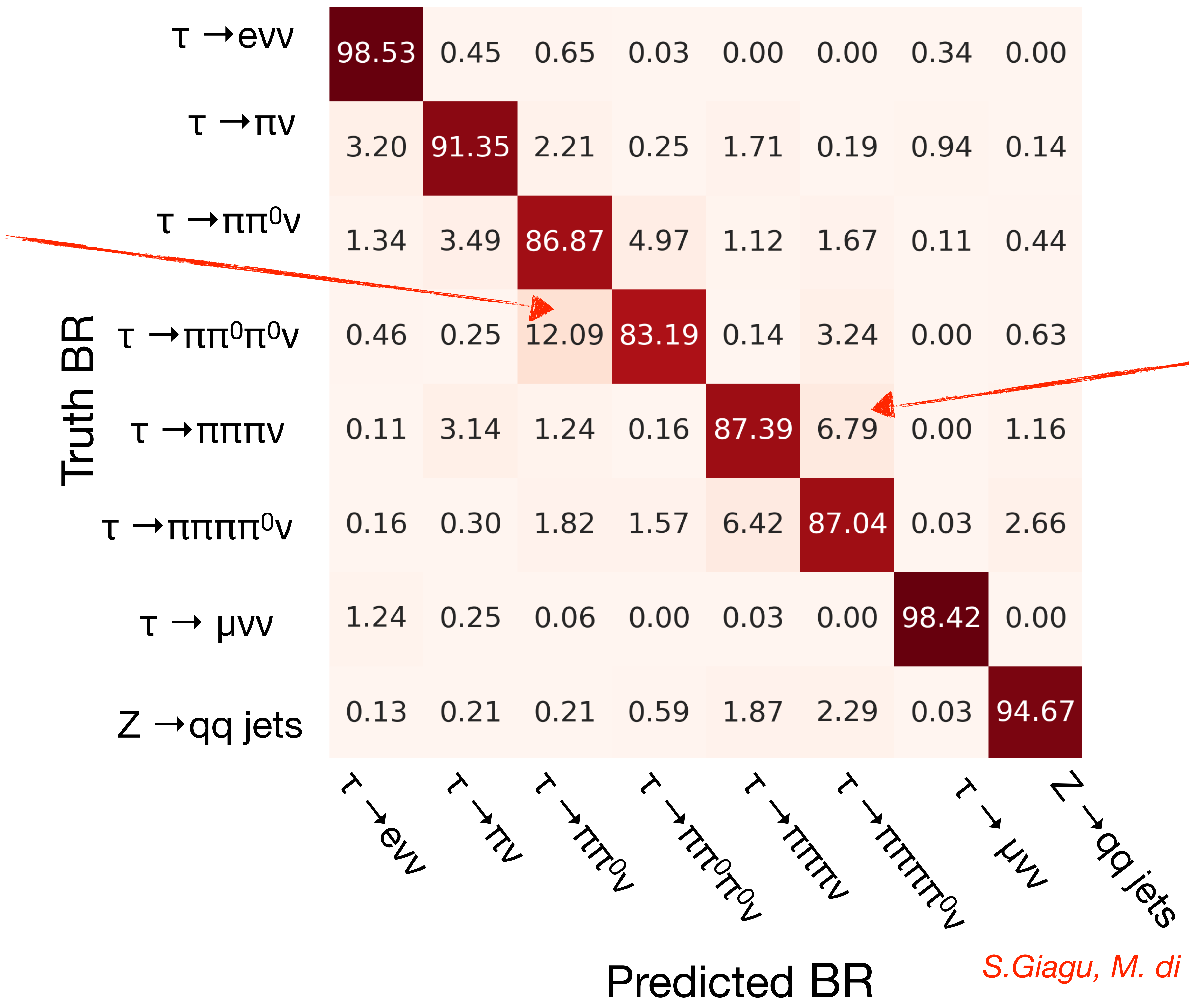
S.Giagu, M. di Filippo, L.Torresi, Front. Phys., Volume 10 - 2022

IDEAL PERFORMANCE ON TEST SET

8-class classification task w/ DGCNN

input features: **coordinates, type** of fibre (S/C), and **#of photo-electrons** in each fibre

small confusion only within topologically similar decays



	$\tau \to e\nu\nu$	$\tau \to \pi\nu$	$\tau \to \pi\pi^0\nu$	$\tau \to \pi\pi^0\pi^0\nu$	$\tau \to \pi\pi\pi\nu$	$\tau \to \pi\pi\pi\pi^0\nu$	$\tau \to \mu\nu\nu$	$Z \to qq$ jets
$\tau \to e\nu\nu$	98.53	0.45	0.65	0.03	0.00	0.00	0.34	0.00
$\tau \to \pi\nu$	3.20	91.35	2.21	0.25	1.71	0.19	0.94	0.14
$\tau \to \pi\pi^0\nu$	1.34	3.49	86.87	4.97	1.12	1.67	0.11	0.44
$\tau \to \pi\pi^0\pi^0\nu$	0.46	0.25	12.09	83.19	0.14	3.24	0.00	0.63
$\tau \to \pi\pi\pi\nu$	0.11	3.14	1.24	0.16	87.39	6.79	0.00	1.16
$\tau \to \pi\pi\pi\pi^0\nu$	0.16	0.30	1.82	1.57	6.42	87.04	0.03	2.66
$\tau \to \mu\nu\nu$	1.24	0.25	0.06	0.00	0.03	0.00	98.42	0.00
$Z \to qq$ jets	0.13	0.21	0.21	0.59	1.87	2.29	0.03	94.67

average accuracy:
90.8%

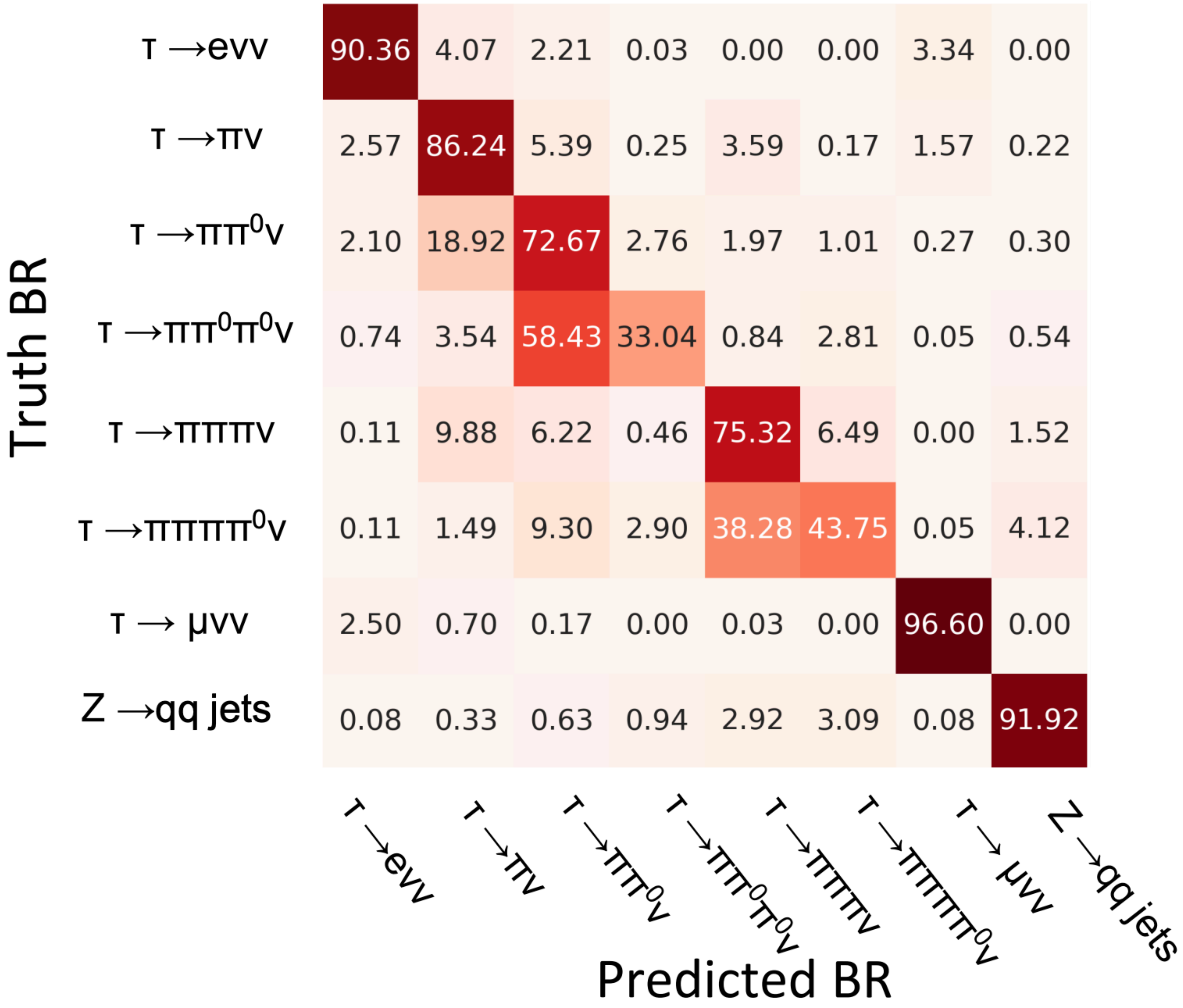
small confusion only within topologically similar decays

