

# FINAL REPORT

Name of the Student	Emin YÜKSEL
Name of the Advisor	Ali Murat GÜLER
Project Title	Study of $D_s$ Decays with Machine Learning Techniques

# Contents

1	INTRODUCTION . . . . .	2
2	METHOD . . . . .	3
3	RESULTS AND DISCUSSION . . . . .	6
	3.1 PYTHIA Monte Carlo data analysis and signal data production	7
	3.2 FLUKA Monte Carlo data analysis and background data pro- duction . . . . .	8
	3.3 Particle separation on TMVA . . . . .	11
4	CONCLUSION . . . . .	13

---

# 1 INTRODUCTION

Measuring the cross-section of the neutrinos is crucial to understand Lepton Universality, which is an important principle for the Standard Model. In this experiment, proton-nucleus interaction is used as a Tau neutrino source, but Tau neutrinos are not produced directly after scattering occurs. Instead after proton-nucleus interaction, D-mesons ( $D$ ,  $D_s$  and  $D^0$ ) are produced, and  $D_s$  meson decays into Tau and Tau neutrino (Figure 1). Therefore, to calculate cross-section of the Tau neutrino cross-section of  $D_s$  also must be known. [1]

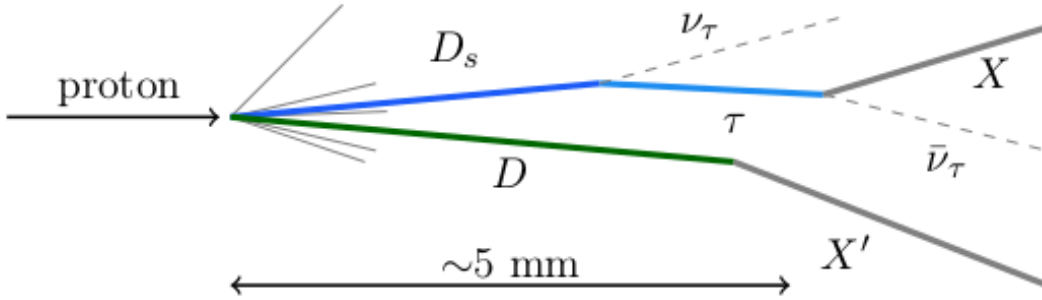


Figure 1: Tau neutrino production diagram. [1, p. 4]

Neutrinos are leptons like electrons and, there are six leptons (electron, muon, tau, and their neutrino counterparts). Since neutrinos are nearly massless and neutral unlike charged leptons, neutrino detection is a challenging process. Moreover, detecting Tau neutrino is even tougher than the others because most of the time Tau neutrinos are produced by the oscillation of the other neutrinos. Therefore, cross-section measurement of the Tau neutrinos is considerably less accurate than the other neutrinos 2. So, the point of this experiment is to decrease the uncertainty in the measurement of the cross-section of Tau neutrinos.

## 2 METHOD

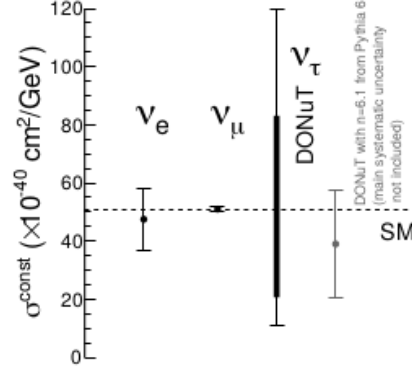


Figure 2: Cross-Section of three neutrino flavors. [1, p. 2]

The flight length and kink angle of  $D_s$  are important characteristics that can be measured accurately in the DsTau experiment. The detector of DsTau is designed to fulfill this property. The DsTau experiment uses nuclear emulsion technology that presents sub-micron and a few mrad spatial and angular resolutions respectively.

These two parameters have been studied by means of the Monte Carlo simulation. Using the Pythia event generator, proton-nucleon interactions at 400 GeV are generated to study Ds decays and Below flight length and kink angle distributions can be seen for each charmed hadron 3.

