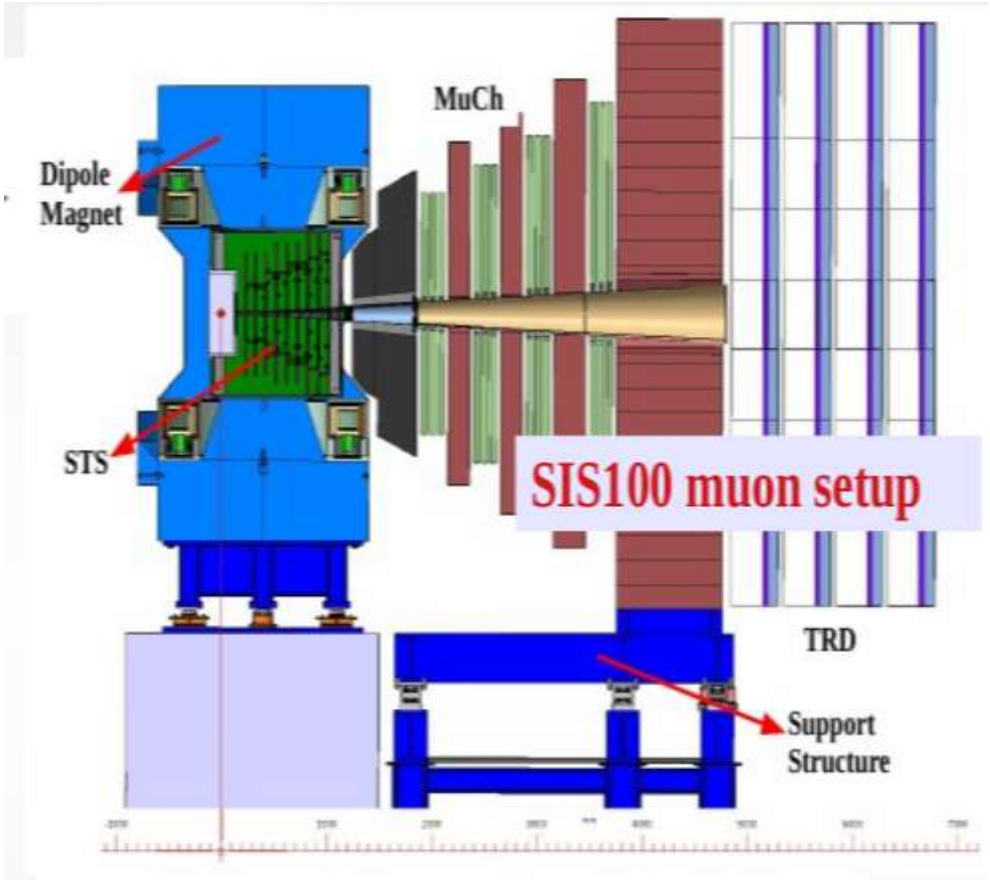
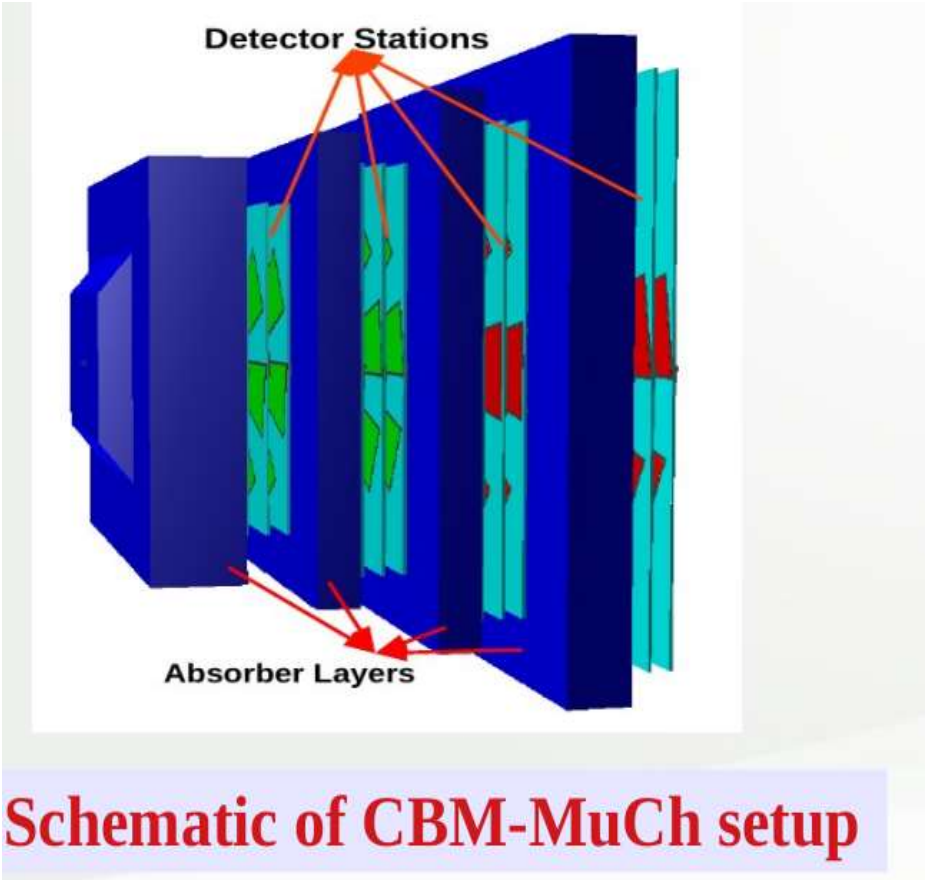
Dimuon reconstruction for Low mass vector mesons (LMVM) using Machine Learning Technique

Pawan Kumar Sharma (VECC Kolkata)
Abhishek Sharma (AMU, Aligarh)
RaktimMukerjee (Summer student, VECC)
Partha Partim Bhaduri (VECC Kolkata)

MuCh @ CBM



CBM at FAIR GSI



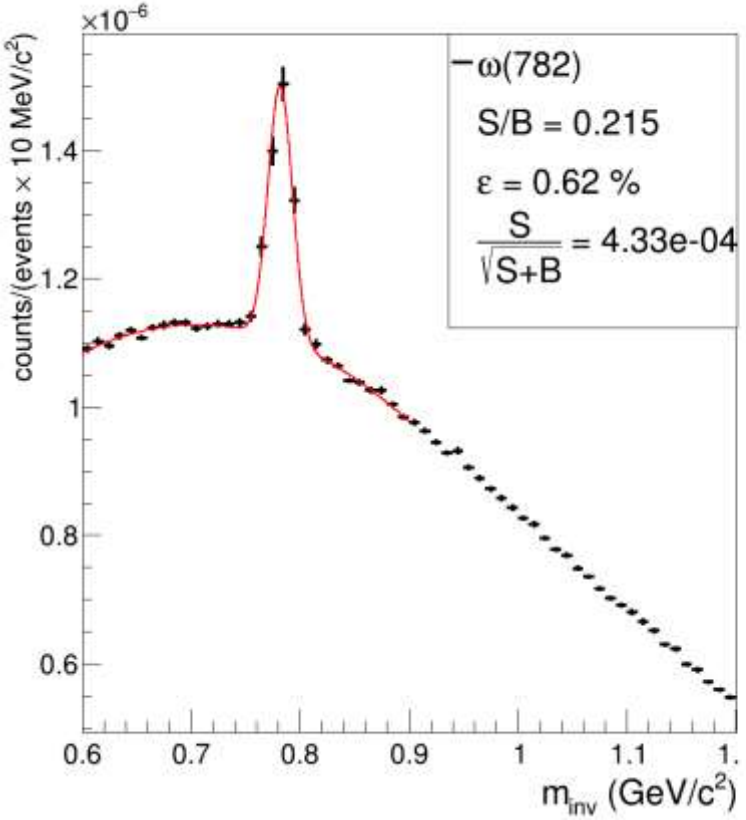
Basic Simulation Info

- **Geometry : v23a**
- **Generator : URQMD & Pluto**
- **Beam & Target system : AuAu**
- **Energy : 8 AGeV**
- **Centrality : central**
- **Setup : SIS100**
- **No of Events: 100k**
- **FairSoft - Apr22p2 & FairRoot – 18.6.7**

Omega meson($\omega \rightarrow \mu^+ + \mu^-$)

Invariant mass spectra of omega (with manual cuts)

- **Simulation for 8AGeV Central AuAu collision in SIS100 setup**



Cuts:

