New proposal for the B $\to$ K*$\mu\mu$ selection

Marcin Chrzaszcz[1,2], Tatiana Likhomanenko[3,4],

Andrey Ustyuzhanin[3,4,5]

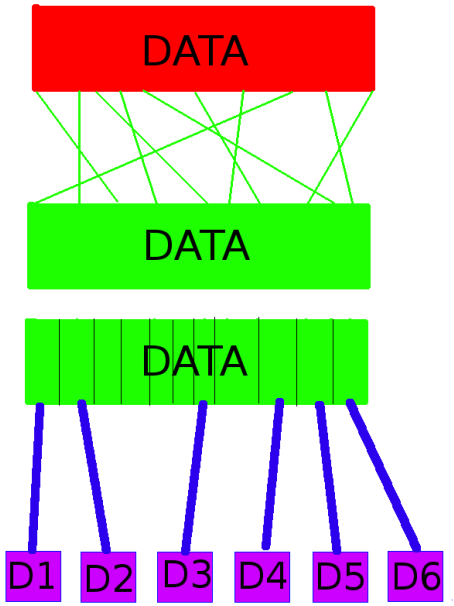
November 27, 2013

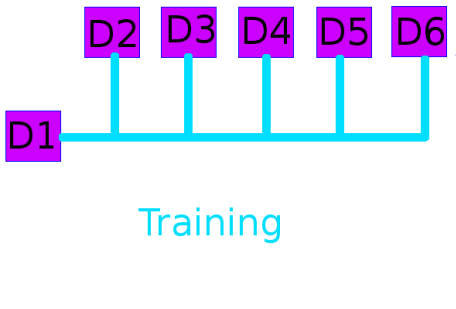
[1] University of Zurich, [2] Institute of Nuclear Physics, [3] Yandex, [4] Kurchatov Institute, [5] Imperial College

University of Zurich[UZH]

1. **Chopping technique**
   - General idea
   - Performance gain
2. **Proposal of new variables**
   - MC/DATA comparison
   - Performance gain (ROC, 1:1 comparison last BDT)
3. **Conclusions**

Chopping Data Set, How to





1. Reshuffling the events to guarantee the uniformity of the data.

2. Chopping in sub-samples.

3. Training using n-1 sub-samples and applying the result on the remaining one (iteratively)
Increase in the statistics used in the training (more stable MVA response), no bias in the result :-)

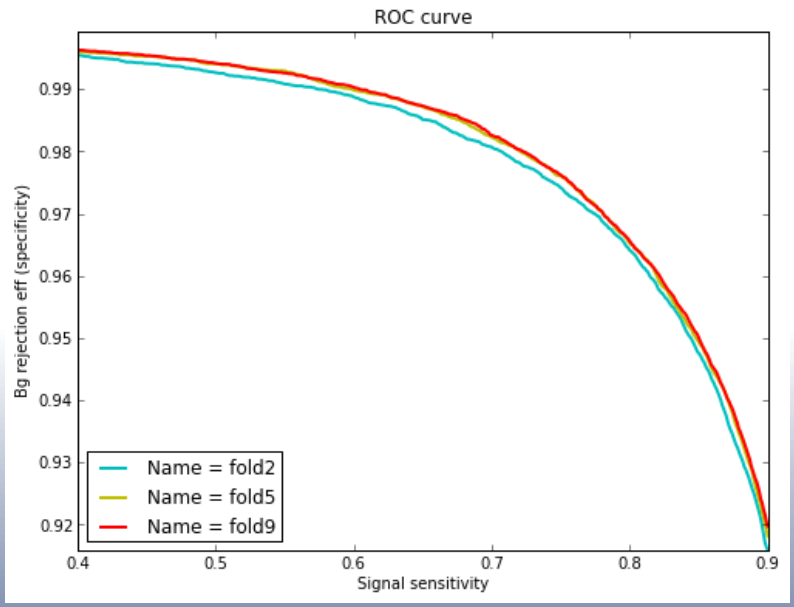
Chopping performance (I)

- Chopping technique studied with MatrixNet and the BDT (TMVA)
- Comparison of ROC curves for different sampling in the chopping procedure
- Comparison repeated for different input variable configurations