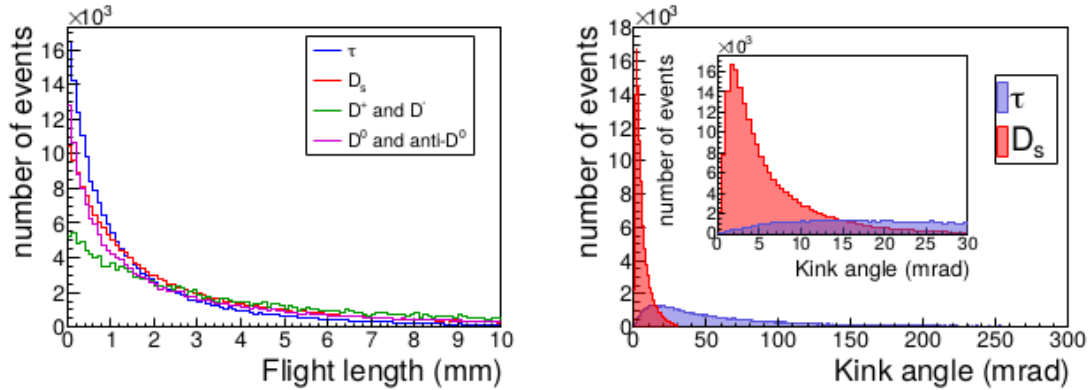


Figure 3: Estimated flight length (left) and kink angle (right) distributions. [1, p. 5]

As shown in Figure 4 detector module consist of two parts. The first part is the decay module which is made of tungsten and emulsion film. First, proton-nucleus



---

DONuT experiment [3] total cross-section formula will be used.

$$\sigma_{\nu_\tau} = \sigma_{\nu_\tau}^{const} \cdot E \cdot K(E) \quad (1)$$

Where  $\sigma_{\nu_\tau}^{const}$  is energy independent cross-section,  $E$  is energy of the neutrino and,  $K(E)$  is kinematic effect due to lepton mass. Because of the uncertainty in the  $D_s$  differential cross-section,  $\sigma_{\nu_\tau}^{const}$  value is can be expressed in terms of  $n$  (longitudinal dependence parameter).

$$\sigma_{\nu_\tau}^{const} = 7.5(0.335n^{1.52}) \times 10^{-40} cm^2 GeV^{-1} \quad (2)$$

Moreover, by using  $n = 6.1$  which, is obtained from the PYTHIA simulation [4] for the DONuT experiment,  $\sigma_{\nu_\tau}^{const}$  will be as

$$\sigma_{\nu_\tau}^{const} = (0.39 \pm 0.13 \pm 0.13) \times 10^{-38} cm^2 GeV^{-1} \quad (3)$$

---

### 3 RESULTS AND DISCUSSION

From the computational reconstruction of the experiment, two results can be deduced. The first result is, with the event generator PYTHIA, kink angle and flight length of the  $D_s$  and  $\tau$  particles gathered. As seen on the figure 6 mean of the kink angle of  $\tau$  is around 6 mrad and the mean of the flight length figure 6 of  $D_s$  is around 3.5 mm. Hence, the accuracy of the emulsion films and the dimension of the films and plastic sheets must be chosen according to these results to take efficient measurements. The second result is, the method of separation of particles from each other and background noise. This method can be considered in three different sections, PYTHIA data analysis and signal data production, FLUKA data analysis and background data production, and particle separation on TMVA.