N of STS hits ≥ 7

N of MUCH hits ≥ 11

N of TRD hits ≥ 1

N of Tof hits ≥ 1

$\chi^2_{vertex} \leq 2.5$

$\chi^2_{STS} \leq 2.0$

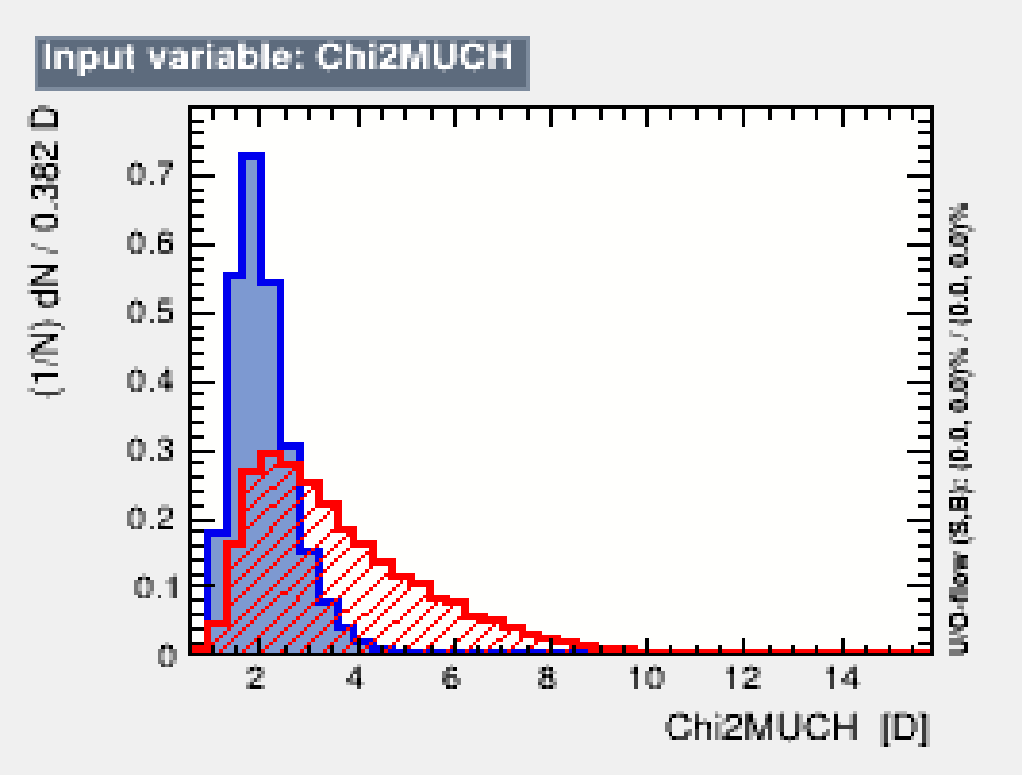
$\chi^2_{MUCH} \leq 3.0$

2σ cut in TOF

ML : Training data

Signal and background have following variables

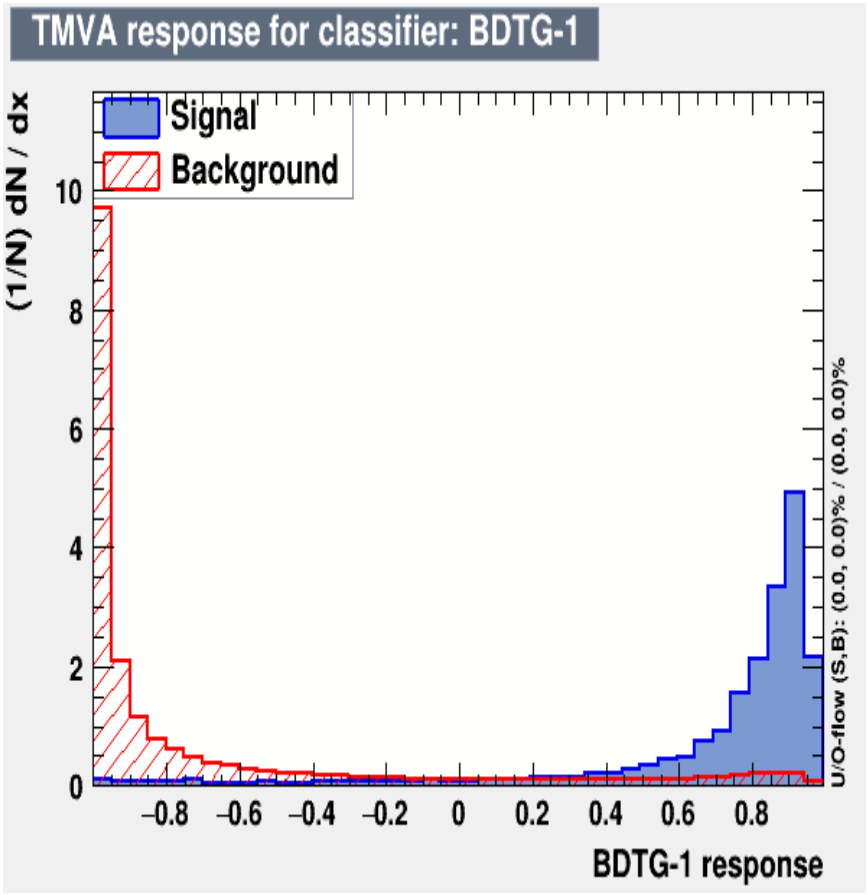
- Mass
- Momentum
- No of MuCh Hits
- No of STS Hits
- No of TRD Hits
- No of Tof Hits
- Chi2Much(χ^2_{Much})
- Chi2STS (χ^2_{STS})
- Chi2Vertex(χ^2_{Vertex})

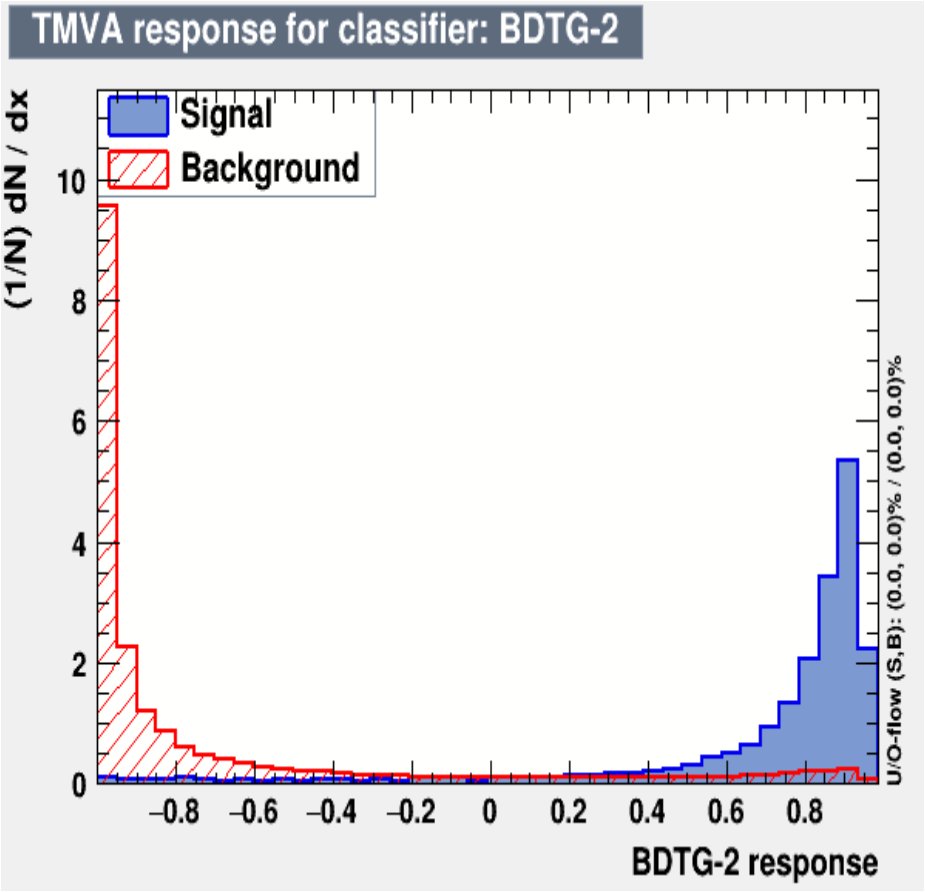


- Various ML model like BDTG, KNN, Hmatrix, MLP have been tested. Among them BDTG model performs better.

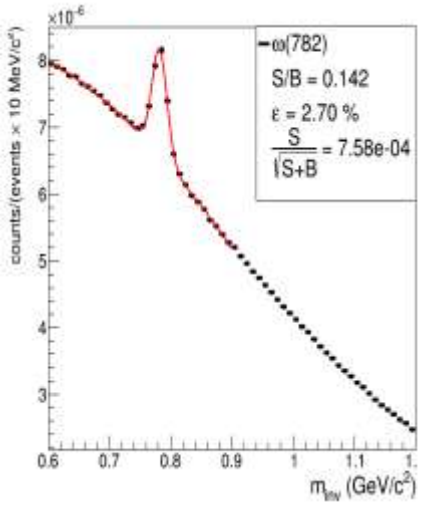
Gradient Boosted Decision Trees (BDTG)

Seperation of Signal & Background

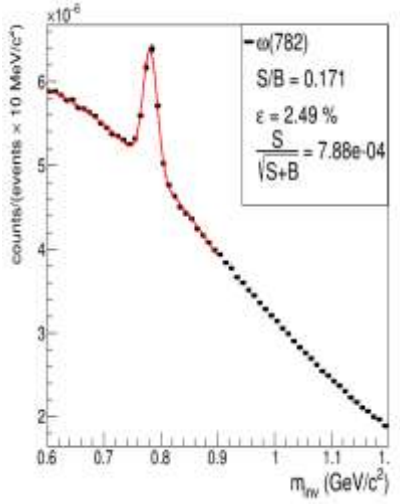




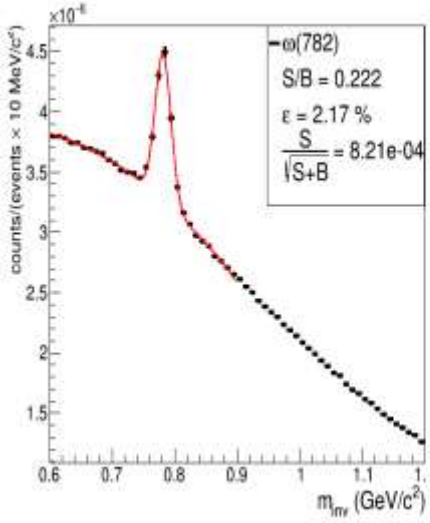
Invariant Mass Spectra (BDTG-1)



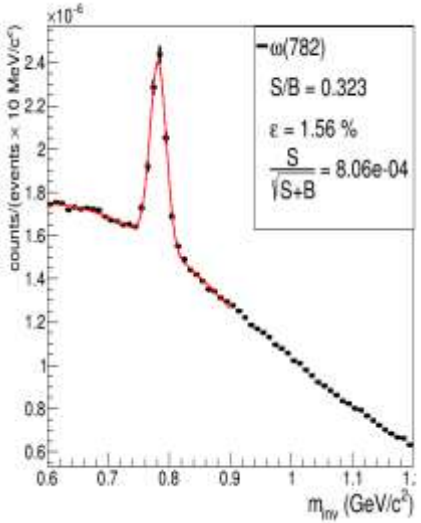
Cuts:
N of STS hits ≥ 4
N of MUCH hits ≥ 8
N of TRD hits ≥ 1
N of TOF hits ≥ 1
$\chi^2_{vertex} \leq 20.0$
$\chi^2_{STS} \leq 20.0$
$\chi^2_{MUCH} \leq 20.0$
2σ cut in TOF
BDTG-1ML Model with cut 0.5



Cuts:
N of STS hits ≥ 4
N of MUCH hits ≥ 8
N of TRD hits ≥ 1
N of TOF hits ≥ 1
$\chi^2_{vertex} \leq 20.0$
$\chi^2_{STS} \leq 20.0$
$\chi^2_{MUCH} \leq 20.0$
2σ cut in TOF
BDTG-1ML Model with cut 0.6

