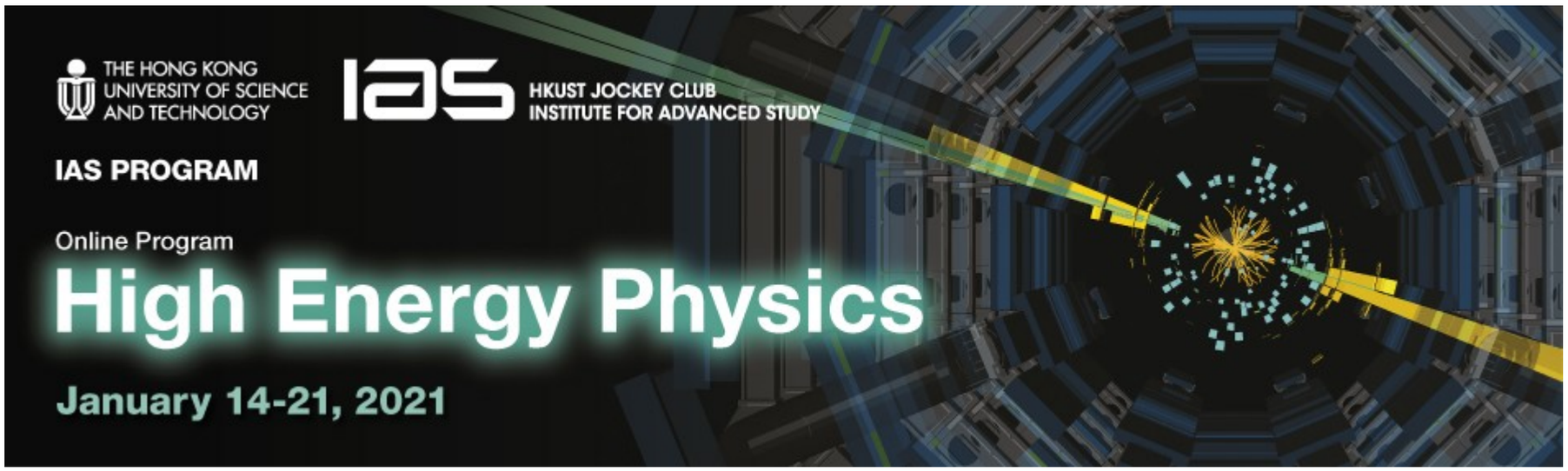
PID with Dual-Readout Calorimeters

Roberto Ferrari
on behalf of the IDEA proto-collaboration

January 15th, 2021



contents

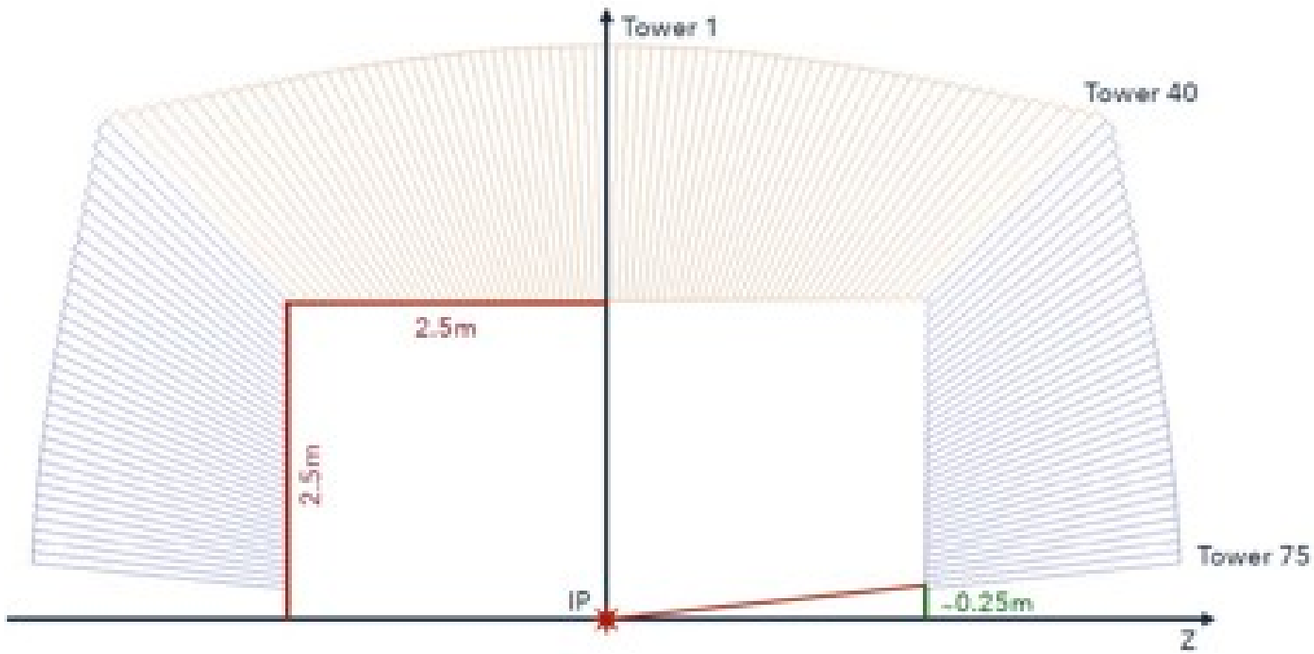
- *dual-readout fibre-sampling calorimetry*
- *e/π discrimination à la RD52*
- *τ-decay identification*
- *quark/gluon jet discrimination*

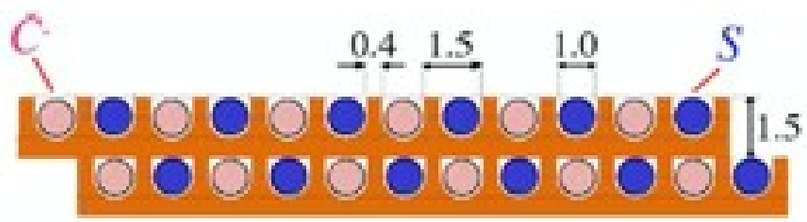
dual-readout fibre-sampling calorimetry

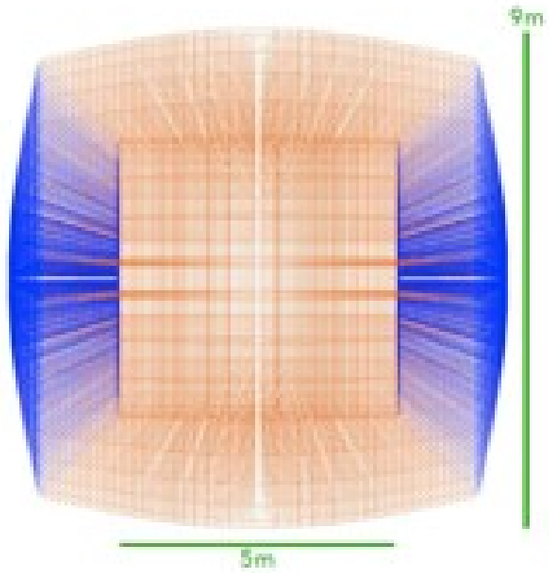
may provide:

- C & S information
- high granularity readout in transverse plane
- timing information for longitudinal development

