IDEA Calorimeter Simulation and Performance - Status and Plans

Lorenzo Pezzotti, **Iacopo Vivarelli**

University of Pavia & INFN Pavia, University of Sussex

on behalf of the IDEA Calo Group

IDEA Collaboration Meeting, 16/17 February 2021

2019-2021 Activities

- Development of a standalone Geant4 application to fully simulate the IDEA dual-readout projective calorimeter.
- Detailed study on the response to single particles ($e^-, \pi^-, p^-, n, k^+, \mu^-$): calibration, χ factor universality, energy resolution, position/angle measurements.
- Simulation interfaced to standard event generators (through HepMC format): $Z \to jj$ events, $Z/W/H$ mass measurements via $2j$ final states, $H \to \gamma\gamma$ events. Now investigating $4j$ final states.
- PID studies: e^-/π^- discrimination, γ/π^0 discrimination, $\tau^\pm$ decay identification and jet rejection.
- SiPM digitization studies: signal saturation, time properties and impact on energy resolution.
- Dual-readout crystal integration: studying the possibility to adopt a crystal EM section in the simulation.
- Integration within modern SW tool: DD4HEP and EDM4HEP integration.

2019-2021 Activities - Today's topics

- Development of a standalone Geant4 application to fully simulate the IDEA dual-readout projective calorimeter.
- Detailed study on the response to single particles ($e^-, \pi^-, p^-, n, k^+, \mu^-$): calibration, χ factor universality, energy resolution, position/angle measurements.
- Simulation interfaced to standard event generators (through HepMC format): $Z \rightarrow jj$ events, $Z/W/H$ mass measurements via $2j$ final states, $H \rightarrow \gamma\gamma$ events. Now investigating $4j$ final states.
- PID studies: e^-/π^- discrimination, γ/π^0 discrimination, $\tau^\pm$ decay identification and jet rejection.
- SiPM digitization studies: signal saturation, time properties and impact on energy resolution.
- Dual-readout crystal integration: studying the possibility to adopt a crystal EM section in the simulation.
- Integration within modern SW tool: DD4HEP and EDM4HEP integration.

SiPM Digitization Studies

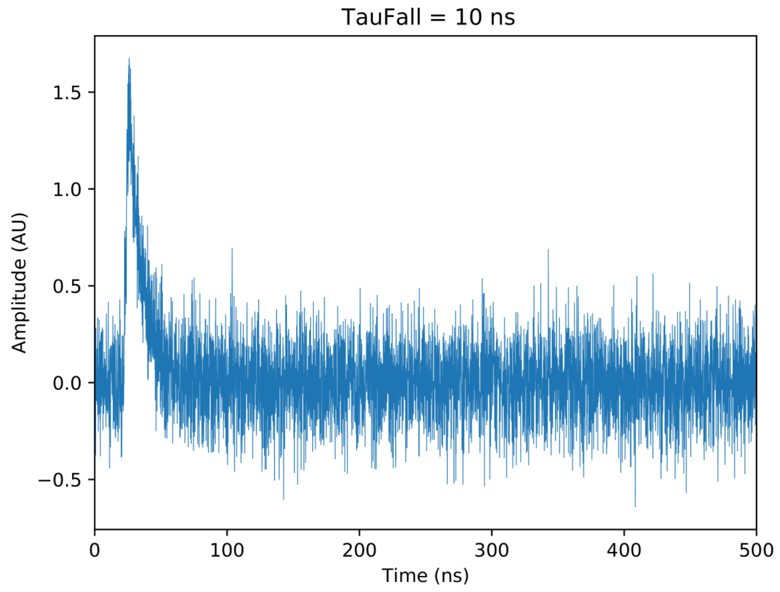
- A SW to simulate the SiPM transfer function has been developed at the University of Insubria [Presentation] and extensively tested at the University of Pavia [Presentation].
- The digitization SW was tested by taking the output of the G4 simulation as timestamps of each p.e. and outputting the SiPM digitized signals.

Signal Length: 500 ns
Sampling: 0.1 ns

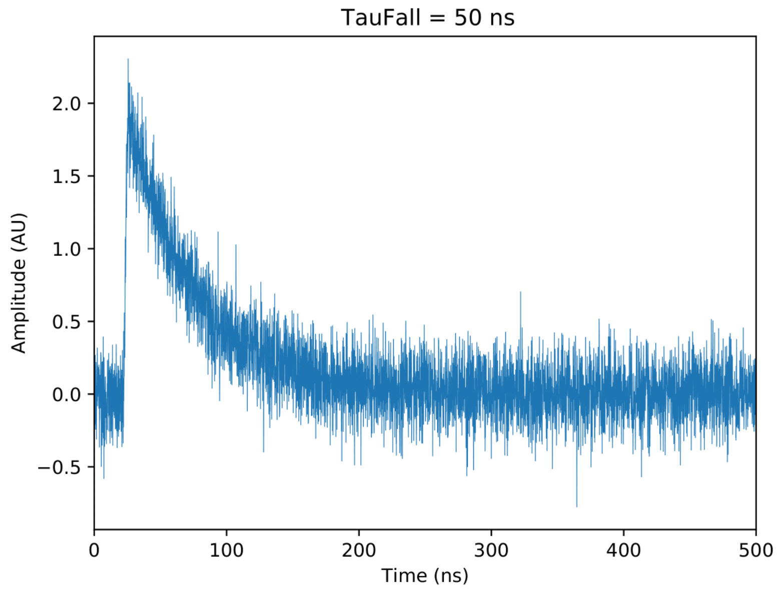
SiPM Size: 1 x 1 mm^2
Cell Size: 25 x 25 μm^2
Dark Count Rate: 200 kHz
CrossTalk: 1 %
After Pulse: 3 %

