Deep learning based event reconstruction for Limadou HEPD-02 on board the CSES satellite

Workshop sul Calcolo nell'INFN, 22-26 maggio 2023

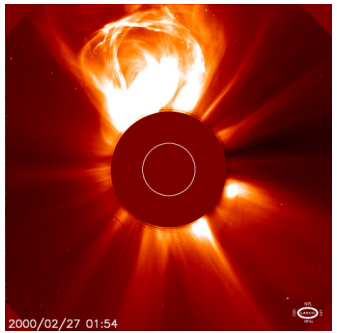
F.M.Follega on behalf of the CSES-Limadou collaboration
University of Trento and INFN-TIFPA

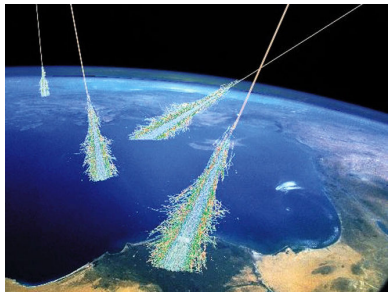
The CSES scientific mission

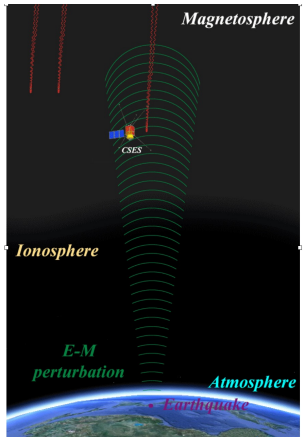
The **CSES-Limadou program** is a collaboration between Italy and China.

Scientific goals:

- **Investigate the ionosphere** and gather world-wide data;
- Measure the particles and plasma **perturbations of the atmo/iono/magnetosphere** (natural sources and anthropic emitters);
- Study solar-terrestrial interactions and solar physics phenomena: CMEs, SEPs, solar flares;
- Study **low energy cosmic rays;**



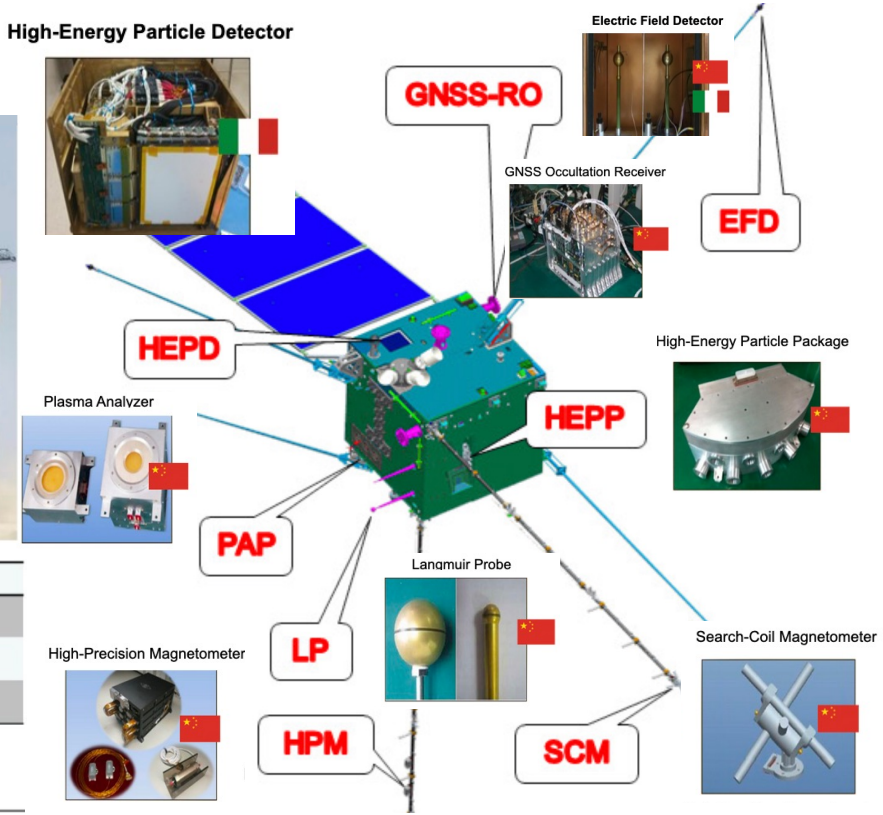




The CSES-01 Satellite

The China Seismo-Electromagnetic Satellite is a sophisticated multi-channel space observatory.

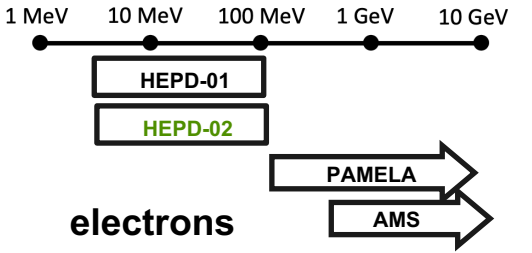
- **Launched** on the **2nd of February 2018**;
- Sun-Synchronous orbit at 500 km;
- Equipped with 9 instruments, among them the **High-Energy Particle Detector (HEPD-01)**
- Payload operation range -65°/65° lat