IDEA calorimeter in Geant4

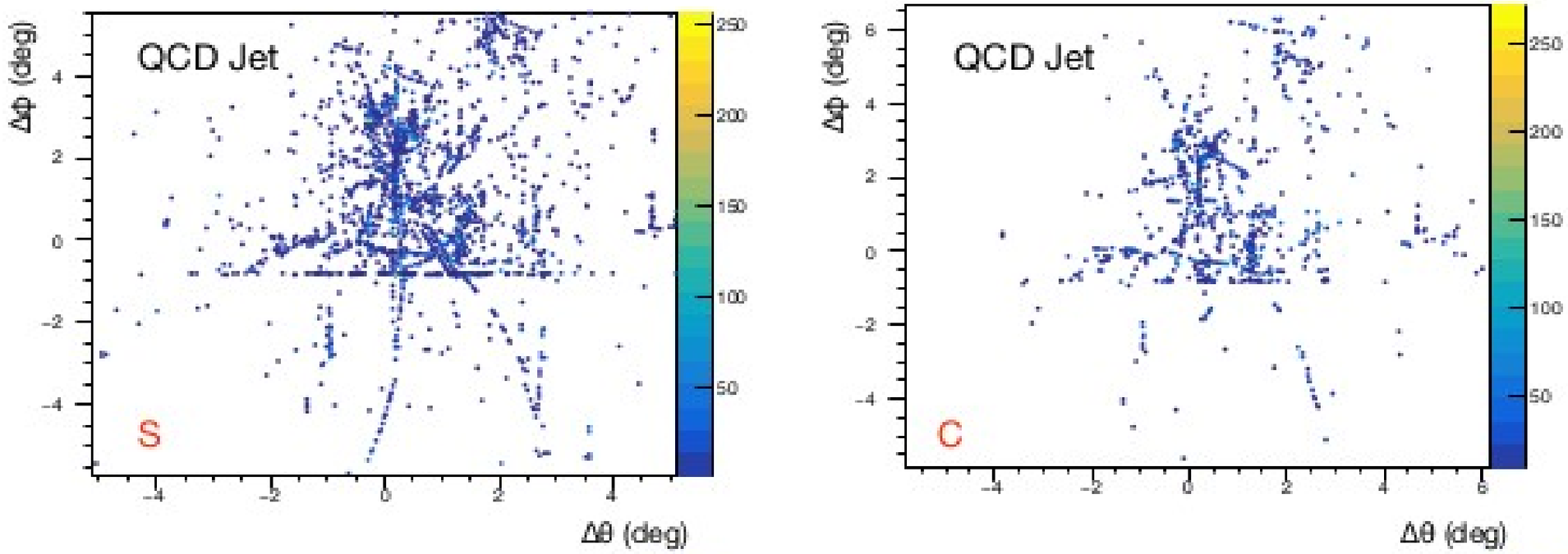




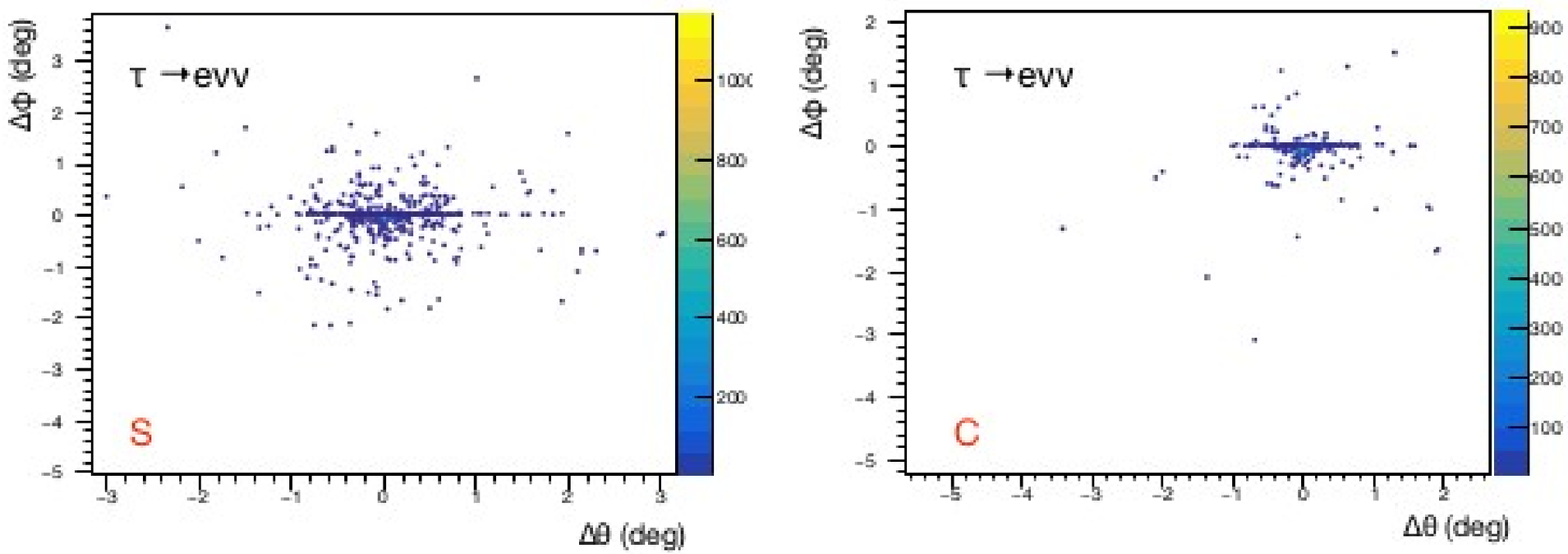


examples of events @ full granularity

Z→qq Jets

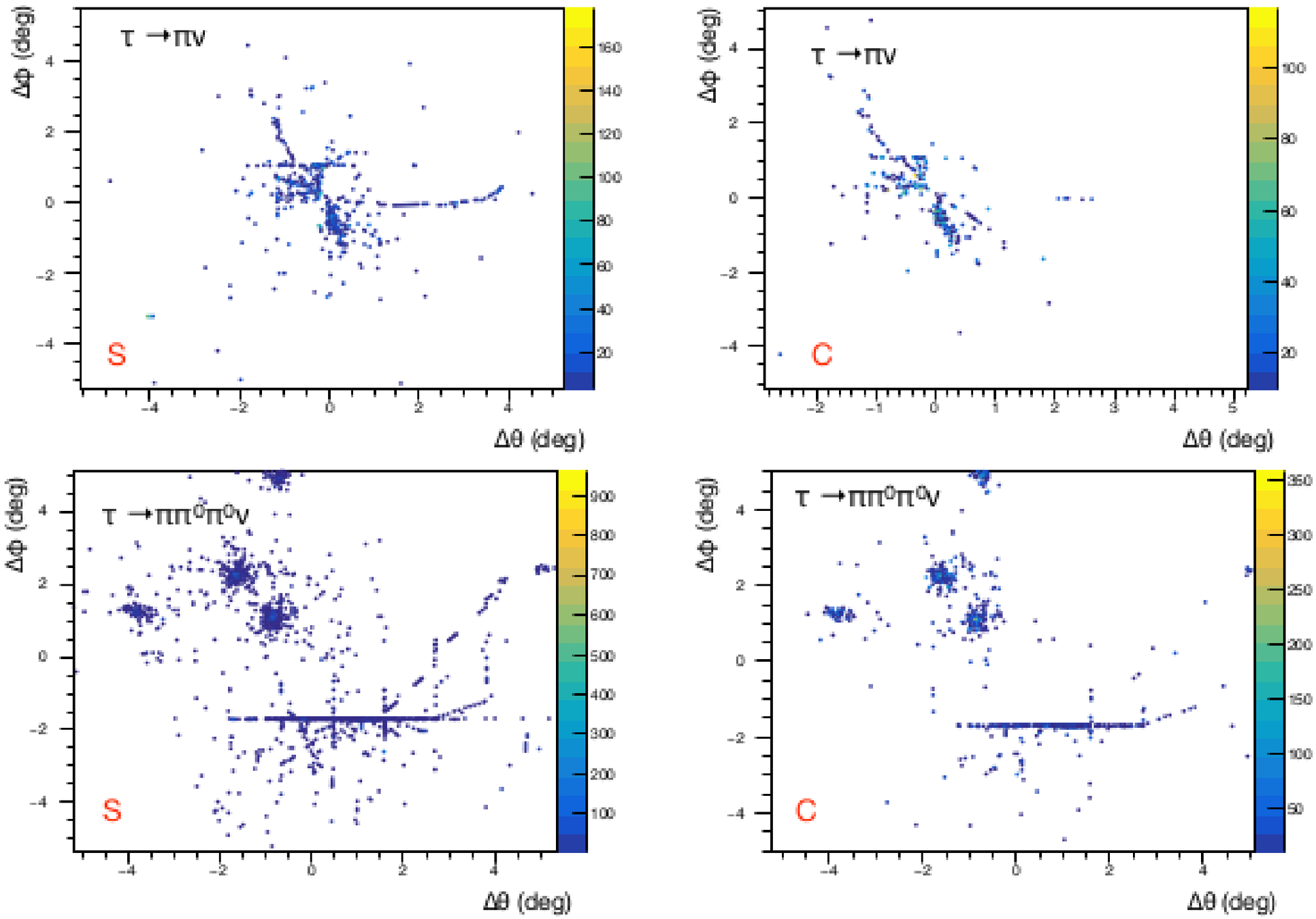


leptonic
Z→ττ



examples of events @ full granularity

hadronic
$Z \to \tau\tau$



hadronic
$Z \to \tau\tau$

simulation - light (p.e.) generation

Full G4 simulation of IDEA* DR calorimeter up to:

*1 X_0 in front, no B field

- (Birks' corrected) energy lost in S fibres
  → S photons → 400 Spe/GeV + decay time

- Čerenkov light production in C fibres
  → C photons → 100 Cpe/GeV

correct timing for emission angle + fibre refractive index