Decay Time Constant: 50 or 10 ns
Rise Time Constant: 1 ns

Integration Gate Start Time: 5 ns
Integration Gate Time Window: 300 ns



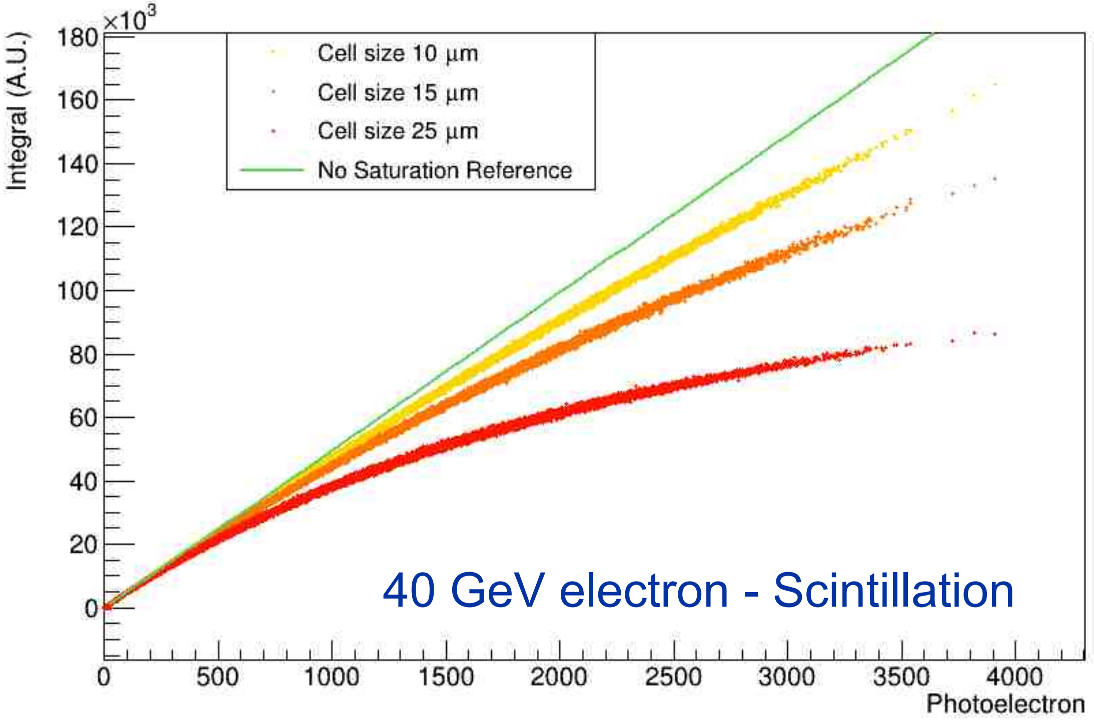
10 ns decay time

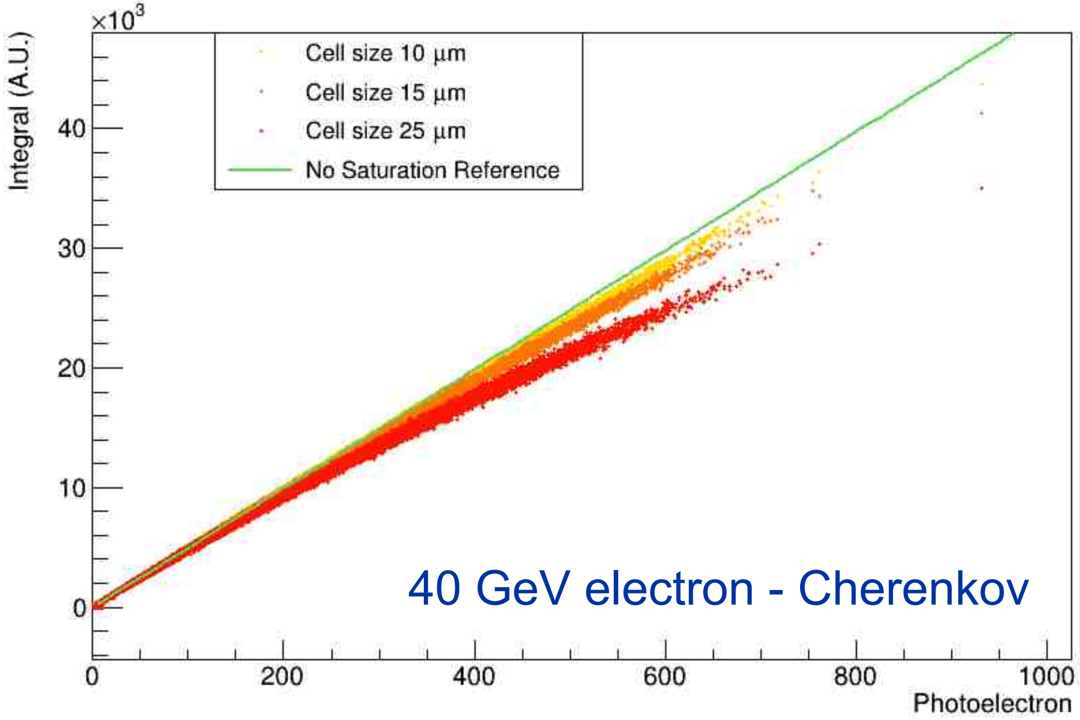


50 ns decay time

SiPM Occupancy Saturation Studies (1/3)

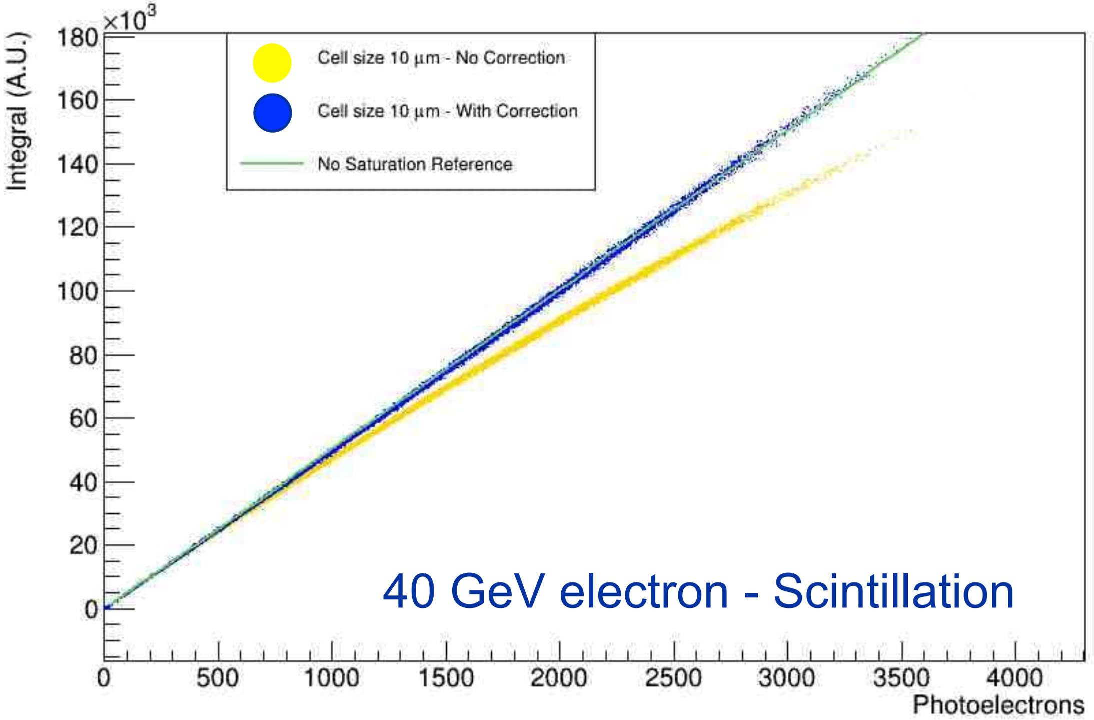
- The relation between the signal integral (a.u.) and the number of p.e. was studied by measuring both the contribution of the DCR and the one related to a single p.e.
- By simulating within G4 light yields of 400 Sp.e./GeV and 100 Cp.e./GeV (at the em scale), the occupancy saturation effect was studied for different cell sizes.

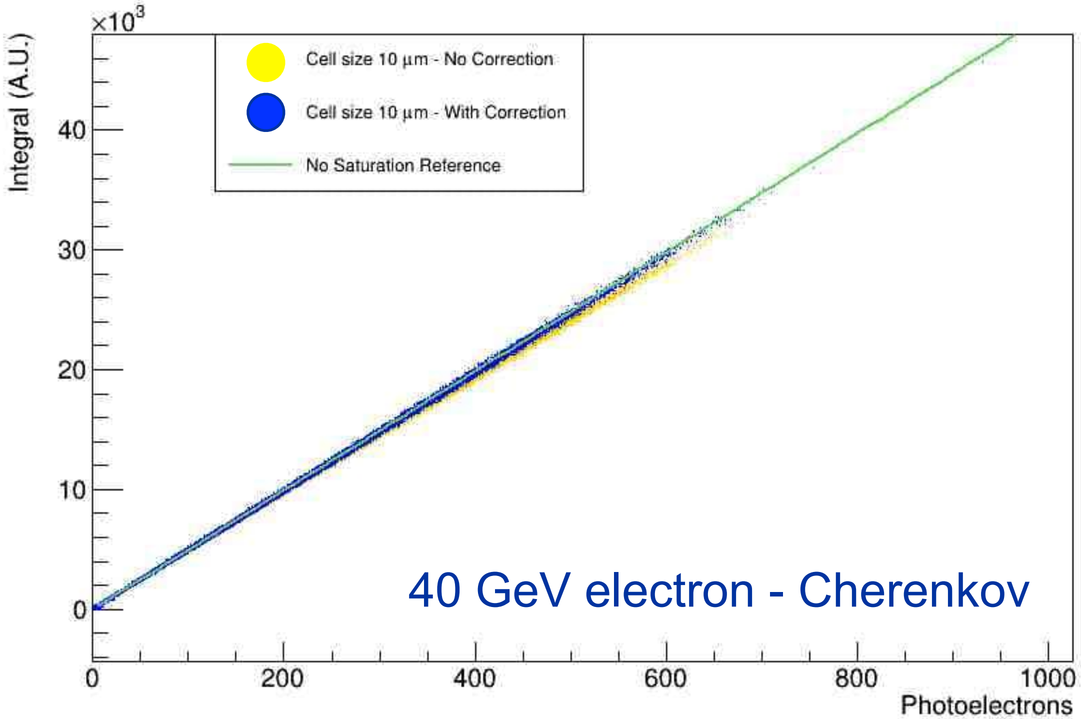




SiPM Occupancy Saturation Studies (2/3)

- Taking the 10-μm pitch as a reference, we applied an analytical correction given by $N_{FC} = N_{tot} \cdot \left(1 - e^{-\frac{N_\gamma PDE}{N_{tot}}}\right)$.
- The correction almost completely fixed the saturation issue.