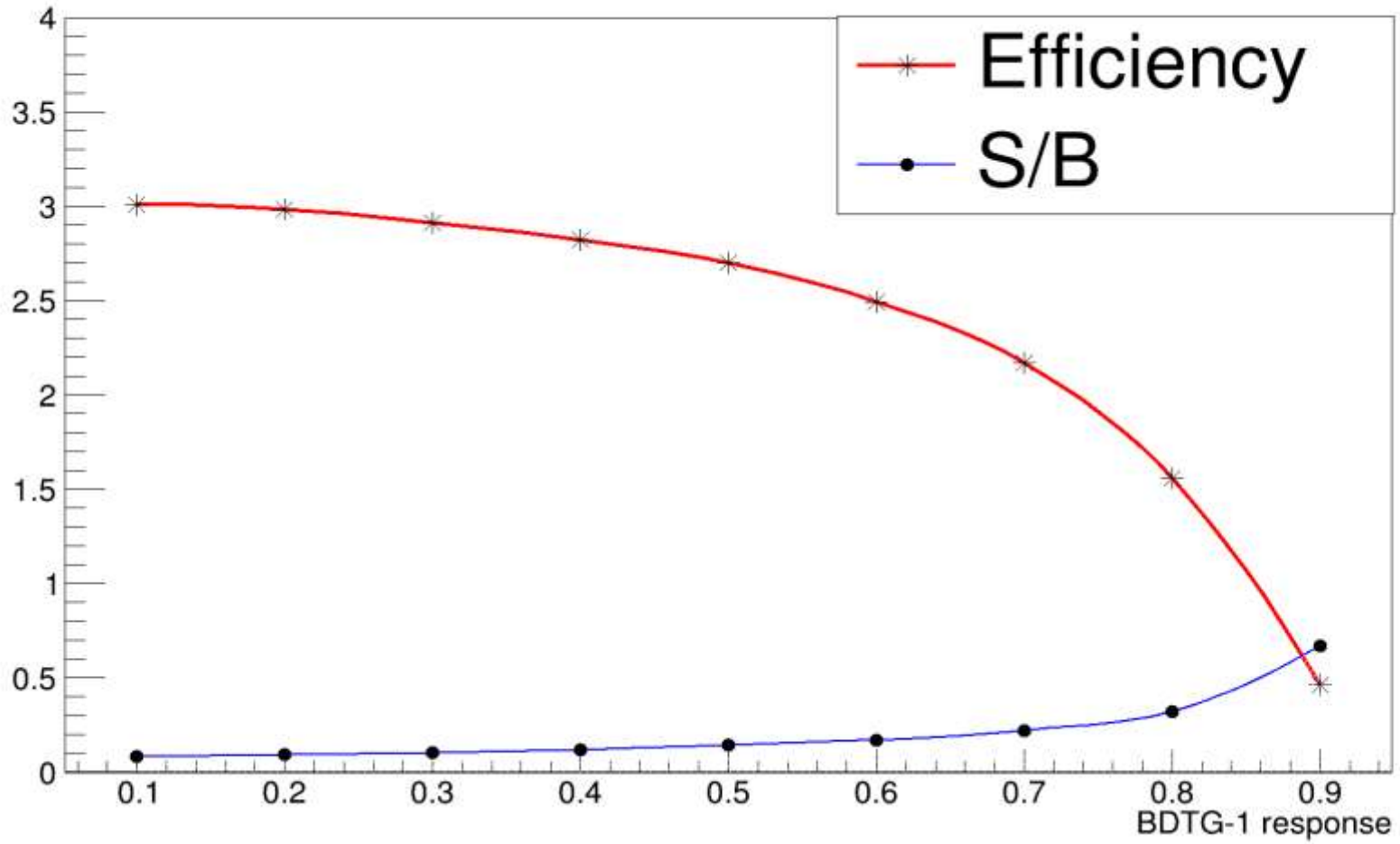


Cuts:
N of STS hits ≥ 4
N of MUCH hits ≥ 8
N of TRD hits ≥ 1
N of TOF hits ≥ 1
$\chi^2_{vertex} \leq 20.0$
$\chi^2_{STS} \leq 20.0$
$\chi^2_{MUCH} \leq 20.0$
2σ cut in TOF
BDTG-1ML Model with cut 0.7

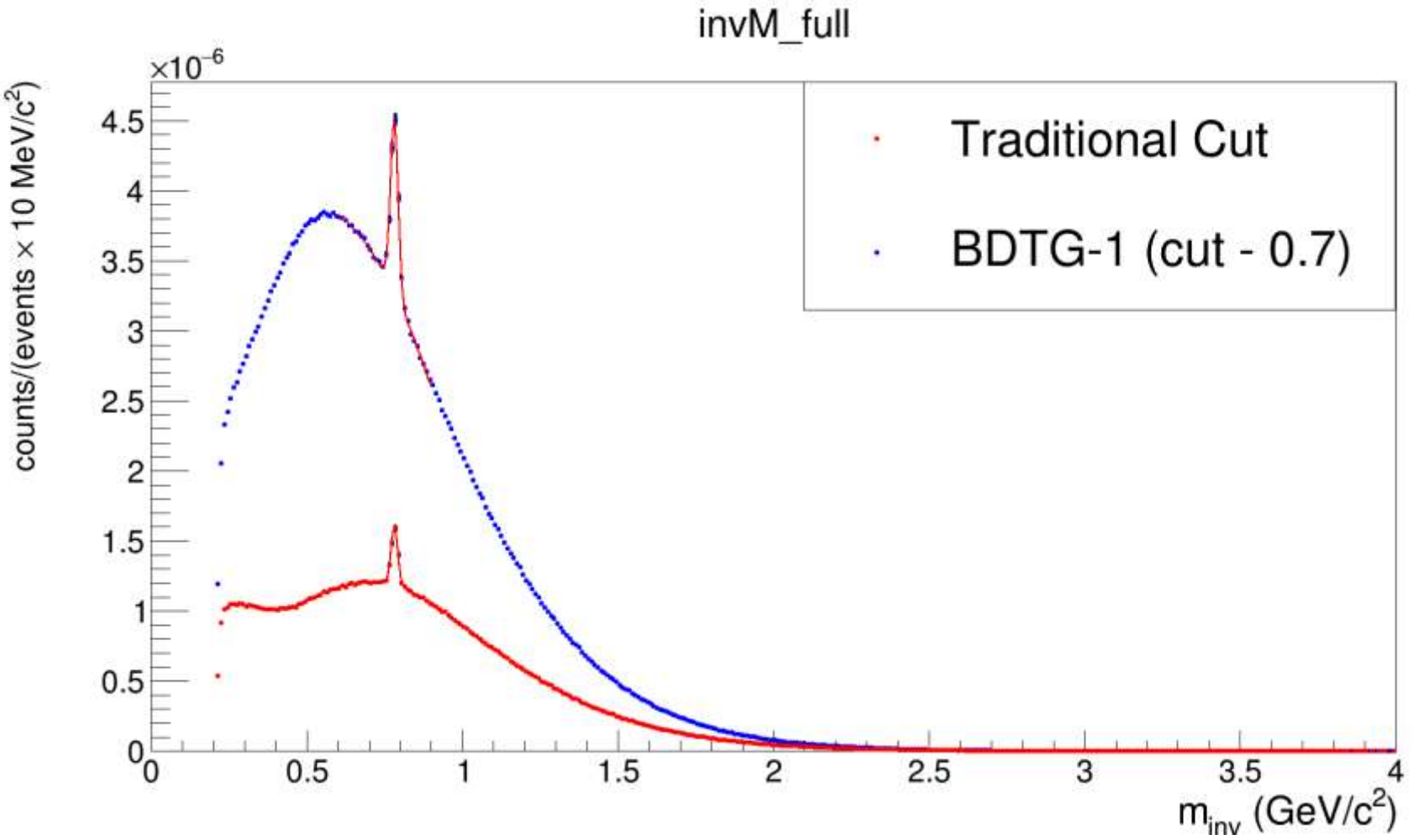


Cuts:
N of STS hits ≥ 4
N of MUCH hits ≥ 8
N of TRD hits ≥ 1
N of TOF hits ≥ 1
$\chi^2_{vertex} \leq 20.0$
$\chi^2_{STS} \leq 20.0$
$\chi^2_{MUCH} \leq 20.0$
2σ cut in TOF
BDTG-1ML Model with cut 0.8

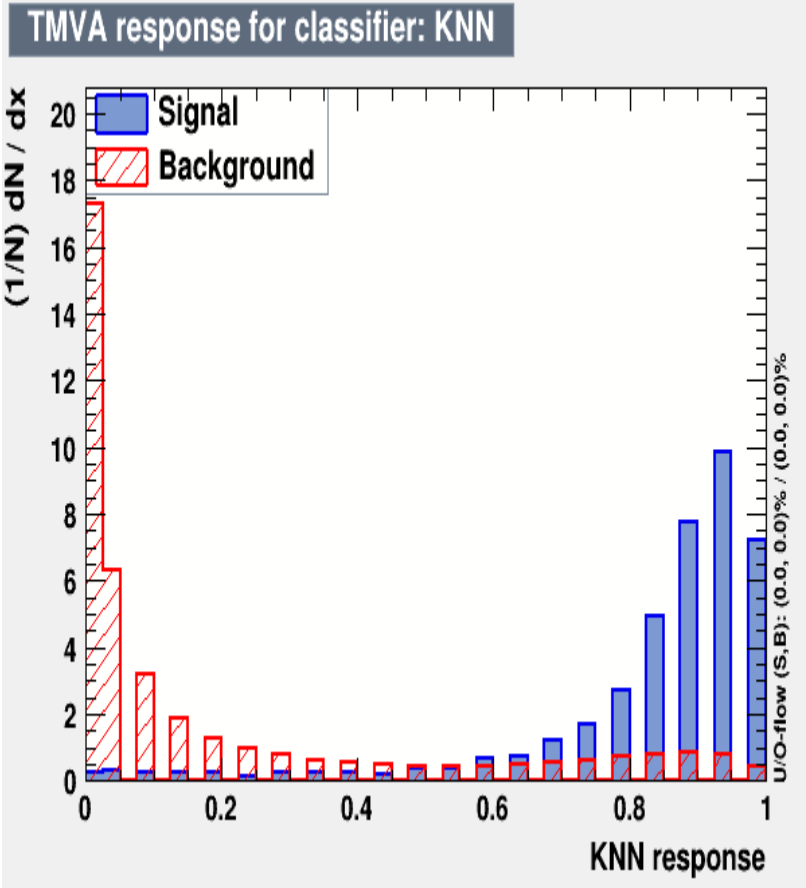
Efficiency & S/B vs BDTG-1 output variable

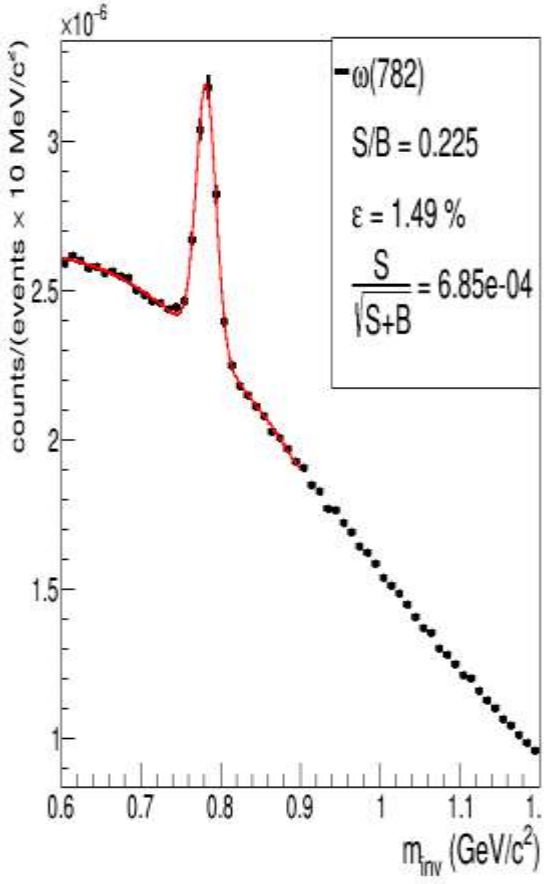


Comparision of ML with traditional cuts



Invariant Mass Spectra (KNN)





