Available online at www.sciencedirect.com

ScienceDirect

Nuclear Physics A 982 (2019) 192–194

NUCLEAR PHYSICS A

www.elsevier.com/locate/nuclphysa

Muon Identification using Neural Networks With the Muon Telescope Detector at STAR

James D. Brandenburg (for the STAR Collaboration)

Rice University, Houston, TX, USA

Abstract

The installation of the Muon Telescope Detector (MTD) at STAR allows a measurement of the dimuon ($\mu^+\mu^-$) production in heavy-ion collisions over a large invariant mass range for the first time. Data has been collected with the MTD from Au+Au collisions at $\sqrt{s_{NN}}$= 200 GeV and from p+p collisions at $\sqrt{s}$= 200 GeV. These two datasets allow for new opportunities to measure the dimuon invariant mass spectra at STAR. Before any dimuon measurements can be made, muons must be identified. This contribution presents muon identification employing deep neural networks (DNN) and compares it with other multi-variate techniques. Applications of the DNN technique for data-driven purity measurements are discussed.

Keywords: Muon, Neural Network, Muon Telescope Detector, STAR, Dimuon

1. Introduction

Dileptons (l^+l^-) provide an ideal probe of the hot and dense medium of strongly interacting matter produced in high energy heavy-ion collisions like those produced by the Large Hadron Collider (LHC) and the Relativistic Heavy-Ion Collider (RHIC) [1–4]. Since leptons are inert to the strong force, they carry pristine information about the whole evolution of the created system. Different production mechanisms related to different stages of the system can be distinguished through the invariant mass of produced pairs.

The Muon Telescope Detector was installed to allow for identification of muons over a large momentum range for the first time at STAR. The plethora of hadrons produced in high-energy hadronic interactions makes hadron punch-through a significant source of background for muon identification. Identifying primordial $\mu^+\mu^-$ pairs is further obscured by secondary muons resulting from weak decays of charged pions and kaons. The precise timing and position measurements of the MTD provide information that is useful for identifying muons and rejecting background sources. In this contribution, we report on the training and use of neural networks for muon identification with information from the MTD at STAR.

2. The Muon Telescope Detector

The Muon Telescope Detector (MTD) is a multi-gap resistive plate chamber (MRPC) based detector installed outside the STAR magnet return yoke steel at a radius of ~410 cm [5]. The magnet steel acts as the hadron absorber providing ~5 interaction lengths at its thickest. The MTD is segmented into 30 backlegs

https://doi.org/10.1016/j.nuclphysa.2018.10.036
0375-9474/

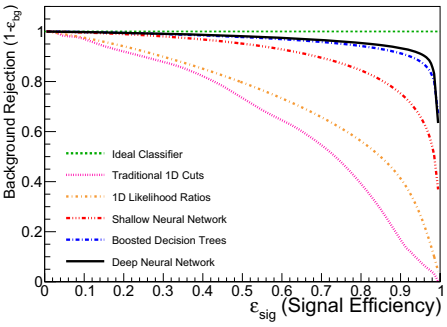

Fig. 1: (color online) The background rejection ($1 - \varepsilon_{bg}$) versus the signal efficiency (ε_{sig}) for several different multivariate classifiers and traditional 1D cuts for tracks with $p_T < 10$ GeV/c.

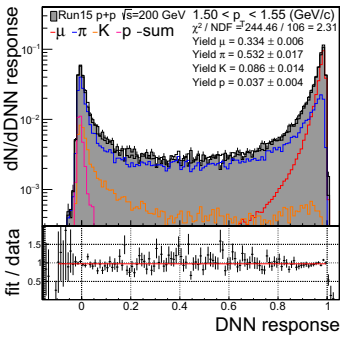

Fig. 2: (color online) Top panel: Extraction of the $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$ and $p/\bar{p}$ yields by fitting the DNN response in data using template shapes from simulation. Bottom panel: The ratio of the total fit result to the data.

in the ϕ direction and 3 to 5 modules in the z (beam line) direction. The MTD provides ~45% coverage in $|\eta| < 0.5$. The precise timing of the MTD, with a resolution of $\sigma_{\mathrm{TOF}} \approx 100$ ps, enables some of the slower hadron-punch through to be rejected. In addition to timing information, the MRPCs also provide ~2 cm hit position resolution in the local Y and local Z directions. Altogether, the MTD and STAR Time Projection Chamber [6] provide 10 quantities that are used in the training of neural networks.

3. Neural Network Details

For these studies, the Toolkit for Multivariate Analysis (TMVA) was used to train the neural networks [7]. Training of this type of neural network classifier is a form of supervised learning which requires a labeled dataset for the training phase. The labeled training samples were produced by simulating individual $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$ and $p/\bar{p}$ trajectories through a GEANT3-based STAR geometry and detector simulation [8]. After simulating the detector responses, the standard event reconstruction software is used to form tracks in the TPC, find the primary collision vertex in the event, and match tracks to simulated MTD hits. The result of the simulation procedure is a separate set of all PID distributions for $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$ and $p/\bar{p}$ that can be used in the neural network training phase.