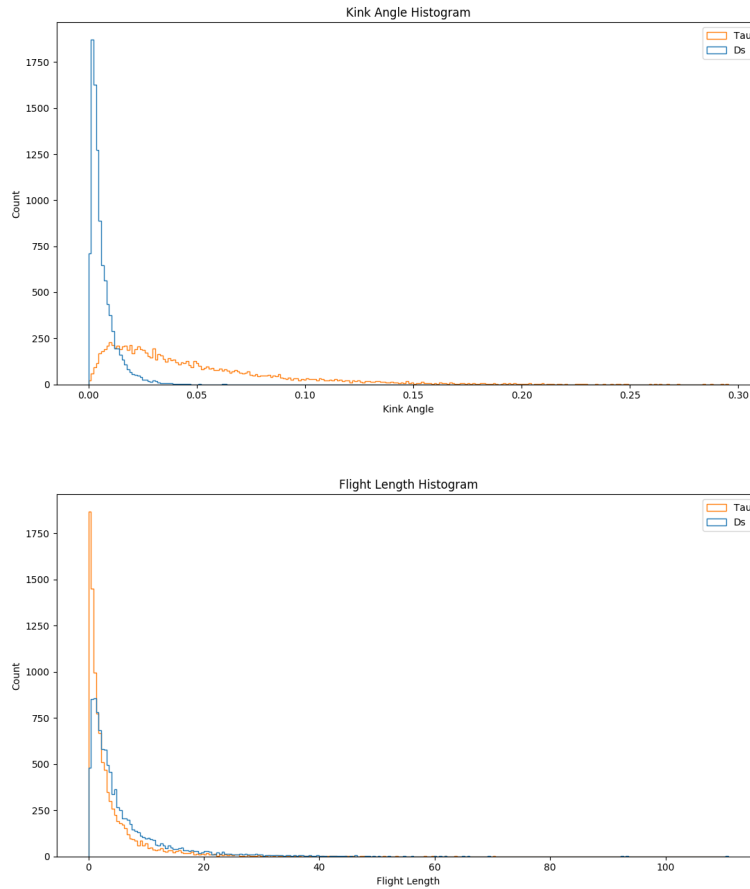


Figure 6: Distribution of kink angle and flight length of  $D_s$  and  $\tau$ .

### 3.1 PYTHIA Monte Carlo data analysis and signal data production

400 GeV proton-nucleus interaction simulated with event generator PYTHIA. PYTHIA can produce only events that decay into desired particles. For this experiment, only  $D_s-$  containing events generated and parameters of the parent ( $D_s-$ ) and daughters ( $\tau$ ,  $\nu_\tau$ , etc.) can be found. Since kink angle and flight length data can be obtained for each particle in PYTHIA, its data can be used while producing signal data.

```

(0)
*****
2 4 6 |
2      | -431(Flight): 0.369220
2      | -431(Kink): 0.068824
-----
      | 15(Kink): 0.018769
      | -16(Kink): 0.161735
-----
4      | 15(Flight): 0.979552
-----
      | 16(Kink): 0.098265
      | -211(Kink): 0.659265
      | -211(Kink): 0.099500
      | 211(Kink): 0.135395
*****

(1)
*****
2 3 2 |
2      | -431(Flight): 0.391904
2      | -431(Kink): 0.045330
-----
      | 15(Kink): 0.008997
      | -16(Kink): 0.142800
-----
3      | 15(Flight): 1.100258
-----
      | 16(Kink): 0.085704
      | 13(Kink): 0.074717
      | -14(Kink): 0.201339
*****

```

Figure 7: PYTHIA data output. Data on the left for event number 0, on the right for event number 1.

Before working on MC production with PYTHIA, must be understood how the PYTHIA data format looks like 7. The PYTHIA data have been modified for the analysis as shown in 7. The first value inside of parenthesis shows the event number and the next three numbers below event number show respectively  $D_s$  daughter,  $\tau$  daughter, and charm particle daughter size. This all process takes place inside the event loop, and there is a second loop for daughters of particles. From first “\*” line to first “-” line  $D_s-$  (-431) parameters produced. From first “-” line to the second “-” line daughters of  $D_s-$  parameters produced. From the second “-” line to third “-” line  $\tau$  (15) parameters produced. From third “-” line to second “\*” line daughters of the  $\tau$  parameters produced. In other words, each “-” line represents interactions in the Feynman diagram of DsTau 1 except the second “-” line since  $\tau$  (15) in the third row is the duplication of  $\tau$  in the second row.

Furthermore, by using the momenta of the particles in each interaction, and with the help of the cosine rule, the kink angle of each particle is calculated. Also, flight length can be calculated straightforwardly by subtracting the decay position



of the particle from the first position where the particle is produced. Then get the magnitude of that displacement vector.