Cuts:

N of STS hits ≥ 4

N of MUCH hits ≥ 8

N of TRD hits ≥ 1

N of TOF hits ≥ 1

$\chi^2_{vertex} \leq 20.0$

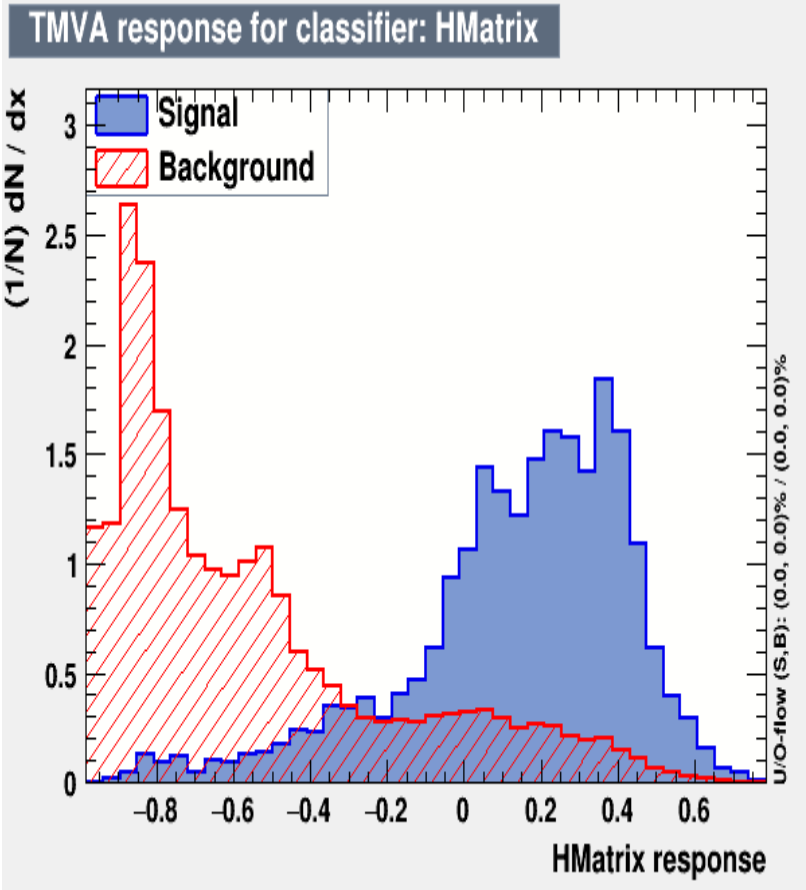
$\chi^2_{STS} \leq 20.0$

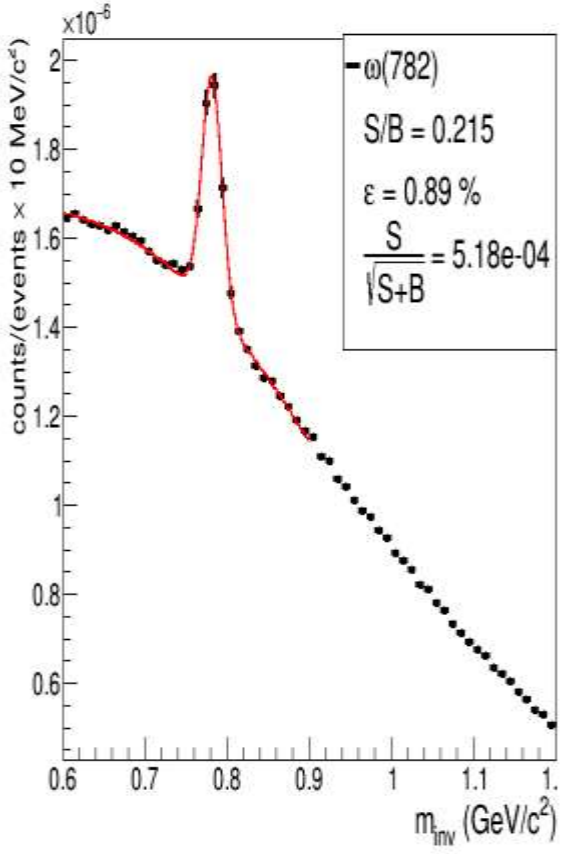
$\chi^2_{MUCH} \leq 20.0$

2σ cut in TOF

KNNML Model with cut 0.9

Invariant Mass Spectra (HMatrix)





Cuts:

N of STS hits ≥ 4

N of MUCH hits ≥ 8

N of TRD hits ≥ 1

N of TOF hits ≥ 1

$\chi^2_{vertex} \leq 20.0$

$\chi^2_{STS} \leq 20.0$

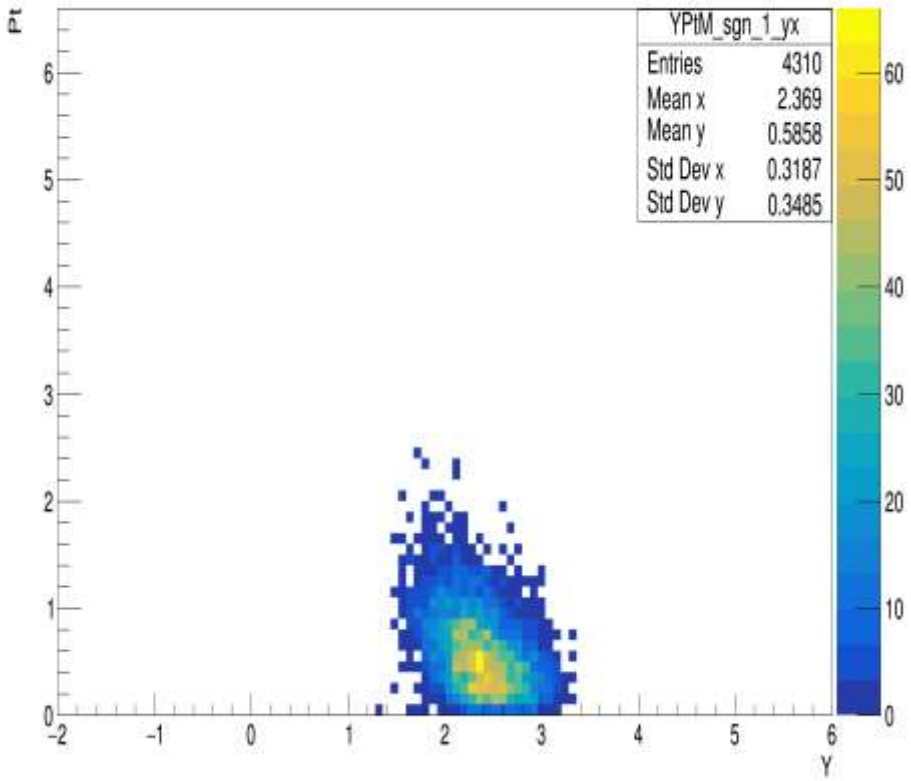
$\chi^2_{MUCH} \leq 20.0$

2σ cut in TOF

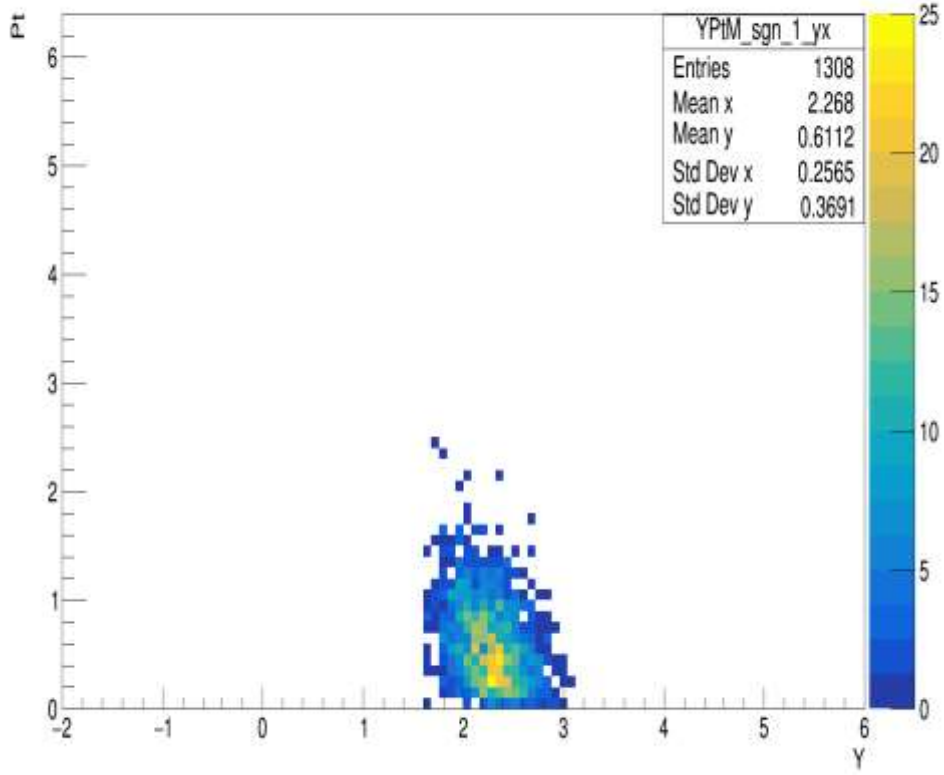
HMatrixML Model with cut 0.2

Y-Pt Spectra

ML: BDTG-1 (cut at 0.7)



Manual cuts



Comparision Table

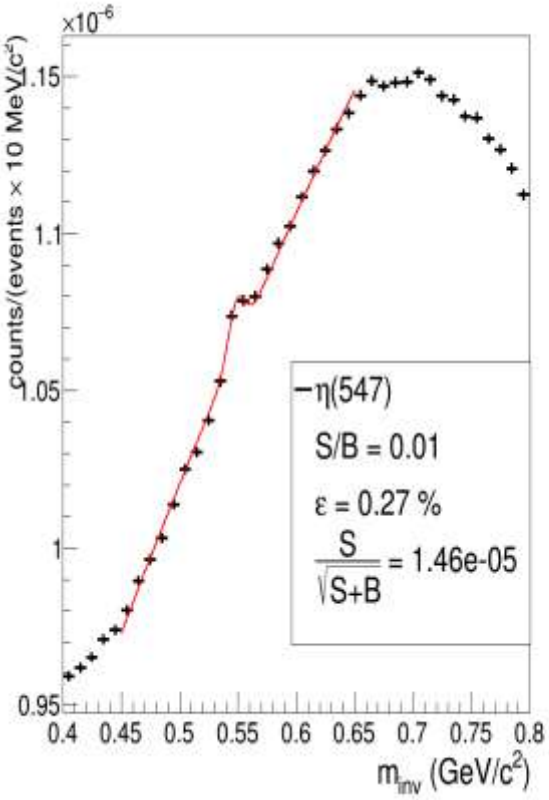
Method	ML output variable	S/B	Efficiency	Normalized Significance (w. r. t. Manual cuts)
Manual Cuts		0.215	0.62	1
BDTG-1	0.7	0.222	2.17	1.89
HMatrix	0.2	0.215	0.89	1.2
KNN	0.9	0.225	1.49	1.58