simulation - SiPM transfer function

SiPM parameters (tunable) with typical values:

duration: 500 ns
sampling: 0.1 ns

SiPM size: 1×1 mm^2
cell size: 10×10 μm^2
dark count rate: 200 kHz
cross talk: 1%
after pulsing: 3%

rise time: 1 ns
decay (recovery) time: 50 ns

integration window: 300 ns

PDE: 100% → included in light yield generation

calibrated with (12-photon) Poissonian source

feature extraction & PID analysis

Simulation output (per fibre): # pe, Q, Pk, ToA, ToT, ToP

After calibration with electrons:

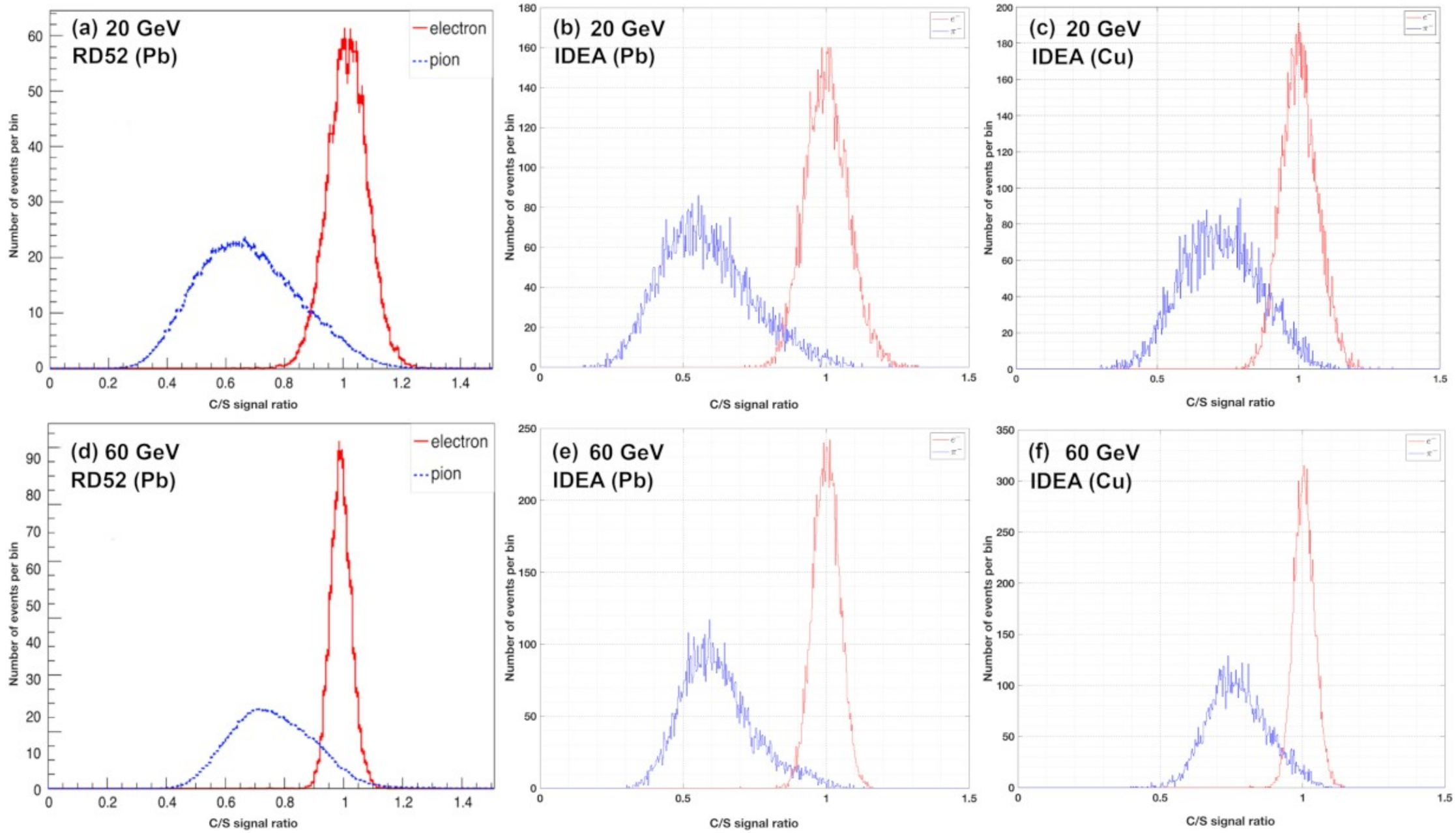
Q (in pe) → E (GeV)

Per-event variables (à la RD52):