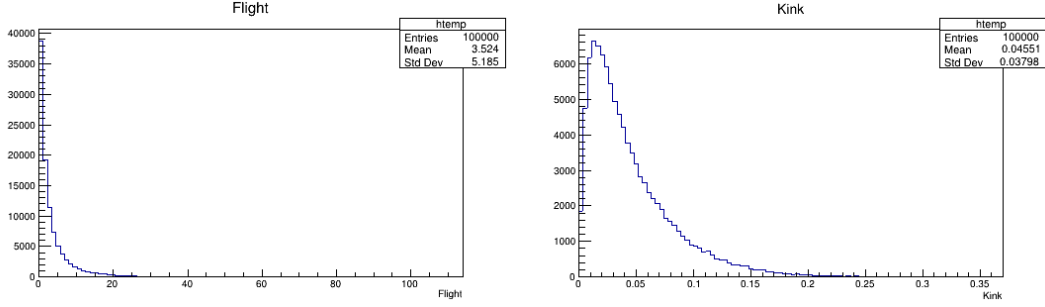


Figure 8: Distribution of  $D_s$  kink angle and flight length that is stored in ROOT file.

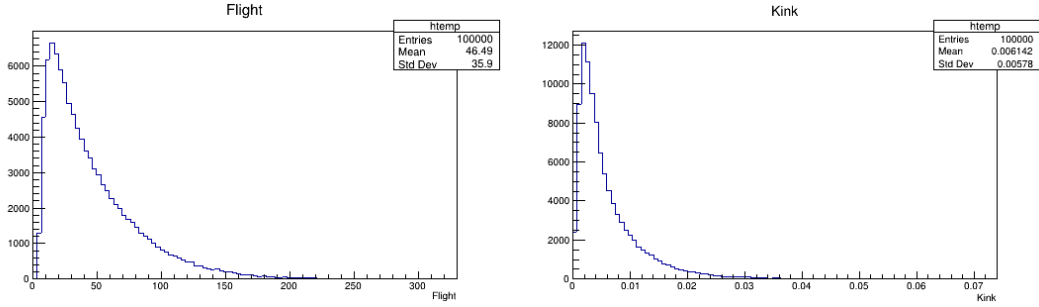


Figure 9: Distribution of  $\tau$  kink angle and flight length that is stored in ROOT file.

After kink angle and flight length data obtained, decay position, particle direction, kink angle, and flight length parameters are stored into a root file that contains  $D_s$  8 and  $\tau$  9 trees.

### 3.2 FLUKA Monte Carlo data analysis and background data production

For background processes, MC data have been generated using the FLUKA event generator. Proton-nucleus interactions are generated and stored in root format. This MC data contains hadron interaction that mimics the hadron decays. Therefore, it constitutes the background for  $D_s$  decays.

Since FLUKA output is stored in the root file, it is easier to work on it and understand this time. FLUKA root file consists of chains. Each chain corresponds

---

```

root [2] t.Show(0)
=====> EVENT:0
Event      = 1
Flag       = 111
ParentId   = 1
Region     = 133
IntInfo    = 1.35275e-05,
            -3.29984e-05, 1, 401.845, 401.844
IntVtx     = 4.69311,
            2.83927, 5.87589
Multip     = 252
Tracks     = 102,
            124, 26
SecondaryId = 7,
            7, 14, 23, 13, 14, 23, 13, 23, 1, 23,
            14, 23, 14, 8, 23, 23, 14, 14, 15
SecondaryInfo = -0.292124,
                0.059967, 0.954499, 1.20582, 0.137682, -0.446902,
                0.883925, 0.918396, 0.378878, -0.252738, 0.890267,
                0.482752, -0.0377614, -0.0337863, 0.998715, 1.36365,
                0.413895, 0.551011, 0.724623, 0.359351

root [3] t.Show(1)
=====> EVENT:1
Event      = 1
Flag       = 333
ParentId   = 8
Region     = 118
IntInfo    = 0.526236,
            0.446142, 0.723902, 0.967801, 0.232066
IntVtx     = 5.46098,
            3.57848, 7.01868
Multip     = 3
Tracks     = 0,
            3, 0
SecondaryId = 8,
            8, 8
SecondaryInfo = -0.922548,
                -0.0199536, -0.385365, 0.00492244, -0.0552343, 0.889759,
                -0.453076, 0.00254142, -0.331081, 0.411219, -0.849285,
                0.000910447

