Cluster Reconstruction in SAND Electromagnetic Calorimeter

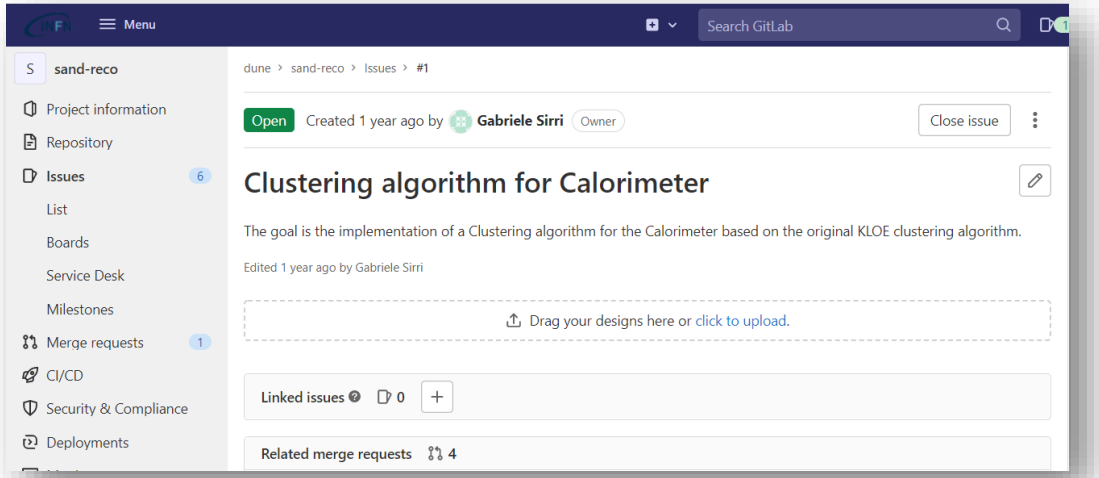
I. Cagnoli, F. Poppi, M. Tenti, G. Sirri
Bologna

DUNE/SAND SW meeting 2022.03.09

A new feature for SAND reconstruction (issue #1)

This development is related to the ISSUE #1 , created 1 year ago

https://baltig.infn.it/dune/sand-reco/-/issues/1



Goal:
implementation of a clustering algorithm for the SAND E.M. calorimenter based on the original KLOE clustering algorithm

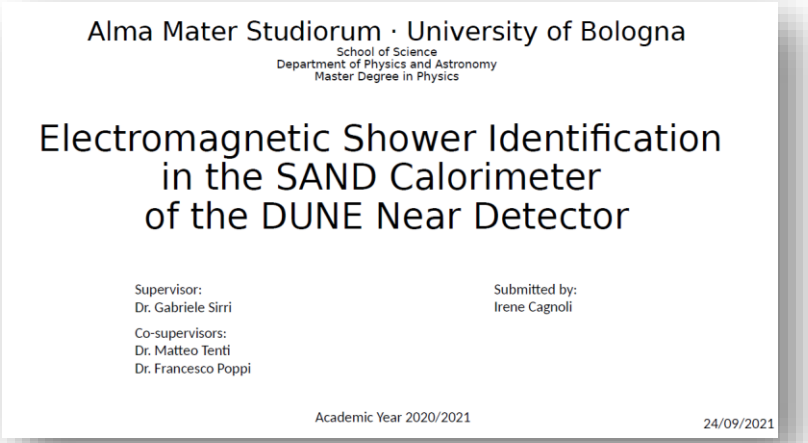
Branch: 1-clustering-algorithm-for-calorimeter (not ready for merge)

Status: **OPEN** (actually, **suspended** since Sept 2021)

people involved:
Francesco Poppi, Irene Cagnoli (master thesis), GS, Matteo Tenti

The present implementation of the algorithm
was exploited by Irene Cagnoli for her master thesis

http://amslaurea.unibo.it/23901/


Alma Mater Studiorum · University of Bologna
School of Science
Department of Physics and Astronomy
Master Degree in Physics