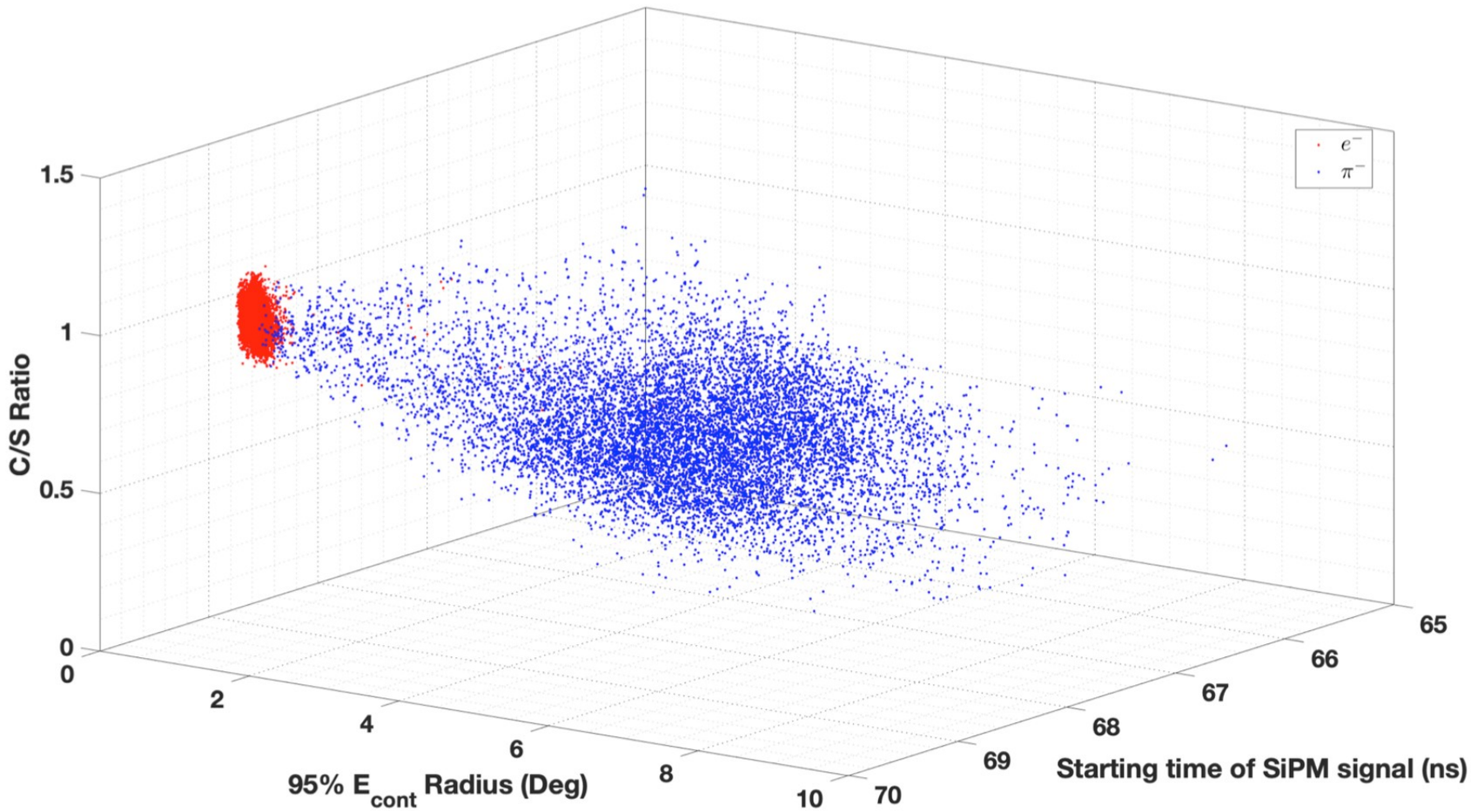
1) C/S ratio
2) R95: radius of cylinder with 95% of total signal
3) ToA for S signals

→ direct comparison with RD52 results

C/S


(a) 20 GeV
RD52 (Pb)
electron
pion
(b) 20 GeV
IDEA (Pb)
(c) 20 GeV
IDEA (Cu)
(d) 60 GeV
RD52 (Pb)
(e) 60 GeV
IDEA (Pb)
(f) 60 GeV
IDEA (Cu)
Number of events per bin
C/S signal ratio

correlations at 60 GeV



	e^-	π^-
corr (C/S, R95):	5%	60%
corr (C/S, ToA):	31%	23%
corr (R95, ToA):	7%	9%

e/π performance

Energy	RD52 e^- ID (%)	RD52 π^- mis-ID (%)	IDEA (Pb) e^- ID (%)	IDEA (Pb) π^- mis-ID (%)	IDEA (Cu) e^- ID (%)	IDEA (Cu) π^- mis-ID (%)
C/S ratio						
20 GeV	99.0	16.9	98.9	7.1	99.0	17.2
60 GeV	98.0	13.9	99.0	4.0	99.0	9.7
Shower profiles						
20 GeV	99.1	6.9	99.4	3.2	99.3	3.8
60 GeV	99.1	5.1	99.8	2.5	98.7	3.2
Signal starting time						
20 GeV	-	-	98.7	17.3	99.0	33.2
60 GeV	99.1	11.3	98.5	3.2	99.0	15.3

Multivariate analysis for 60 GeV events:

	e^- ID	π^- mis-ID
RD52:	99.8%	0.2%
IDEA (Pb):	99.8%	0.4%

τ-decay discrimination

Preliminary analysis with testbeam geometry (brass abs.)
dimensions: 133.2×133.2×250 cm^3

Reduced granularity (1.2×1.2 cm^2, 32 S & 32 C fibres):
→ 111×111 modules

Signals from (S, C) fibres in each module combined

Simulation (detector + SiPM) as in previous analysis

Available features: E(Q), Pk, ToP, ToA, ToT

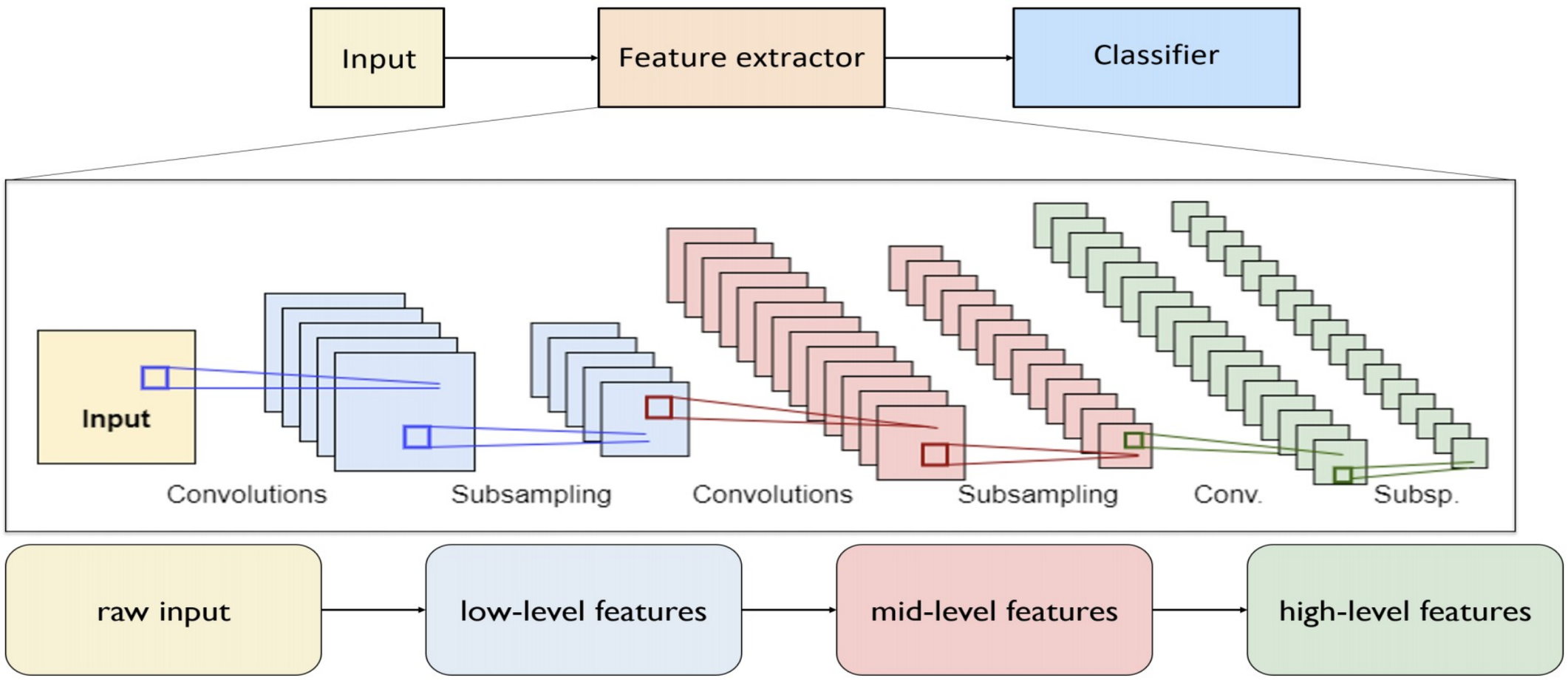
→ each event represented by a 111×111×5×2 tensor

