# Particle Identification with MPD-TPC tracks

MexNICA Collaboration Winter Meeting

Julio César Maldonado González

December 16th, 2020

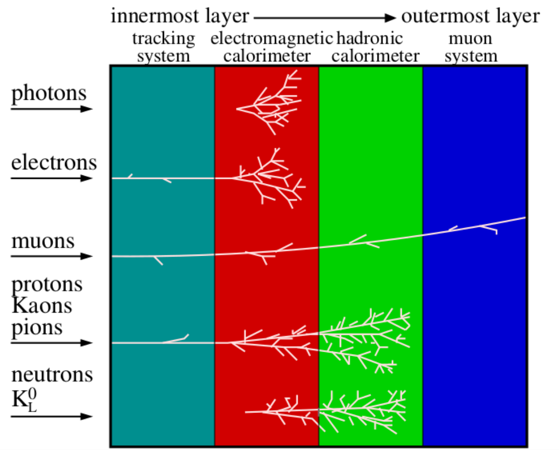
Facultad de Informática Culiacán - Universidad Autónoma de Sinaloa

# Table of contents

# Introduction

# Introduction

- Particles are accelerated and collided at high energies.
- We can only see the information stored in the detectors translated into the form of tracks of the trajectories of the particles. [1]



# Introduction

This project will participate in the particle identification analysis used in the simulation and generation of events to study physics in heavy ion collisions for the Multi-Purpose Detector (MPD) experiment of the Nuclotron Ion Collider fAcility (NICA) at the Joint Institute for Nuclear Research (JINR) in Dubna, Russia.

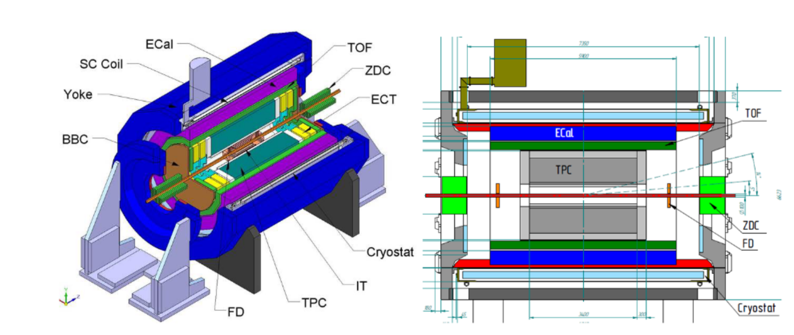



**Figure 1:** MPD sub-detectors at NICA. [3]

# Introduction

- In the experimental high energies physics, it is essential to identify the type of particle that occurs in the Particle Identification (PID) experiment.
- These identification techniques depend on the properties or observables that are obtained from the experiment. [2]
- From reconstruction on the tracks of the particles, you can get the momentum, energy loss, and other features of the particles as they pass through the TPC detector.
- MPD tracks features could be used as input data for machine learning techniques like GLMs models.
- Statistical methods (Bayesian approach) results could be compare with results obtained by machine learning.

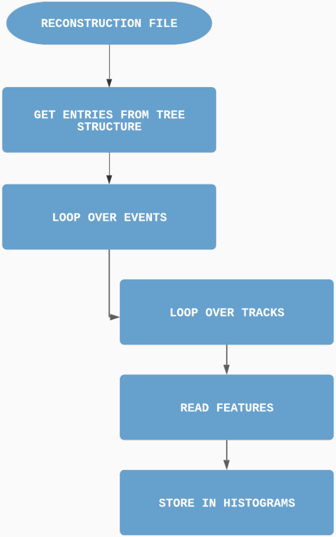
# Reconstruction

# Reconstruction

- The **event reconstruction** consists in finding particles tracks using the Kalman filter technique. [3]
- The **physics analysis** consists in finding the PID from observable signals from detectors through reconstruction data.