Electromagnetic Shower Identification
in the SAND Calorimeter
of the DUNE Near Detector

Supervisor:
Dr. Gabriele Sirri

Co-supervisors:
Dr. Matteo Tenti
Dr. Francesco Poppi

Submitted by:
Irene Cagnoli

Academic Year 2020/2021

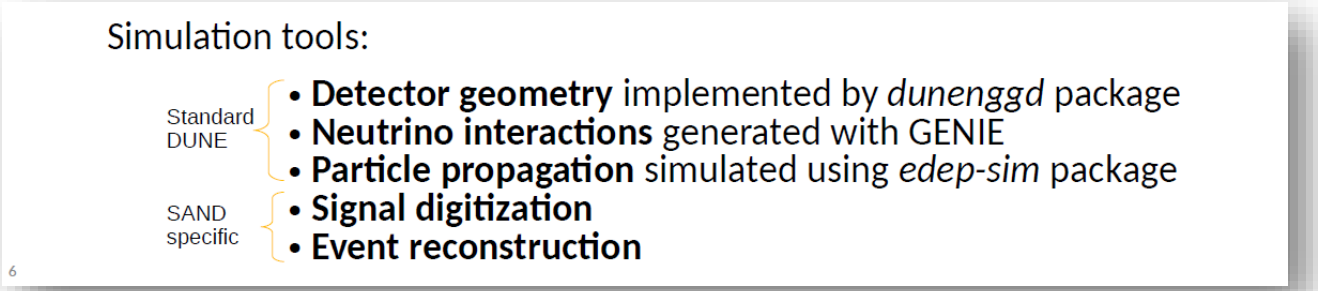
24/09/2021

with the aim of a preliminary
study on the electromagnetic calorimeter (ECAL) capabilities
to discriminate electromagnetic showers and (muon) tracks

Francesco Poppi was largely involved in the algorithm implementation.

Simulation Tools

The simulation chain of the full reconstruction was used.



from Irene Cagnoli master thesis

- Detector Geometry as available in 2021

  nd_hall_kloe_sttonly_v1.gdml

- Event generation, digitization, reconstruction in STT as usual

→ New ROOT macro /src/clustering.cpp

For prototyping and testing this clustering algorithm

Clustering algorithm

The algorithm was developed in C++ and integrated in the SAND simulation framework.

It is inspired by the reconstruction procedure developed by the KLOE experiment, but adapted and tuned for the SAND detector in the DUNE experiment.

in a nutshell:

- selects fired cells of the calorimeter
  i.e. those cells with at least one photon detected by one of the two readout channels.

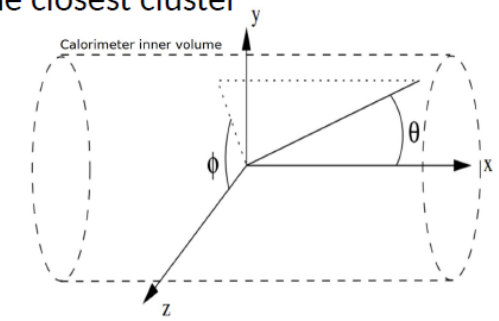
- divides fired cells into
  - complete cells which have photons detected at both sides;
  - incomplete cells, with one side without any photo-signal.

- groups neighboring fired cells into pre-clusters.
  Only complete cells are included in the pre-clusters search and the subsequent steps, while the incomplete ones will be considered only at a later stage.

ECAL reconstruction algorithm

- **Pre-cluster phase**: search for a group of neighboring cells
  cells must have t_{TDC} and Q_{ADC} signals available on both sides
- **Splitting procedure**: for clusters likely produced by the overlapping of different particles trajectories
- **Merging procedure**: for clusters likely originated from the same particle
- **Track fitting**: to each cluster is associated a direction vector
- Cells with missing t_{TDC} or Q_{ADC} signals in the barrel (endcap) are recovered
  if $\Delta\phi$ ($\Delta\theta$) < 3°
  $\Delta\phi$ ($\Delta\theta$): angular distance between the cell and the closest cluster

OUTPUT: **cluster**
- **energy**: sum of energies for all cells
- **position**: energy-weighted averages
- **position variances**
- **average time:** energy-weighted average
- **direction vector**
- **set of cells** belonging to the cluster



from Irene Cagnoli master thesis

Clustering algorithm current status

The current implementation of the clustering algorithm can be used only for preliminary test and validation studies.