CNN models

multilayer networks designed for visual pattern recognition

made of convolution and subsampling layers

→ hierarchical extraction of topological properties



implementation

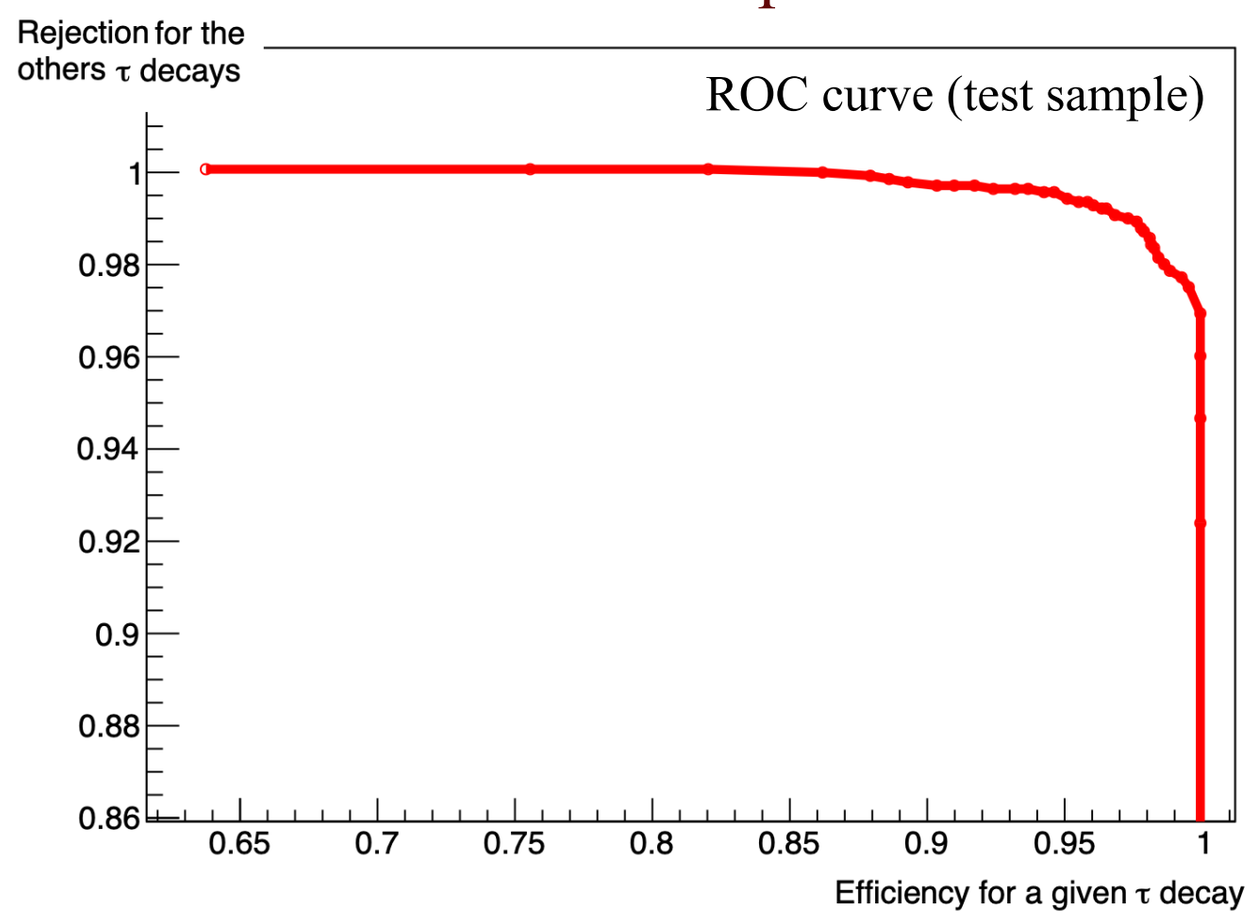
Two DNN architecture variant studied:

- VGG-11 like (VGG = Visual Geometry Group, Oxford Un.)
- Dynamic Graph CNN (DGCNN)

6 event classis (covering ~ 90% of τ decays)
Training set: 6 BR × 2000 evts

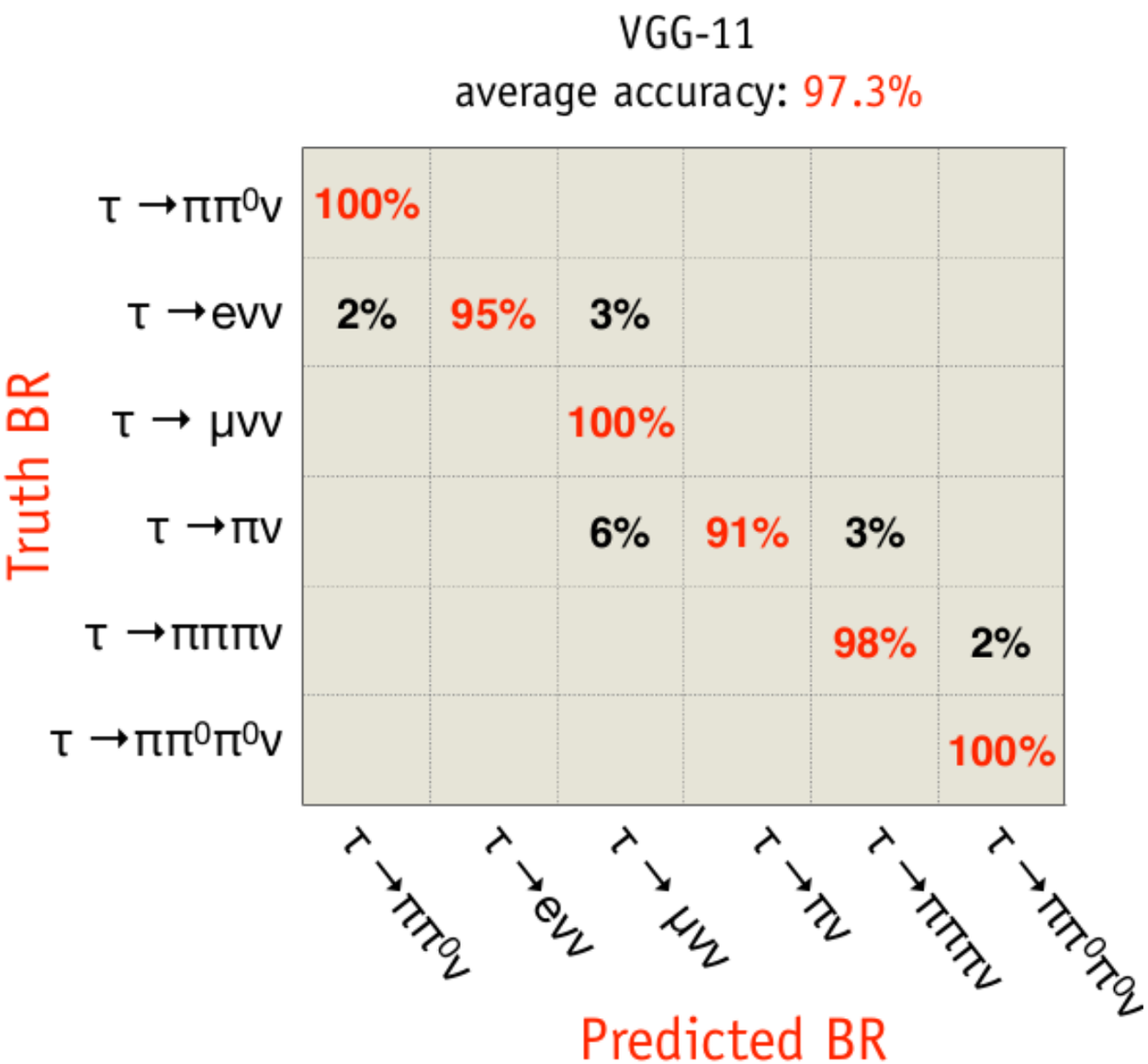
$\tau \to \pi\pi^0\nu$
$\tau \to \mu\nu\nu$
$\tau \to e\nu\nu$
$\tau \to \pi\nu$
$\tau \to \pi\pi\pi\nu$
$\tau \to \pi\pi^0\pi^0\nu$