stat. uncertainty on accuracies ~3÷5%

S.Giagu, M. di Filippo, L.Torresi, Front. Phys., Volume 10 - 2022

PERFORMANCE ONLY USING GEOMETRICAL OR GEOM. + FIBER TYPE INFORMATION

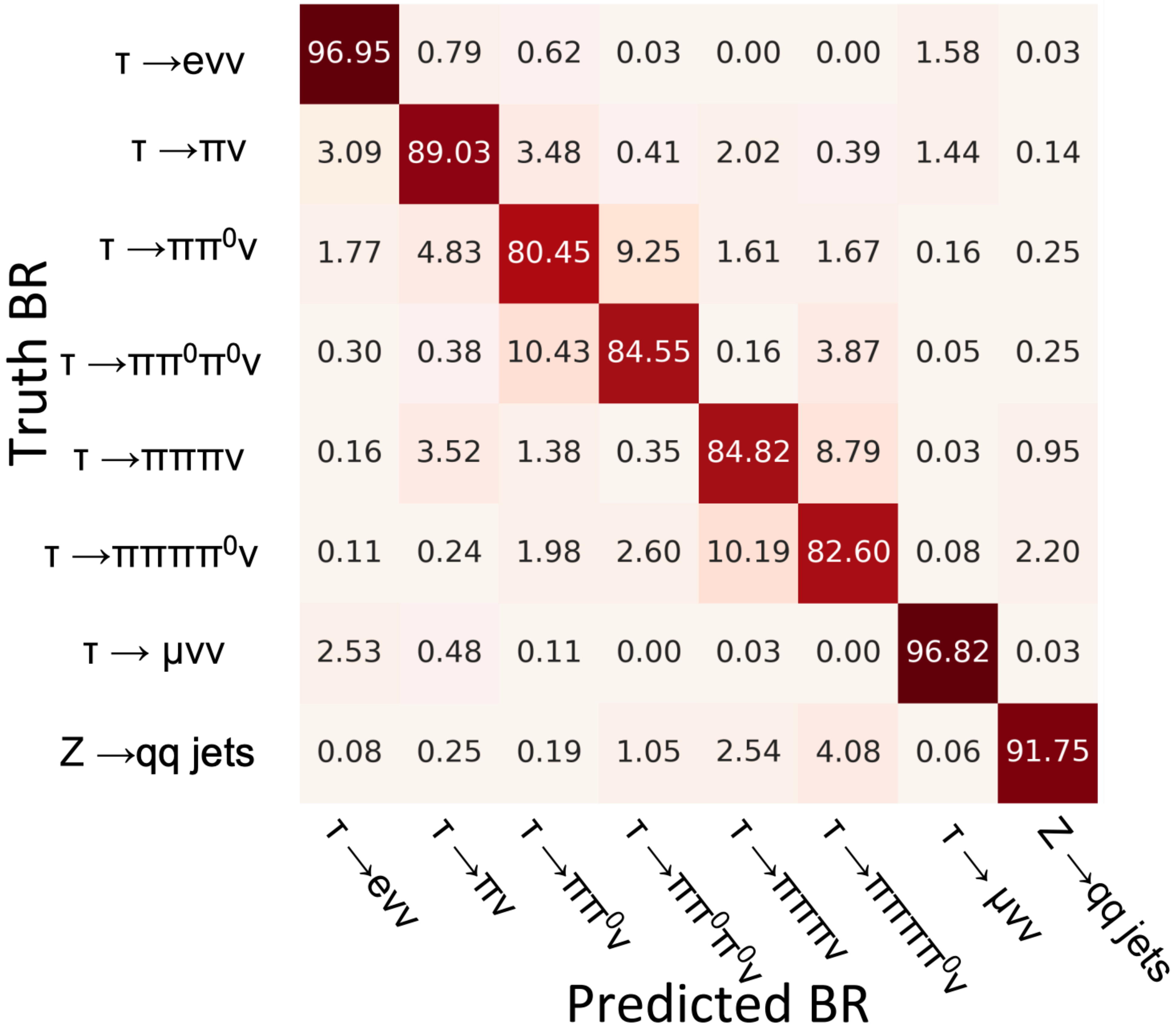
input features: **coordinates only**



	$\tau \rightarrow e\nu\nu$	$\tau \rightarrow \pi\nu$	$\tau \rightarrow \pi\pi^0\nu$	$\tau \rightarrow \pi\pi^0\pi^0\nu$	$\tau \rightarrow \pi\pi\pi\nu$	$\tau \rightarrow \pi\pi\pi\pi^0\nu$	$\tau \rightarrow \mu\nu\nu$	$Z \rightarrow qq$ jets
$\tau \rightarrow e\nu\nu$	90.36	4.07	2.21	0.03	0.00	0.00	3.34	0.00
$\tau \rightarrow \pi\nu$	2.57	86.24	5.39	0.25	3.59	0.17	1.57	0.22
$\tau \rightarrow \pi\pi^0\nu$	2.10	18.92	72.67	2.76	1.97	1.01	0.27	0.30
$\tau \rightarrow \pi\pi^0\pi^0\nu$	0.74	3.54	58.43	33.04	0.84	2.81	0.05	0.54
$\tau \rightarrow \pi\pi\pi\nu$	0.11	9.88	6.22	0.46	75.32	6.49	0.00	1.52
$\tau \rightarrow \pi\pi\pi\pi^0\nu$	0.11	1.49	9.30	2.90	38.28	43.75	0.05	4.12
$\tau \rightarrow \mu\nu\nu$	2.50	0.70	0.17	0.00	0.03	0.00	96.60	0.00
$Z \rightarrow qq$ jets	0.08	0.33	0.63	0.94	2.92	3.09	0.08	91.92

average accuracy: 73.7%

coordinates + type of fibre (S/C)



	$\tau \rightarrow e\nu\nu$	$\tau \rightarrow \pi\nu$	$\tau \rightarrow \pi\pi^0\nu$	$\tau \rightarrow \pi\pi^0\pi^0\nu$	$\tau \rightarrow \pi\pi\pi\nu$	$\tau \rightarrow \pi\pi\pi\pi^0\nu$	$\tau \rightarrow \mu\nu\nu$	$Z \rightarrow qq$ jets
$\tau \rightarrow e\nu\nu$	96.95	0.79	0.62	0.03	0.00	0.00	1.58	0.03
$\tau \rightarrow \pi\nu$	3.09	89.03	3.48	0.41	2.02	0.39	1.44	0.14
$\tau \rightarrow \pi\pi^0\nu$	1.77	4.83	80.45	9.25	1.61	1.67	0.16	0.25
$\tau \rightarrow \pi\pi^0\pi^0\nu$	0.30	0.38	10.43	84.55	0.16	3.87	0.05	0.25
$\tau \rightarrow \pi\pi\pi\nu$	0.16	3.52	1.38	0.35	84.82	8.79	0.03	0.95
$\tau \rightarrow \pi\pi\pi\pi^0\nu$	0.11	0.24	1.98	2.60	10.19	82.60	0.08	2.20
$\tau \rightarrow \mu\nu\nu$	2.53	0.48	0.11	0.00	0.03	0.00	96.82	0.03
$Z \rightarrow qq$ jets	0.08	0.25	0.19	1.05	2.54	4.08	0.06	91.75

average accuracy: 88.3%

double-readout geometry alone allows excellent tau identification

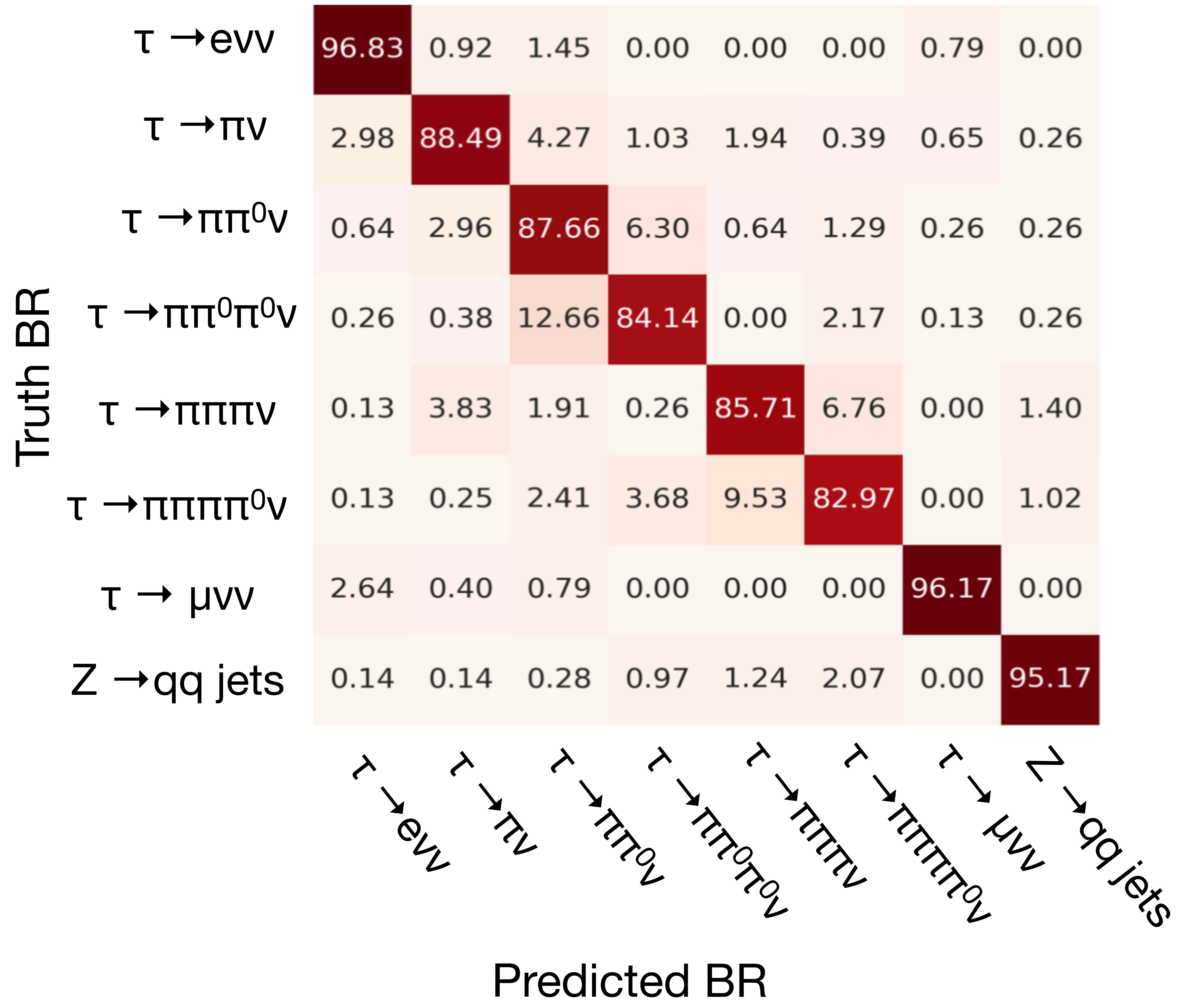
PERFORMANCE USING REALISTIC SiPM READOUT INFORMATION