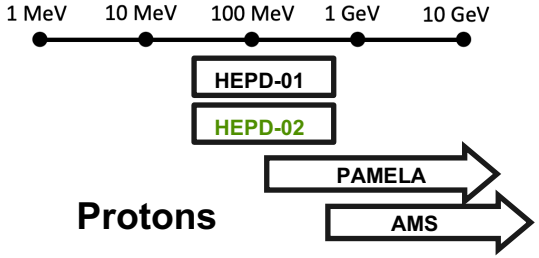


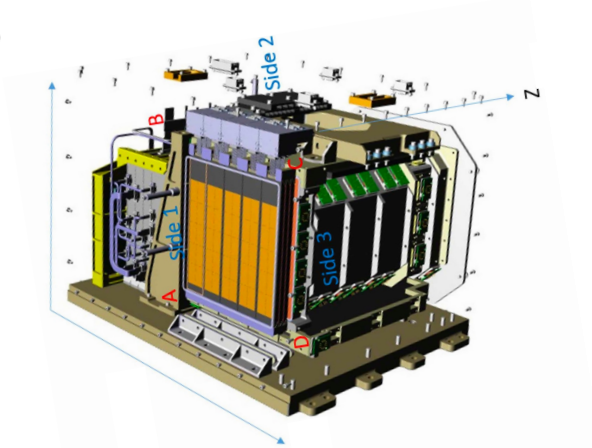
Category	Payload Name	Observation Targets
Electro-Magnetic Field	Electric Field Detector	Electric Field: DC ~ 3.5MHz
	High Precision Magnetometer	Magnetic Field: DC ~ 15Hz
	Search Coil Magnetometer	Magnetic Field: 10Hz ~ 20kHz
Energetic Particle	Italian HEPD(INFN Prod.) High Energy Particle Package	Proton： 2MeV~200MeV Electron： 100keV~100MeV

The High Energy Particle Detector HEPD-02

HEPD-02 is designed to measure fluxes of charged particles: mostly electrons (3-100 MeV) and protons (30-300 MeV) and heavy nuclei in a wide range of energies.

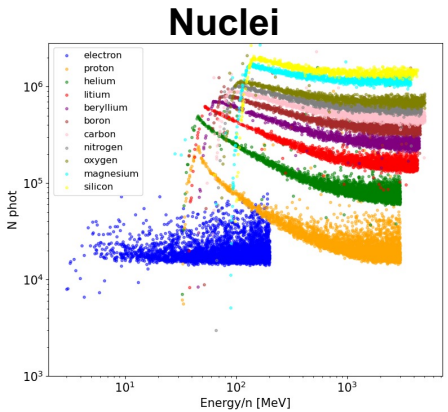


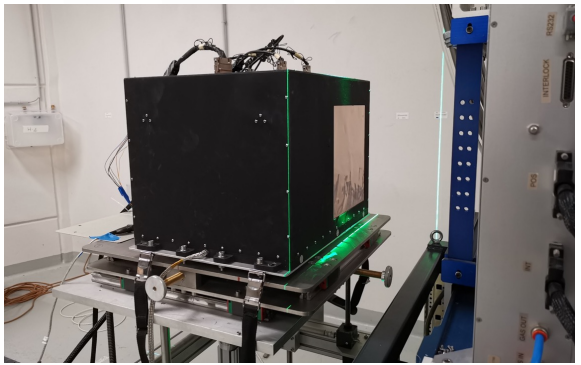




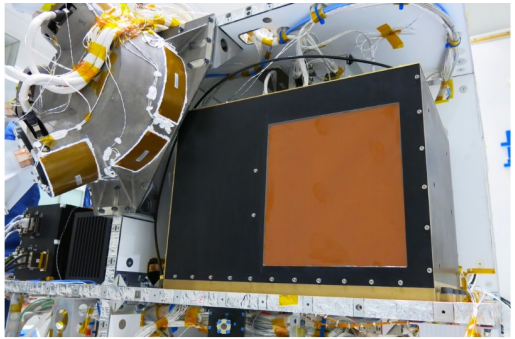
Detector subsystems:

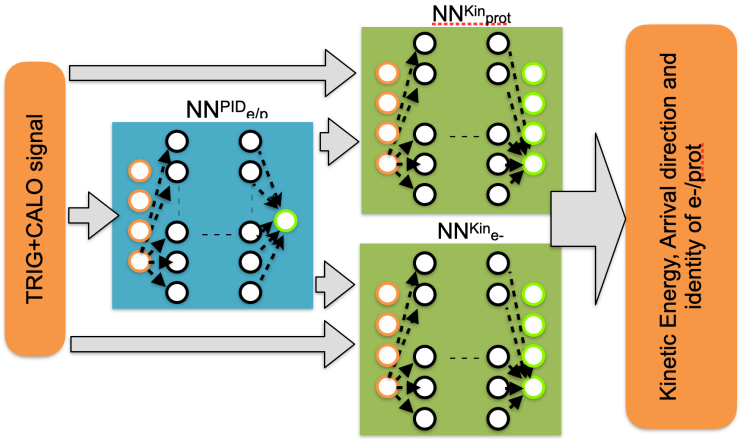
- **the tracker system -** pixel
- **two trigger planes -** EJ200
- **the calorimeter -** plastic scintillator tower (12 planes) + two layers of LYSO inorganic crystals
- **the veto system** - EJ200