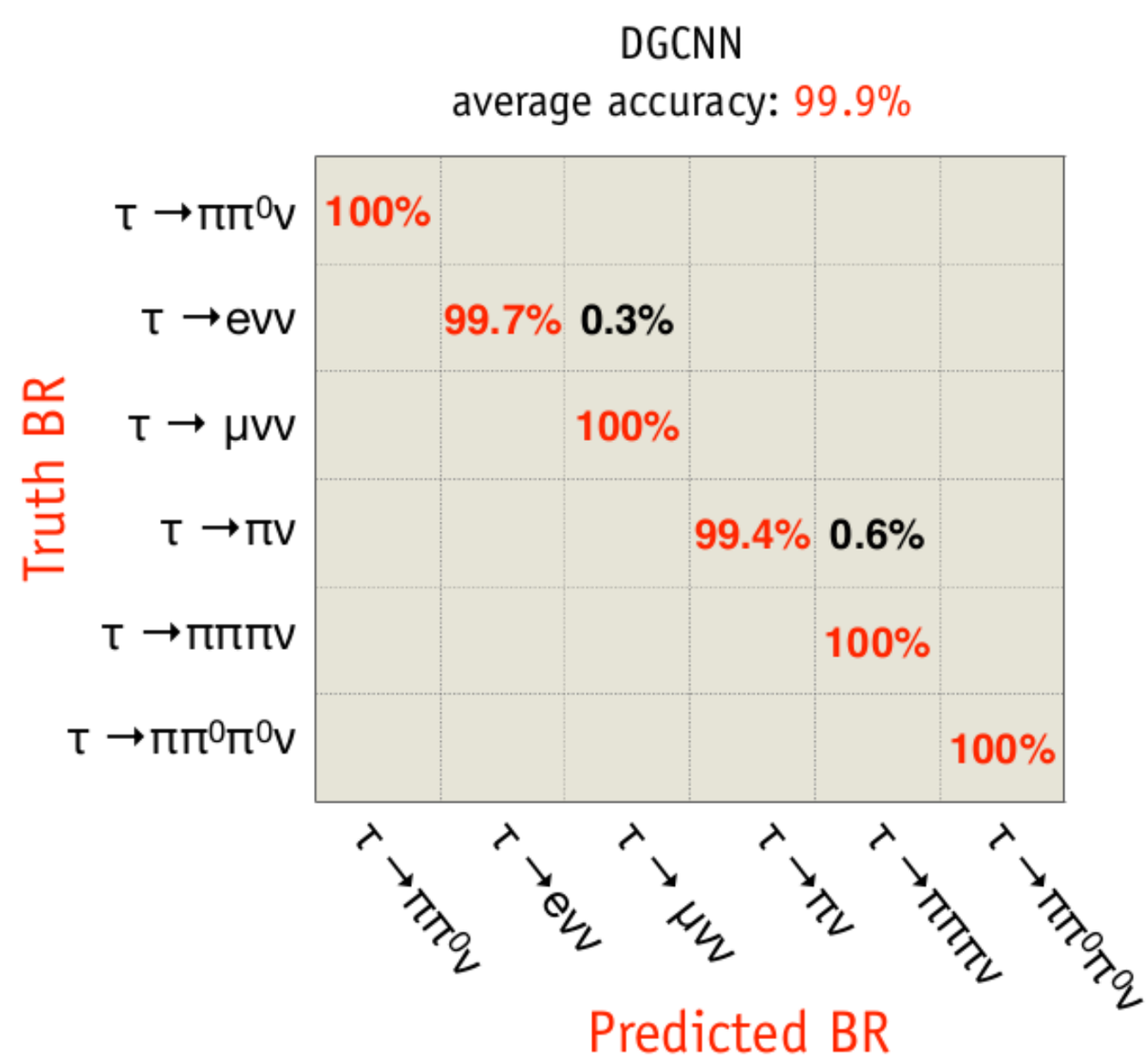
VGG example



performance

Confusion matrix on test set





uncertainty on accuracies ~ 4÷5%

realistic IDEA DR geometry

No SiPM response simulation yet

→ information: fibre signal output (# p.e.)

3-class classification:

τ_{lep}, τ_{had}, QCD jet

8-class classification:

τ_0, τ_1, τ_2, τ_3, τ_4, τ_5, τ_6, QCD jet

[τ from $Z \to \tau\tau$ decays]

3-class label	8-class label	
0	0	$\tau \to \mu\nu\nu$
0	1	$\tau \to e\nu\nu$
1	2	$\tau \to \pi\nu$
1	3	$\tau \to \pi\pi^0\nu$
1	4	$\tau \to \pi\pi^0\pi^0\nu$
1	5	$\tau \to \pi\pi\pi\nu$
1	6	$\tau \to \pi\pi\pi^0\nu$
2	7	$Z \to qq$ jets

DNN models

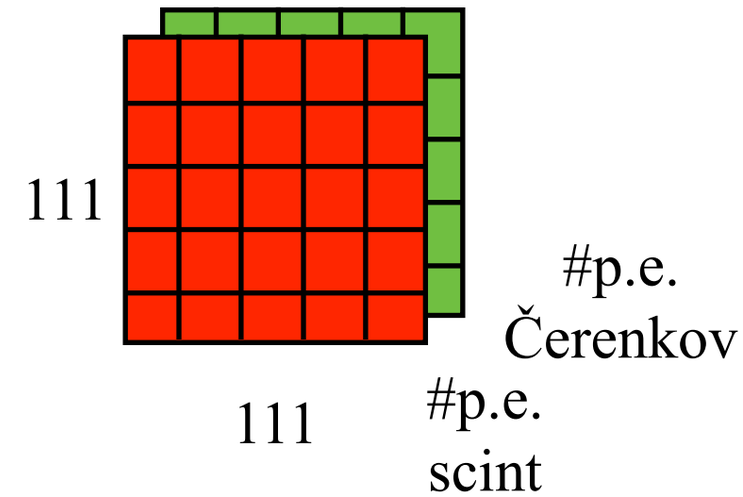
DNN models:

- VGG-like CNN with 3D and 2D convolutions:

jet/τ representation 2-channel 111x111 mesh

reduced granularity as in previous case

1.2×1.2 cm^2, 32 S & 32 C fibres



- DGCNN:

jet/τ representation: 2D point-cloud of spatial coordinates + #p.e. as features

full granularity within cluster cone

Training/validation/test sets: 4k/2k/4k events per class

Data preprocessing:

- simple geometrical clustering
- save fibre signal aroud clusters ($\sqrt{(\Delta\theta^2+\Delta\varphi^2)} < 1$)

preliminary results - 3-class classification

B=0 and no solenoid

Truth BR \ Predicted BR	τ_{lep}	τ_{had}	J_{QCD}
τ_{lep}	97%	2%	1%
τ_{had}	1.5%	94%	4%
J_{QCD}		4%	96%

average accuracy: 96%

CNN

B and material

Truth BR \ Predicted BR	τ_{lep}	τ_{had}	J_{QCD}
τ_{lep}	97%	2%	1%
τ_{had}	3%	91%	6%
J_{QCD}		7%	93%

average accuracy:94%

CNN

Truth BR \ Predicted BR	τ_{lep}	τ_{had}	J_{QCD}
τ_{lep}	98%	2%	
τ_{had}	1%	96%	3%
J_{QCD}		2%	98%

average accuracy:97%

DGCNN

uncertainty on accuracies ~3÷5%

preliminary results - 8-class classification

DGNN architecture (not yet optimised)

uncertainty on accuracies ~3÷5%

Truth BR \ Predicted BR	$\tau \to \mu\nu\nu$	$\tau \to e\nu\nu$	$\tau \to \pi\nu$	$\tau \to \pi\pi^0\nu$	$\tau \to \pi\pi^0\pi^0\nu$	$\tau \to \pi\pi\pi\nu$	$\tau \to \pi\pi\pi\pi^0\nu$	$Z \to qq$ jets
$\tau \to \mu\nu\nu$	97%	2%	1%					
$\tau \to e\nu\nu$	1%	97%	1%	1%				
$\tau \to \pi\nu$	1%	2%	89%	3%		5%		
$\tau \to \pi\pi^0\nu$		1%	3%	74%	9%	3%	9%	
$\tau \to \pi\pi^0\pi^0\nu$				6%	80%	1%	12%	1%
$\tau \to \pi\pi\pi\nu$			7%	5%		80%	7%	1%
$\tau \to \pi\pi\pi\pi^0\nu$			1%	7%	8%	12%	69%	2%
$Z \to qq$ jets				1%	1%	1%		97%

(B=0 and no solenoid material)

average accuracy: 85.3% (CNN: 82-83%)

Truth BR \ Predicted BR	$\tau \to \mu\nu\nu$	$\tau \to e\nu\nu$	$\tau \to \pi\nu$	$\tau \to \pi\pi^0\nu$	$\tau \to \pi\pi^0\pi^0\nu$	$\tau \to \pi\pi\pi\nu$	$\tau \to \pi\pi\pi\pi^0\nu$	$Z \to qq$ jets
$\tau \to \mu\nu\nu$	97%	2%	1%					
$\tau \to e\nu\nu$	1%	97%	1%	1%				
$\tau \to \pi\nu$	1%	3%	87%	4%	1%	2%		1%
$\tau \to \pi\pi^0\nu$		1%	5%	76%	13%	1%	3%	1%
$\tau \to \pi\pi^0\pi^0\nu$				7%	88%		4%	1%
$\tau \to \pi\pi\pi\nu$			5%	2%		75%	15%	2%
$\tau \to \pi\pi\pi\pi^0\nu$			1%	1%	4%	8%	81%	5%
$Z \to qq$ jets					1%	1%	2%	96%

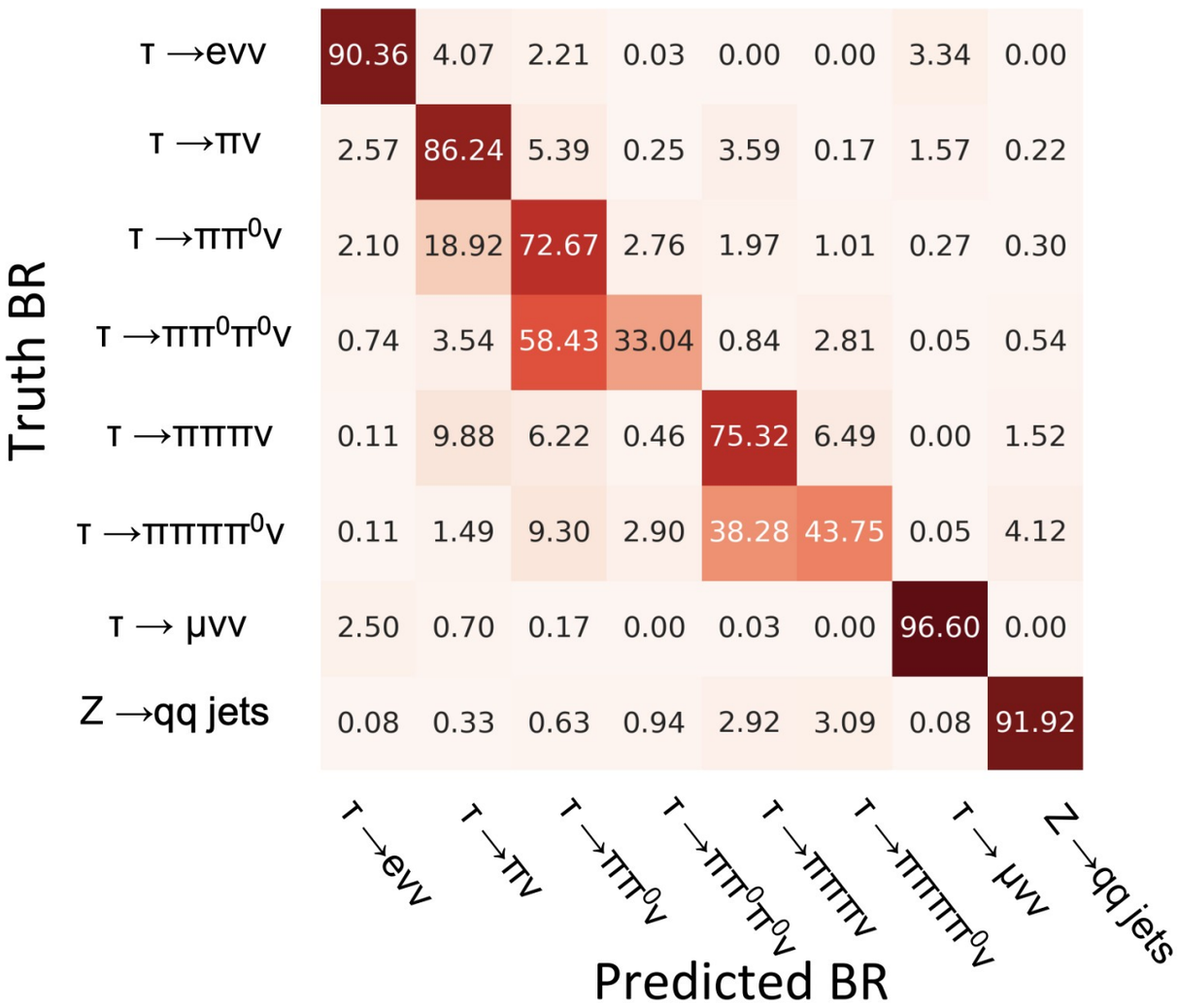
(B field and material)

