PROCEEDINGS OF SCIENCE

Identification of Beam Particles Using Detectors based on Cerenkov effect and Machine Learning in the COMPASS Experiment at CERN

František Voldřich,[a,*] **Marcin Stolarski,**[b] **Martin Zemko,**[a] **Flavio Tosello,**[c] **Josef Nový**[a] **and Miroslav Virius**[a]

[a] *Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering, Brehova 78/7, Prague 115 19, Czech Republic*

[b] *Laboratory of Instrumentation and Experimental Particles Physics, Av. Prof. Gama Pinto 2, 1649-003 Lisbon, Portugal*

[c] *Torino section, National Institute for Nuclear Physics, Via Pietro Giuria, 10125 Torino, Italy*

E-mail: voldrichfrantisek@gmail.com, Marcin.Stolarski@cern.ch

Cerenkov Differential counters with Achromatic Ring focus (CEDARs) in the COMPASS experiment beamline were designed to identify particles in limited intensity beams with divergence below 65μrad. However, in the 2018 data taking, a beam with a 15 times higher intensity and a beam divergence of up to 300μrad was used, hence the standard data analysis method could not be used. A machine learning approach using neural networks was developed and examined on multiple Monte Carlo simulations. Different types of network were tested and their configurations optimized using a genetic algorithm with the best performing model being integrated into the current data analysis software written in C++.

41st International Conference on High Energy physics - ICHEP2022
July 6th - July 13th, 2022
Bologna, Italy

[*] Speaker

PoS(ICHEP2022)244

1. Introduction

The COMPASS experiment at CERN focuses on study of hadron structure and spectroscopy using high intensity muon and hadron beams. Hadron beams are a mixture of particles, e.g. a 190 GeV negative hadron beam contains about 97 % of π^-, 2.5% of K^- and < 1% of $\bar{p}$. CEDAR detectors in the COMPASS experiment beamline were designed to identify particles in limited intensity beams with a divergence below 65 μrad. However, during the 2018 data taking, a beam with 15 times higher intensity was used.

CEDARs were prepared to withstand such conditions by a major upgrade of the frontend electronics and photomultipliers as well as a redesign of the corresponding firmware. In addition, the beam divergence of those runs was up to 300 μrad with only 10-15 % of particles being within the designed divergence range. Hence, the previous method of data analysis using the likelihood approach cannot be used in this case. [1][2]

This work (for details, see [3]) presents a new approach for particle separation based on artificial neural networks developed in Python, their optimization, analyses and integration into C++ framework. In addition, effects aggravating separation are identified and possible improvements for the AMBER experiment replacing the COMPASS experiment in 2023 are suggested. All results presented in this work are based on Monte Carlo simulations.

2. CEDAR detectors

CEDAR is an abbreviation for 'CErenkov Differential counters with Achromatic Ring focus', a name used for a detector based on Cerenkov effect used for π^-, K^- and $\bar{p}$ identification. Two alike CEDARs are positioned approximately 30 meters upstream of the COMPASS target. [1][2][4]

CEDARs detect rings of Cerenkov photons of predefined radius, which is different for different particles (due to different emission angles), using 8 photomultipliers (PMTs). Originally, they were designed as a majority counter, where for highly parallel beam 6-8 PMTs should fire to positively identify a particle. However, not all beam particles traverse CEDARs parallel to their optical axis. Although the beam inclination leads to a low efficiency of the majority approach, a certain PMTs hit pattern is characteristic for different particle species at a given inclination. This phenomenon was utilized to develop a likelihood ansatz where response of each PMT was parameterized individually as a function of radiation angle (for details, see [2]), to better compensate for the aforementioned issues.

2.1 Challenges with 2018 data taking

In 2018, a beam with 15 times higher intensity was used, which results in observing correlated events (additional undetected track). In addition, SI trackers were replaced by FI detectors that can withstand the increased radiation, but have lower angle measurement precision. Because of that, the likelihood method cannot be used in this case and since the AMBER experiment will also use a beam of the same composition, it will not be applicable for its data taking as well.