Eta meson($\eta \rightarrow \mu^+ + \mu^-$)

Invariant mass spectra of Eta

- **Simulation for 8AGeV Central AuAu collision in SIS100 setup**

with manual cuts



Cuts:

N of STS hits ≥ 7

N of MUCH hits ≥ 11

N of TRD hits ≥ 1

N of Tof hits ≥ 1

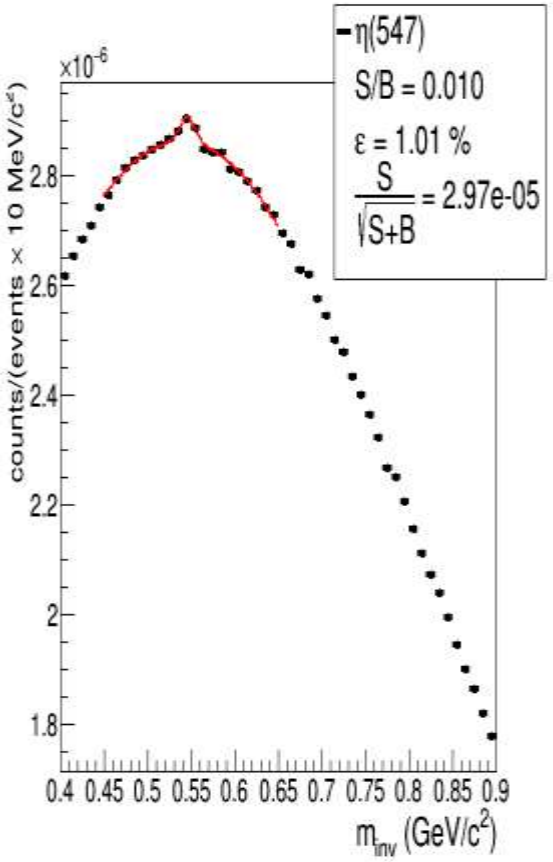
$\chi^2_{vertex} \leq 3.0$

$\chi^2_{STS} \leq 2.0$

$\chi^2_{MUCH} \leq 3.0$

2σ cut in TOF

ML:KNN



Cuts:

N of STS hits ≥ 4

N of MUCH hits ≥ 8

N of TRD hits ≥ 1

N of TOF hits ≥ 1

$\chi^2_{vertex} \leq 20.0$

$\chi^2_{STS} \leq 20.0$

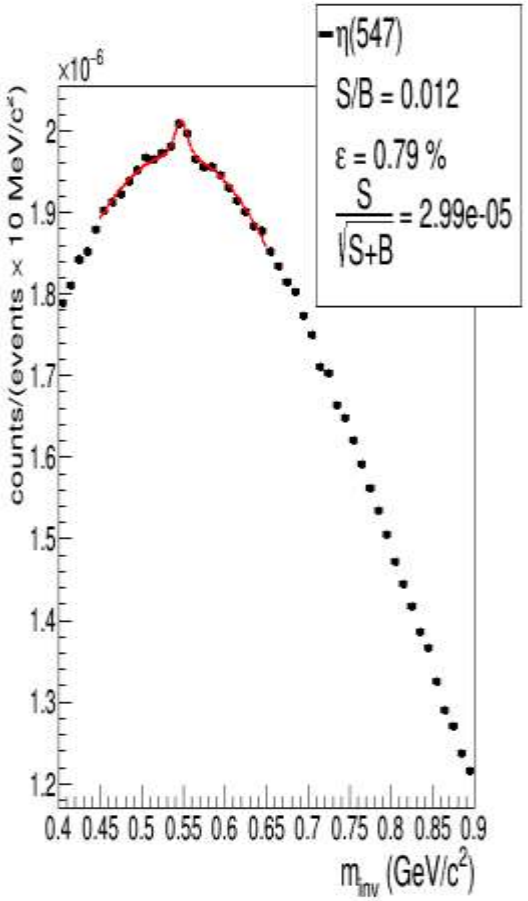
$\chi^2_{MUCH} \leq 20.0$

2σ cut in TOF

KNNML Model with cut 0.9

Invariant mass spectra

BDTG-1



Cuts:

N of STS hits ≥ 4

N of MUCH hits ≥ 8

N of TRD hits ≥ 1

N of TOF hits ≥ 1

$\chi^2_{vertex} \leq 20.0$

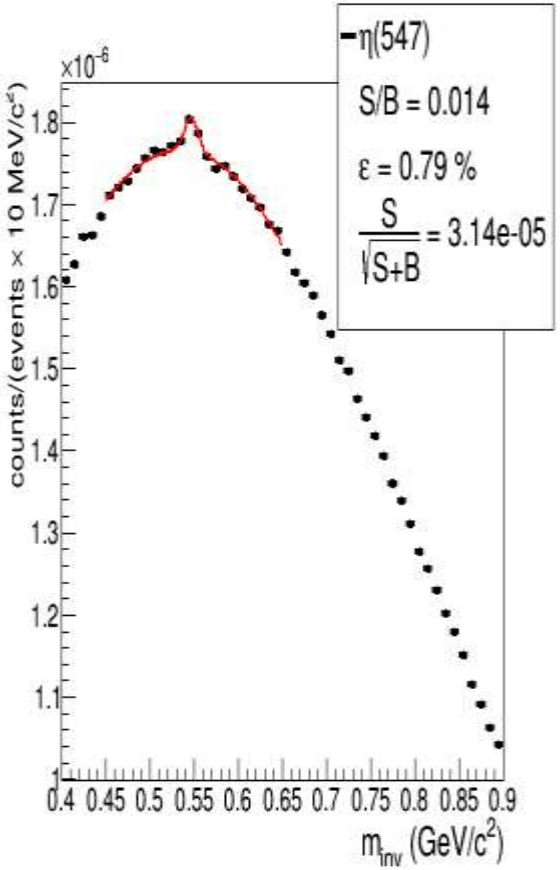
$\chi^2_{STS} \leq 20.0$

$\chi^2_{MUCH} \leq 20.0$

2σ cut in TOF

BDTG-1ML Model with cut 0.8

BDT-2



Cuts:

N of STS hits ≥ 4

N of MUCH hits ≥ 8

N of TRD hits ≥ 1

N of TOF hits ≥ 1

$\chi^2_{vertex} \leq 20.0$

$\chi^2_{STS} \leq 20.0$

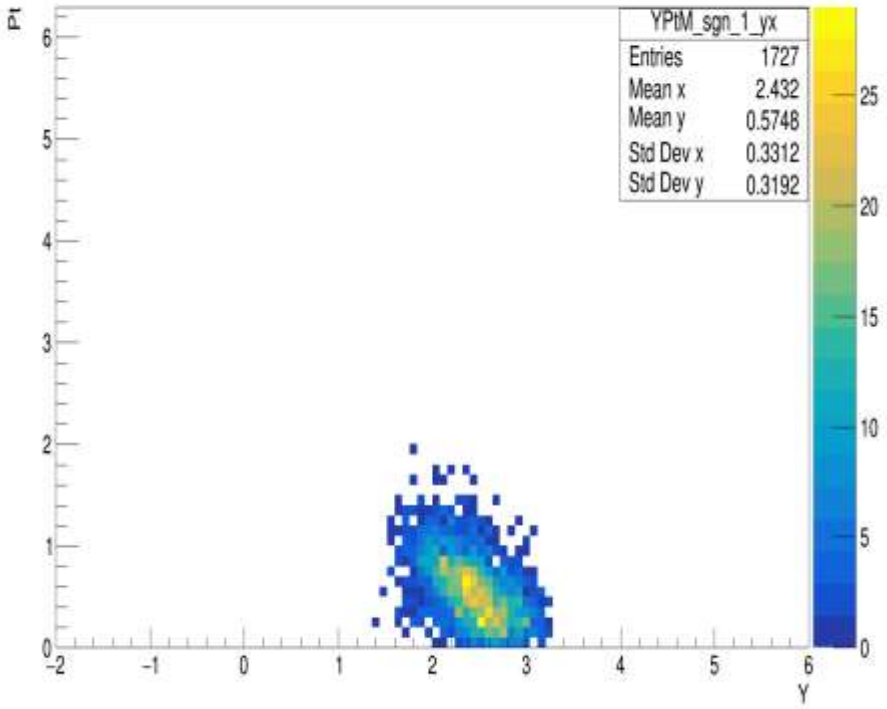
$\chi^2_{MUCH} \leq 20.0$

2σ cut in TOF

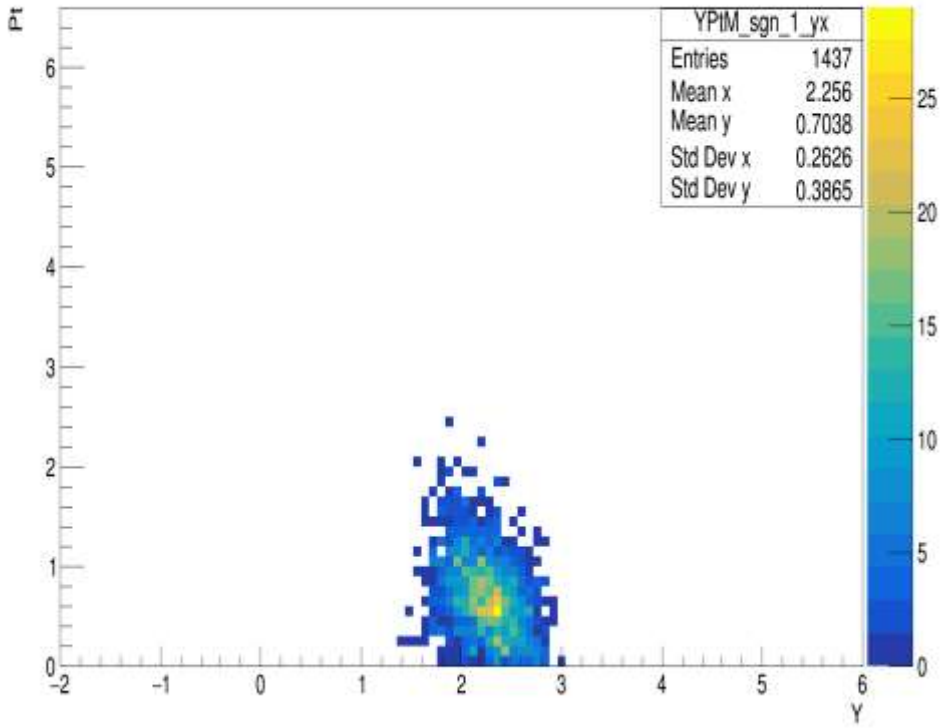
BDT-2ML Model with cut 0.1

Y-Pt Spectra

ML: BDTG-1 (cut at 0.8)