using only geometry and Integral/Peak of the signal

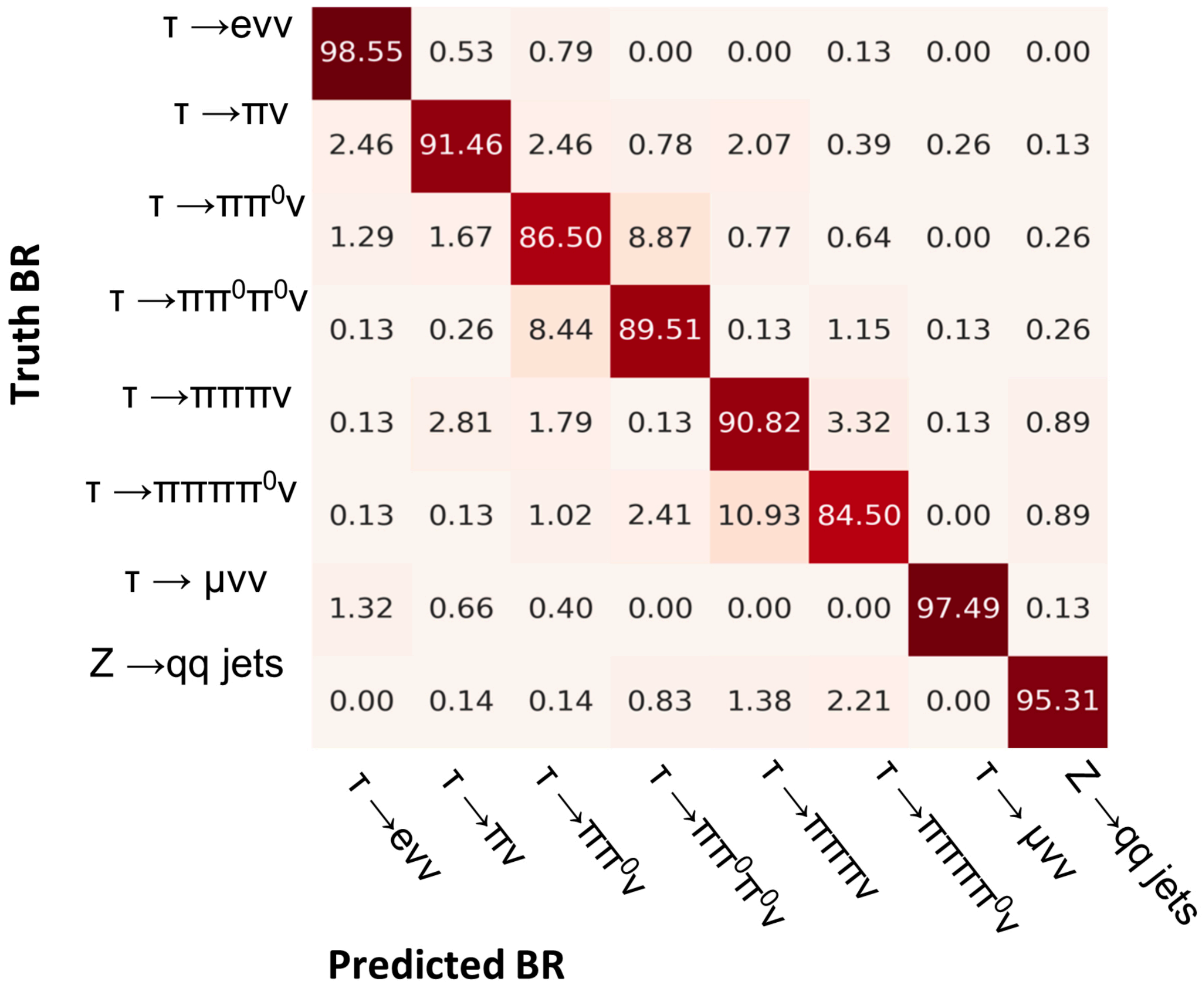
average accuracy: **88.8%**

adding also SiPM timing information

average accuracy: **91.4%**



	$\tau \rightarrow e\nu\nu$	$\tau \rightarrow \pi\nu$	$\tau \rightarrow \pi\pi^0\nu$	$\tau \rightarrow \pi\pi^0\pi^0\nu$	$\tau \rightarrow \pi\pi\pi\nu$	$\tau \rightarrow \pi\pi\pi\pi^0\nu$	$\tau \rightarrow \mu\nu\nu$	$Z \rightarrow qq$ jets
$\tau \rightarrow e\nu\nu$	96.83	0.92	1.45	0.00	0.00	0.00	0.79	0.00
$\tau \rightarrow \pi\nu$	2.98	88.49	4.27	1.03	1.94	0.39	0.65	0.26
$\tau \rightarrow \pi\pi^0\nu$	0.64	2.96	87.66	6.30	0.64	1.29	0.26	0.26
$\tau \rightarrow \pi\pi^0\pi^0\nu$	0.26	0.38	12.66	84.14	0.00	2.17	0.13	0.26
$\tau \rightarrow \pi\pi\pi\nu$	0.13	3.83	1.91	0.26	85.71	6.76	0.00	1.40
$\tau \rightarrow \pi\pi\pi\pi^0\nu$	0.13	0.25	2.41	3.68	9.53	82.97	0.00	1.02
$\tau \rightarrow \mu\nu\nu$	2.64	0.40	0.79	0.00	0.00	0.00	96.17	0.00
$Z \rightarrow qq$ jets	0.14	0.14	0.28	0.97	1.24	2.07	0.00	95.17



	$\tau \rightarrow e\nu\nu$	$\tau \rightarrow \pi\nu$	$\tau \rightarrow \pi\pi^0\nu$	$\tau \rightarrow \pi\pi^0\pi^0\nu$	$\tau \rightarrow \pi\pi\pi\nu$	$\tau \rightarrow \pi\pi\pi\pi^0\nu$	$\tau \rightarrow \mu\nu\nu$	$Z \rightarrow qq$ jets
$\tau \rightarrow e\nu\nu$	98.55	0.53	0.79	0.00	0.00	0.13	0.00	0.00
$\tau \rightarrow \pi\nu$	2.46	91.46	2.46	0.78	2.07	0.39	0.26	0.13
$\tau \rightarrow \pi\pi^0\nu$	1.29	1.67	86.50	8.87	0.77	0.64	0.00	0.26
$\tau \rightarrow \pi\pi^0\pi^0\nu$	0.13	0.26	8.44	89.51	0.13	1.15	0.13	0.26
$\tau \rightarrow \pi\pi\pi\nu$	0.13	2.81	1.79	0.13	90.82	3.32	0.13	0.89
$\tau \rightarrow \pi\pi\pi\pi^0\nu$	0.13	0.13	1.02	2.41	10.93	84.50	0.00	0.89
$\tau \rightarrow \mu\nu\nu$	1.32	0.66	0.40	0.00	0.00	0.00	97.49	0.13
$Z \rightarrow qq$ jets	0.00	0.14	0.14	0.83	1.38	2.21	0.00	95.31

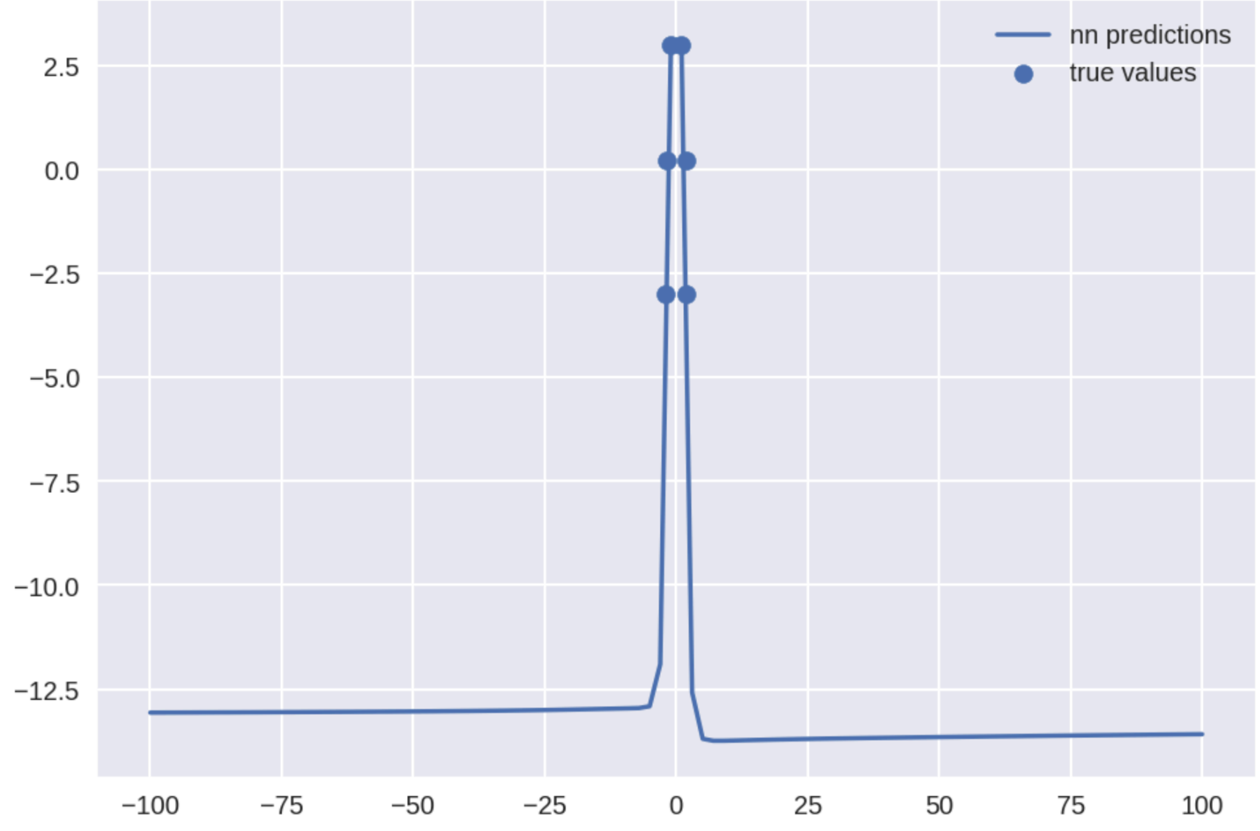
comparable identification performance w/r the ideal case

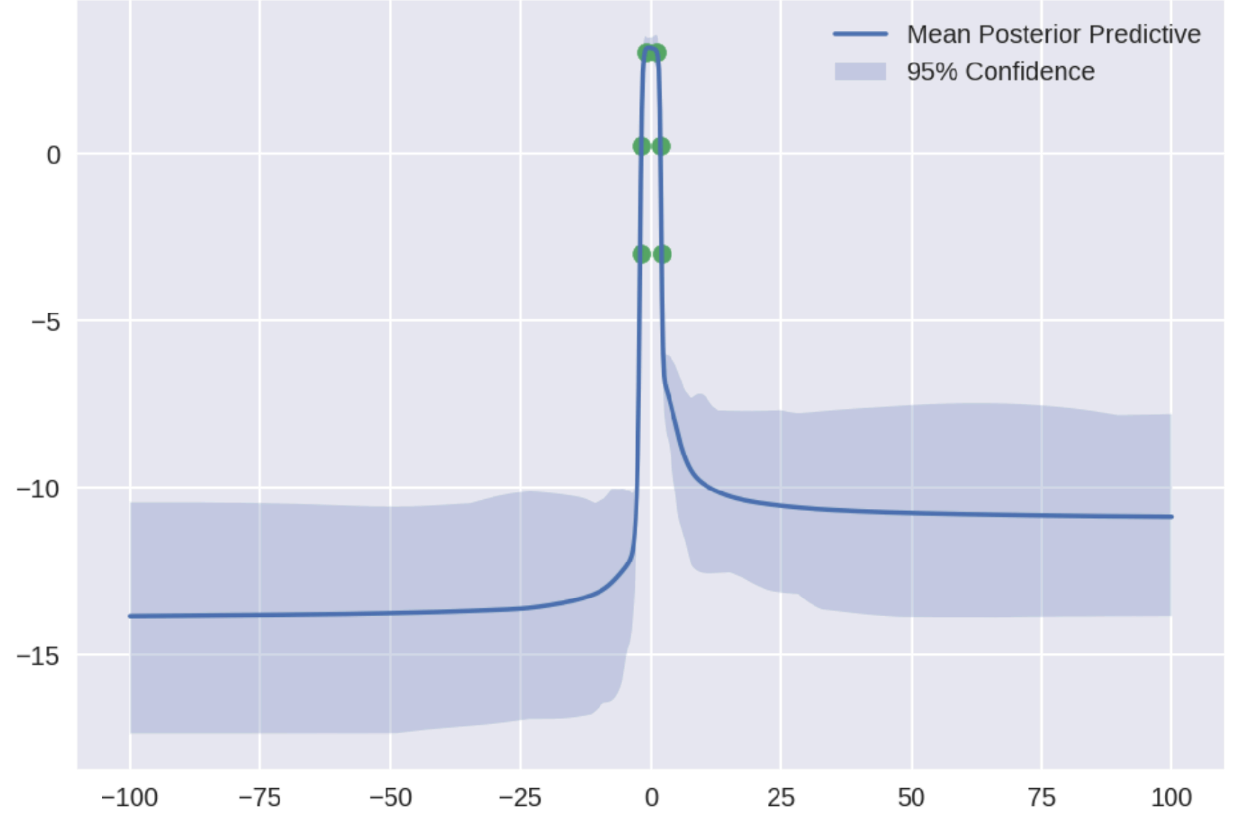
CHECK OF POSSIBLE BIAS ON ENERGY

no bias observed over distribution of total energy per event

UNCERTAINTY IN THE CLASSIFICATION: BAYESIAN-DGCNN

- Neural networks based on point values for weights may suffer of overconfidence when analysing new data especially for predictions in regions with few or w/o examples in the training set
- Bayesian neural networks mitigate the problem by introducing probability distributions over the weights and **predicting distributions instead of point values**
  - a Bayesian-NN learns a variational approximation of the true posterior distribution P(w|D), and predict an estimate of the expected value $E_{P(w|D)}[P(y|x,w)]$ → since the weights are random variables, each predictions is a random variable too
  - allows to measure uncertainty, identify outliers in the input, regularise the whole model
- Designed and implemented in pytorch a full Bayesian version of a DGCNN (leveraging the *Bayes by Backprop* algorithm (https://arxiv.org/abs/1505.05424):





$$p(y\,|\,x, D) = \int p(x\,|\,y, w)p(w\,|\,D)dw \quad \Rightarrow \quad p(w\,|\,D) \approx q_\theta(w\,|\,D)$$

bayesian inference

generally intractable via MC integration

replace it with a variational (eg optimisation) problem

approximate p with a more tractable parametric distribution q (eg. gaussian) w/ learnable parameters (eg a NN)