Deep Learning event reconstruction HEPD-01





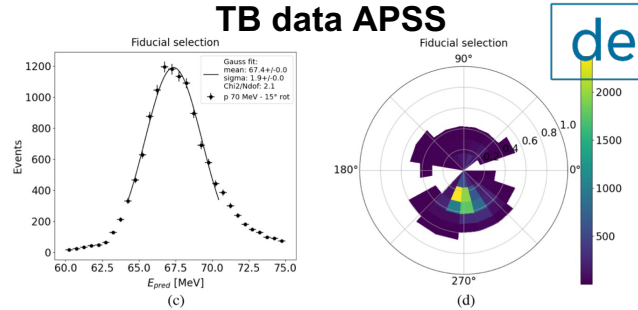


FIG. 13. Kinetic energy and arrival direction predictions from FCNN_{kin} for beamed protons. In plots (a) and (b) protons $E_{\text{beam}} = 154$ MeV and $(\theta, \phi) = (0°, 0°)$). In plots (c) and (d) protons have $E_{\text{beam}} = 70$ MeV and $(\theta, \phi) = (15°, 270°)$. The color specifies the number of events.

Phys. Rev. D 105, 022004

Flight data

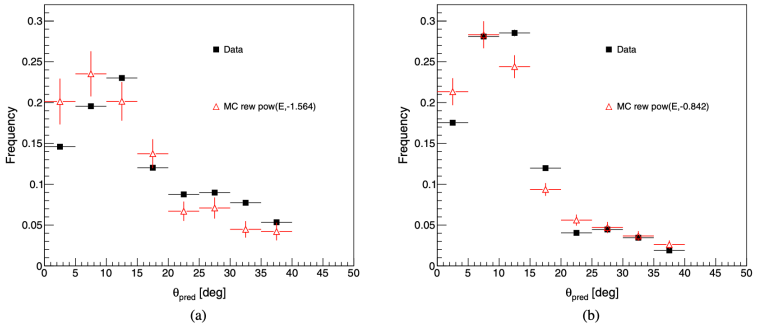


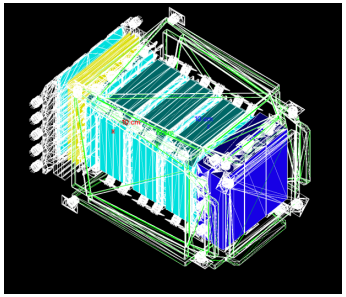
FIG. 15. Distribution of reconstructed polar angle θ for protons (a) and electrons (b). Monte Carlo events (red triangles) are reweighted as described in the text. Error bars represent statistical uncertainty.

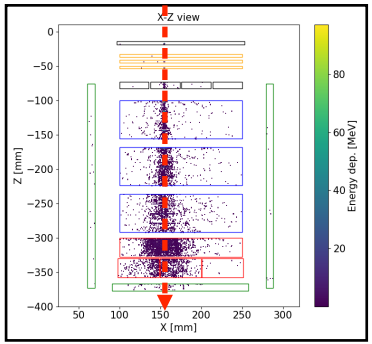
Training of Deep Neural networks

Typically event reconstruction is a set of supervised task: PID classification, energy and direction regression

Training set (MC based)

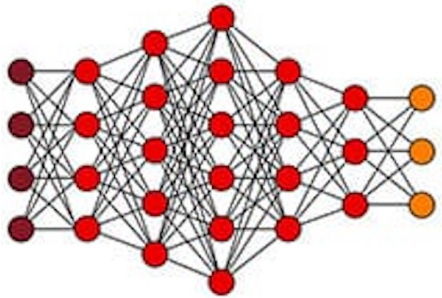
HEPD02 - G4 model

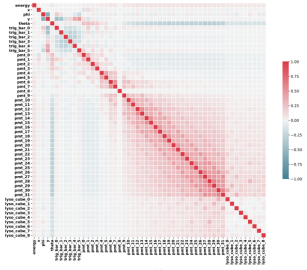




Training on CPus/**GPUs**

PyTorch TensorFlow



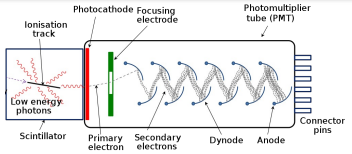


Identity/class of the impinging particle

Energy/momentum of the primary particle

Incoming direction of the primary particle

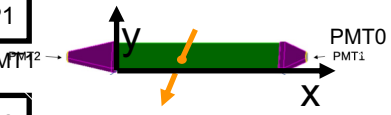
HEPD-02 simulation scheme



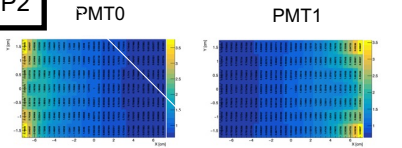
Digitization

(x,y,z,Edep) of all the hits on TR1

STEP1

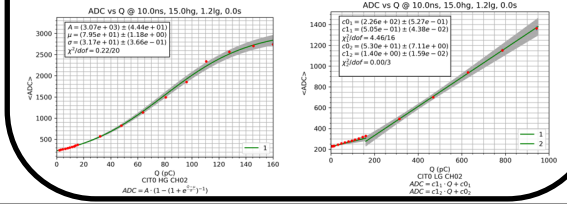


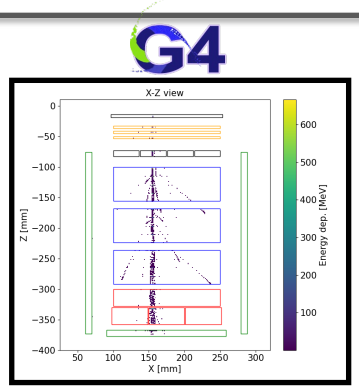
STEP2



$$NPE_{hit} = F(E_{dep}^{hit}) \times fact(x^{hit}, y^{hit})$$

$$NPE_{tot}^{PMT0} = \sum_{i=1}^{N_{hits}} Poiss(NPE_{hit_i}^{PMT0}) \qquad NPE_{tot}^{PMT0} = \sum_{i=1}^{N_{hits}} Poiss(NPE_{hit_i}^{PMT0})$$