```

Figure 10: FLUKA data output. Data on the left for chain number 0, on the right for chain number 1.

to a particle. As seen in figure 10 the first EVENT gives the number of the generated events, and the second event only shows the chain number in the events. For the flag, there are three options 111, 222, and 333. If the flag is 111, it means that the chain shows the first interaction, if the flag is 222, the chain shows the decay, and if the flag is 333, the chain shows the secondary interaction. “ParentId” is the particle that is shown in the current chain. The first three elements of the “IntInfo” is the 3-vector that gives energy and momentum of the parent. “IntVtx” gives the 3-vector of the interaction vertices. “Multip” shows the number of particles (daughters) produced after the interaction. “SecondaryId” is the FLUKA id of the daughters. “SecondaryInfo” shows the information of the daughters, and each daughter has four elements, and the first three of them are 3-vector of the direction vector. Therefore, “SecondaryId” and “SecondaryInfo” are written respectively.

Before calculations of kink angle and flight length, neutral particles must be filtered out from the background data. Because in the real experiment, only charged particles will be detected on the emulsion film, neutral particles in the FLUKA data is filtered out.

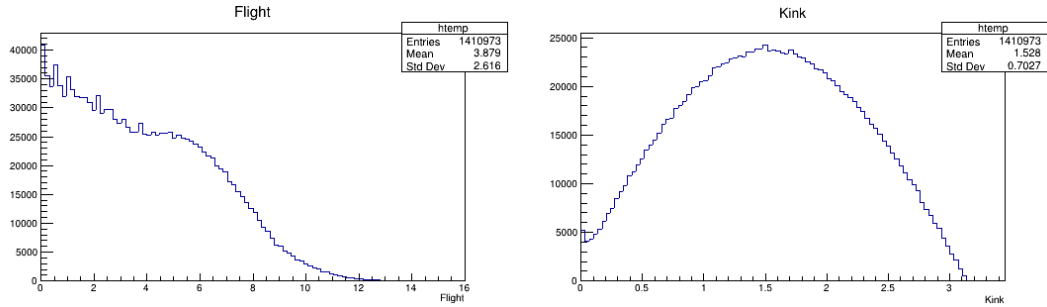


Figure 11: Graph of particles that is produced on DsTau by event generator FLUKA.

The same procedure to obtain kink angle and flight length on PYTHIA also can be applied to the FLUKA. However, unlike PYTHIA data, FLUKA loops over chains so, interactions will be determined by the flags. To find flight length, “IntVtx” of flag 333 particles will be subtracted and find displacement vector, and the rest of the process is the same. To find a kink angle each particle needs to be looped over multiplicity and find the angle between the direction of the particle in the “IntInfo” and the direction of the daughters in the “SecondaryInfo”. With this method kink angle in the first and second interactions can be calculated. However, since there is no interaction vertex data, it is not possible to calculate the flight length of the daughters of the FLUKA data that worked on. After kink angle and flight length are calculated, both stored in a root file 11 as background data.

---

### 3.3 Particle separation on TMVA

At this part of the analysis, signal data is separated from the background events using machine learning techniques. For this purpose, the TMVA package has been used. As a main discriminator or variable, flight length and kink angle are used. For signal events PYTHIA and for background events FLUKA event generator is used.

Firstly a factory constructed on the TMVA, and as background data, background root trees are added. For the signal data, the first  $D_s$  tree in the signal root file is added. Kink angle and flight length branches inside trees are added as a variable. Then, with the help of PYTHIA data 7, found a point where the kink angle and flight length becomes zero and assumed them as the cut. For  $D_s$  flight length cut:  $flight < 40$ , kink angle cut:  $kink < 0.05$ . Finally, added Likelihood, Fisher, MLP MVA methods. After that, the same procedure applied to the  $\tau$  tree in the signal root, and cuts are found as flight length cut:  $flight < 40$ , kink angle cut:  $kink < 0.02$ . Then, the macro is compiled and run.