BAYES BY BACKPROP

find optimal $\theta*$ by minimising the
Kullback-Leibler divergence between p and q

$$\theta* = \arg\min_{\theta} \mathrm{KL}[q_\theta(w \mid D) \| p(w \mid D)]$$

$$\mathrm{KL}[q_\theta(w \mid D) \| p(w \mid D)] = \int q_\theta(w \mid D) \log \frac{q_\theta(w \mid D)}{p(w \mid D)} dw$$

we have another integral here, but now q is more
tractable and we can approximated it via MC sampling

➡

$$\theta* \approx \arg\min_{\theta} \sum_{i=1}^{n} \log q_\theta(w^{(i)} \mid D) - \log p(w^{(i)}) - \log p(D \mid w^{(i)})$$

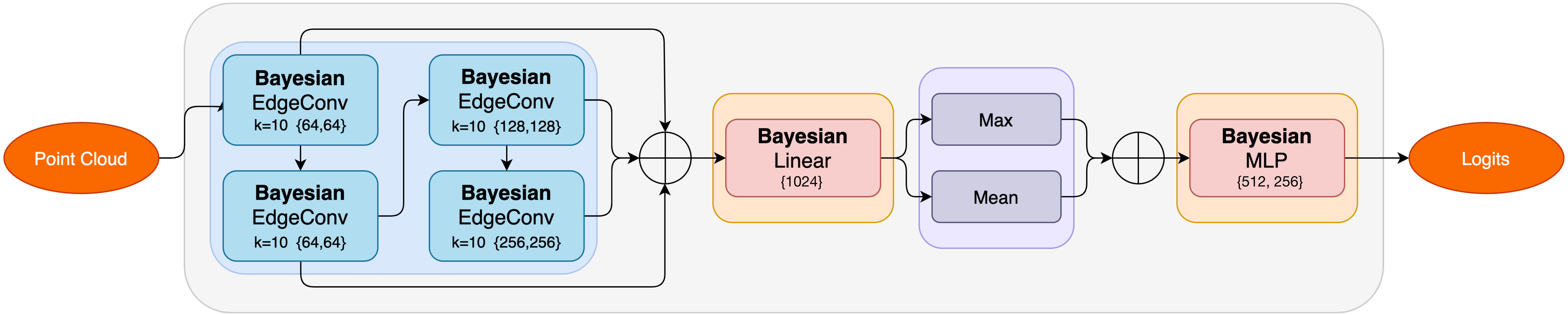
with $w^{(i)}$ samples sampled from $q_\theta(w \mid D)$

in practice as q_θ is an ANN, the Kingma local reparameterization trick is used to make the whole expression differentiable

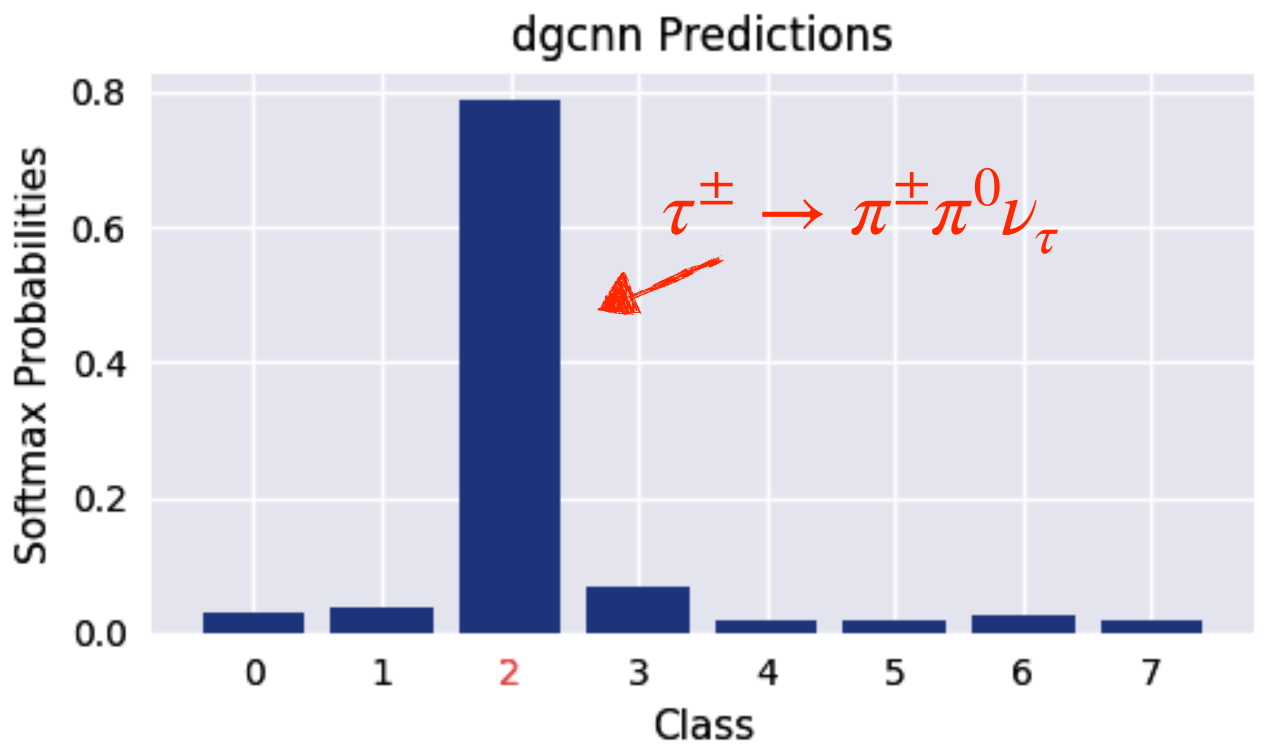
Blundell et al., arXiv:1505.05424

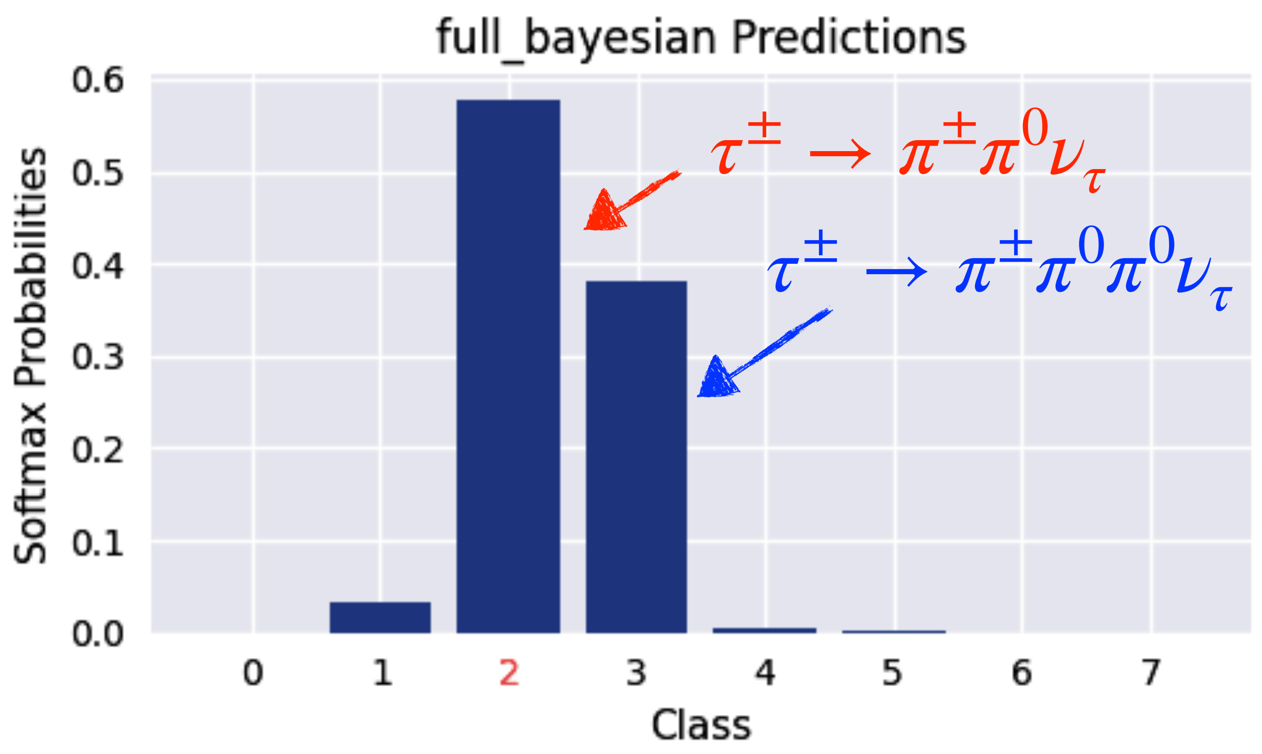
S.Giagu, M. di Filippo, L.Torresi, Front. Phys., Volume 10 - 2022