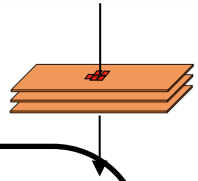


Truth level information:

- Hit position (x,y,z);
- Hit Edep;
- Detector ID

Digitized information:

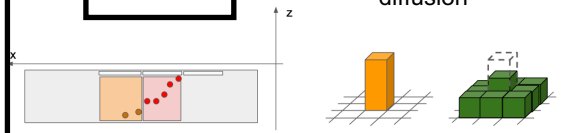
- PMTs ADC signal;
- Pixels and clusters



TROPix (pixels) + Mapping code

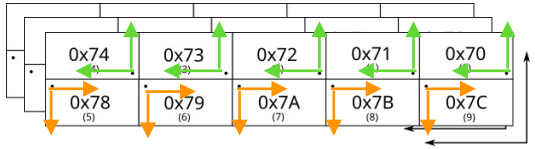
STEP1

Pixel generation + charge diffusion



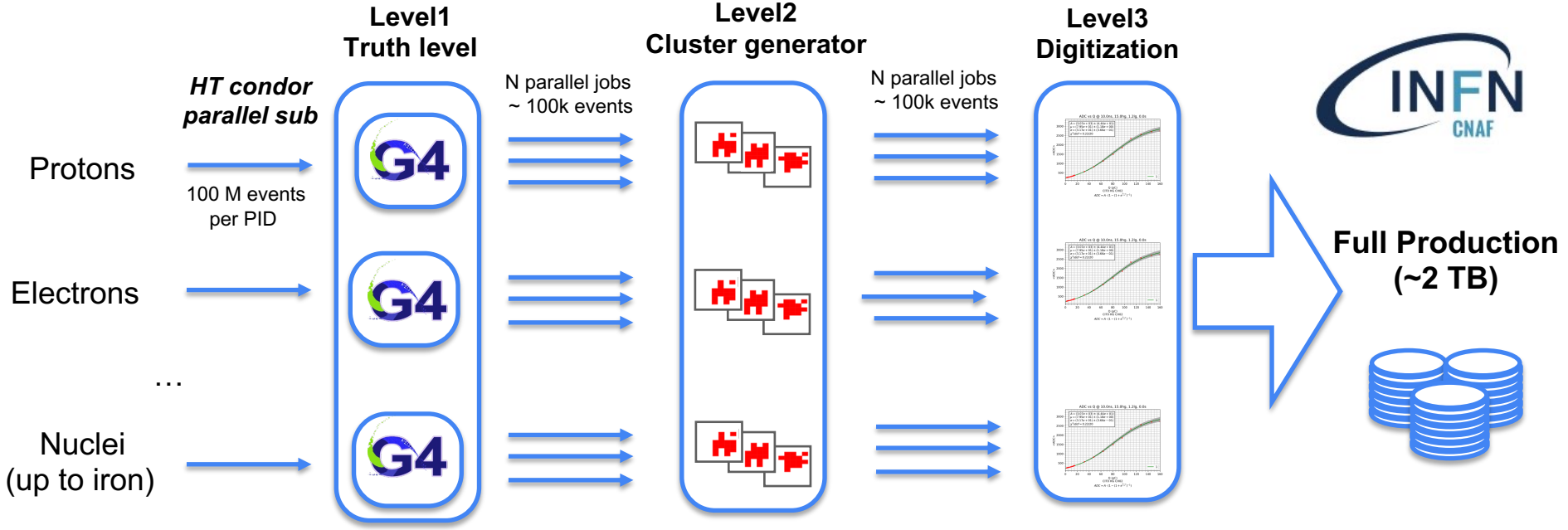
STEP2

Remapping to HEPD-02 tracker pixel coordinates



HEPD-02 simulation production

All steps of the HEPD-02 simulations are performed at CNAF and all the simulation outputs are stored in the storage area



Simulation scheme designed to be robust to reprocessing and resource effective (heaviest part at the beginning) to avoid multiple reprocessing at each small change of the chain

Preprocessing and selections

Singularity image

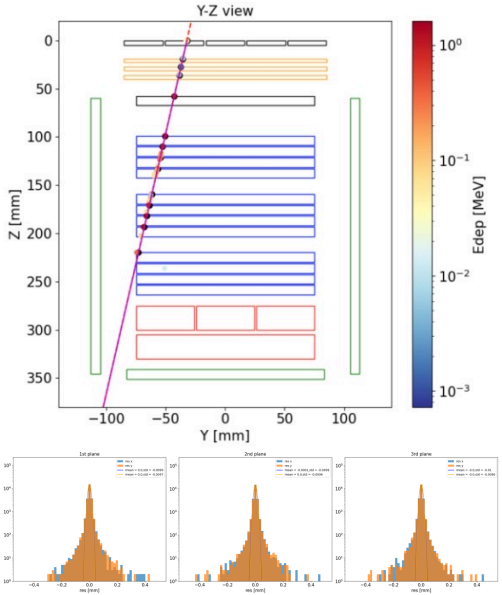
Preprocessing of training MC samples is performed @ CNAF