# Reconstruction

A brief example of a MPDROOT macro to read reconstruction files:



# Data for simulation and reconstruction

- Detector: MPD (TPC)
- Input file: reconstruction.root file (DST)
- Event generator: UrQMD
- Bi-Bi a 11 GeV (MB)
- Number of events: 10k
- Macro base: CompareSpectra.C, anaDST.C

Cuts suggested in the TPC documentation [7]

- Eta cut ($\eta < 1.2$)
- $P_T$ cut ($1.0 > P_T > 0.1$)
- Primary and secondary particles ($MotherID \leq 0$, $MotherID > 0$ )

# Bayesian method

# Bayesian method

The probability of a particle $i$, if a signal $s$ is observed,

$$\omega(i|s) = \frac{r(s|i)C_i}{\sum_k r(s|k)C_k}$$

- $r(s|i)$ Conditional probability density function of the signal observed $s$ of a detector, if a particle $i(e, \mu, \pi, K, p, ...)$is detected. It reflects properties of the detector.
- $C_i$ is the probability of finding a particle type in the detector (frequency). It does not depend on the detector.

# Probability density functions

Using the TPC, $s$ is the signal $dE/dx$ assigned to each track from the reconstruction.

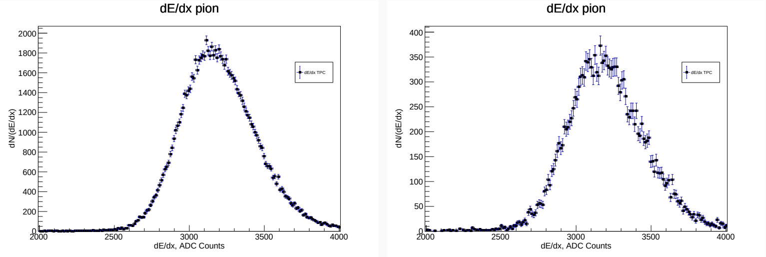
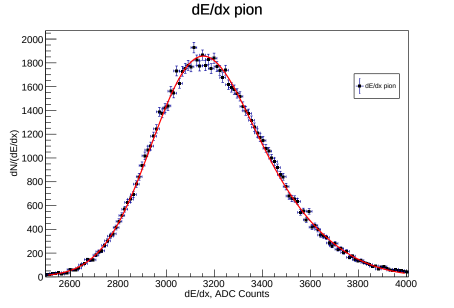



**Figure 2:** dE/dx from TPC for 10k events taking the momentum range of 0.28 to 0.32 GeV/c in [6] for primary and secondary particles.

# Fit and probability density function

Using a Gaussian probability density function to fit data,

$$r(s|\pi) = A\exp\left(-\frac{(s-\mu)^2}{2\sigma^2}\right)$$




Obtaining the result:

- $FCN = 68.3756$
- Constant
  $A = 1.84022e+03$,
  $ERROR = 1.17648e+01$
- Mean
  $\mu = 3.13593e+03$,
  $ERROR = 3.19178e+00$
- Sigma
  $\sigma = 2.00574e+02$,
  $ERROR = 1.67078e+00$

# Probabilities a priori $C_i$

The probabilities $C_i$ can be estimated with Time-of-Flight measurements,

$$m = \frac{p}{\beta\gamma} = p\sqrt{\frac{c^2 t^2}{l^2} - 1}$$

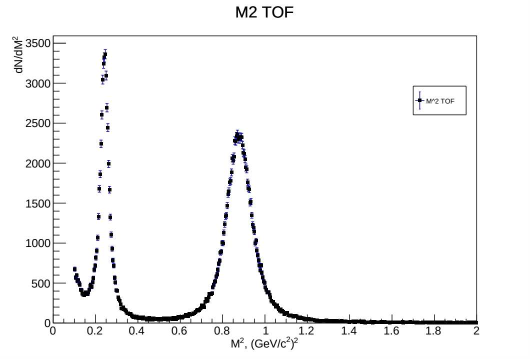
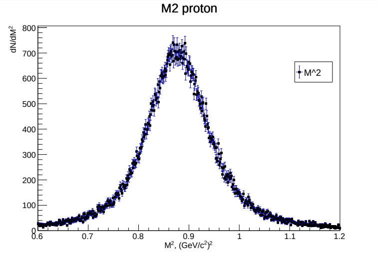
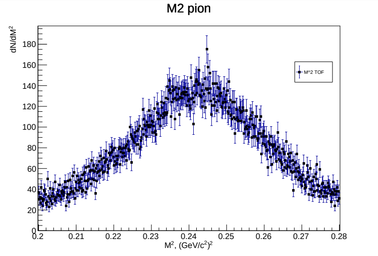



**Figure 3:** $M^2$ TOF for 10k events for primary particles.

# Probabilities a priori $C_i$

Taking each peak in an histogram, we have $C_p = 741$ and $C_\pi = 175$



# Data Science Techniques

# Classification models

In high-energy physics, statistical techniques are traditionally used to determine a mathematical function $f(x)$.