Several steps still to be finalized:

- Software code not peer reviewed yet,
  i.e. not examined by a developer other than its author(s)

- Some features still missed,
  see ISSUE #4, https://baltig.infn.it/dune/sand-reco/-/issues/4
  For each event, the current clustering takes into account only the first photo signal of a sensor
  Improve the algorithm to consider the full sequence of photo signals generated by each sensor.

- Integration with other algorithms which are (or may be) available in the SAND-RECO SW not done
  i.e. not possible to select between different reconstruction algorithms and compare them

Preliminary Performance Evaluation

- A preliminary evaluation of the performance was done in the framework of the Irene Cagnoli's master thesis

  for the discrimination between electromagnetic (e.m.) showers and muon tracks

The classication analysis was operated in two separated steps:

(a) multivariate algorithms are trained and validated using a dedicated dataset obtained by simulating particle guns of electrons and muons originated close to the ECAL;

(b) the selected optimal classifiers are tested with simulated neutrino interactions in the SAND detector considering the case of a ν_μ-dominated beam in neutrino mode.

(a) simulated dataset: e^- and μ^- particle guns

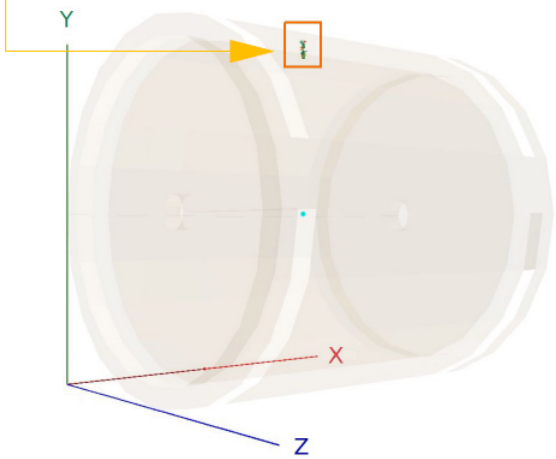
PRELIMINARY

e-/μ- training dataset

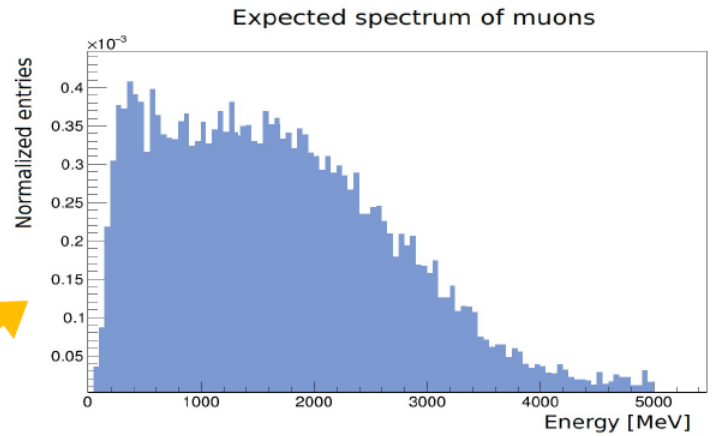
particle guns

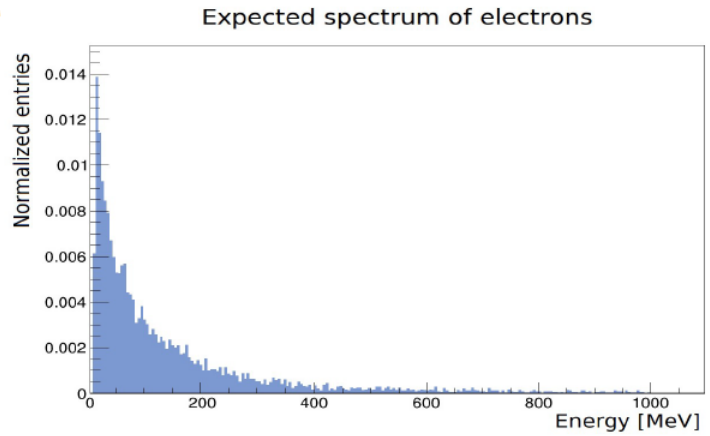
- originated within SAND, 1 cm from the ECAL, at the center of the ECAL module
- perpendicular to the ECAL module

- 410k e- and 410k μ- events
- energy spectra as from νμ interactions in the SAND fiducial volume



Event display of an e.m. shower in the ECAL originated by a 1 GeV e-





9

from Irene Cagnoli master thesis

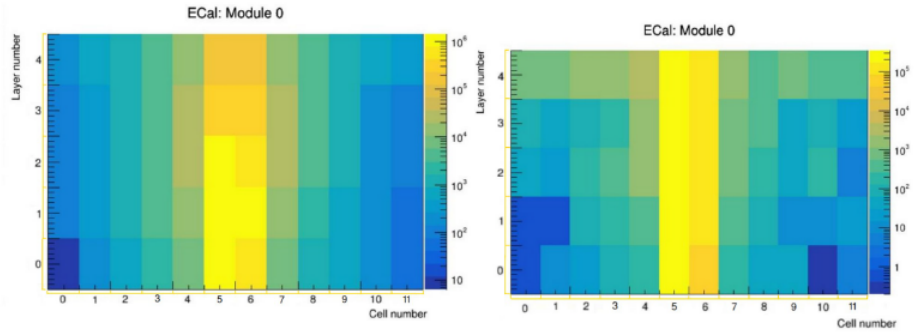
(a) Multivariate analysis (Fisher, MLP, BDT)

PRELIMINARY

Multivariate Analysis

using the TMVA package:

- **Classes:**
  signal: e- clusters
  background: μ- clusters
- **Weights:** signal and background with w= 1.0
- **Classification methods:**
  - Fisher method
  - MLP (Neural network)
  - BDT (Boosted Decision Tree)
- **10 input variables:**
  5 → visible energy deposited in a ECAL layer
  5 → #cells fired in ECAL layer
- **Preliminary cut:** only events with just 1 cluster reconstructed



Distribution of energy released in the calorimeter by an electron (left) and a muon (right)