FULL MC simulation

electrons and protons

~ 500 GB

First level reconstruction



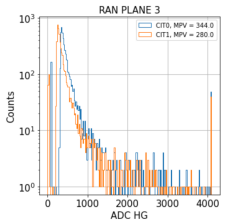
Skimming

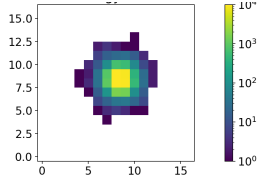
StdScaling

~ 5 GB

RAW feature from calo

HL features from tracker + cluster images

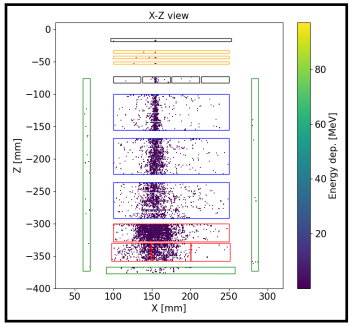




Fiducial selection for good quality events

Deep learning event reconstruction for HEPD02

Simulation&Preprocessig
@CNAF



G4 + TROPix

Cluster Size + angles

Input from Tracking algo

Data from pixel tracker as images

Layer 1

Layer 2

Layer 3

Trigger bar

Calo plane

Lyso bar

CNN cluster feature extractor

FCNN ADC signal extractor

fcl

fcl

θ

φ

Output

θ

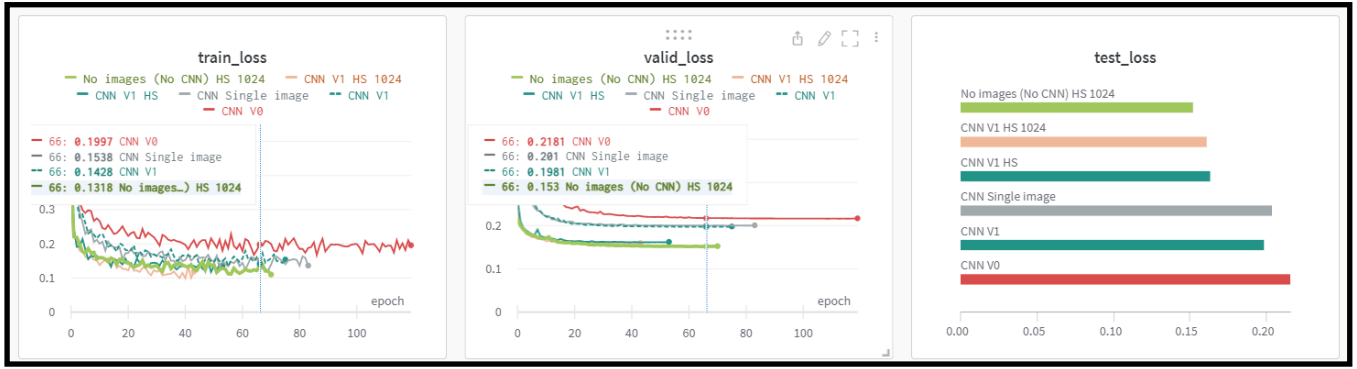
φ

PID

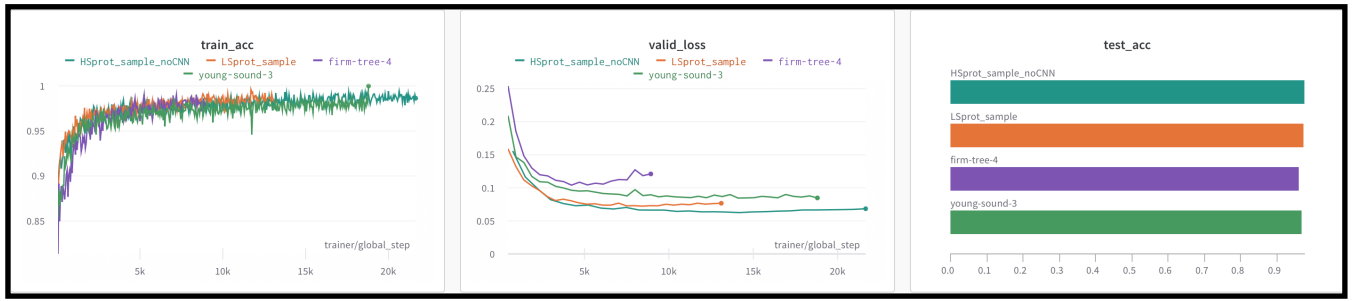
E

Training of models @ RECAS

Kinetic energy reconstruction - regression



Particle identification - classification



+ PyTorch

Usual request (one job):
- 16 CPUs
- 2 GPUs (NVIDIA A100)