TMVA results show  $D_s$  separated with the 0.68 signal efficiency by the Likelihood method. In addition, as seen on the figure 12 results consistent with signal data 6. Moreover, ROC Curve 13 is drawn by TMVA Gui. Also, ROC Integral values of Likelihood is 0.773, MLP is 0.747, and Fisher is 0.698. That is, Likelihood is the best method to separate  $D_s$  decay from the background.

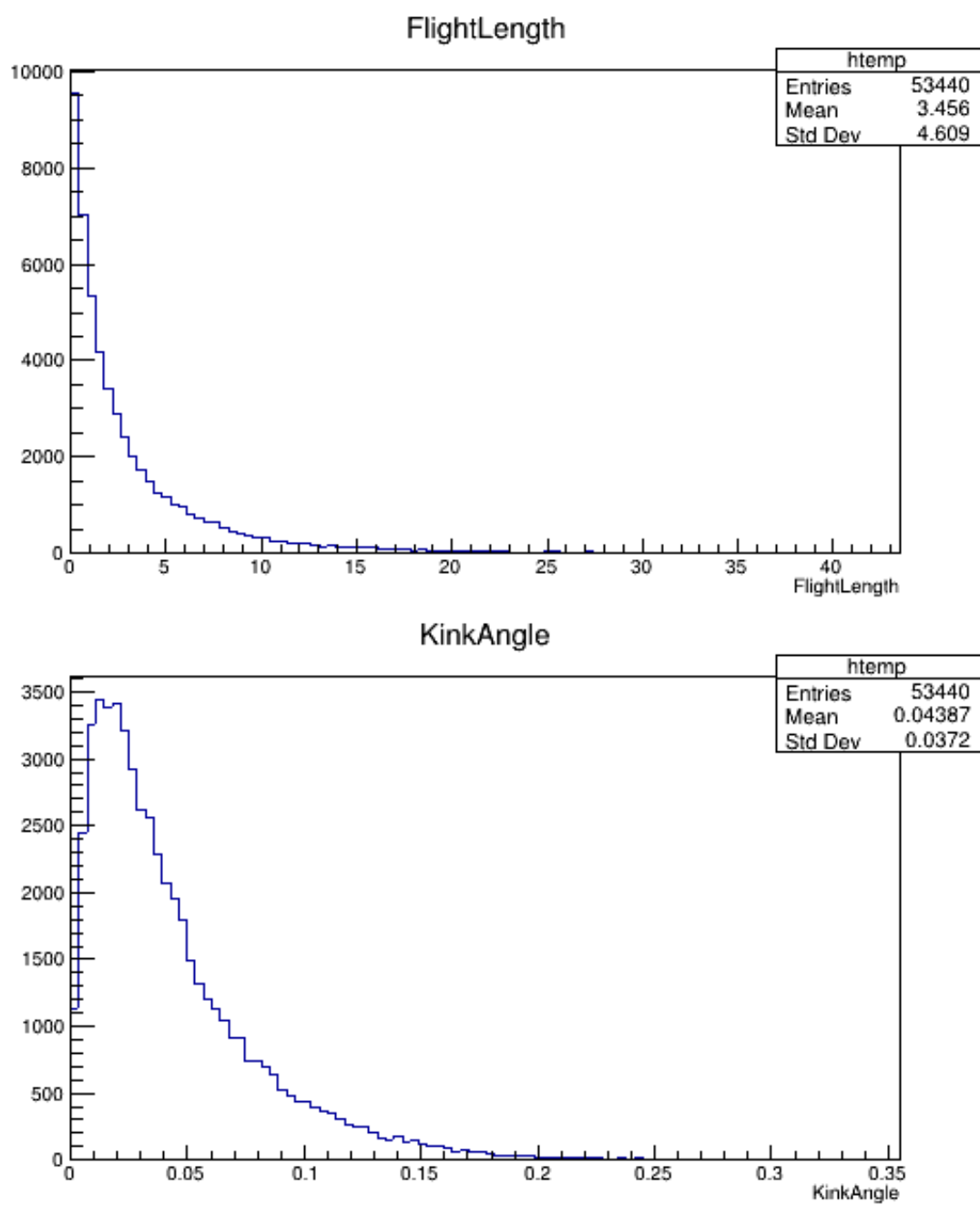


Figure 12: TMVA results after background separation.