	Signal (e-)	Background (μ-)
Total n. of clusters used	222 157	316 148

10

from Irene Cagnoli master thesis

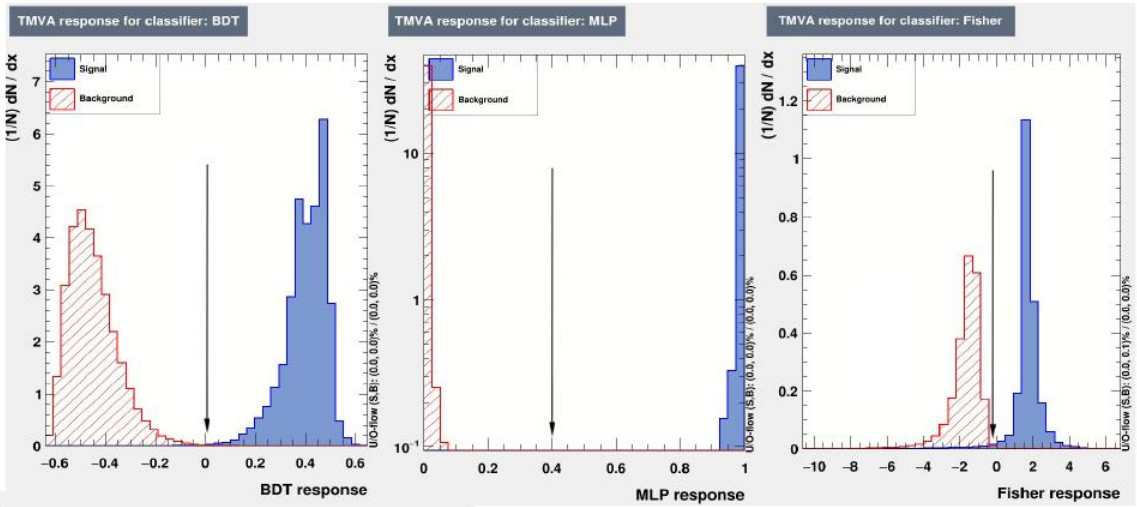
(a) e^- and μ^- discrimination

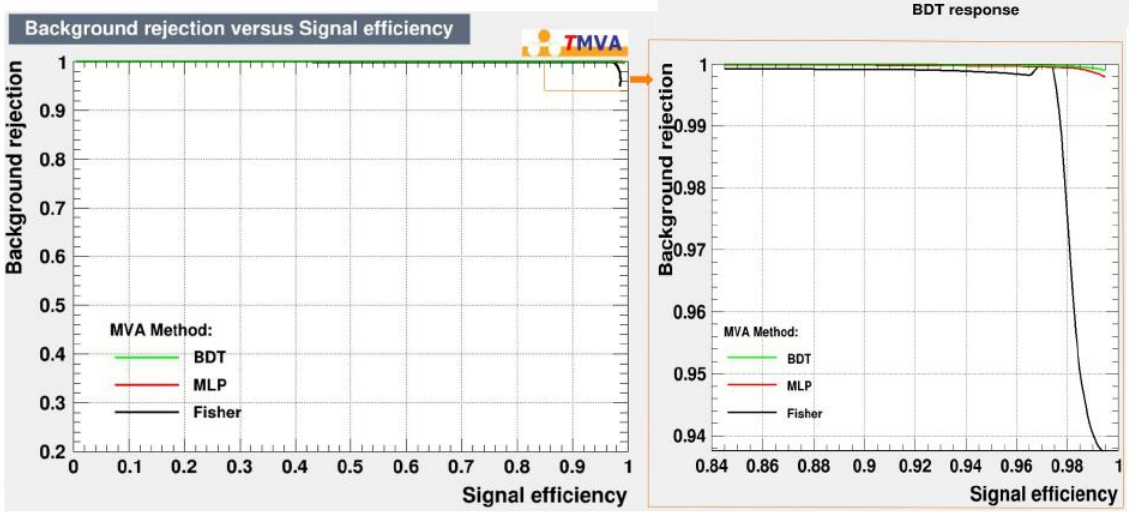
PRELIMINARY

Training and validation

Classificator outputs

e.g.	Cut	Signal efficiency	Background rejection