BAYESIAN-DGCNN

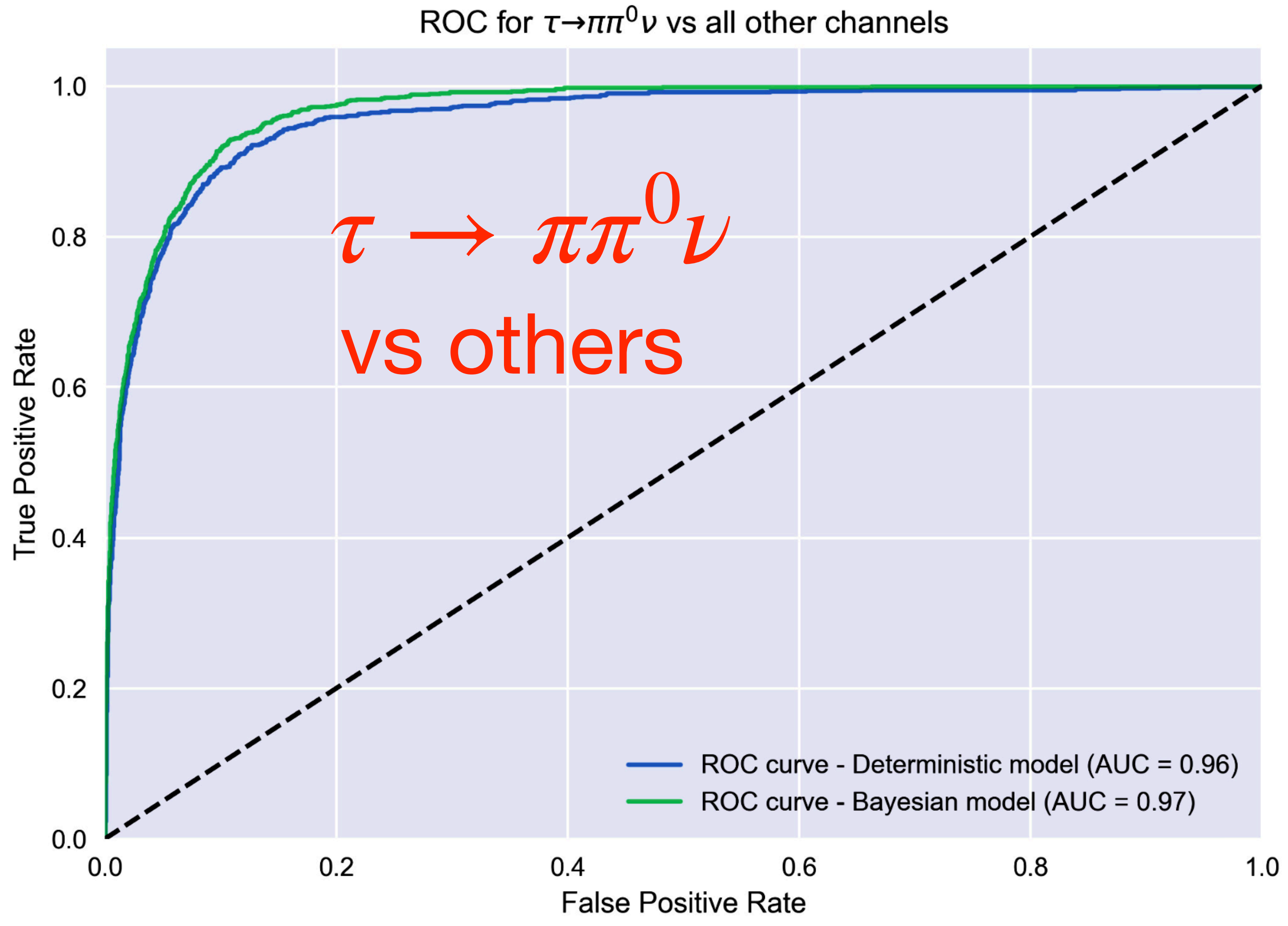


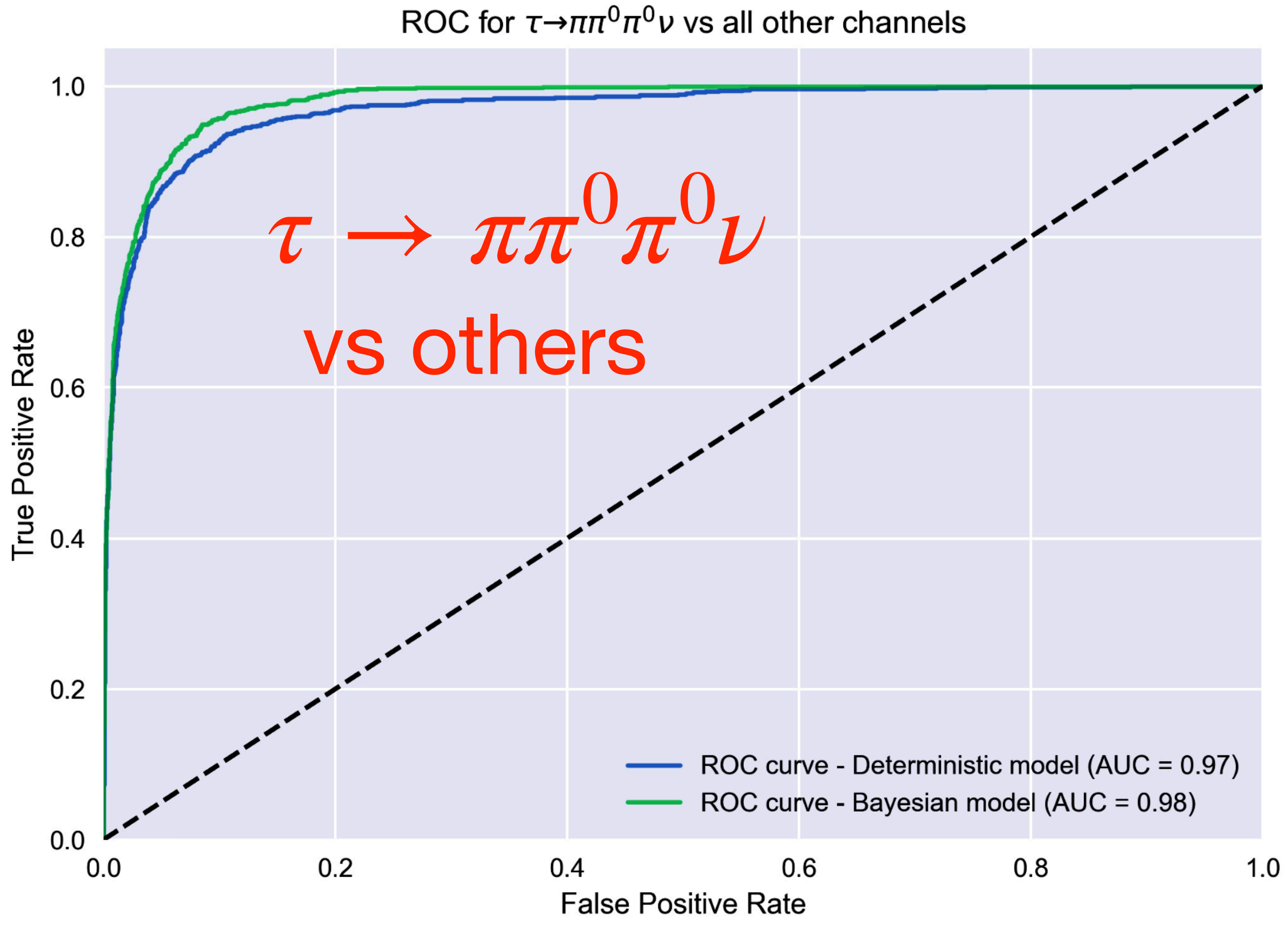
- all bayesian layers (EdgeConv, MLP, etc.), w/ gaussian priors (uncorrelated between layers and neurons)
- better classification performance wrt the point DGCNN
- class probabilities better aligned with physics expectations