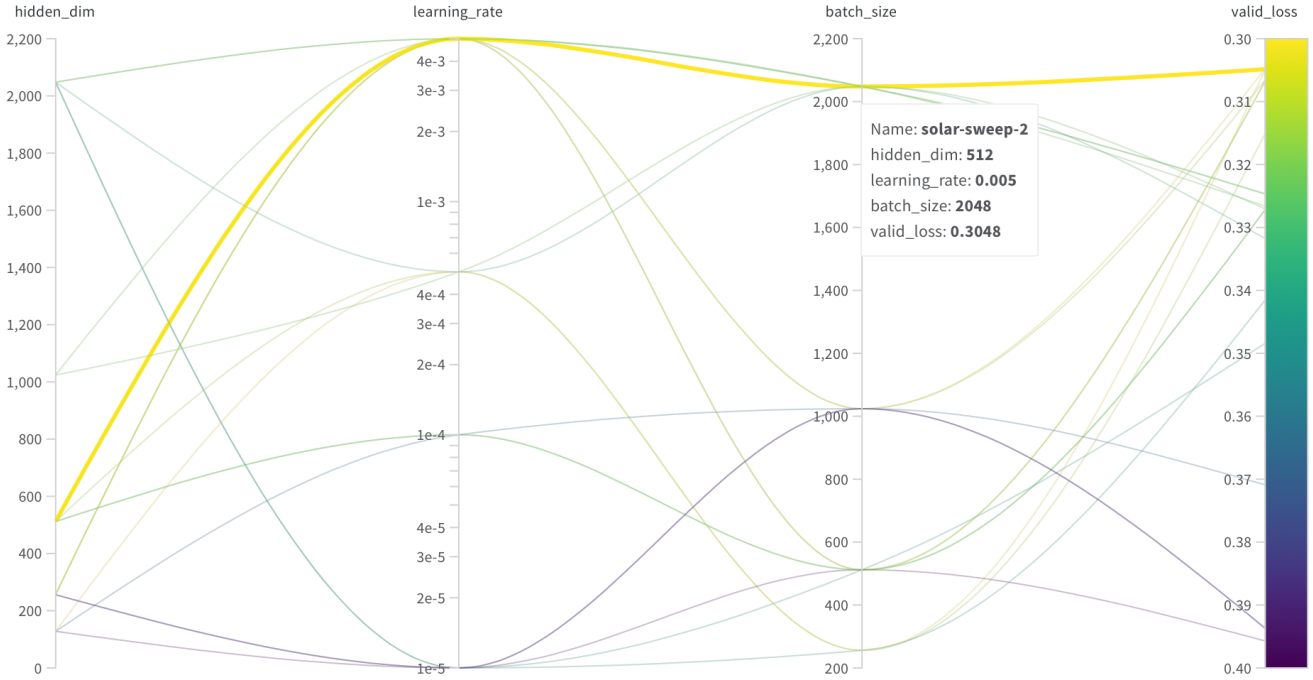
```
+-----------------------------------------------------------------------------+
| NVIDIA-SMI 520.61.05    Driver V
|-------------------------------+----------------------+----------------------+
| GPU  Name        Persistence-M|
| Fan  Temp  Perf  Pwr:Usage/Cap|
|                               |
|===============================+======================+======================|
|   0  NVIDIA A100-PCI...  Off  |
| N/A   24C    P0    33W / 250W |
|                               |
+-------------------------------+----------------------+----------------------+
|   1  NVIDIA A100-PCI...  Off  |
| N/A   24C    P0    33W / 250W |
|                               |
+-------------------------------+----------------------+----------------------+
```

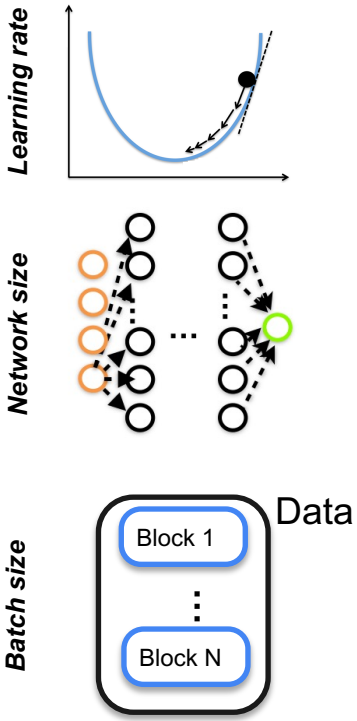
Online monitoring of model performance and storage of trained models

Weights & Biases

HP optimization of DNN

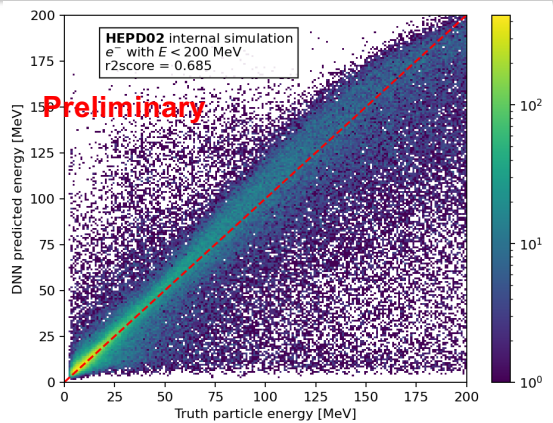
We run a grid search to optimize the deep neural network architecture/training