A grid-search strategy was used to determine the optimal multi-layer perceptron neural network architecture. Shallow neural networks, with only a single hidden layer, and deep neural networks (DNN) with two or more hidden layers were considered. Neural network architectures were scored based on their signal versus background rejection power, their complexity (number of neurons) and the shape of the discriminator output. The highest scoring architecture was a neural network with 2 hidden layers containing N+10 and N+12 neurons each (N=10 is the number of input variables).

The signal versus background discrimination power was analyzed through receiver operating characteristic (ROC) curves by plotting the background rejection ($1 - \varepsilon_{bg}$) versus the signal efficiency (ε_{sig}). An ideal classifier is able to perfectly distinguish between signal versus background by rejecting 100% of the background while keeping 100% of the signal and has an ROC curve equal to 1. Figure 1 shows a comparison of shallow and deep neural network classifiers compared to several other multivariate algorithms and optimized 1D cuts. The likelihood ratio and boosted decision tree (BDT) classifiers were trained using the TMVA package using default options (more details in Ref. [9]). The cuts used in the traditional 1D cut-based PID were optimized on the J/ψ peak in p+p collisions at $\sqrt{s}$= 200 GeV.

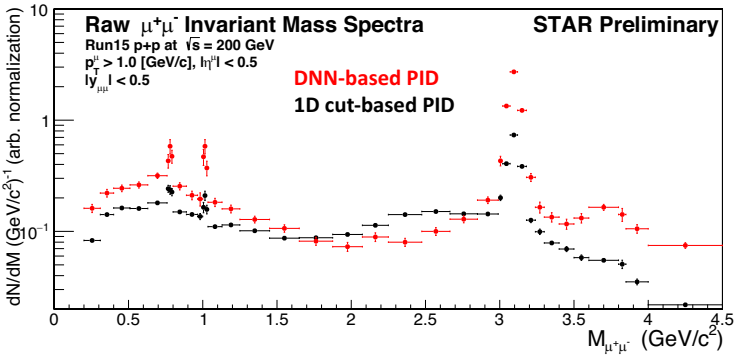

Fig. 3: (color online) Comparison of the raw $\mu^+\mu^-$ invariant mass distribution in p+p collisions at $\sqrt{s}$=200 GeV for traditional 1D muon identification cuts versus the deep neural network based muon identification.

4. Results and Application

An analysis of the ROC curves of several muon identification algorithms in Fig. 1 shows that deep neural networks performed best on simulated data. Figure 3 shows a comparison of the the raw $\mu^+\mu^-$ invariant mass spectra in p+p collisions at $\sqrt{s}$= 200 GeV using optimized 1D cut-based muon identification and the DNN-based muon identification. The distributions are normalized in $1.2 < M_{\mu\mu} < 2.5$ GeV/c^2 to make comparison easier. The DNN-based PID provides higher significance ($S/\sqrt{S+B}$) and S/B ratios of the ω and ϕ. Additionally, the DNN-based PID greatly improves the significance of the ψ(2S) at higher masses, making its analysis possible. Our paper containing significantly more details will be available shortly.

In addition to muon identification, the trained DNN can also assist in measuring the muon purity in data. Template shapes for the signal and each background contribution can be determined by evaluating the trained neural network on the simulated data. These shapes can then be used to fit the DNN response in data to determine the yield contributions from each source. Figure 2 shows the DNN response for tracks in the data in the range $1.5 < p_T < 1.55$ GeV/c along with the $\mu^\pm$, $\pi^\pm$, K$^\pm$, and $p/\bar{p}$ contributions resulting from a maximum likelihood fit.

5. Summary

In this contribution, we reported muon identification techniques employing neural networks with the information from the MTD at STAR. An analysis of the the signal versus background separation power shows that the deep neural network performs better than all other multivariate algorithms investigated. The efficacy of the deep neural network based muon identification is tested in p+p collisions at $\sqrt{s}$= 200 GeV and found to be significantly superior to optimized 1D cut-based PID. Finally, a technique for measuring the muon purity in data by leveraging the improved signal versus background separation power of the deep neural network is presented.

References

[1] J. Adams *et al.*, (BRAHMS Collaboration), Nucl. Phys. A. 757 (2005) 102.
[2] B.B. Back *et al.*, (PHOBOS Collaboration), Nucl. Phys. A. 757 (2005) 28.
[3] I. Arsene *et al.*, (STAR Collaboration), Nucl. Phys. A. 757 (2005) 1.
[4] K. Adcox *et al.*, (PHENIX Collaboration), Nucl. Phys. A. 757 (2005) 184.
[5] C. Yang et al., Nucl. Instrum. Meth. A. 762 (2014) 1.
[6] M. Anderson, *et al.*, Nucl. Instrum. Meth. A. 499 (2003).
[7] A. Hoecker, P. Speckmayer, J. Stelzer, *et al.*, PoS A CAT 040 (2007)
[8] S. Agostinelli, et al., Nucl. Instrum. Meth. A. 506(3), 250 (2003)
[9] T.C. Huang et al., Nucl. Instrum. Meth. A. 833 (2016) 88.