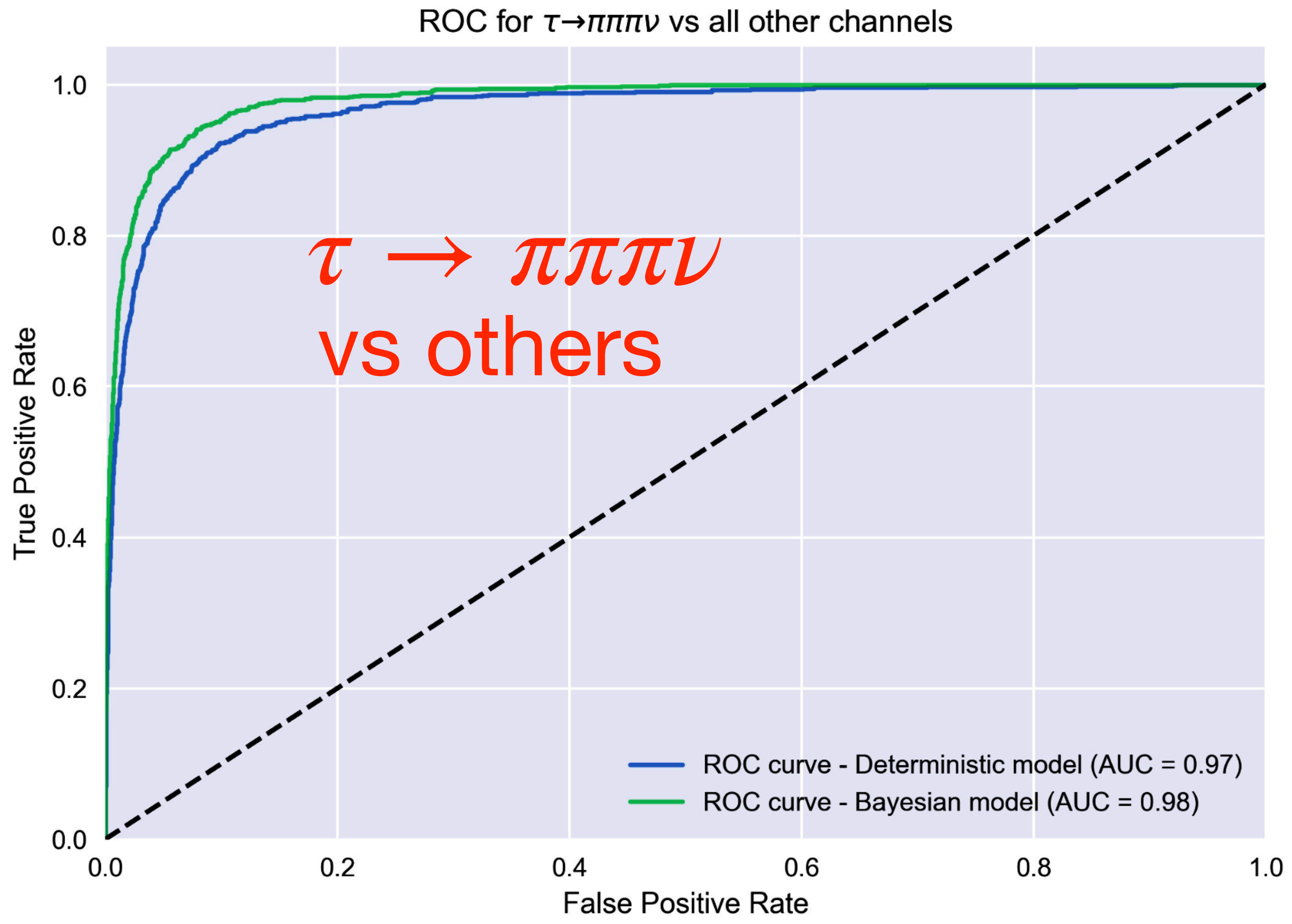


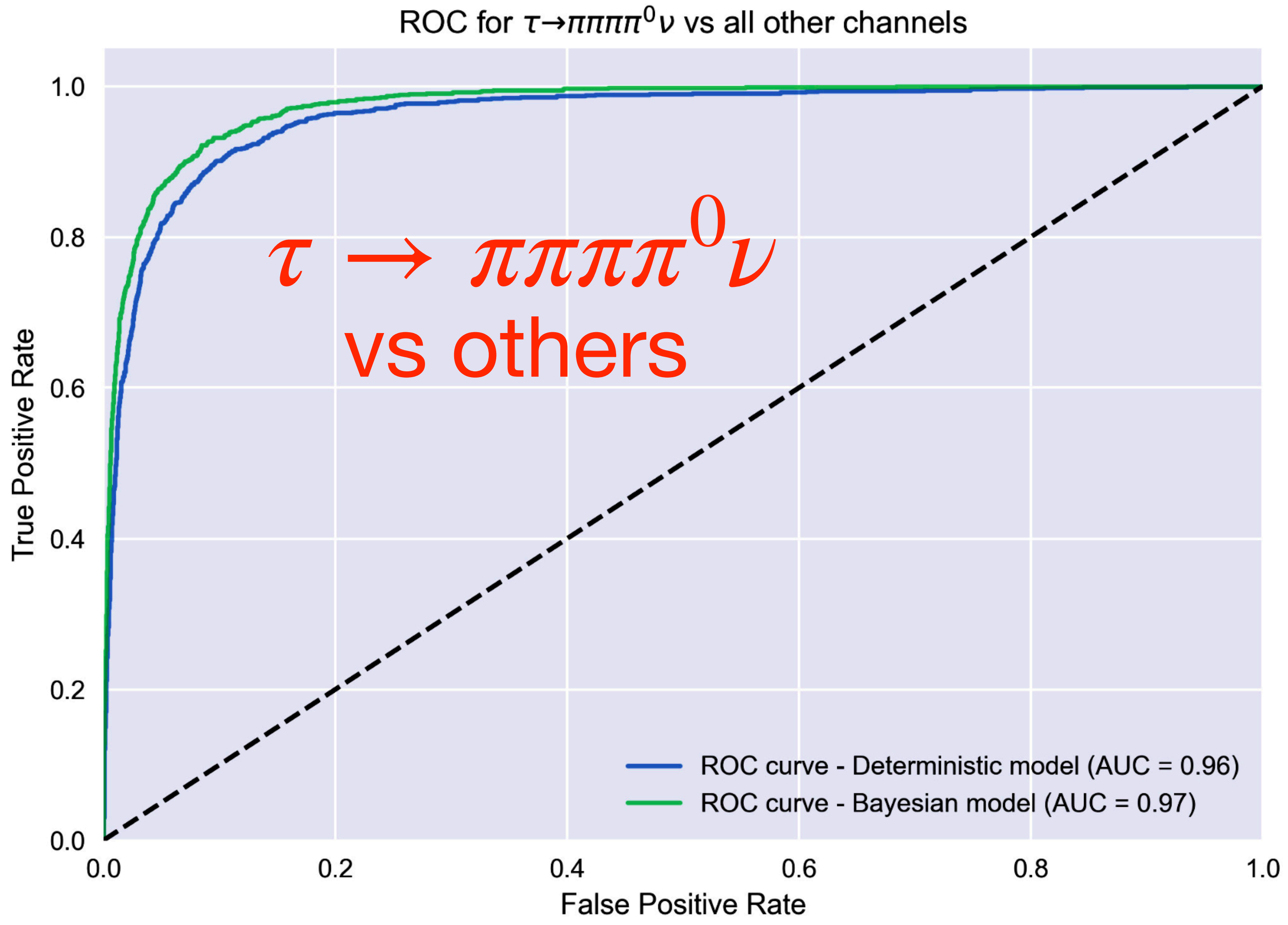


BAYESIAN-DGCNN


ROC for $\tau\to\pi\pi^0\nu$ vs all other channels

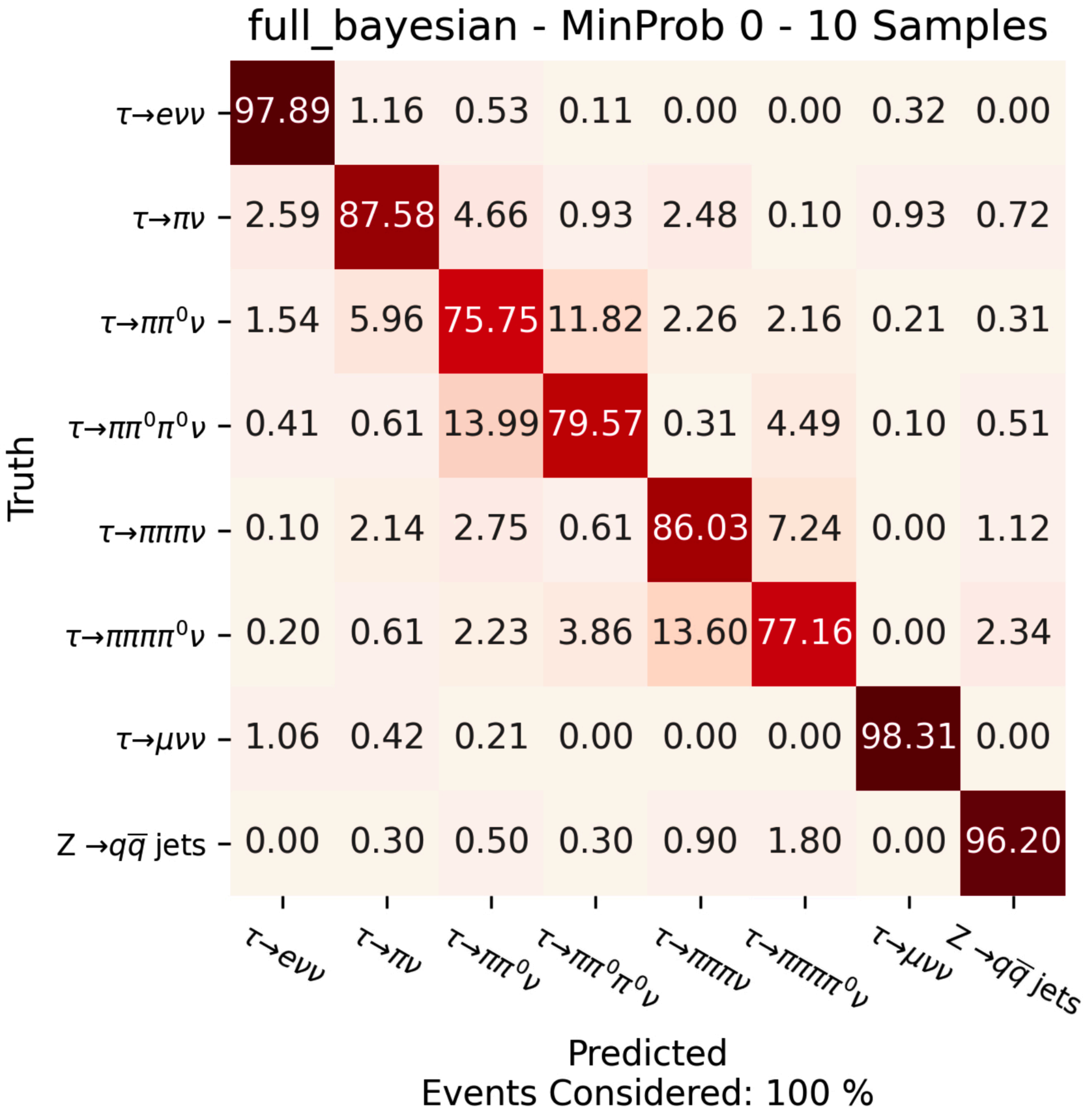

ROC for $\tau\to\pi\pi^0\pi^0\nu$ vs all other channels


ROC for $\tau\to\pi\pi\pi\nu$ vs all other channels


ROC for $\tau\to\pi\pi\pi\pi^0\nu$ vs all other channels

ROC / AUC

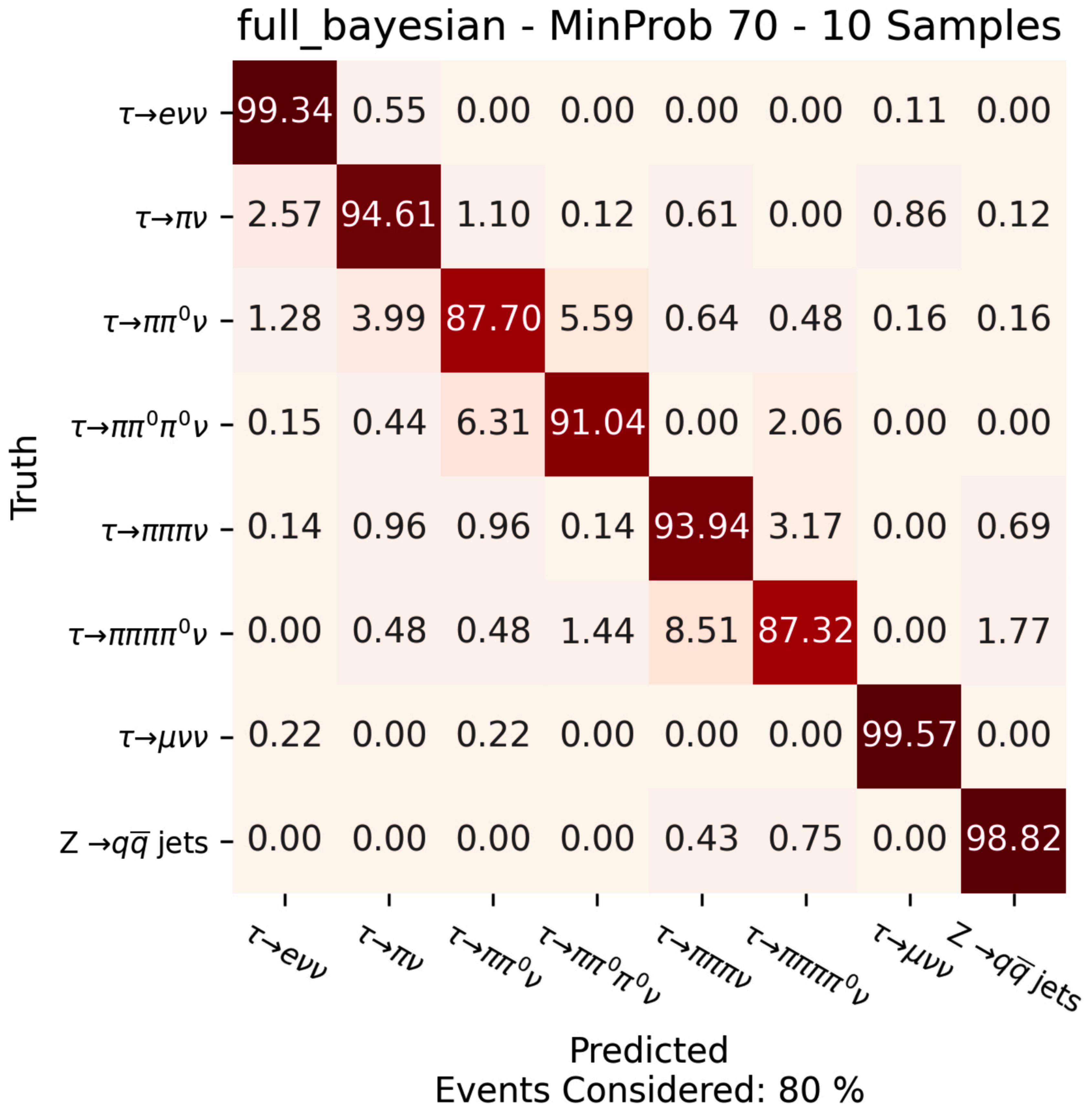
BAYESIAN-DGCNN


full_bayesian - MinProb 0 - 10 Samples

Truth	$\tau\to e\nu\nu$	$\tau\to\pi\nu$	$\tau\to\pi\pi^0\nu$	$\tau\to\pi\pi^0\pi^0\nu$	$\tau\to\pi\pi\pi\nu$	$\tau\to\pi\pi\pi\pi^0\nu$	$\tau\to\mu\nu\nu$	$Z\to q\bar{q}$ jets
$\tau\to e\nu\nu$	97.89	1.16	0.53	0.11	0.00	0.00	0.32	0.00
$\tau\to\pi\nu$	2.59	87.58	4.66	0.93	2.48	0.10	0.93	0.72
$\tau\to\pi\pi^0\nu$	1.54	5.96	75.75	11.82	2.26	2.16	0.21	0.31
$\tau\to\pi\pi^0\pi^0\nu$	0.41	0.61	13.99	79.57	0.31	4.49	0.10	0.51
$\tau\to\pi\pi\pi\nu$	0.10	2.14	2.75	0.61	86.03	7.24	0.00	1.12
$\tau\to\pi\pi\pi\pi^0\nu$	0.20	0.61	2.23	3.86	13.60	77.16	0.00	2.34
$\tau\to\mu\nu\nu$	1.06	0.42	0.21	0.00	0.00	0.00	98.31	0.00
$Z\to q\bar{q}$ jets	0.00	0.30	0.50	0.30	0.90	1.80	0.00	96.20