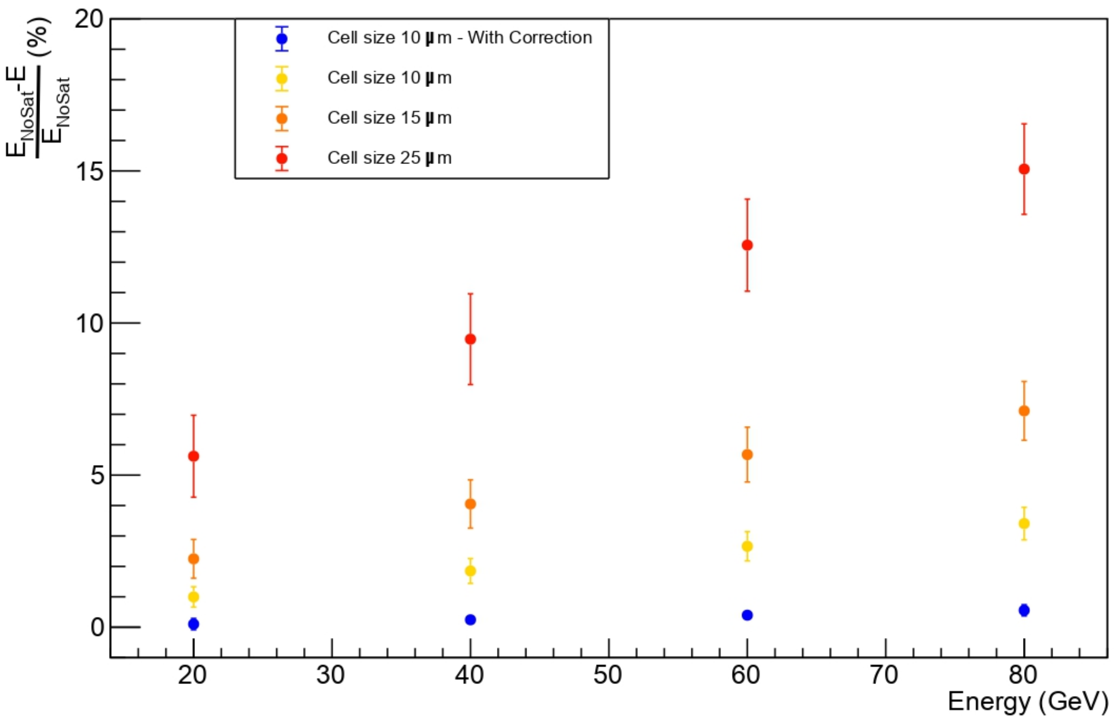




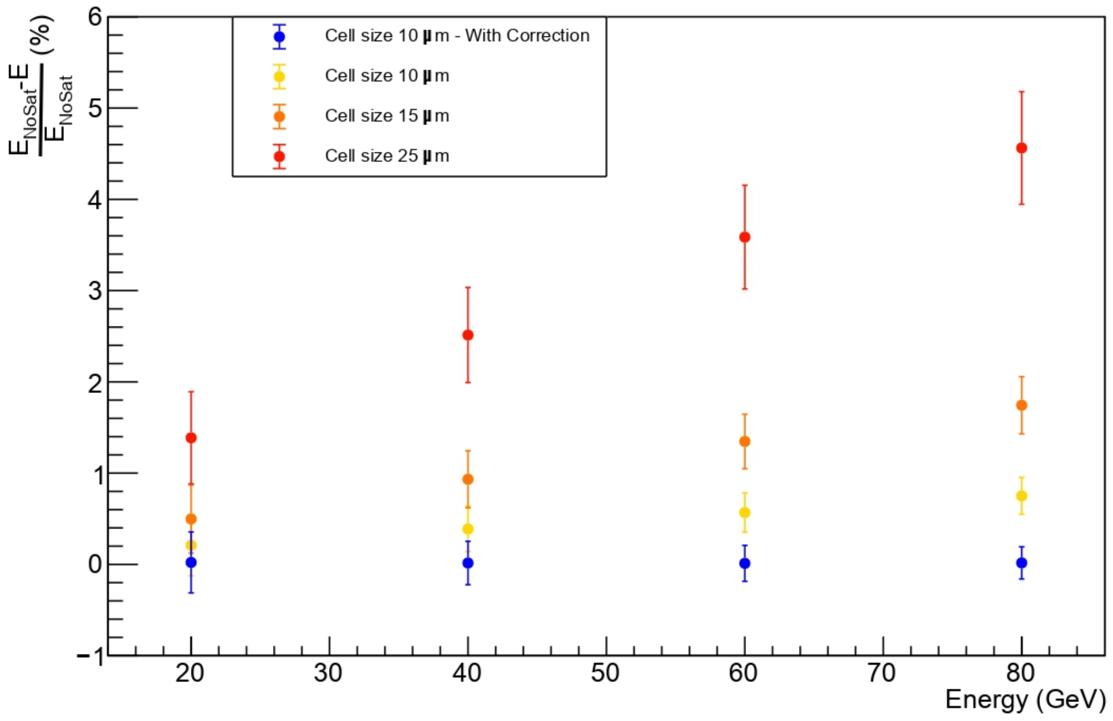
SiPM Occupancy Saturation Studies (3/3)

- The work demonstrated that 10-μm-pitch sensors could be considered the baseline for the simulation. This sets the goal of future sw/hw activities.
- Indeed, the calorimeter can reach $\pm 1\ \%$ linearity in the FCCee expected energy range.

20 - 80 GeV electron - Scintillation



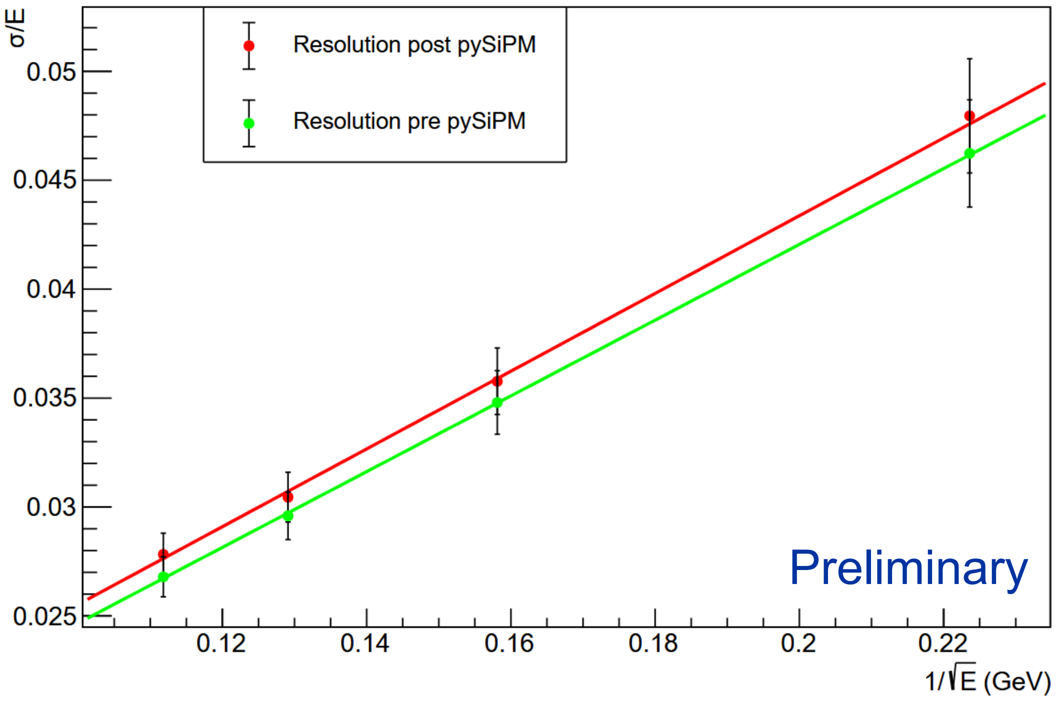
20 - 80 GeV electron - Cherenkov



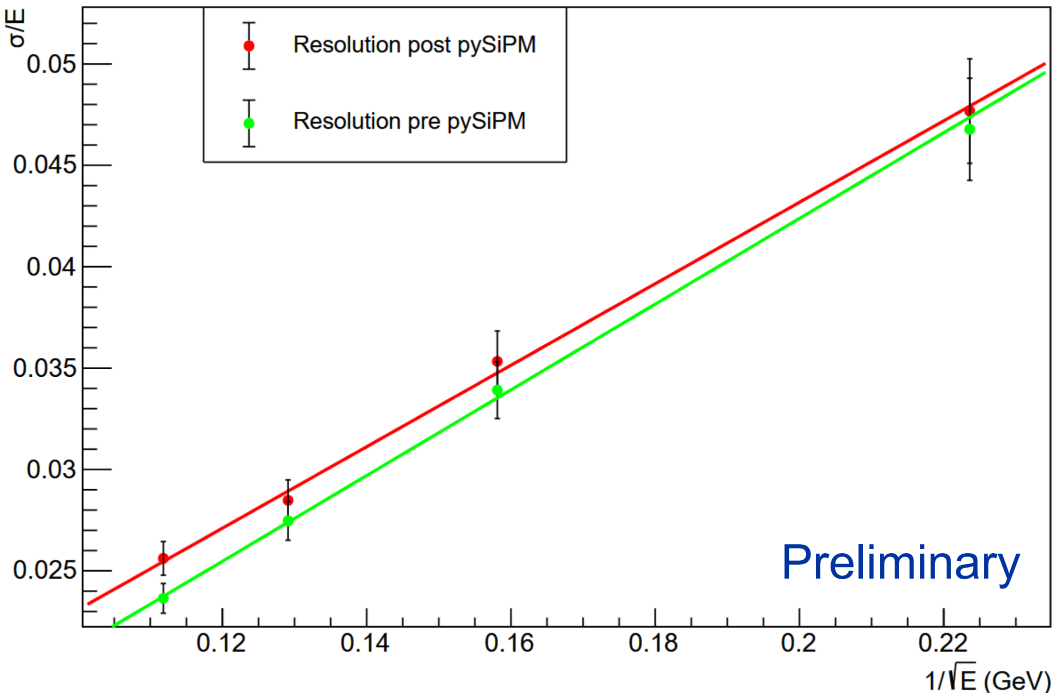
Digitization impact on energy resolution

- By applying calibration constants (estimated with a known source of light), it is possible to estimate the impact of the digitization on the energy resolution.
- Results for electron events with a 1 p.e.-suppression applied show no significant degradation of the energy resolution.