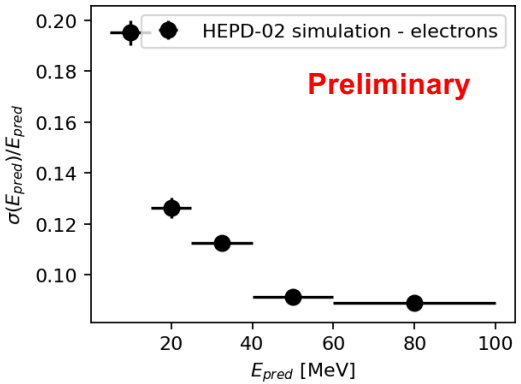




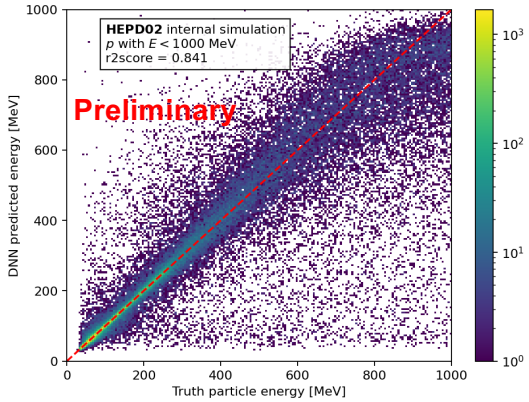
Performance in primary energy reconstruction