Predicted
Events Considered: 100 %

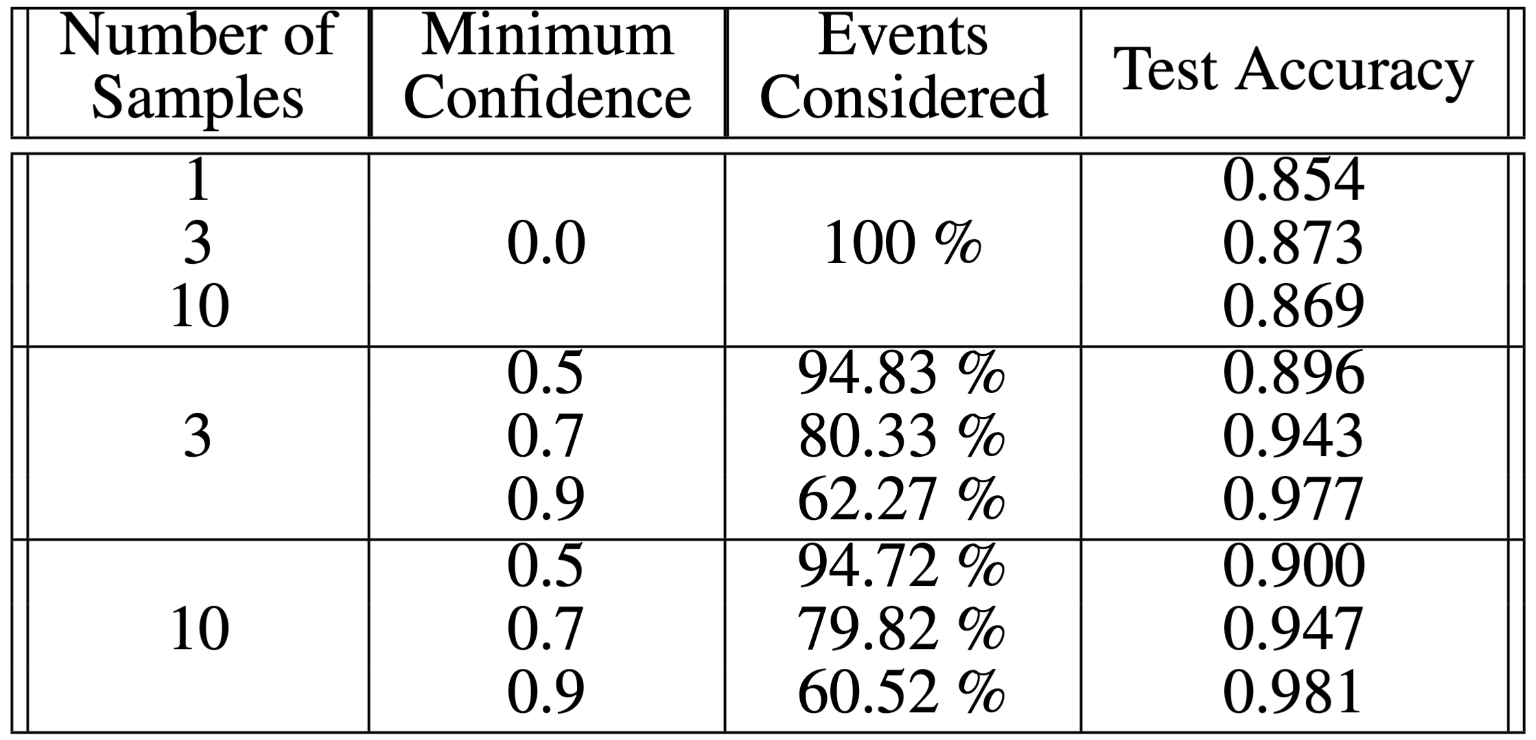
B-DGCNN:10 samples per prediction
no threshold on minimum confidence


full_bayesian - MinProb 70 - 10 Samples

Truth	$\tau\to e\nu\nu$	$\tau\to\pi\nu$	$\tau\to\pi\pi^0\nu$	$\tau\to\pi\pi^0\pi^0\nu$	$\tau\to\pi\pi\pi\nu$	$\tau\to\pi\pi\pi\pi^0\nu$	$\tau\to\mu\nu\nu$	$Z\to q\bar{q}$ jets
$\tau\to e\nu\nu$	99.34	0.55	0.00	0.00	0.00	0.00	0.11	0.00
$\tau\to\pi\nu$	2.57	94.61	1.10	0.12	0.61	0.00	0.86	0.12
$\tau\to\pi\pi^0\nu$	1.28	3.99	87.70	5.59	0.64	0.48	0.16	0.16
$\tau\to\pi\pi^0\pi^0\nu$	0.15	0.44	6.31	91.04	0.00	2.06	0.00	0.00
$\tau\to\pi\pi\pi\nu$	0.14	0.96	0.96	0.14	93.94	3.17	0.00	0.69
$\tau\to\pi\pi\pi\pi^0\nu$	0.00	0.48	0.48	1.44	8.51	87.32	0.00	1.77
$\tau\to\mu\nu\nu$	0.22	0.00	0.22	0.00	0.00	0.00	99.57	0.00
$Z\to q\bar{q}$ jets	0.00	0.00	0.00	0.00	0.43	0.75	0.00	98.82

Predicted
Events Considered: 80 %

B-DGCNN:10 samples per prediction
minimum threshold on confidence 0.7

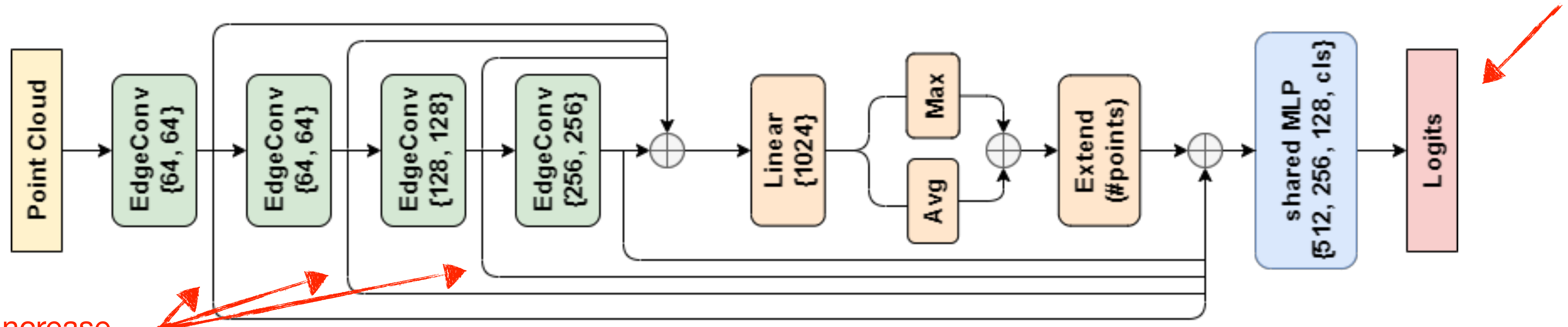


Number of Samples	Minimum Confidence	Events Considered	Test Accuracy
1 3 10	0.0	100 %	0.854 0.873 0.869
3	0.5 0.7 0.9	94.83 % 80.33 % 62.27 %	0.896 0.943 0.977
10	0.5 0.7 0.9	94.72 % 79.82 % 60.52 %	0.900 0.947 0.981

JET CONSTITUENTS ID

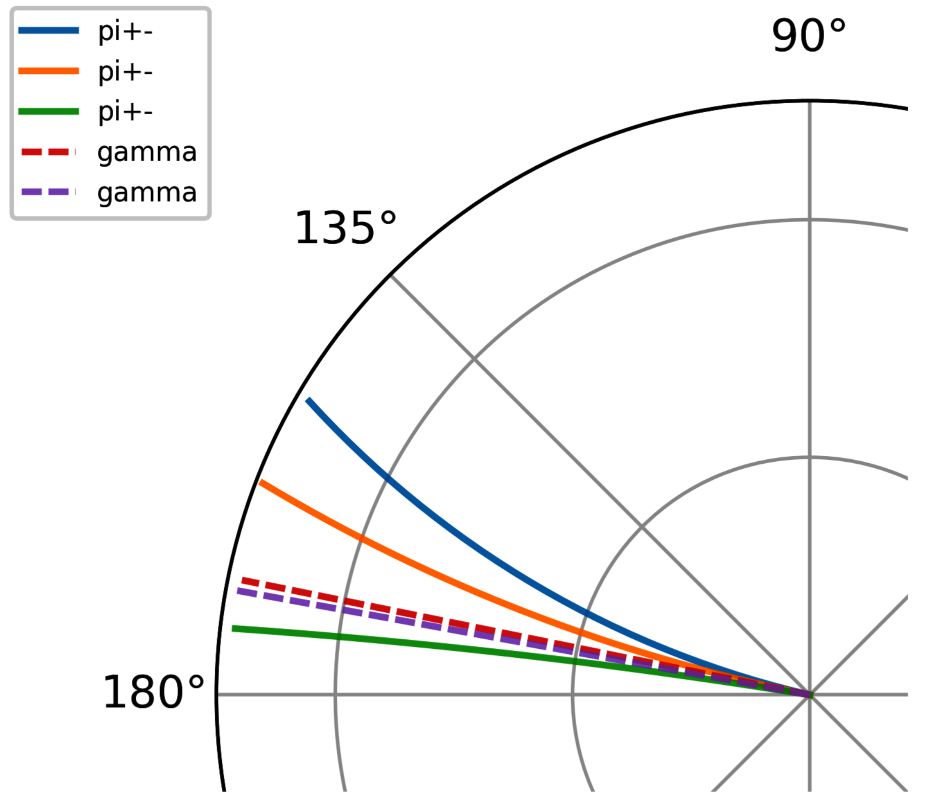
- DGCNN and dual-readout calorimeter high granularity can be exploited for object (particle) detection inside taus and jets
  - a proto-step for a particle flow algorithms
  - can be implemented with a similar approach as in segmentation tasks (eg pixel/node/fiber-level classification)
- DenseNet like modification of the DGCNN architecture for a segmentation task:

logits provided per-fibre (segmentation task)



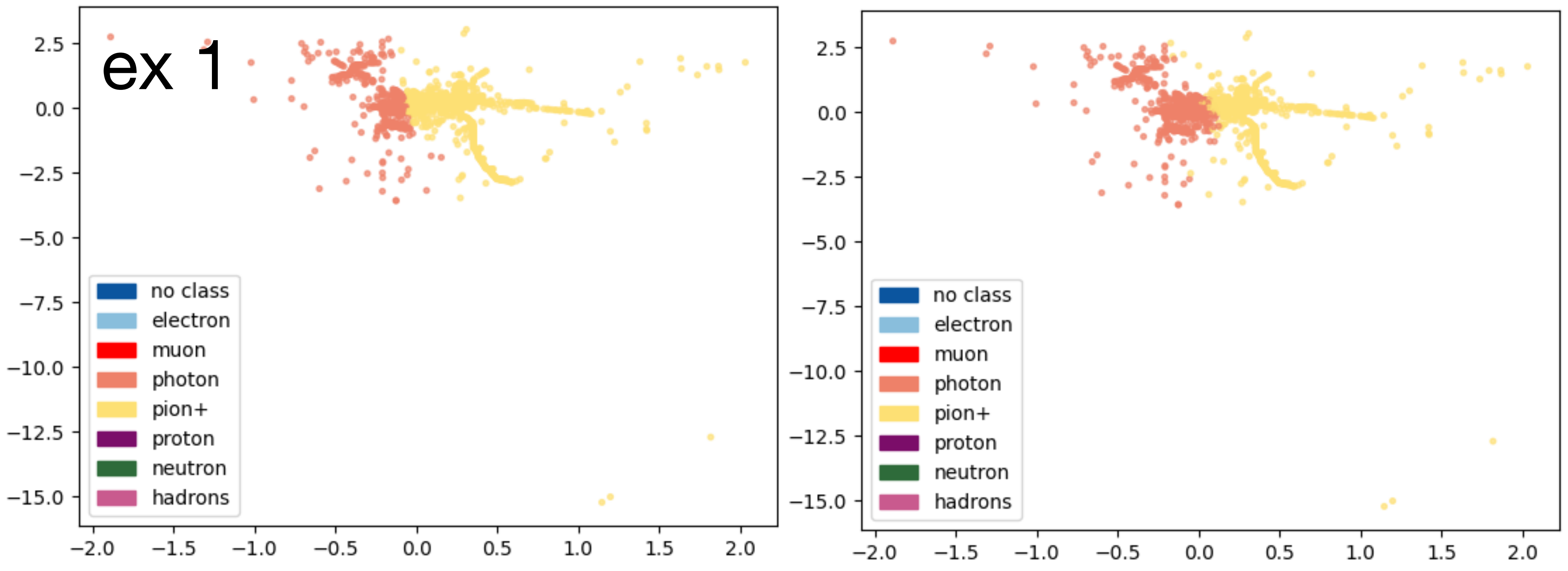
skip connections to increase sensitivity to multiscale features …

- identify the particle associated to the larger energy deposit in each fibre
- label each fibre by extrapolating Monte Carlo truth particles from production to the DRC into the IDEA magnetic field
- train the DGCNN to predict the label associated to each fibre
- **Initial study:** using point DGCNN only