20 - 80 GeV electron - Scintillation



20 - 80 GeV electron - Cherenkov



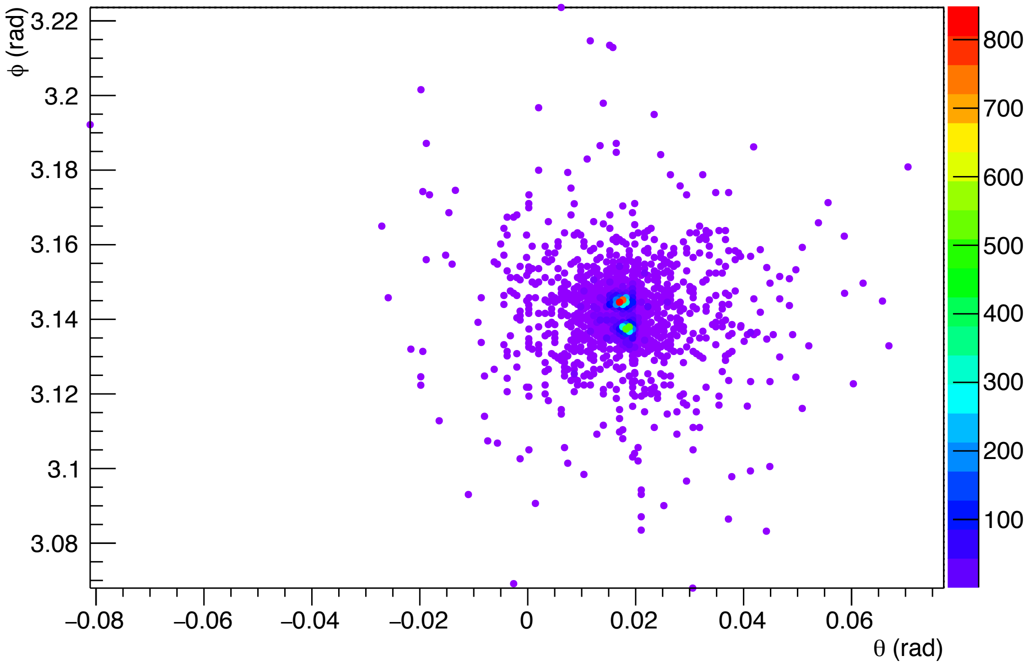
Digitization studies recap

- A dedicated SiPM digitization tool has been developed and tested against the G4 output.
- It indicates that, in the FCCee energy range, the IDEA Calorimeter can reach a signal linearity of $\pm 1\,\%$ by adopting 10-μm-pitch sensors, considering 400 Sp.e./GeV and 100 Cp.e./GeV light yields. This sets the scale for a proper light smearing in our present and future simulations.
- By considering 1p.e.-suppressed events and fully digitizing the G4 output, we found minimal worsening on the em energy resolution with respect to previous results.
- Not discussed here: basic timing information from the signals has already been studied and confirms the known/expected properties of a dual-readout fiber calorimeter. More advanced studies can be performed if needed.

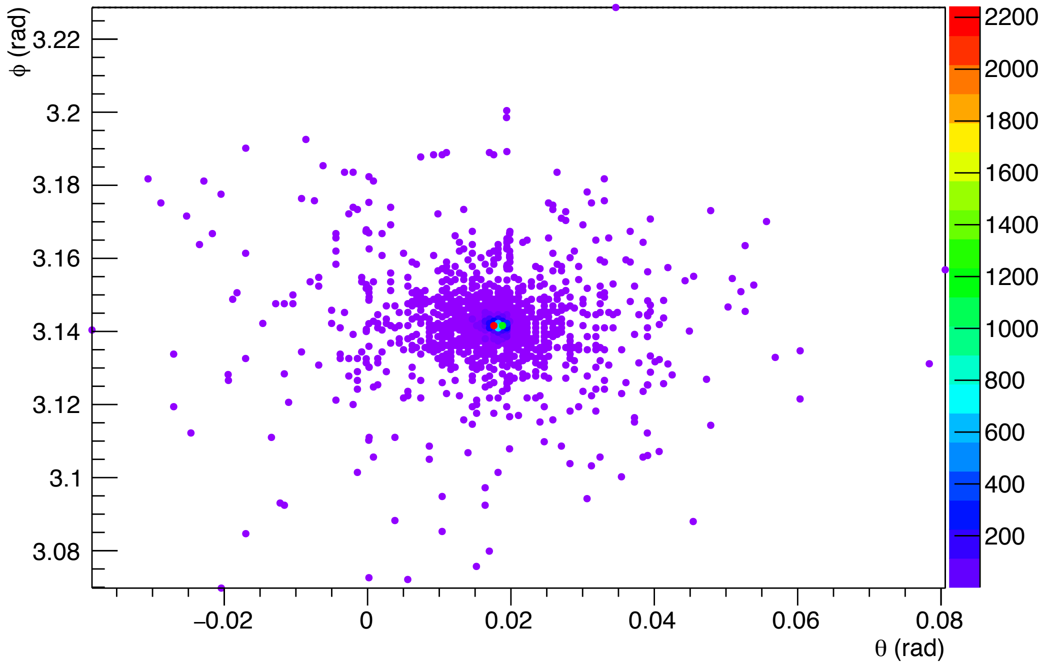
Particle Identification: γ/π^0 (1/2)

- The extremely high 2D-granularity of the IDEA Calorimeter brings to some spectacular results. An example is the π^0 identification from two γ-initiated showers.

40 GeV π^0 - scintillation



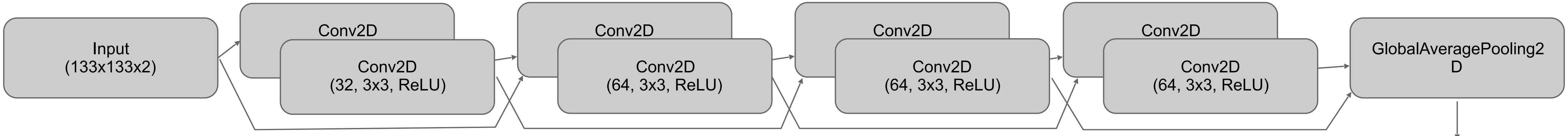
40 GeV γ - Cherenkov

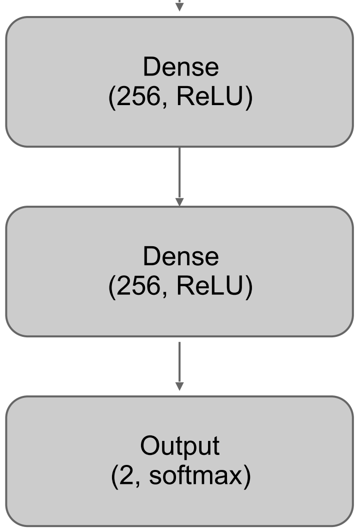


- A quantitative analysis on the possibility to distinguish between π^0 and γ was performed using a convolutional neural network. Results on events with no selection and fully digitized.

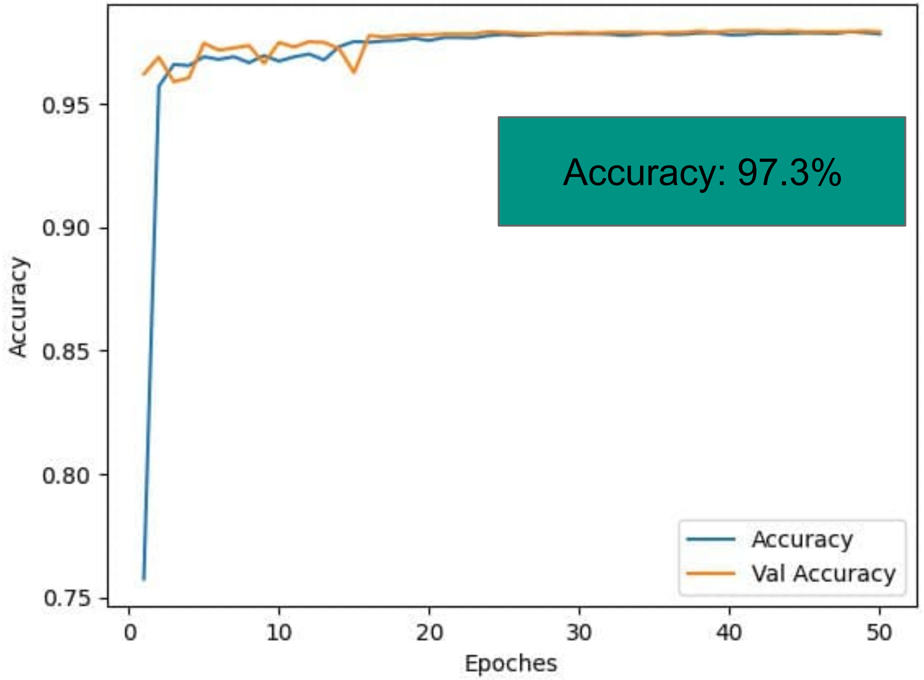
Particle Identification: γ/π^0 (2/2)

The residual network:

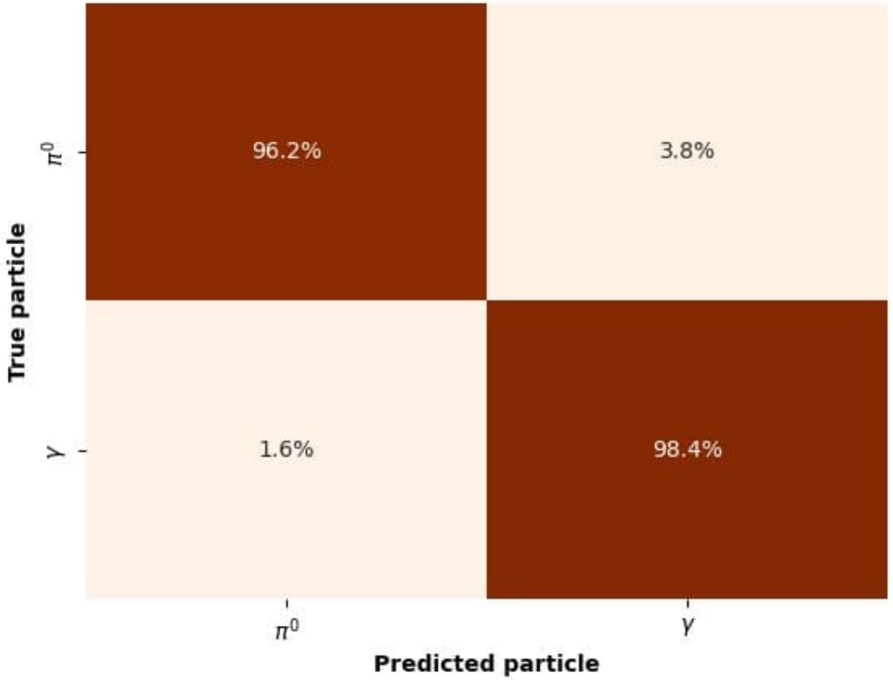




Accuracy during 50 training epochs



Confusion Matrix



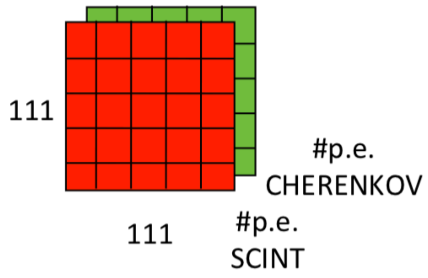
Now estimating the accuracy using a VGG network…

$\tau^{\pm}$ Identification with ML (1/2)

- A detailed study to identify tau decays and discriminate $\tau/jets$ events was performed, using the fiber-by-fiber calorimeter information and two convolutional NN architectures.
Refer to [Presentation].

Data preparation

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



DNN Selection

- VGG-like CNN with 3D and 2D convolutions: data representation 2-channel 111x111 mesh
- DGCNN: 2D point-cloud including 2000 fiber coordinates, including fiber type and signals (p.e.)

Results including B field and solenoid

	τ_{lep}	τ_{had}	J_{QCD}
τ_{lep}	98%	2%	
τ_{had}	1%	98%	
J_{QCD}		7%	93%

avg accuracy: 96.4%

CNN

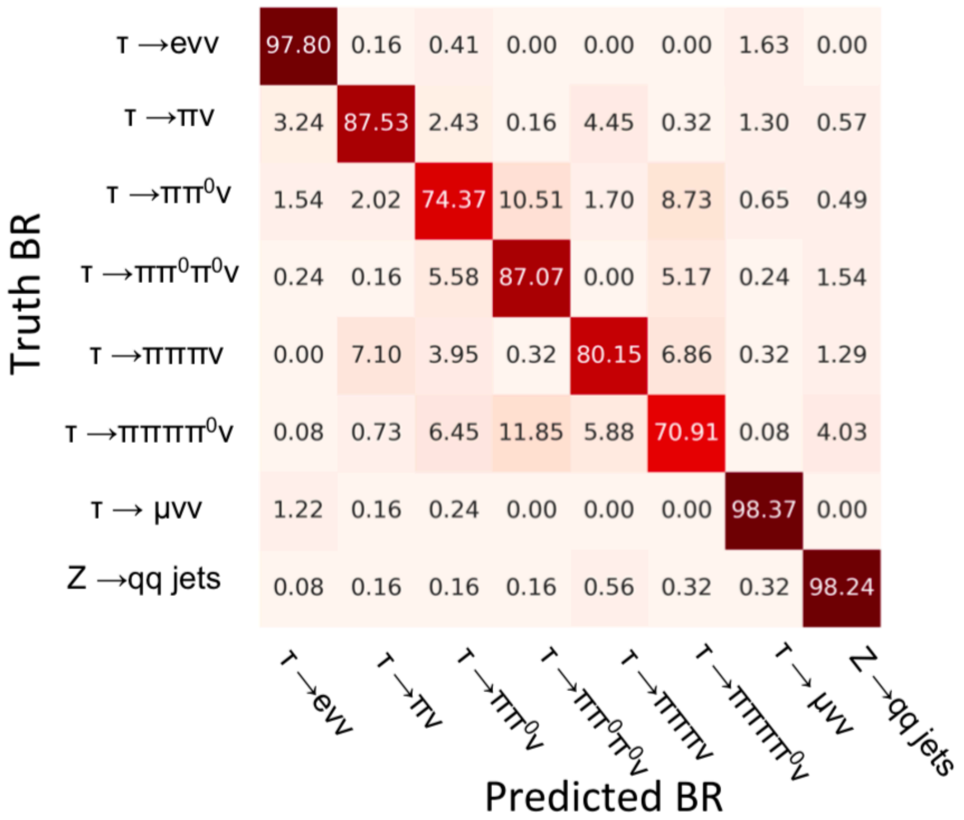
	τ_{lep}	τ_{had}	J_{QCD}
τ_{lep}	99%	1%	
τ_{had}	1%	97%	2%
J_{QCD}		4%	96%

avg accuracy: 97%

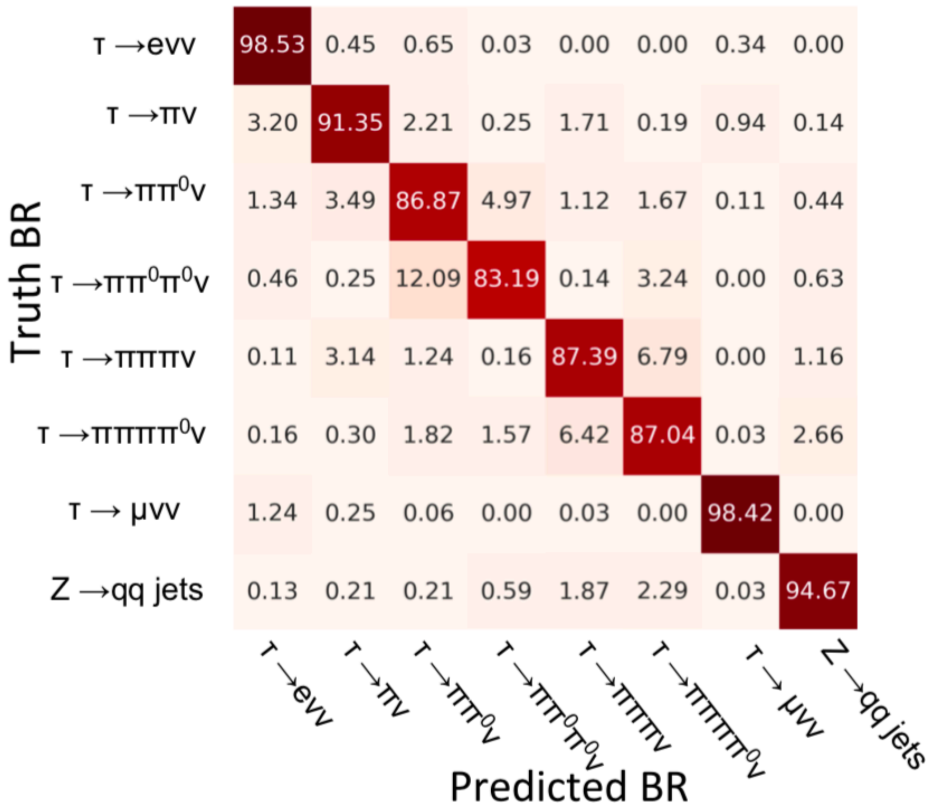
DGCNN

$\tau^{\pm}$ Identification with ML (2/2)

Example: 8-class classification task with DGCNN architecture without energy information