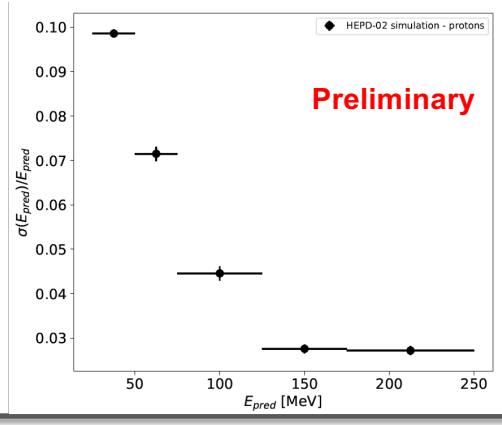
Electrons



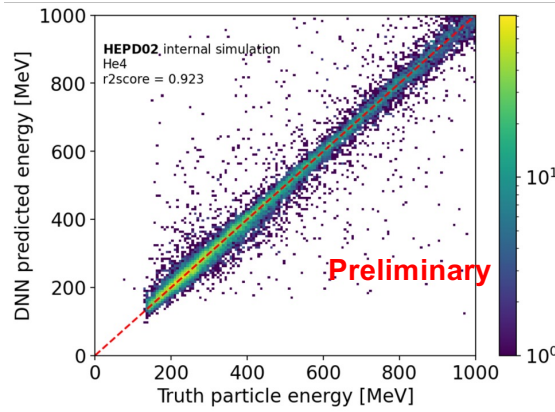


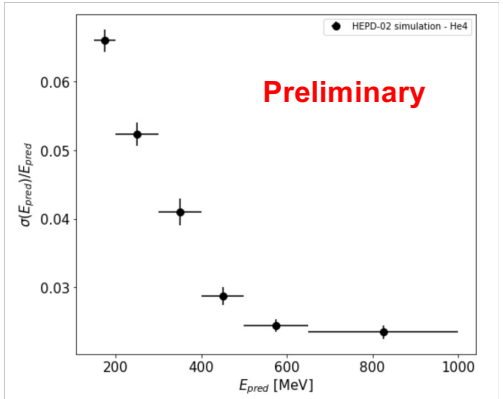
Protons



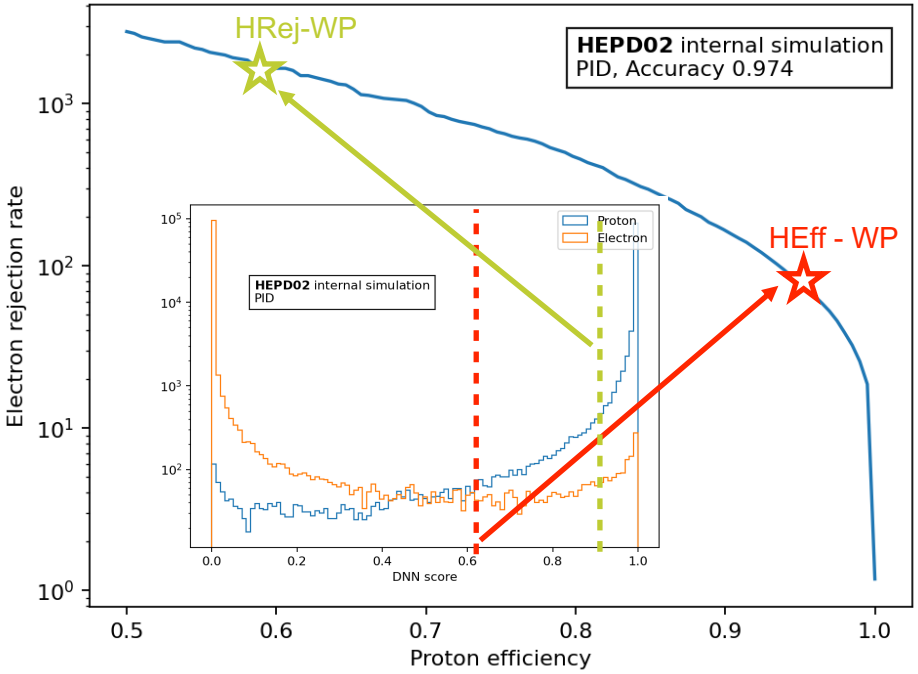


Helium 4

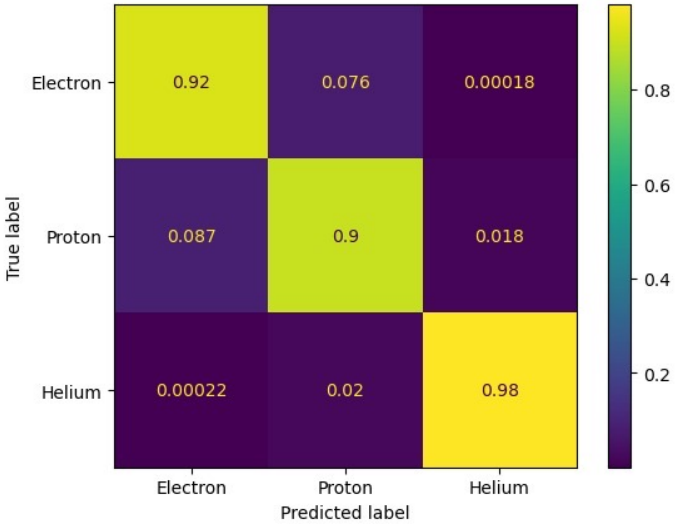




Performance particle identification



3 class classifier (e, p, heavy nuclei)



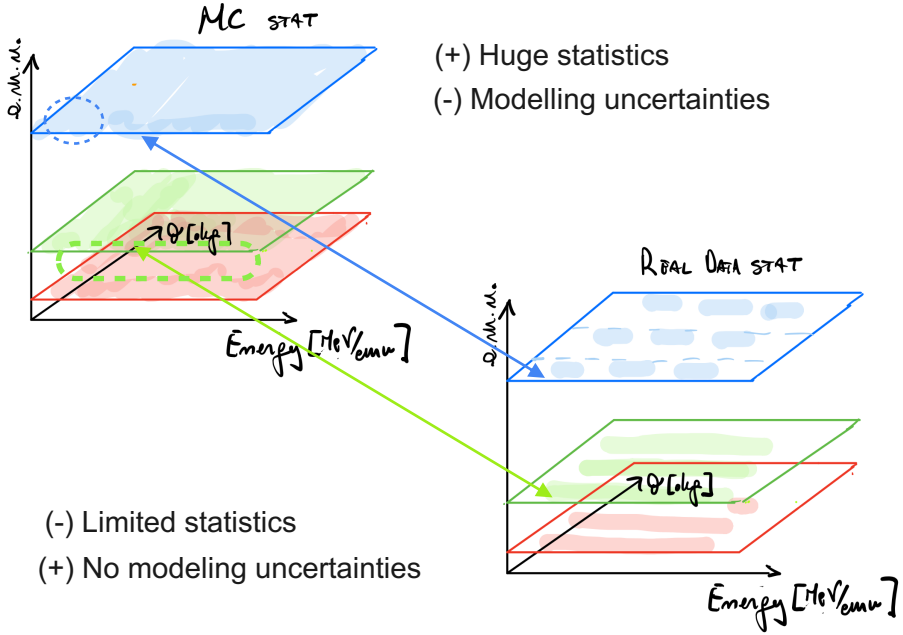
WP optimization on going!

$$\text{Signal efficiency} = \frac{Signal > TH}{Signal}$$

$$\text{Background rejection rate} = \frac{Background}{Background > TH}$$

Training on real data: domain adaptation

DL-algorithms are trained on large samples of MC data and tested on real data, **Deep Neural Networks need to be inser**

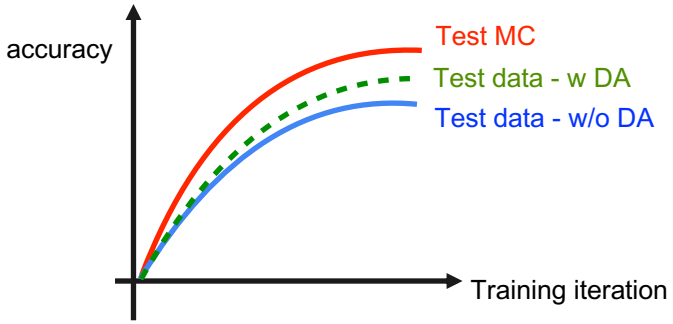


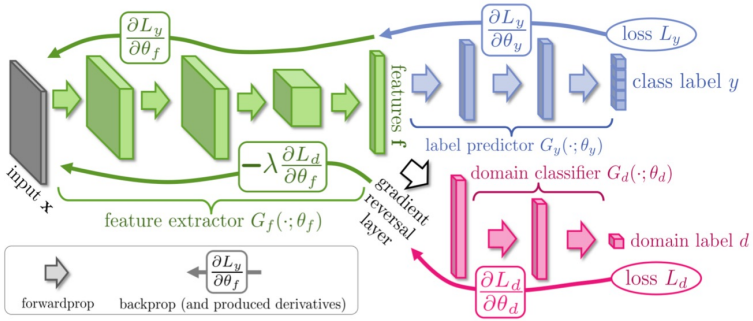
(+) Huge statistics

(-) Modelling uncertainties

(-) Limited statistics

(+) No modeling uncertainties





Summary and conclusion

DL research activity for Limadou HEPD is carried out with and thanks to INFN computing infrastructure:

- **Simulation and data preprocessing** are carried out **@CNAF**;
- **Training of DNN** and test on simulated data **@RECAS**

  - recently established a possibility of direct connection CNAF<->ASI/RECAS via a storage area webdav.

Next steps will be performed in the next months:

- Moving the model **deployment phase on shared INFN cloud machines**: VM with sing/docker images or Jupyter Notebooks.
- Validation and training of the DL evReco using test beam data: Proton Therapy (Trento), BTF (Frascati), CNAO (Pavia).