Manual cuts



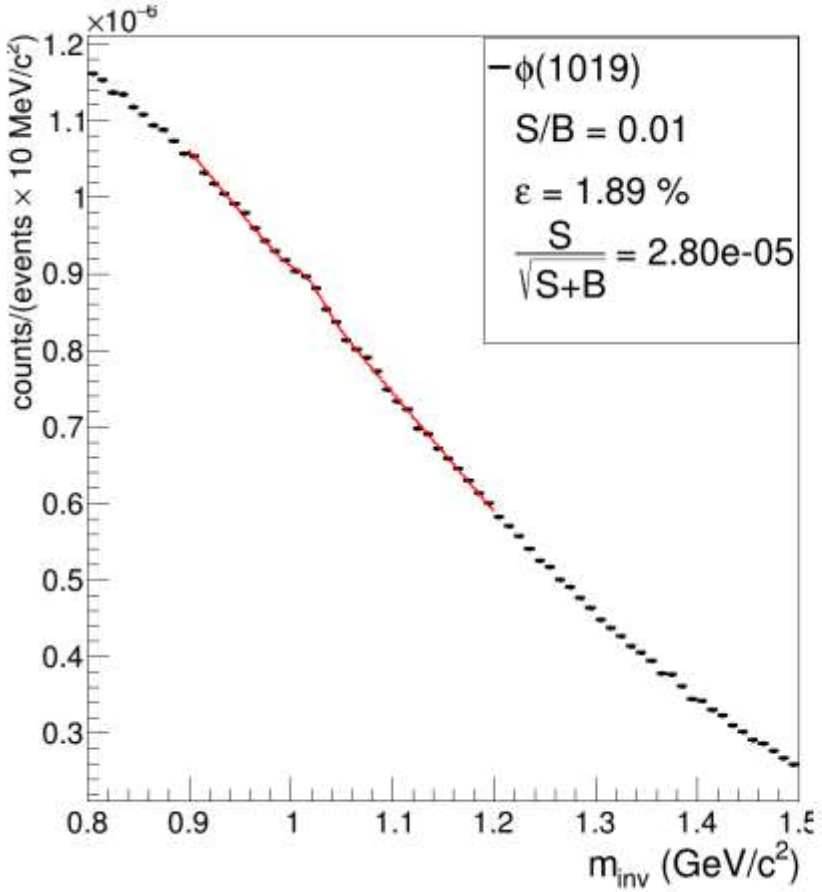
Comparision Table

Method	ML output variable	S/B	Efficiency	Normalized Significance
Manual Cuts		0.01	0.27	1
BDTG-1	0.8	0.012	0.79	2.05
BDT-2	0.1	0.014	0.79	2.15
KNN	0.9	0.010	1.01	2.03

Phi Meson ($\phi \rightarrow \mu^+ + \mu^-$)

Invariant mass spectra of Phi (with manual cuts)

- **Simulation for 8AGeV Central AuAu collision in SIS100 setup**



Cuts:

N of STS hits ≥ 7

N of MUCH hits ≥ 11

N of TRD hits ≥ 1

N of Tof hits ≥ 1

$\chi^2_{vertex} \leq 3.0$

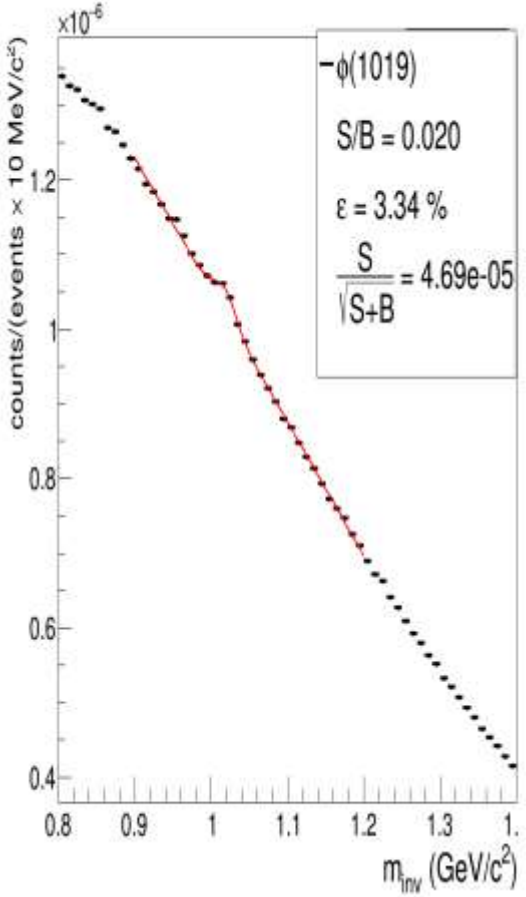
$\chi^2_{STS} \leq 2.0$

$\chi^2_{MUCH} \leq 3.0$

2σ cut in TOF

Invariant mass spectra of Phi

BDTG-1



Cuts:

N of STS hits ≥ 4

N of MUCH hits ≥ 8

N of TRD hits ≥ 1

N of TOF hits ≥ 1

$\chi^2_{vertex} \leq 20.0$

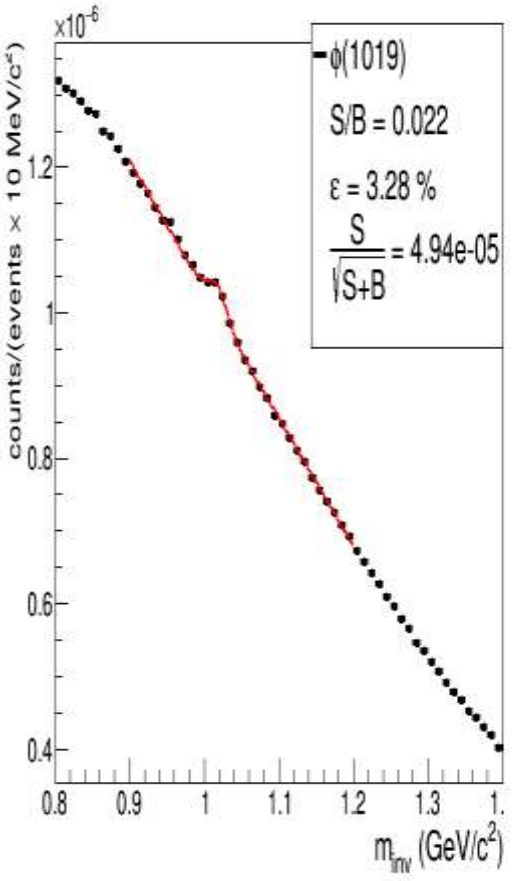
$\chi^2_{STS} \leq 20.0$

$\chi^2_{MUCH} \leq 20.0$

2σ cut in TOF

BDTG-1ML Model with cut 0.8

BDTG-2



Cuts:

N of STS hits ≥ 4

N of MUCH hits ≥ 8

N of TRD hits ≥ 1

N of TOF hits ≥ 1

$\chi^2_{vertex} \leq 20.0$

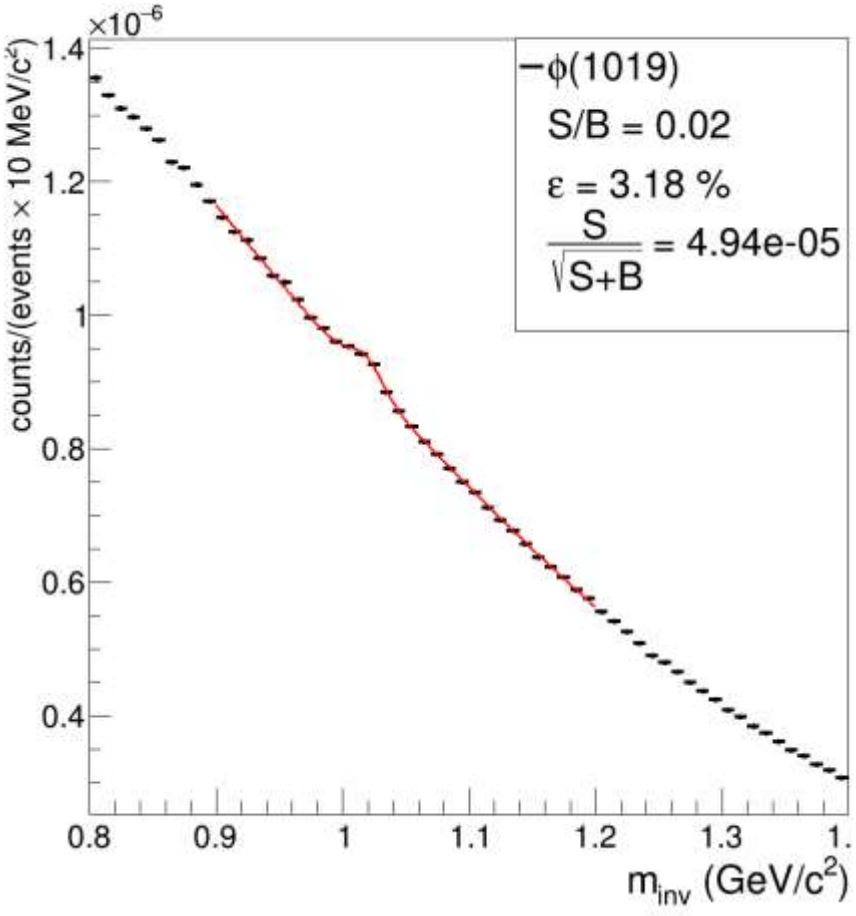
$\chi^2_{STS} \leq 20.0$

$\chi^2_{MUCH} \leq 20.0$

2σ cut in TOF

BDTG-2ML Model with cut 0.8

Invariant mass spectra (BDTG-1)*



Cuts:

N of STS hits ≥ 4

N of MUCH hits ≥ 8

N of TRD hits ≥ 1

N of TOF hits ≥ 1

$\chi^2_{vertex} \leq 20.0$

$\chi^2_{STS} \leq 20.0$

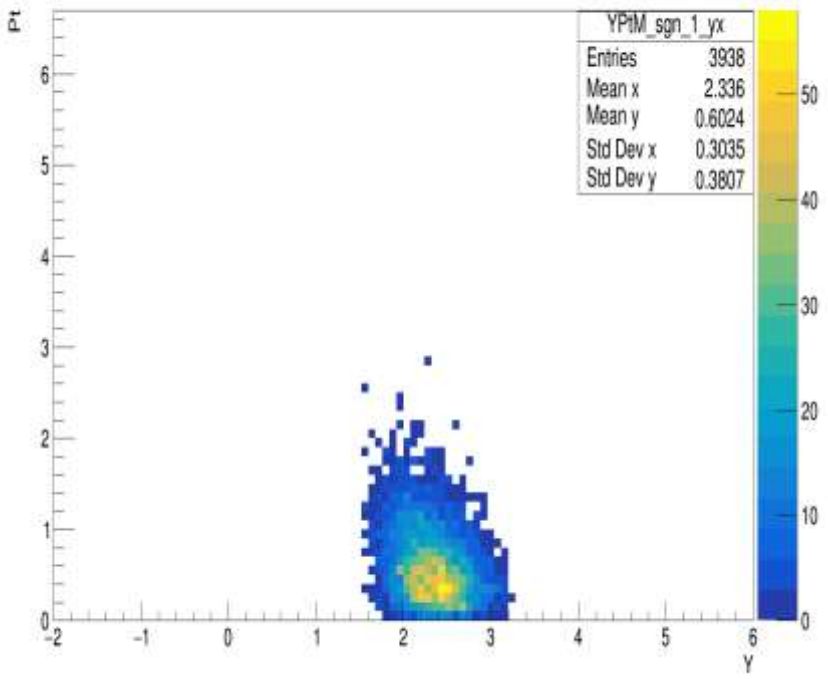
$\chi^2_{MUCH} \leq 20.0$

2σ cut in TOF

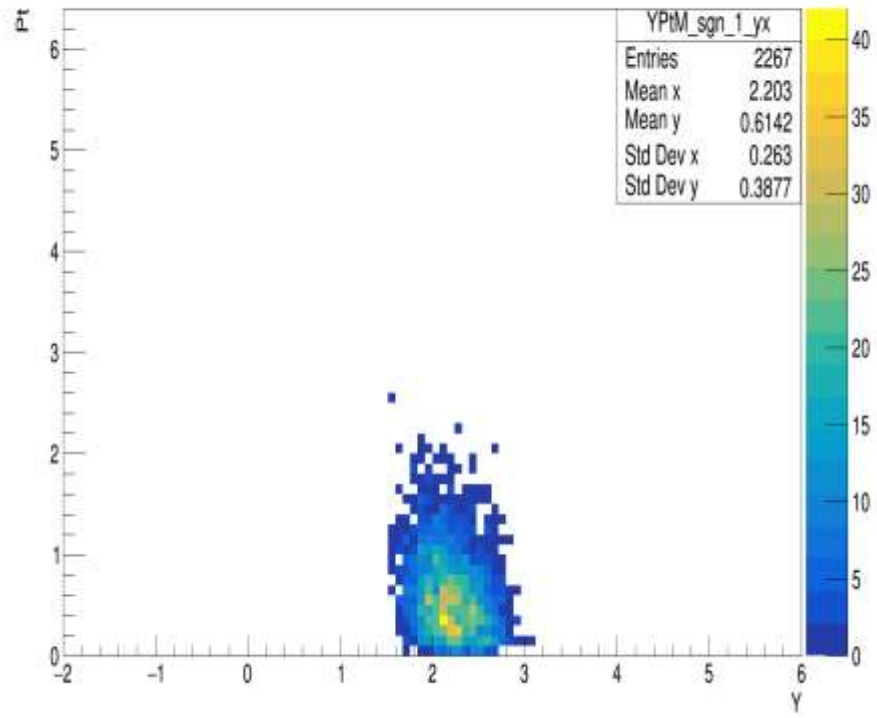
* using weight file of cocktail data

Y-Pt Spectra

ML: BDTG-1 (cut at 0.8)



Manual cuts

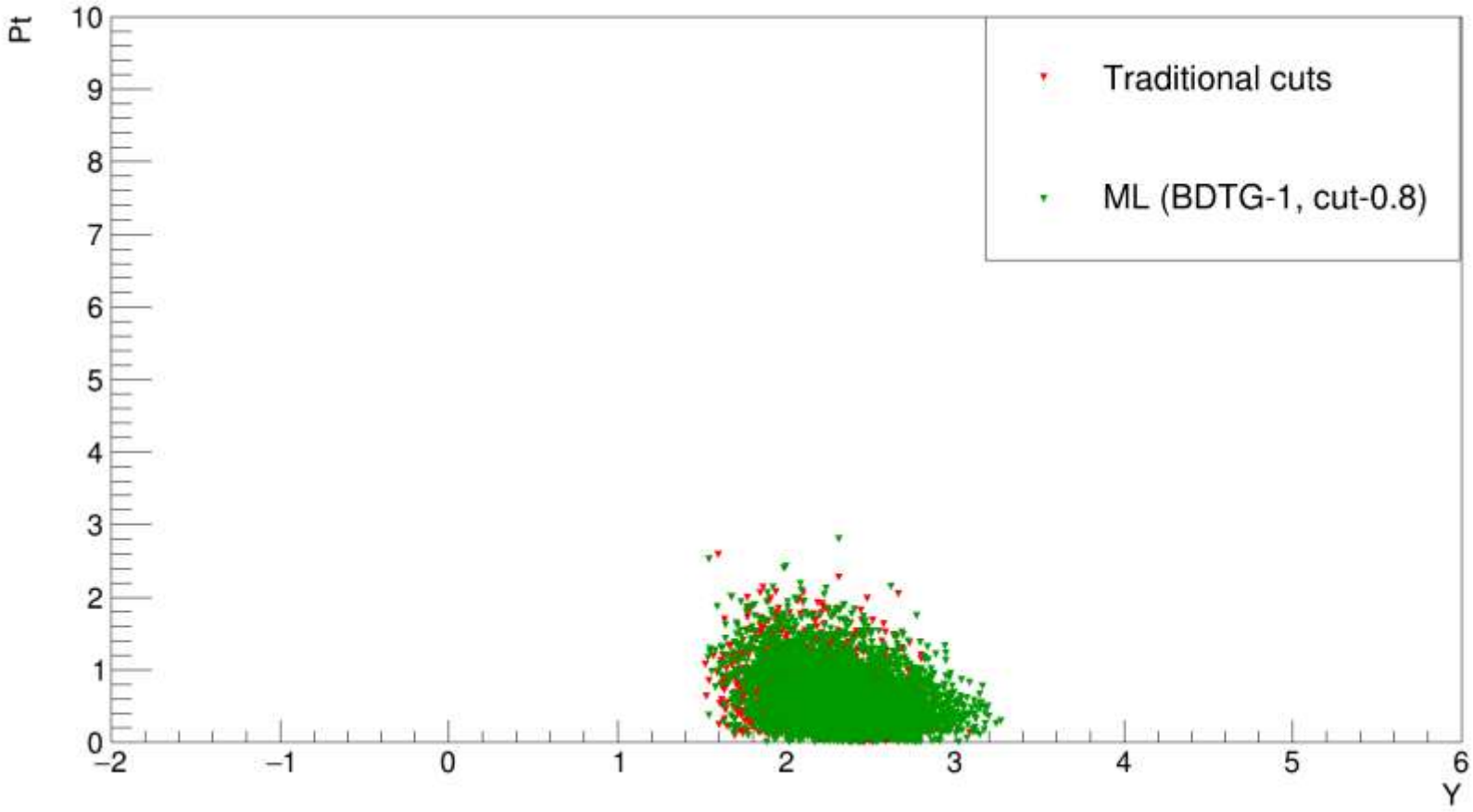


Result and Future Plan

- There are improvement in Dimuon performance through ML models.
- Based on requirement, we can select ML output variable.
- We are trying to obtain cocktail spectra of LMVMs using ML models.
- Some additional variables like Opening angle, rapidity of muon pairs can be used to train data and there is still scope of improvements.

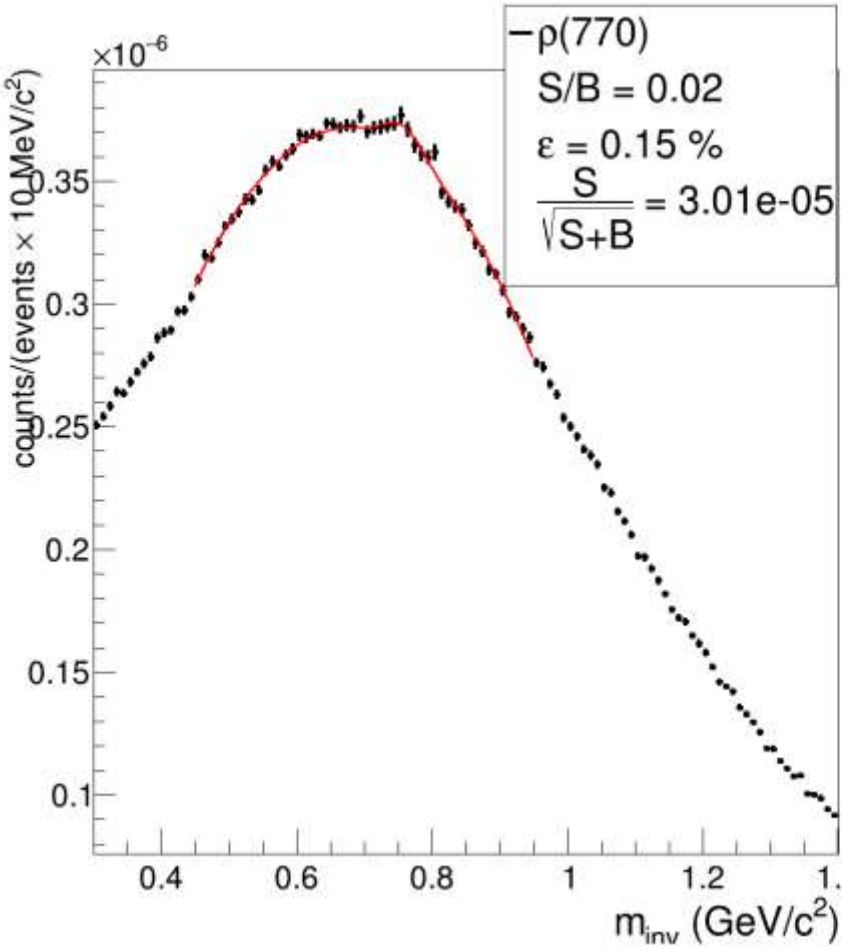
Backup

Y_Pt Spectra


Pt
10
9
8
7
6
5
4
3
2
1
0
−2
−1
0
1
2
3
4
5
6
Y
Traditional cuts
ML (BDTG-1, cut-0.8)

Rho Meson

Invariant mass spectra of rho (BDTG-1, cut – 0.875)



Cuts:

N of STS hits ≥ 4

N of MUCH hits ≥ 8

N of TRD hits ≥ 1

N of TOF hits ≥ 1

$\chi^2_{vertex} \leq 20.0$

$\chi^2_{STS} \leq 20.0$

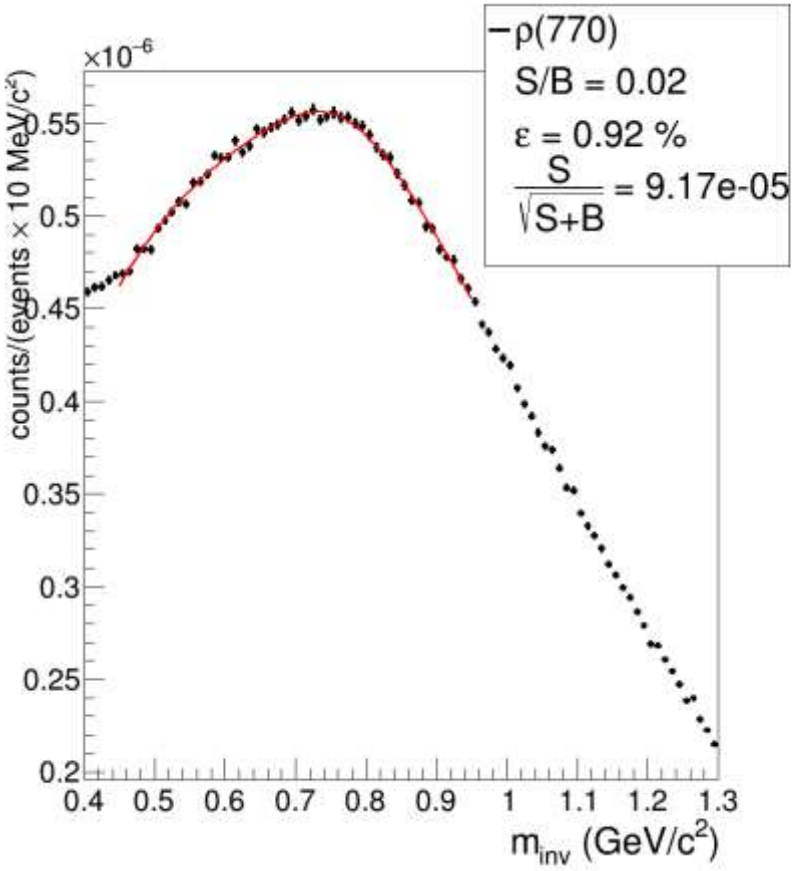
$\chi^2_{MUCH} \leq 20.0$

2σ cut in TOF

- For 100k events

Invariant mass spectra of rho (with manual cuts)

- **Simulation for 8AGeV Central AuAu collision in SIS100 setup**



Cuts:

N of STS hits ≥ 7

N of MUCH hits ≥ 11

N of TRD hits ≥ 1

N of Tof hits ≥ 1

$\chi^2_{vertex} \leq 2.0$

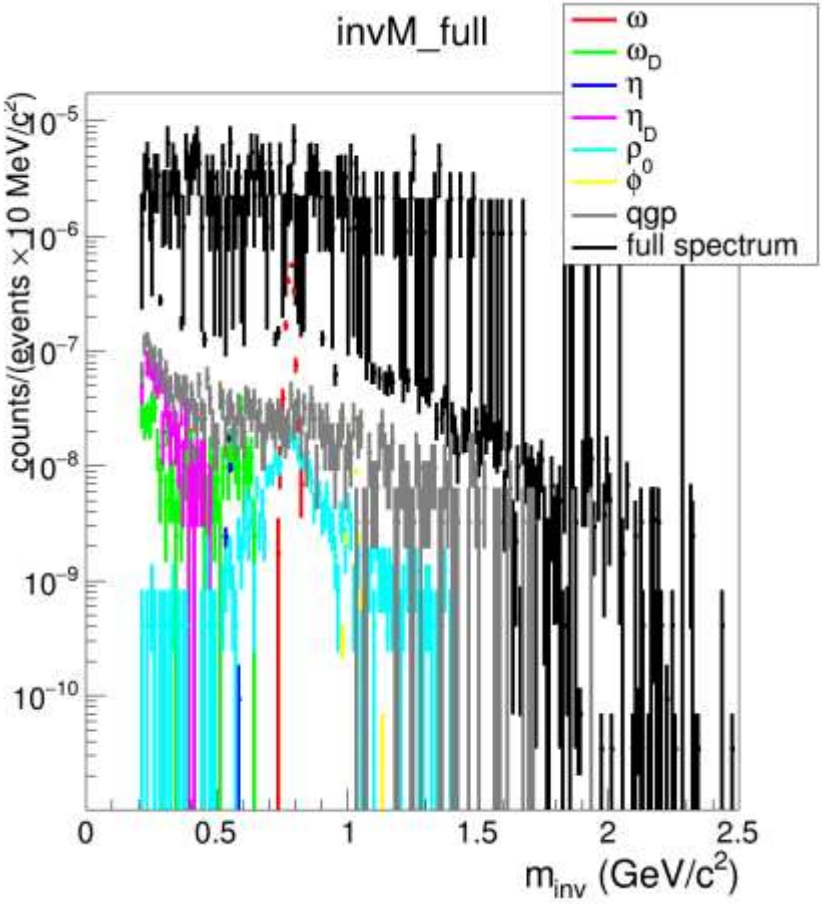
$\chi^2_{STS} \leq 2.0$

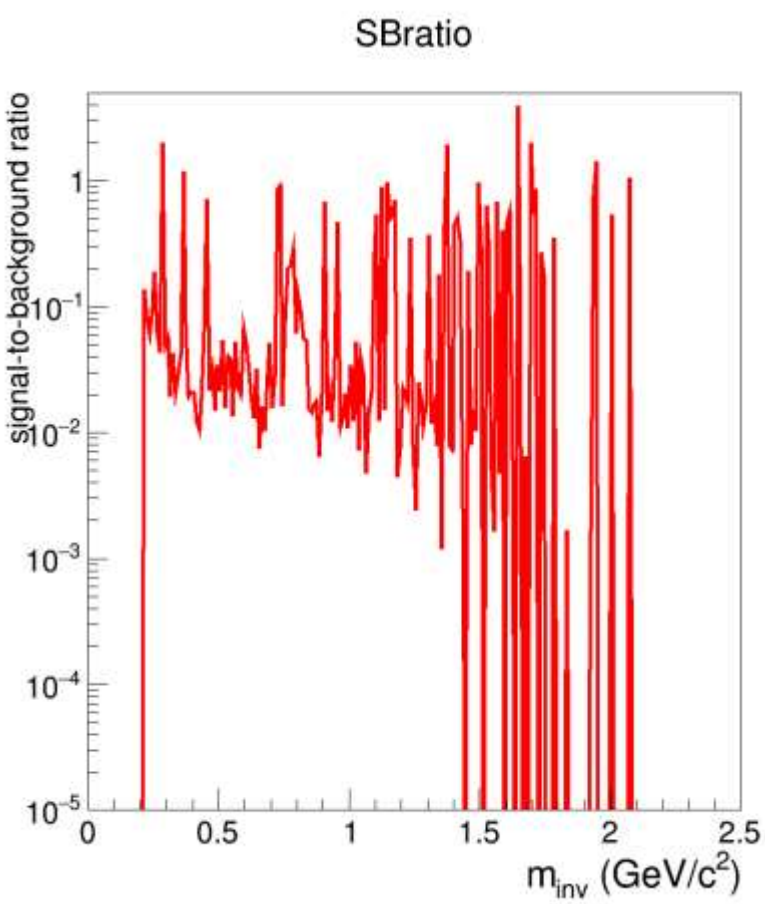
$\chi^2_{MUCH} \leq 2.5$

2σ cut in TOF

- For 100k events

Cocktail


invM_full
ω
ω_D
η
η_D
ρ_0
φ
qgp
full spectrum
counts/(events ×10 MeV/c²)
10^-5
10^-6
10^-7
10^-8
10^-9
10^-10
0
0.5
1
1.5
2
2.5
m_inv (GeV/c²)


SBratio
signal-to-background ratio
1
10^-1
10^-2
10^-3
10^-4
10^-5
0
0.5
1
1.5
2
2.5
m_inv (GeV/c²)

Backup

Machine Learning Concept	Description
Linear Regression	A statistical method for modeling the relationship between a dependent variable and one or more independent variables.
Clustering	The task of grouping a set of objects in such a way that objects in the same group are more similar to each other than to thos other groups.
Decision Tree	A decision support tool that uses a tree-like model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility.
Neural Networks	A set of algorithms, modeled loosely after the human brain, designed to recognize patterns and interpret data by labeling or clustering raw input.
Reinforcement Learning	A type of machine learning where an agent learns to behave in an environment by performing actions and seeing the results.
Logistic Regression	A statistical model that uses a logistic function to model a binary dependent variable, although more complex extensions exist
Naive Bayes	A simple probabilistic classifier based on applying Bayes' theorem with strong (naive) independence assumptions between th features.
Supervised Learning	A type of machine learning where the model is trained on a labeled dataset, meaning the data is already paired with the corre answers.
Support Vector Machine	A set of supervised learning methods used for classification, regression, and outliers detection.
Probability	The measure of the likelihood that an event will occur, often used as the basis for statistical models in machine learning.
Random Forest	An ensemble learning method for classification, regression, and other tasks that operates by constructing a multitude of decis trees at training time.
Variance	A measure of the dispersion of a set of data points around their mean value, crucial in statistical modeling and machine learni
Evaluation Metrics	Standards or methods used to measure the effectiveness of a machine learning model (accuracy, precision, recall, F1 score,
Bagging	A method in ensemble learning that improves the stability and accuracy of machine learning algorithms by combining multiple models.
Data Wrangling	The process of cleaning, structuring, and enriching raw data into a desired format for better decision making in less time.
Dimensionality Reduction	The process of reducing the number of random variables under consideration by obtaining a set of principal variables.
K-nearest Neighbors Algorithm	A non-parametric method used for classification and regression, where the input consists of the k closest training examples in feature space.
Programming	The process of creating a set of instructions for a computer to perform tasks, often an essential skill in implementing machine learning models.
Regularization	A technique used to reduce the error by fitting a function appropriately on the given training set and avoiding overfitting.
Statistics	The science of using data to make decisions, fundamental to machine learning for interpreting data and results.
Binomial Distribution	A probability distribution that summarizes the likelihood that a value will take one of two independent states under a given set parameters.
Bootstrap Sampling	A statistical method for estimating the sampling distribution of an estimator by sampling with replacement from the original sa
Exploratory Data Analysis	An approach to analyzing data sets to summarize their main characteristics, often using visual methods.
Data Collection	The process of gathering and measuring information on variables of interest in a systematic fashion, enabling one to answer relevant questions.