[0] In reality, the PMTs are shifted by 22.5 clockwise.

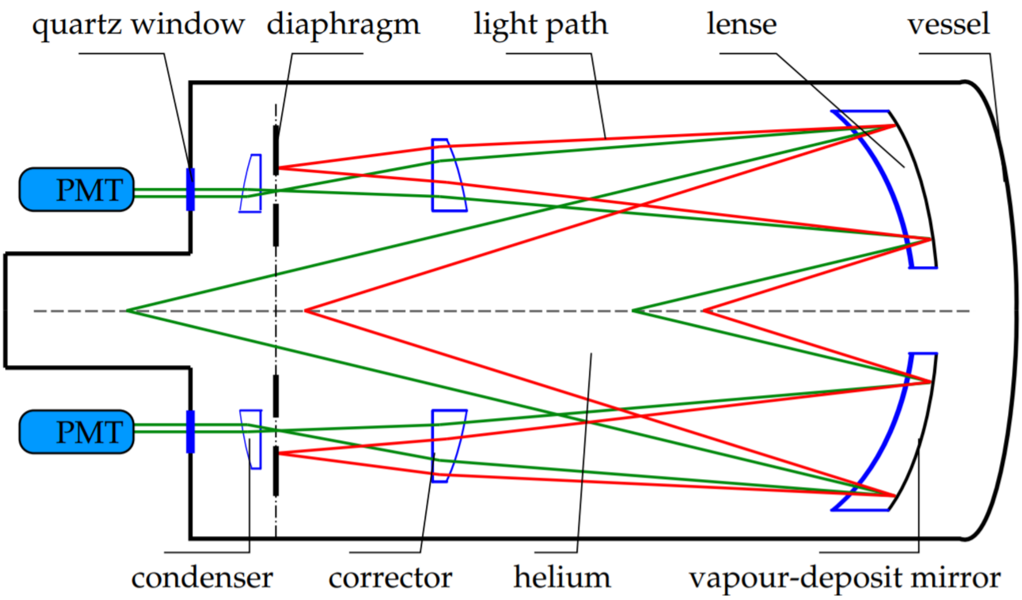

Figure 1: CEDAR schematic - two particles with the same momentum but with different masses, represented by red and green lines, radiate Cherenkov light at different angles. This results in rings with different radii. A diaphragm selects the rings from the required particle type (green line), while the radiation from the other particle does not go through the diaphragm to trigger signal from the PMTs. [1]

3. New approach

There are three methods using artificial neural networks (ANNs) proposals to examine with the goal of selecting kaons (signal) and discarding pions (background). While the first one tries to directly predict the particle type, the other two formulate and compare two hypothesis to get the final likelihood.

1. **Method 1: ANN as a direct binary classifier**
   For input we use beam position and angle at spectrometer together with responses from PMTs of each CEDAR. The network has only one output neuron that can be interpreted as a likelihood (due to the imbalance between kaons and pions in the sample) of observing a signal event.
2. **Method 2: ANN as a PMTs response predictor**
   The network is used to predict a probability that a certain PMT fires assuming the event is induced by kaon or by pion. One model is trained on signal sample, the other on background sample. The input consists of position and angle at spectrometer and output contains 8 neurons representing each PMT of a CEDAR. We then calculate both of these likelihoods for a given event and take their ratio. This method could show problematic as it does not take account correlation between responses of the PMTs.
3. **Method 3: ANN as a CEDAR response predictor**
   The aim of this method is to take into account the above mentioned correlation between PMTs responses, so the ANN predicts the exact pattern of PMTs assuming the type of the particle that induced it. The likelihoods are then aggregated similarly to the previous method. Hence,

we have 4 input neurons and 256 output neurons[1] for each CEDAR. This network might suffer from instability.

3.1 Methods comparison

As shown in fig. 2, method 2 performed significantly worse than the other two methods, probably due to the correlation between PMTs responses. Method 1 and 3 performed similarly and yet no problems with instability in method 3 was observed. However, since method 1 is notably faster and easier to work with, it was used for further analysis.