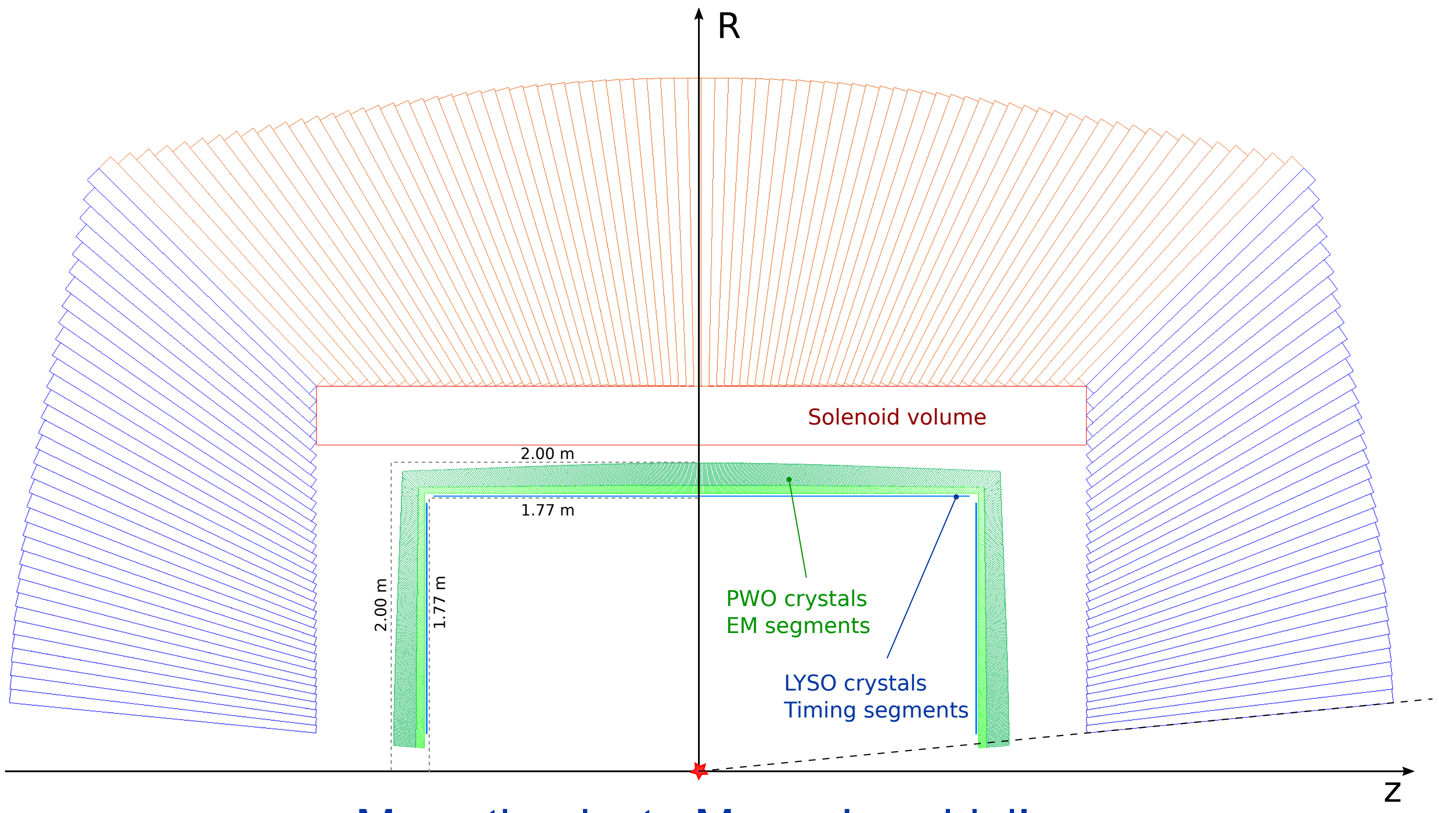
No B field and solenoid
Average accuracy 86.8% (CNN: 85.4%)



With B field and solenoid
Average accuracy 90.8% (CNN: 87.3%)

A Crystal Option for IDEA?

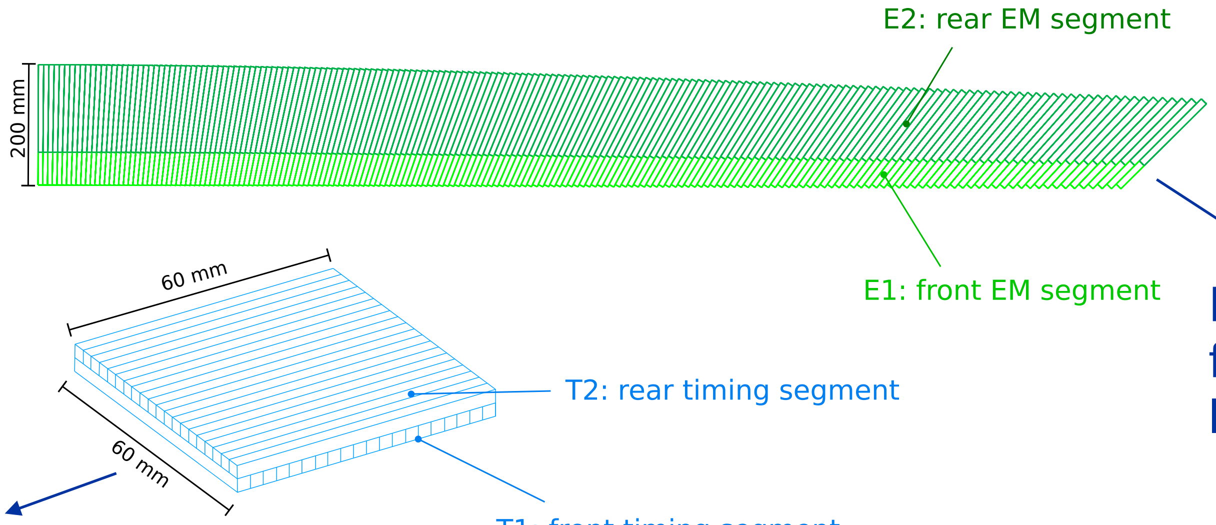
- Recently proposed in [Article], a dual-readout crystal em calorimeter could be adopted in the IDEA calorimeter system.
- Already integrated in the existing Geant4 Calorimeter application.
- Goal: maintain the key capability to correct for fluctuations of the electromagnetic fraction in hadronic showers while boosting the energy resolution for photons and electrons to about $3\,\%/\sqrt{E}$.



Many thanks to Marco Lucchini!

Crystal section geometry

- The EM section is made of four longitudinal layers:
  - two thin and highly-segmented layers (LYSO crystals - $1X_0$) for the time tagging of mips with 20 ps resolution,
  - two thicker layers (PWO crystals $6X_0 + 16X_0$)
- Located inside the solenoid with R between 1.8 and 2.0 m.



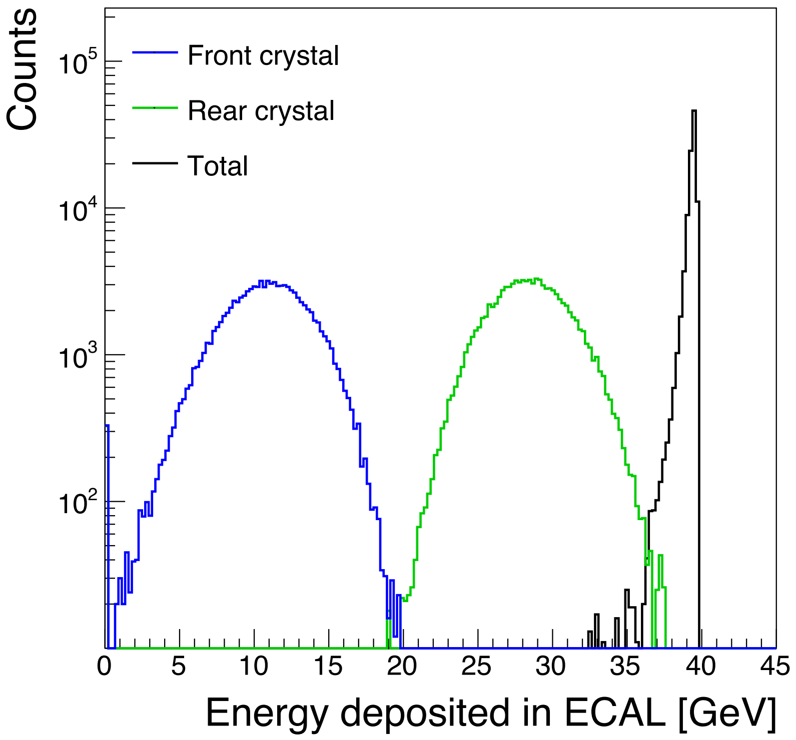
High refractive index (n=2.2) for a high Cherenkov light yield

$90°$ rotated for a cell size of $3 \times 3\ mm^2$

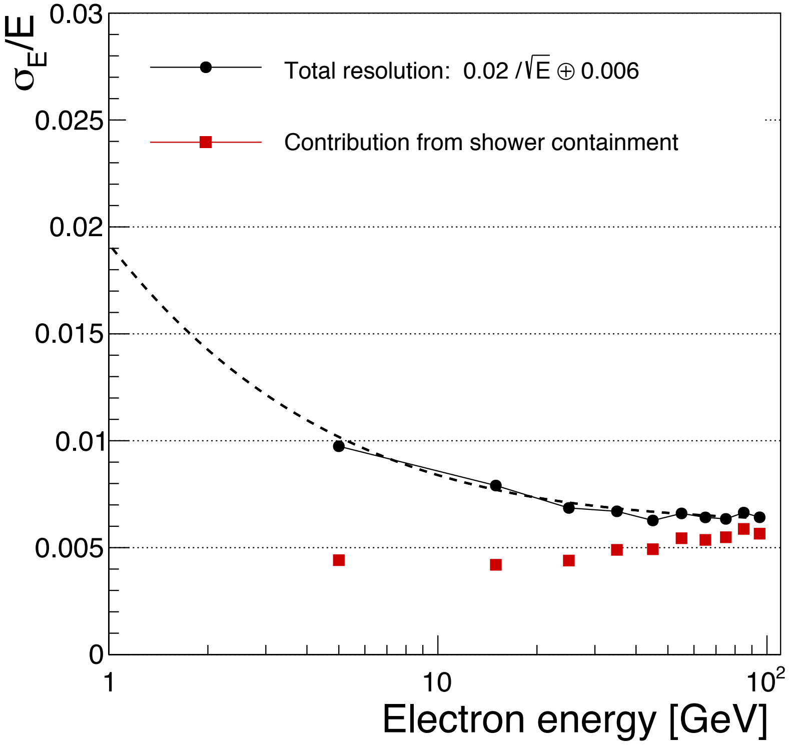
Geant4 results on the crystal section (1/3)

- Currently tuning the simulation to 2000 pe/GeV (S) and 160 pe/GeV (C) at the em scale. Take care: no light saturation (Birks' law) is needed for the crystal description.