BDT	0.001	0.998	0.998
MLP	0.400	0.997	0.997
Fisher	-0.300	0.980	0.996





ROC curve

11

from Irene Cagnoli master thesis

Using the simulated dataset of e and μ particle guns, we obtained an Area Under the ROC Curve very high for both BDT and MLP

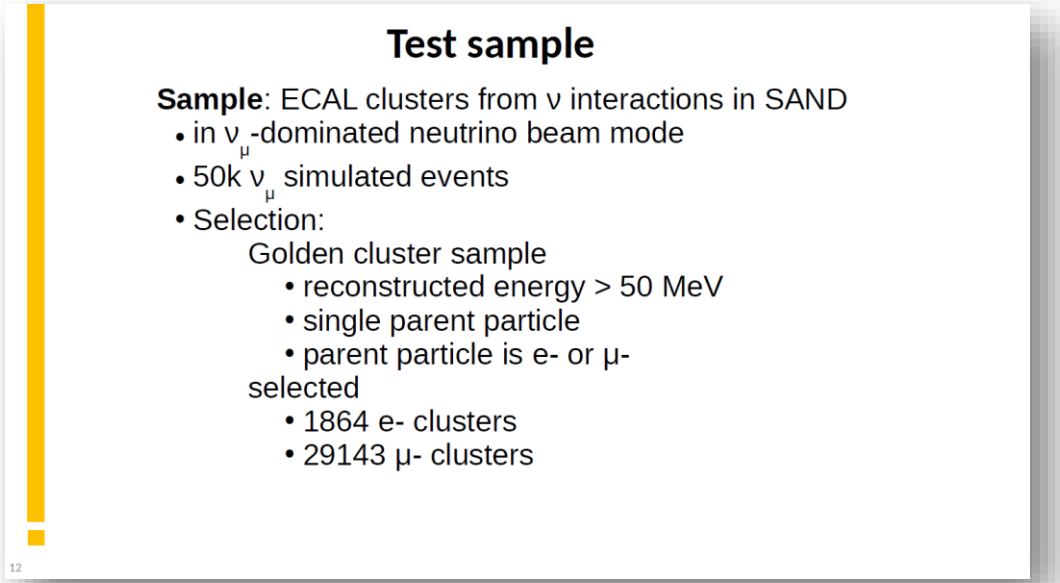
(b) Test with ν_μ interactions simulated in SAND