3.2 Meta parameters optimisation

Three different network types were impemented: Multi Layer Perceptron (MLP), Radial Basis Function and Random Vector Functional Network. Performance of these networks with hyper parameters selected by hand was compared with MLP being slightly better than the other two and hence selected for further work.

For hyper parameters optimisation of the MLP network such as number of hidden layers and neurons, learning rates, activation functions etc., a differential evolution heuristic was used. Its results were comparable to structures selected by hand, which implies that the problem is insensitive to network structure (likely due to its simplicity).

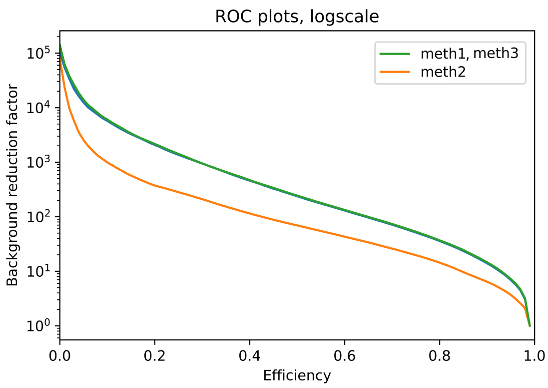

Figure 2: ROC curves of the new methods.

3.3 Dataset size

Monte Carlo used simulations contain ≈80 million events. Analysis and training over such datasets takes a long time. By iteratively halving the training dataset and batch size, we determined that there is a limit of improving results with the use of more data. It appears the improvements are only significant to ≈300 thousand events, which accounts for only ≈6.5 thousand of kaons. This may prove to be important especially when obtaining a kaon sample from measured data.

3.4 Inference in C++

Aforementioned models were developed using TensorFlow and Keras. Using the frugally-deep library, the best performing trained model was exported to a binary file, which was then loaded using the mentioned library in a C++ program. This program can execute inference with no dependence on the Python Interpreter. [5]

A class for handling the loading and inference was created. For using the model in PHAST, a C++ framework used for COMPASS data analysis, one must include it in `lib` or `user` directory to be automatically compiled and linked by PHAST upon its compilation, because it is not part of the standard installation. [6]

[1]Each of the 8 PMTs can output 2 values, i.e. $2^8 = 256$ combinations.

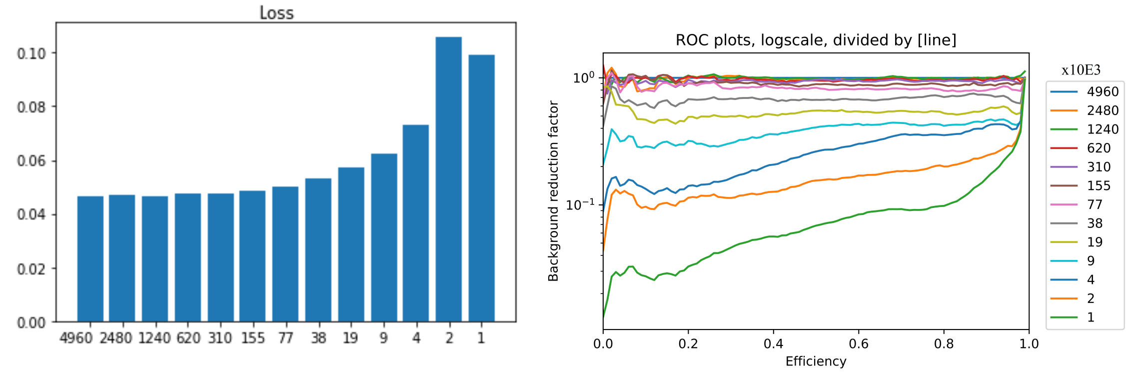

Figure 3: Values of loss functions and ROC curves of different dataset sizes (divided by curve of the largest dataset) used for training.