Energy deposited in the timing sections and in the whole crystal section - 40 GeV e-



EM Energy resolution



$\sigma/E = 2.5\,\%/\sqrt{E} \oplus 0.6\,\%$

- Take care: when the upstream material budget will be included, the stochastic term is expected to be 3%.
- Standalone fiber em energy resolution: $\sigma/E = 14\,\%/\sqrt{E} \oplus 0.6\,\%$

Geant4 results on the crystal section (2/3)

Hadron energy reconstruction in the combined ECAL + HCAL system:

$$E_{HCAL} = \frac{S_{HCAL} - \chi_{HCAL} C_{HCAL}}{1 - \chi_{HCAL}}$$

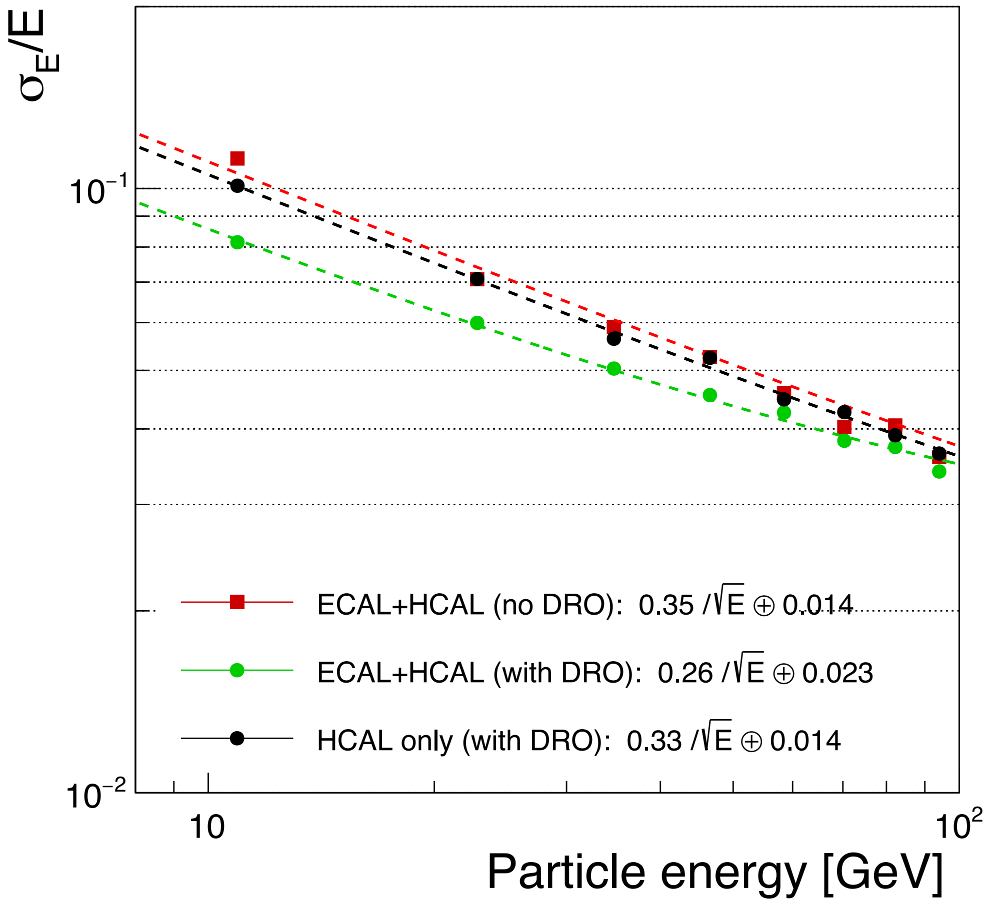
$$E_{ECAL} = \frac{S_{ECAL} - \chi_{ECAL} C_{ECAL}}{1 - \chi_{ECAL}}$$

$$E_{total} = E_{HCAL} + E_{ECAL}$$

Note:

$$\chi_{HCAL} \simeq 0.41 - 0.43$$
$$\chi_{ECAL} \simeq 0.37$$

Preliminary result for k_L^0

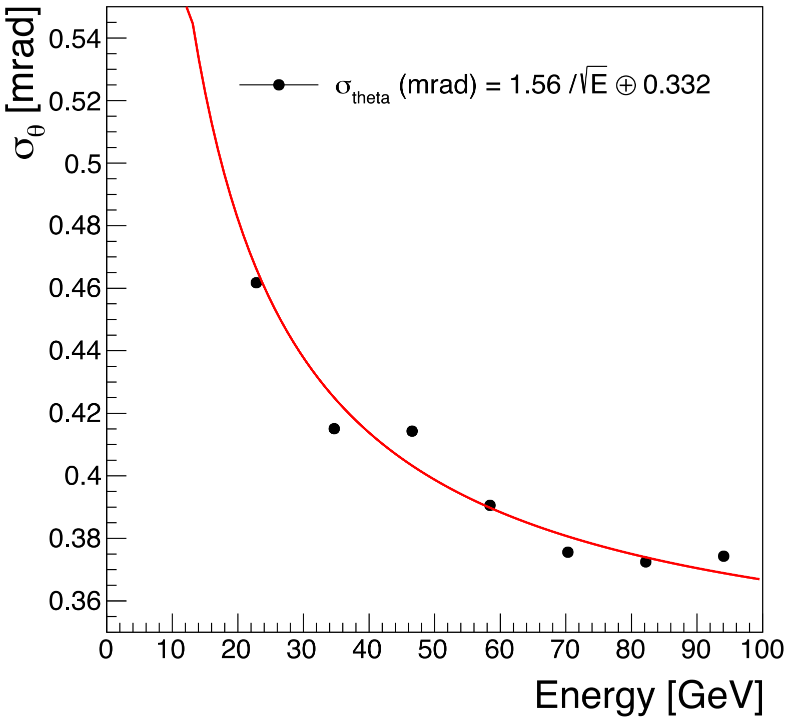


- Very good agreement found with the standalone fiber calorimeter performance.

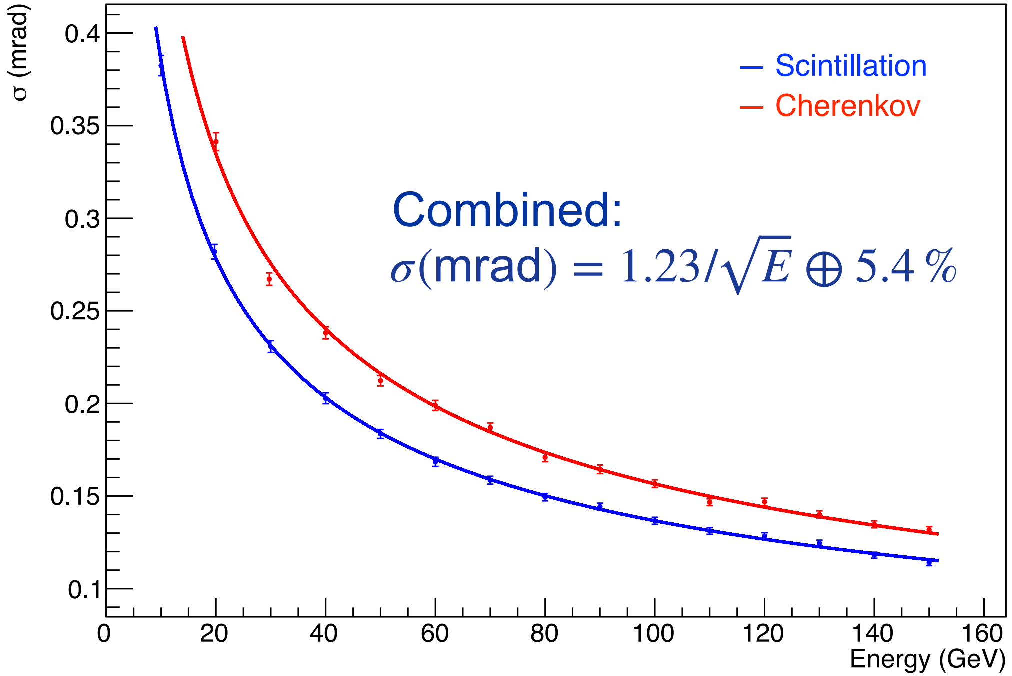
Geant4 results on the crystal section (3/3)

- The angular resolution for em showers was studied with both the standalone fiber calorimeter and the combined crystal+fiber calorimeters.
- Excellent angular resolutions are expected with both configurations.