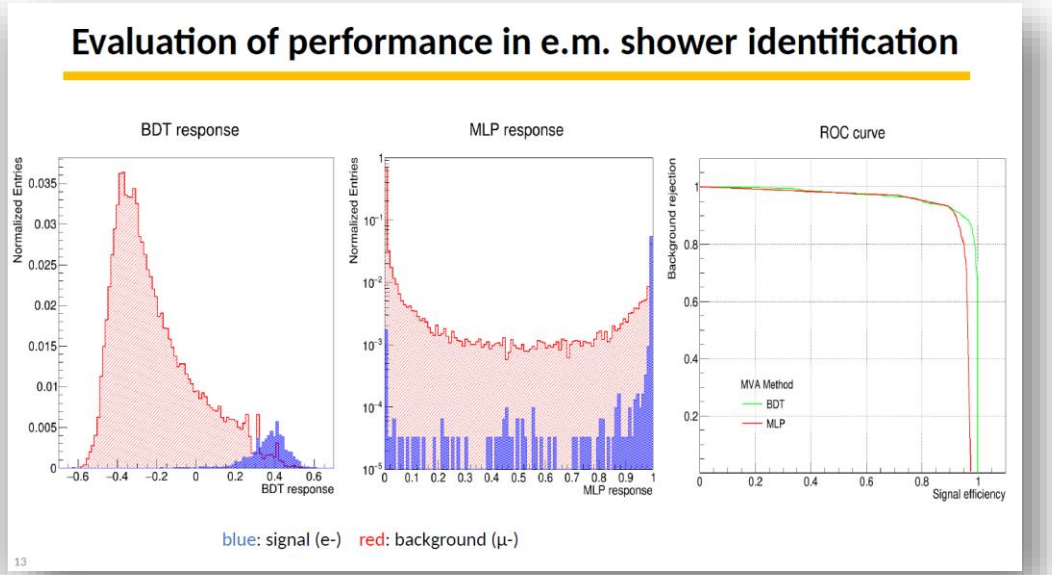
We apply the multivariate analysis (as trained with particle guns) to see the capabilities of ECAL to discriminate shower/tacks in a sample of neutrino interactions simulated inside SAND

This task opens requires a major step: a strategy to define a true primary track for each reconstructed cluster must be implemented in the analysis; otherwise is not possible to evaluate the success rate of the classification.

We started the effort (https:/baltig.infn.it/dune/cluster-analysis) but the task is quite challenging.

At the end, we just apply some selection to the neutrino simulated sample to obtain preliminary numbers but these results cannot be considered conclusive.





from Irene Cagnoli master thesis

Conclusions

- A new reconstruction algorithm has been developed and validated for the SAND ECAL
- A MVA is implemented to study the capabilities of the ECAL as a stand-alone detector in e.m. showers and muon tracks discrimination
- Promising results have been obtained
- Next steps
  - Training with neutrino events
  - Combine data from all SAND sub-detectors
  - Multiclass classification
  - ...

from Irene Cagnoli master thesis