4. Possible improvements for the AMBER experiment

Because the AMBER Drell-Yan experiment will use a beam of the same composition, the results are applicable for it as well. For this reason, possible improvements (see the performance of the model in tab. 1) of the new method were examined, which required identification of the effects aggravating separation.

	predicted 0	predicted 1
true 0	4 819 955	19 124
true 1	60 667	60 666

Sensitivity	0.5000
Specificity	0.9961
Accuracy	0.9839
Background reduction	253.04

Table 1: Confusion matrix and some statistics for 50 % working point.

Four issues needed to be taken into account in Monte Carlo simulations to better match data:

1. MC-1xxx: additional undetected track (correlated noise)
2. MC-x1xx: additional random noise
3. MC-xx1x: PMTs inefficiency
4. MC-xxx1: angle smearing

Combinations of different effects were examined to identify the main factors complicating separation. The biggest improvement can be achieved by removing angle smearing.

4.1 Further improvements

In reality, each PMT consists of four pads giving signal individually. Replacing binary response of PMTs with number of active pads improves the separation almost as much as removing correlated noise, using the response of each pad is expected to perform even better.[2]

[2]These simulations were unfortunately not available at the time of writing this work.

In addition, pads give signal for 10ns. This time can be reduced and signal off time hits caused by undetected tracks discarded.

5. Conclusion

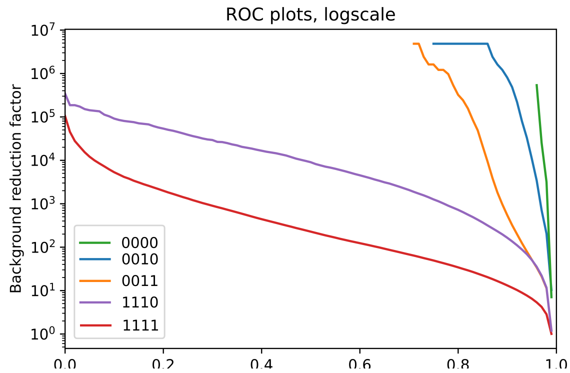

Figure 4: The biggest improvements achievable by removing 1, 2, 3 and all problems.

Three new methods and tools for their performance analysis were developed and the underlying ANN was optimized using genetic heuristic. The best performing model was exported and integrated to a C++ framework PHAST with no dependency on Python interpreter for inference. In addition, issues aggravating separation were identified and improvement achievable by resolving them was quantified, the most significant being angle measurement precision. This can be accomplished by replacing FI detectors with radiation hard SI trackers. Meanwhile, correlated noise can be improved by fastening electronic, i.e. shortening signal length, random noise can be reduced by improving shielding and efficiency by enlarging the diaphragm opening (in this moment, this would also increase correlated noise). Further opportunities for improvement were recognized by using PMTs pads individually.

References

[1] ABBON, P. The COMPASS experiment at CERN. In: *Nuclear Instruments and Methods in Physics Research A577*. 2007, s. 455-518.

[2] Wallner, S., Bicker K., Friedrich J.M., Grube B., Huber S. a Krämer M. *CEDAR PID using the Likelihood Approach for the Hadron-Beam: 2008 Data Set*, COMPASS Note 2017-1, [online]. In: . 2017 [cit. 2021-7-17]. Available at: https://wwwcompass.cern.ch/compass/notes_public/2017-1.pdf

[3] VOLDŘICH, František. *Identiffcation of Beam Particles Using CEDAR Detectors and Machine Learning in the COMPASS Experiment at CERN*. Prague, 2022. Diploma thesis. Czech Technical University. Supervisor: Martin Zemko, Consultant: Marcin Stolarski.

[4] WEISROCK, Tobias. Likelihood Methods for Beam Particle Identification at the COMPASS Experiment. Prague, 2012. Availible at: https://wwwcompass.cern.ch/compass/publications/conf_proc/proc/t2012/weisrock_praha2012_proc.pdf

[5] HERMANN, Tobias. *Frugally-deep* [online]. In: . 2022 [cit. 2022-04-15]. Available at: https://github.com/Dobiasd/frugally-deep

[6] GERASIMOV, Sergei. *PHysics Analysis Software Tools* [online]. 2021 [cit. 2021-7-17]. Available at: http://ges.web.cern.ch/ges/phast/