In machine learning we have the approximate function $f(x, w)$ from a test data set, $\{x, y\}$, with $x$ the feature vector, and $y$ the relation between those features.

We seek to obtain the model parameters $\mathbf{w}$.

# Generalized Linear Model (GLM)

In a generalized linear model, the response $y_i$,

$$y_i = \beta_0 + \beta_1 x_{1i} + ... + \beta_p x_{pi} + \epsilon_i$$

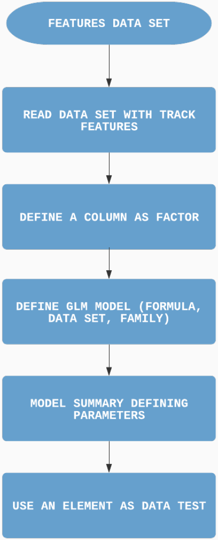
with variables $x_j$ and an error term $\epsilon_i$.

GLMs can be applied with R using the function glm,

$$y \sim x1 + x2$$

- $y$ is the response (type of particle: "proton particule" or "pion particles").
- $x1, x2$ are numeric feature vectors (momentum "P" and energy deposition "dEdx".)

# Generalized Linear Model (GLM)



# Generalized Linear Model (GLM)

We use a data set for training (defining parameters model) with two classes: pions (211) and protons (2212).

| P | dEdx | PDGID |
|---|---|---|
| 0.377411 | 2782.666 | 211 |
| 0.633236 | 2715.193 | 211 |
| 0.404283 | 2710.947 | 211 |
| 0.749252 | 3044.496 | 211 |
| 0.423279 | 13775.852 | 2212 |
| 0.353684 | 2907.477 | 211 |

# Generalized Linear Model (GLM)

Coefficients:

| | Estimate | Std. Error |
|---|---|---|
| (Intercept) | -7.51945 | 12.06997 |
| dEdx | 0.00377 | 0.00365 |
| P | 8.06809 | 9.61674 |

The discriminant function can be write as follow,

$$y = -7.51954 + (0.00377)dEdx + (8.06809)P$$

A test data element is defined as followed, obtaining a probability from 0 to 1 for both classes (protons (0) and pions(1)),

```
newdata = data.frame(P = 0.377411, dEdx = 2782.666)
predict(model, newdata, type="response")
```

# Further work

# Further work

- Implement $r(s|i)$ probability density functions and $C_i$ obtain from histograms to calculate bayesian conditional probability $\omega(i|s)$.
- Compare with GLM model results implementing in R for a binomial class data set.
- Expand GLM model to a more complex model made by several GLM individual models.
- Using larger data set made from larger numbers of events for protons, pions, kaons and electrons.

# References

[1] C. Lippmann (2011) Particle identification. Nucl. Instrum. Meth., vol. A666, pp. 148–172, 2012. DOI:10.1016/j.nima.2011.03.009.

[2] E. Garutti, Particle Identification (PID). Distinguishing particle types http://www.desy.de/ garutti/LECTURES/ParticleDetectorSS12/

[3] O. Rogachevsky (2020) Purposes of the MpdRoot framework – MPD EXPERIMENT http://mpd.jinr.ru/mpdroot-start-guide/

[4] Dryablov, D., Gudima, K., Kapishin, M., Litvinenko, E., Musulmanbekov, G., Zheger, V. (2013). *Event Centrality Determination and Reaction Plane Reconstruction at MPD*.

[5] Kryshen, E. Ivanishchev, D Kotov, Dmitry Malaev, M Riabov, V Ryabov, Yu. (2019). *Study of neutral meson production with photon conversions in the MPD experiment at NICA*. Journal of Physics: Conference Series. 1400. 055055. 10.1088/1742-6596/1400/5/055055.

[6] Mudrokh, Alexander. (2019). Prospects for the study of

event-by-event fluctuations at MPD/NICA project. EPJ Web of Conferences. 204. 07014. 10.1051/epjconf/201920407014.

[7] A. Averyanov et al. (2017). TPC status for MPD experiment of NICA project. JINST 12 C06047