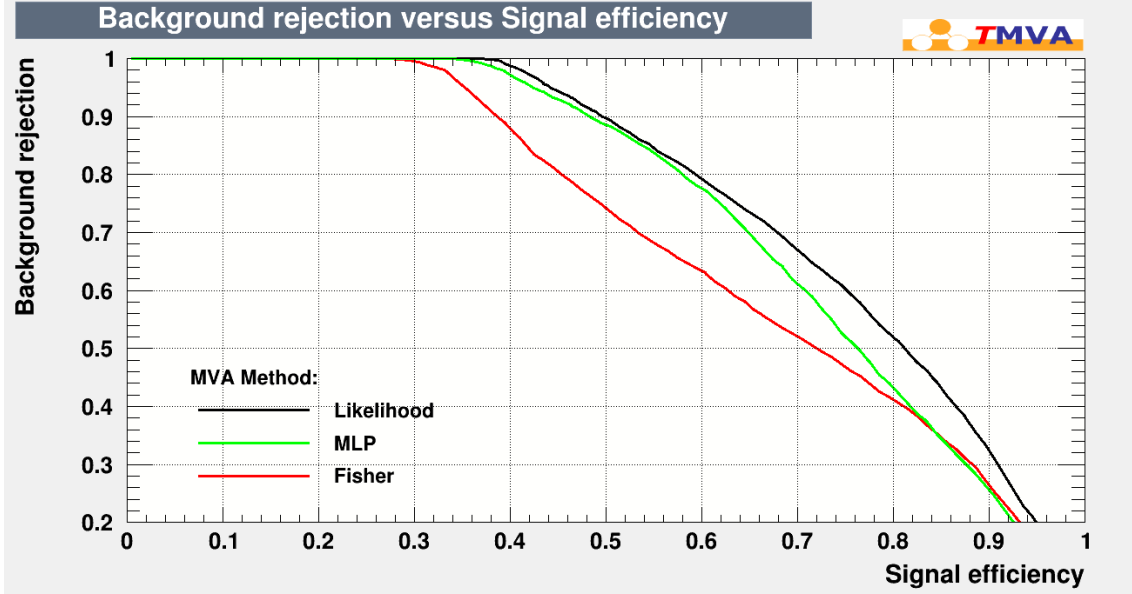


Figure 13: ROC curve of TMVA results.

## 4 CONCLUSION

Due to the lack of information on  $D_s$  production, the cross-section of the  $\nu_\tau$  cannot be measured precisely. Since looking for  $\nu_\tau$  that are produced by  $D_s$  meson interaction,  $\nu_\tau$  cross-section also dependent on  $D_s$  cross-section. Therefore, to get information about  $D_s$  interactions and get an idea of the properties of the instruments that will be used in the DsTau experiment, this computational experiment conducted.

Using the ROOT package TMVA and two discriminator variables signal is separated from the background data with 68% efficiency by using the Likelihood method.

# Bibliography

- [1] Shigeki Aoki, Akitaka Ariga, Tomoko Ariga, Sergey Dmitrievsky, Elena Firt, Dean Forshaw, Tsutomu Fukuda, Yuri Gornushkin, Ali Murat Guler, Maria Haiduc, and et al. Dstau: Study of tau neutrino production with 400 gev protons from the cern-sps, Jun 2019.
- [2] N. Agafonova et. al. (OPERA Collaboration). Momentum measurement by the multiple coulomb scattering method in the OPERA lead-emulsion target. *New Journal of Physics*, 14(1):013026, jan 2012.
- [3] K. Kodama, N. Ushida, C. Andreopoulos, N. Saoulidou, G. Tzanakos, P. Yager, B. Baller, D. Boehnlein, W. Freeman, B. Lundberg, and et al. Final tau-neutrino results from the donut experiment, Apr 2016.
- [4] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. A brief introduction to pythia 8.1, Oct 2007.