Crystals option



Fiber calorimeter only



A New DR Calo Simulation EDM (1/2)

- Up to now results obtained with a standalone Geant4 simulation and an ad hoc customized EDM.
- Now time to study "busy" events, i.e. events where the parton-to-cluster association is not obvious (e.g. $ZH \rightarrow 4j$, highly requested by Franco 🙂).
- Need to systematically include the fiber-by-fiber information, including time stamps of each p.e., in an dimension-optimized file.
- Need to apply digitization tool(s) in an automatic step on top of the data output.

↓

Investigating the possibility to adopt EDM4HEP as the next EDM for any future study.

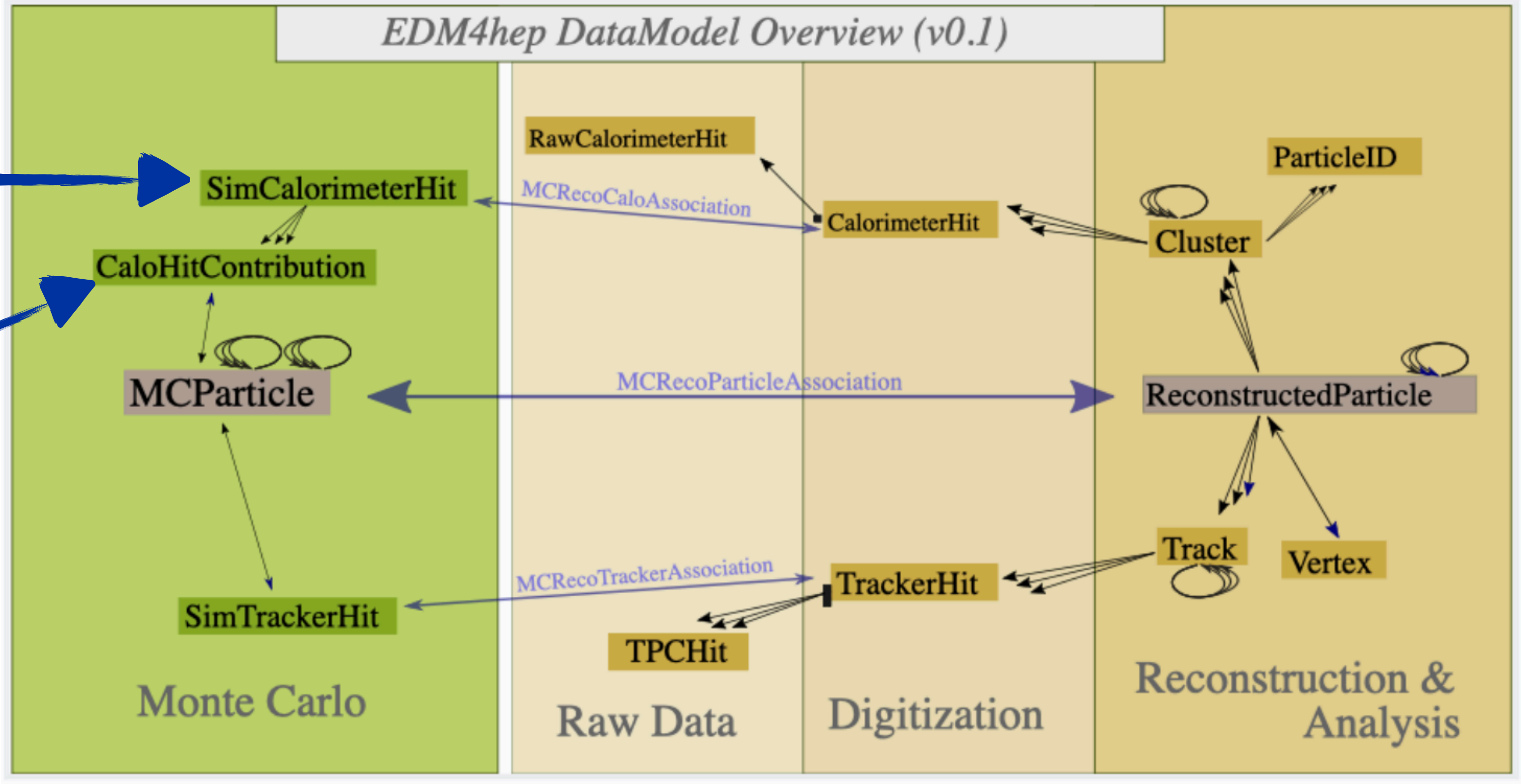
A New DR Calo Simulation EDM (2/2)

- EDM4HEP integration already at an advanced stage. Existing code available at [EDM4HEPbranch].
- What does it imply? Code will be using only environmental variables from the key4hep stack.

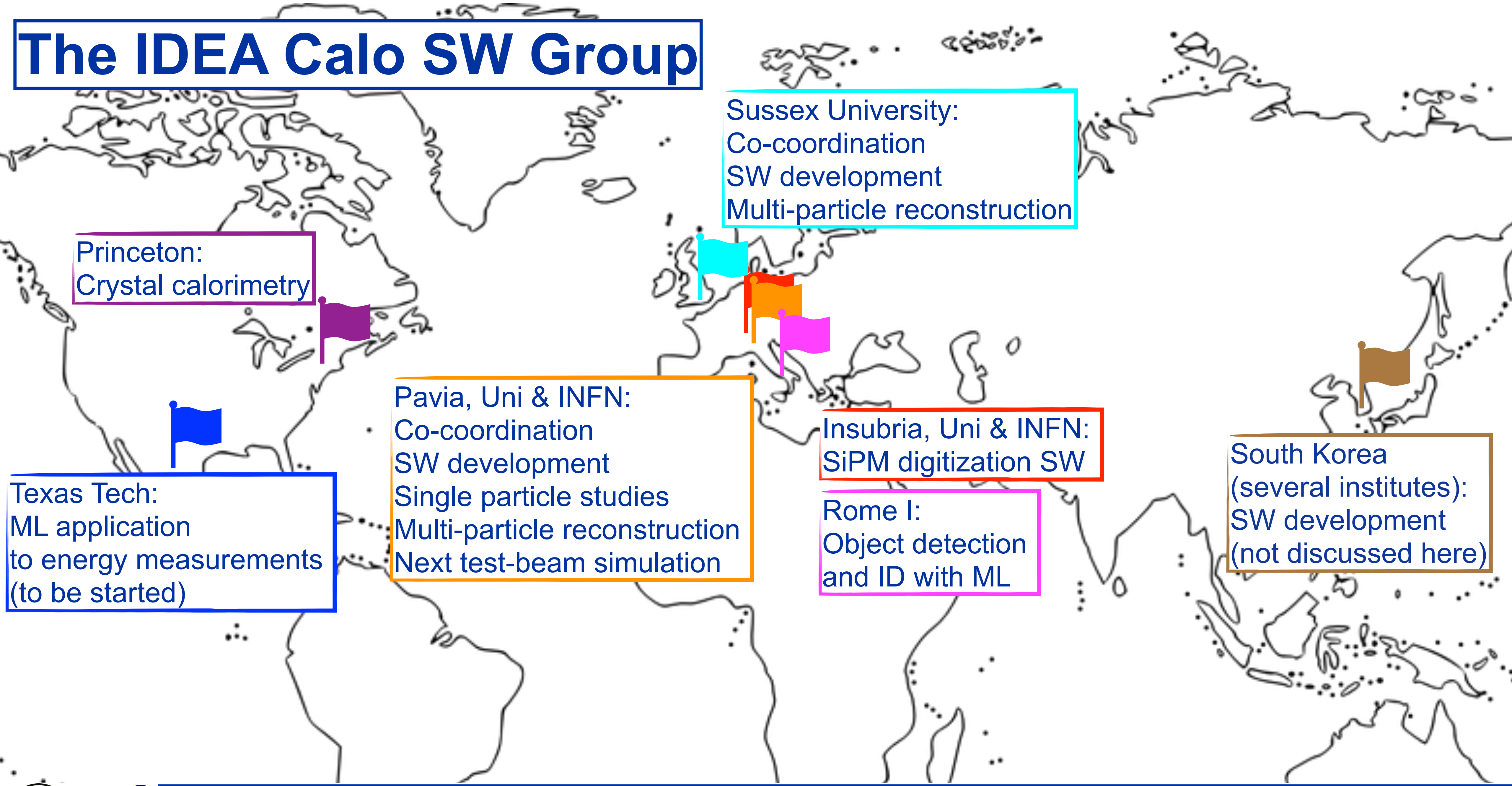
Information stored in pre-defined classes:

- Fiber info now stored as a SimCalorimeterHit including type, ID, (X,Y,Z) and integrated signal
- P.e. timestamps included as CaloHitContribution classes referring to each SimCalorimterHits.

Instruction to use this version of the code will soon be publicly available…



The IDEA Calo SW Group



Conclusions

- Several tasks completed in the 2019-2021 period, including code design, single particle performances, jet reconstruction and ML applications (mostly on PID), all with and without crystals.
  Today I could only report on the most recent results, but much more has been done.

- Mostly important to collect all the results in a common updated reference. Document already exists and about ten of us are currently contributing to it (many other authors will be added when a semi-complete document will be internally distributed).

- Three master thesis completed/ongoing (Pavia/Sussex/Roma I).

Contents

*lorenzo.pezzotti@cern.ch

1

If interested, join our bi-weekly IDEA Dual-Readout Calorimetry Meeting, please subscribe to idea-dualreadout@cern.ch and refer to our [INDICO].