average accuracy: **87%**

using geometrical information only

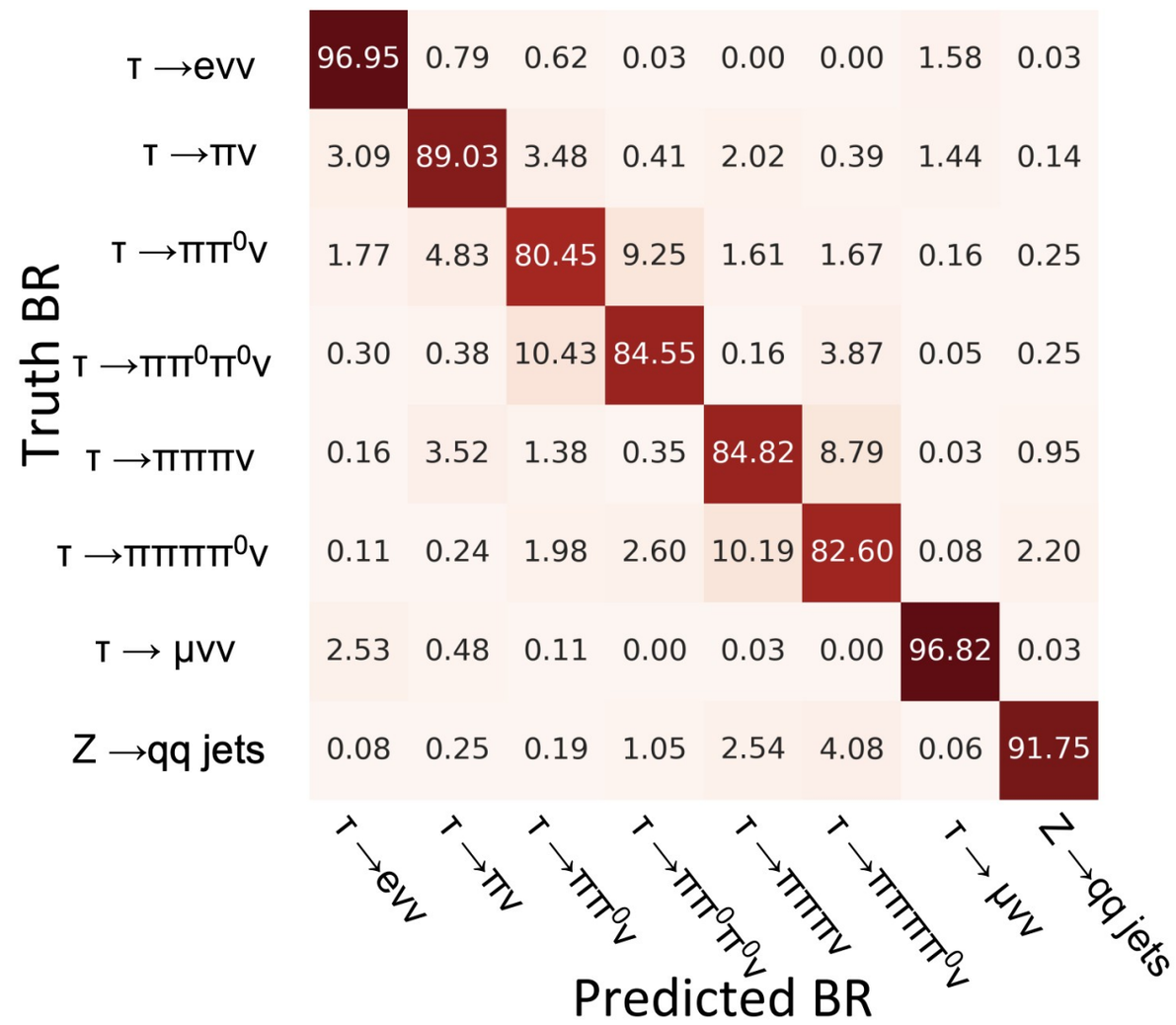
DGCNN optimised but w/o #pe as input feature

B field and material in



no fibre type tag

avg accuracy: 73.7%

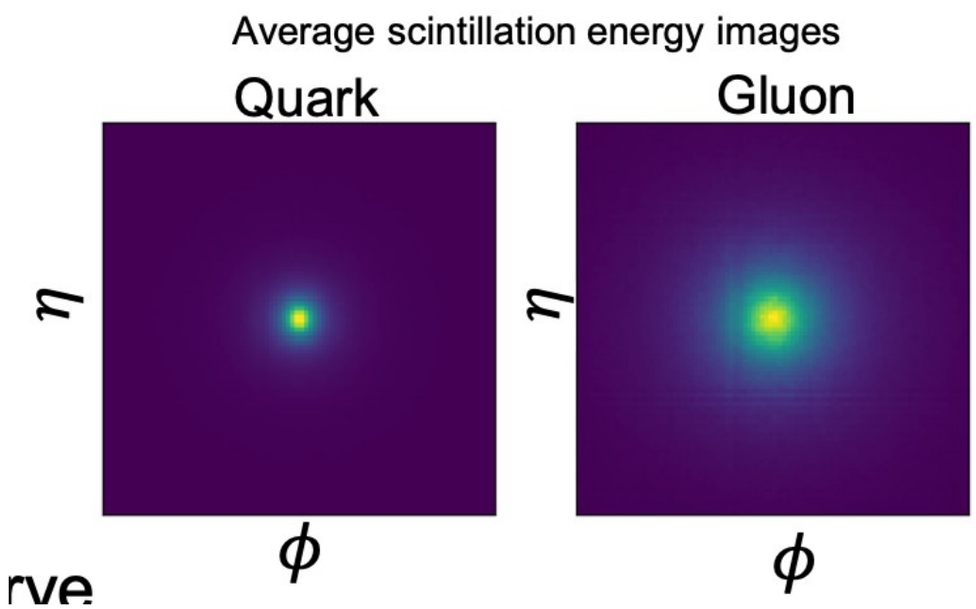


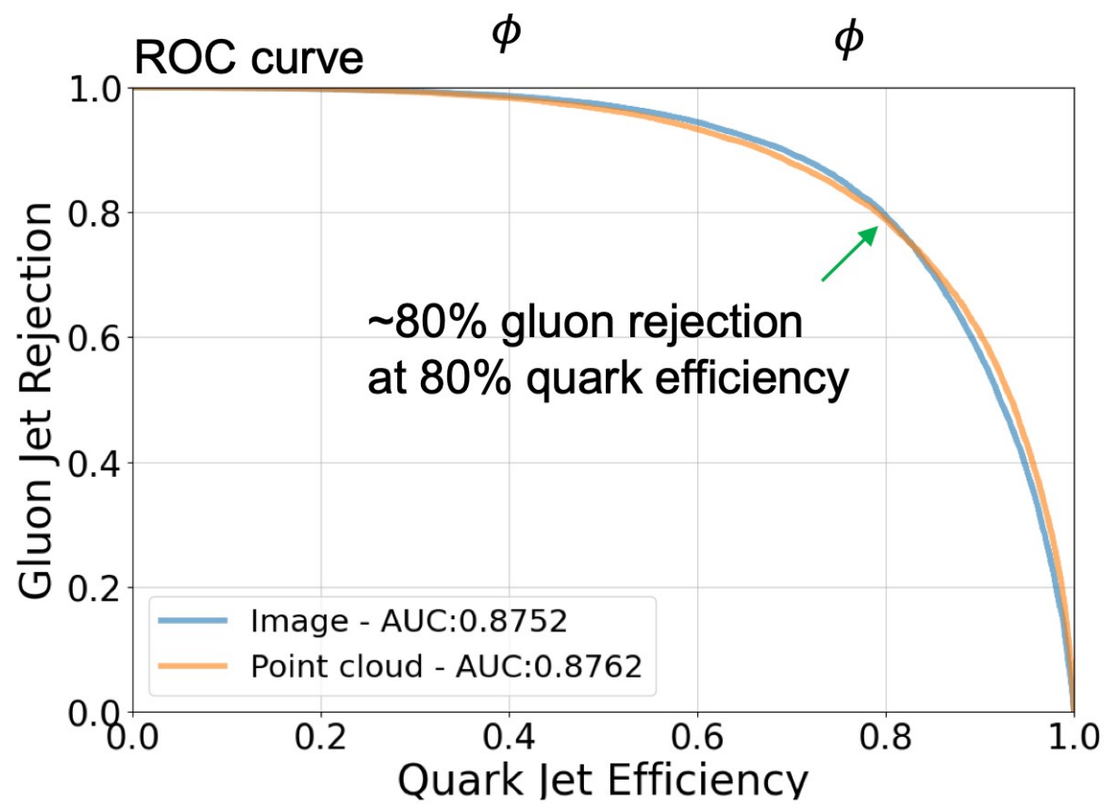
w/ fibre type tag

avg accuracy: 88.3% (w/#p.e. 90.8%)

last but not least … very preliminary

quark/gluon jet identification:





conclusions

Preliminary studies have been presented for:

- (isolated) e/pi discrimination
- classification of tau-decay modes and QCD jets
- quark/gluon jet discrimination (very preliminary)

all based on reduced sets of (calorimeter only) information

Dual-readout information coupled with high-granularity SiPM readout and ML show very promising performance in all cases

Next step: implement full simulation chain and feature extraction in IDEA detector