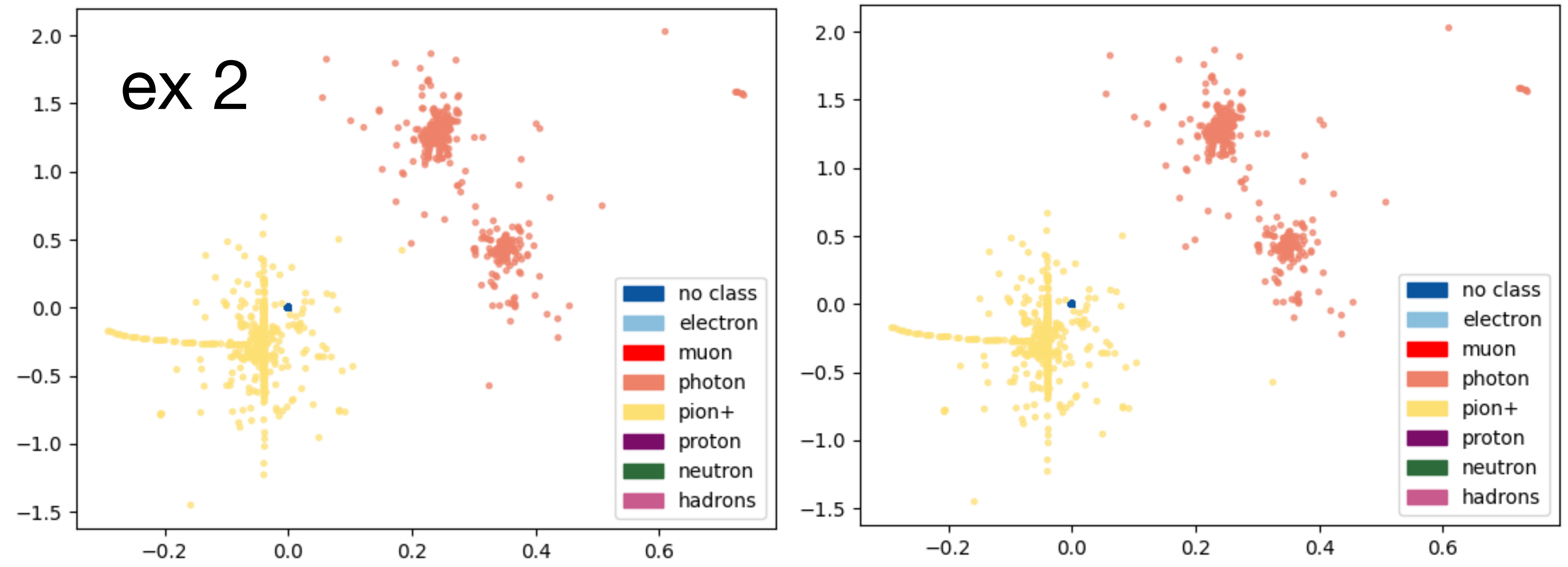
RESULTS SEGMENTATION

Two nice examples: $\tau^{\pm} \rightarrow \pi^{\pm}\pi^{0}\nu_{\tau} \rightarrow \pi^{\pm}\gamma\gamma\nu_{\tau}$



Ground Truth

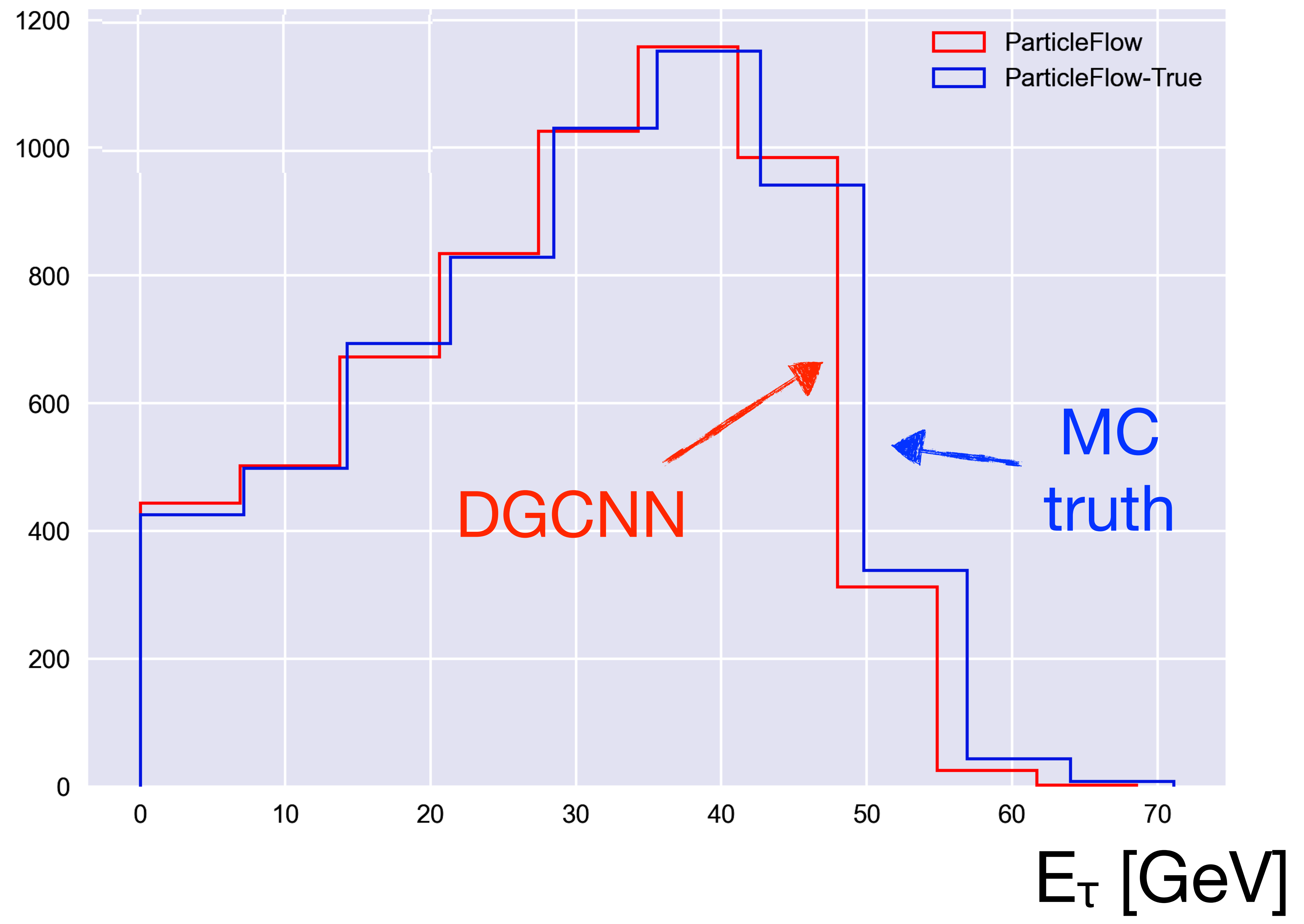
Reconstructed



tau visibile energy reconstructed using:

- DRC for photons
- MC truth for other particles

comparison of the visible energy distributions obtained when photons are identified by the DGCNN and when using the MC truth



RESULTS SEGMENTATION

Two “less nice” results:

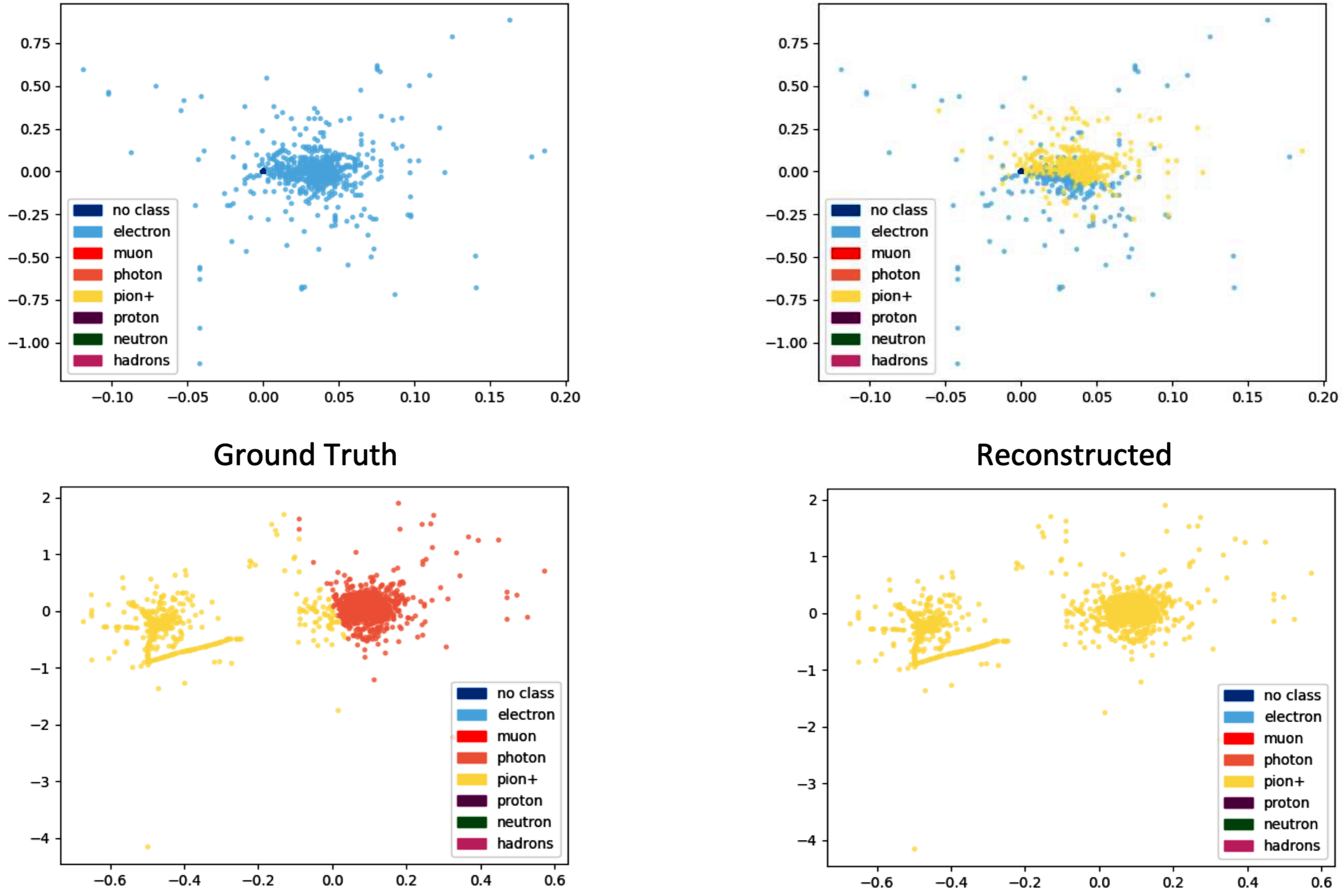
ex 1

$\tau \rightarrow e\nu_e\nu_\tau$

ex 2

$\tau \rightarrow \pi\pi^0\nu_\tau \rightarrow \pi\gamma\gamma\nu_\tau$

w/ overlapping photons



ONGOING DEVELOPMENTS

- improve jet constituents ID by moving from DGCNN segmentation to Graph-Transformer object detection
- hybrid architecture: GNN → high-level representation → Transformed encoder/decoder → bounding box predictions
  - GNN encode the graph for the transformed architecture and extracts compact representation of the global graph structure
  - the transformer encoder process this data to produce an embedding context representing the whole graph
  - this embedding is passed to a transformer decoder that takes as input a small fixed of learned positional embeddings (object queries) and attends to the encoder output
  - the self-attention and the encoder-decoder attention over the embedding and the object queries allows the model to analyse all objects together using pair-wise relations between them, and to independently decode it into box coordinates and class labels by a FFN head
  - the FFN head is a shared MLP that predict class and bounding box for each object
- results will be ready in time for the fall meetings/conferences …