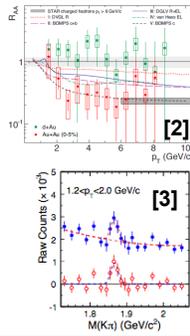


Identification of charmed particles using Multivariate analysis in STAR experiment

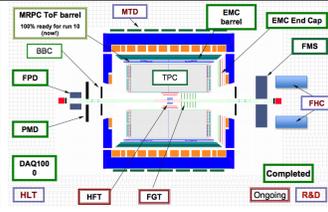
Jonathan Bouchet*, for the STAR collaboration
*Kent State University, Ohio, USA

MOTIVATIONS

- Hot/dense matter created in HIC with strong collectivity[1].
- Heavy quarks → ideal to probe the medium.
- Elliptic flow v_2 → indication of thermalization.
- Nuclear modification factor sensitive to energy loss mechanisms.
- Measurements :
 1. Indirect method via semi leptonic decays[2].
 2. Direct method via statistical combination of hadronic decays[3].



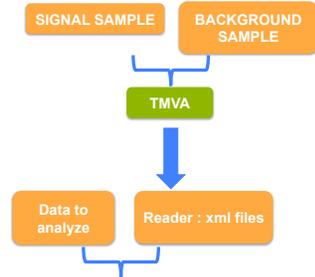
STAR[4] EXPERIMENT @ RHIC



- a large acceptance **Time Projection Chamber** provides particle identification and track momentum.
- Silicon vertex detectors :
- SVT+SSD → **Pointing resolution ~ 280 μm @ 1GeV/c**
- Upgrade HFT(>Y2013) → low mass detector. → **Pointing resolution ~ 40 μm @ 1GeV/c**

Toolkit for MultiVariate data Analysis

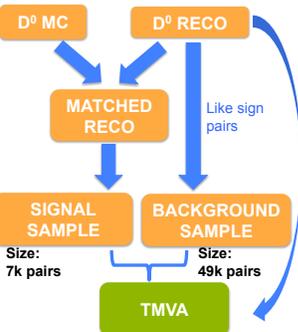
TMVA[5] is a package of several classification methods for multivariate data analysis.



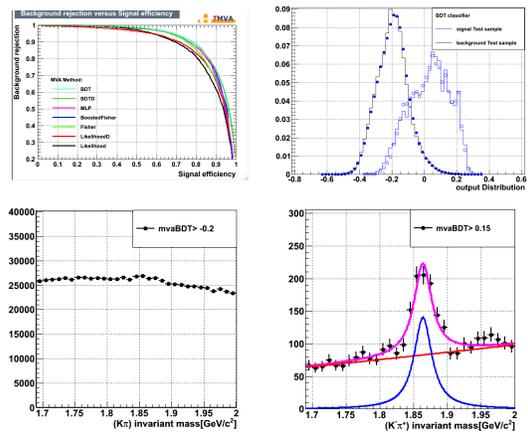
- 1. Training Phase :**
 1. Signal and background input : (X_1, X_2, \dots, X_N) variables
 2. Mapping $(X_1, X_2, \dots, X_N) \rightarrow$ MVA output is written to weight files for each classifier.
- 2. Application Phase :**
 1. Classification based on cut of MVA output.

Results using embedded data

• Root version 5.30 ; TMVA 4.1.2
• Data input : Charmed particle $D^0 \rightarrow K\pi$ embedded into real data with flat transverse momentum.



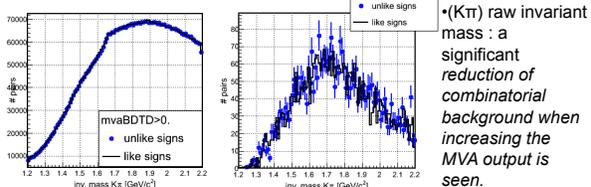
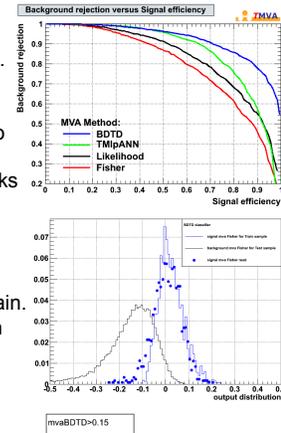
- Classifiers studied : Fisher discriminant (Fisher), Boosted decision Tree (BDT), Projective Likelihood (Likelihood), neural network (MLP)[5].
- 16 variables trained.
- Other classifiers were tested and differences in efficiencies between them were not significant.
- Sophisticated classifiers (Support Vector Machine, RuleFit) were found to be CPU intensive therefore not used in this example.
- correlated variables were removed from the training, although it did not highly affect the results of the classifiers.



- Top left : Background rejection versus signal efficiency ("ROC curve") for a set of classifiers.
- Top right : classifier output distributions for signal (blue line) and background (black line) training samples for BDT classifier. Symbols are the output distributions when reading the signal sample (open) and data (filled). A good agreement is seen, making a decision based on the MVA output possible.
- bottom : $(K\pi)$ raw invariant mass for 2 different cuts on MVA output for BDT classifier.

Preliminary look at the real data

- Data from Au+Au collisions during run7 @RHIC have been analyzed (~10M Minbias events).
- Event vertex and tracks quality cuts applied.
- Daughters tracks are required to have 3 or more silicon hits.
- D^0 reconstructed by pairing tracks identified as kaon and pion.
- Secondary vertex fit computes signed decay length.
- Signal sample : ~8k $D^0 \rightarrow K\pi^+$ through STAR reconstruction chain.
- Background sample : same sign $(K\pi)$ from real data.
- Classifiers : Fisher, BDT, MlpANN, Likelihood.
- 13 variables trained.



CONCLUSIONS and PERSPECTIVES

- TMVA toolkit provides several classifications methods to separate a signal from a given background.
- Classifiers were tested on simulation for the identification of secondary particles.
- Additional classifiers tuning is ongoing to provide an optional signal in real data.

REFERENCES

- [1] Adams et al., STAR Coll., Nucl. Phys. **A757**, 102(2005).
- [2] Abelev et al., STAR Coll., Phys. Rev. Lett. **98** 192301 (2007)
- [3] Adamczyk et al., STAR Coll., arXiv:nucl-ex/1204.4244
- [4] Ackerman et al., STAR Coll., Nucl. Instrum. Meth. **A 499**, 624 (2003)
- [5] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, TMVA - Toolkit for Multivariate Data Analysis, arXiv:physics/0703039