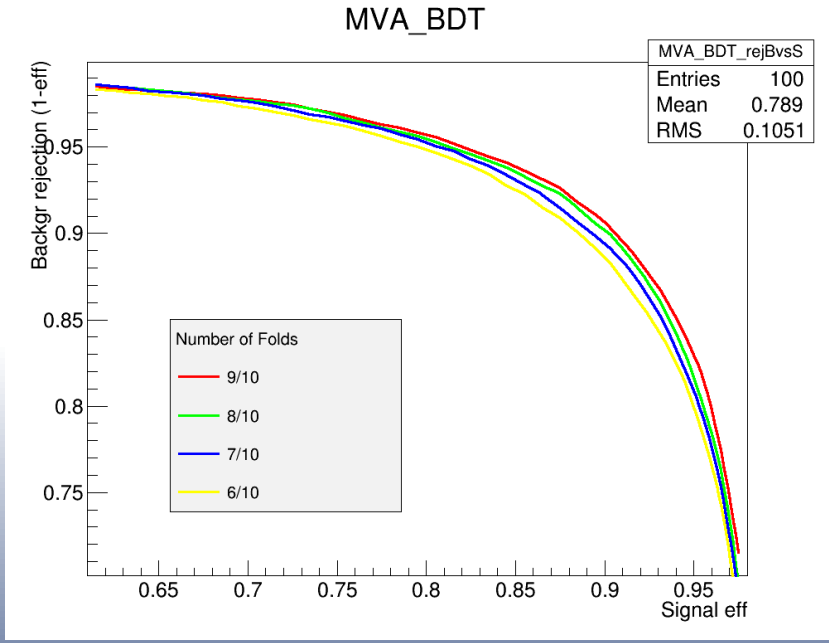
Chopping performance (II)

Comparison performed using as input: var5 + hadron DLL = BASE

MatrixNet



TMVA



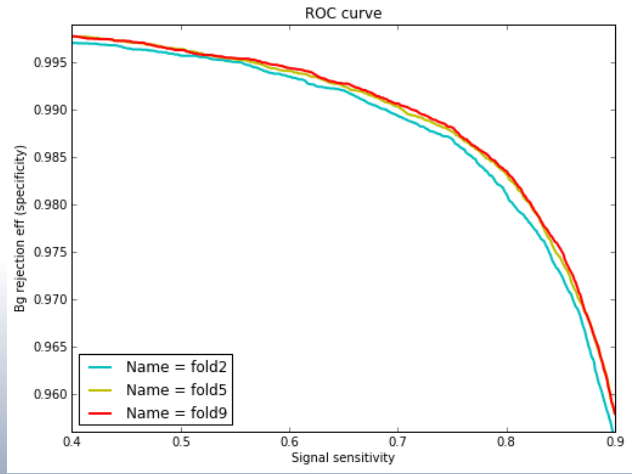
Better performance with 9 sub-samples for training.

Chopping performance (III)

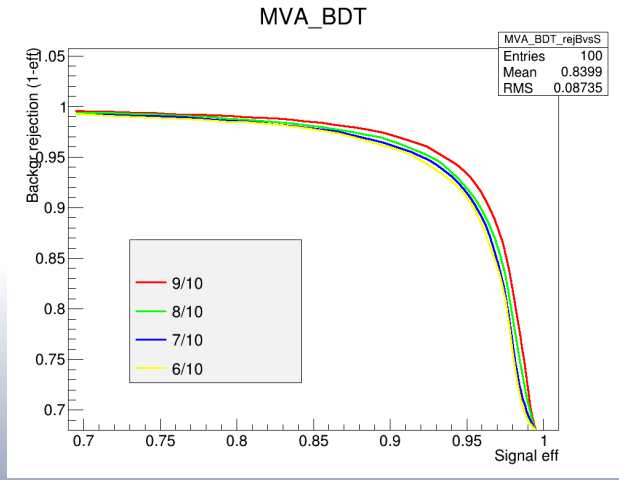
Comparison performed using as input: base + ISO + TAUERR + MUPID

- Data-MC comparison and single variable effect in the BDT performance in few slides.

MatrixNet



TMVA



Better performance with 9 sub-samples for training: result more evident here than with less variables as expected from statistics

Implication for the analysis

- The tuple to be analyzed will have only 1 branch which contains the different MVAs for the different subsamples.
- The MC for the acceptance correction will contain one branch that is the average of the different BDTs.
- Thanks to the reshuffling the BDT response over many sample is the same as the average of the BDTs.

Complication is only in the training phase and preparing the ntuples, after that the analysis goes exactly in the same way as having only 1 BDT.

Chopping conclusions

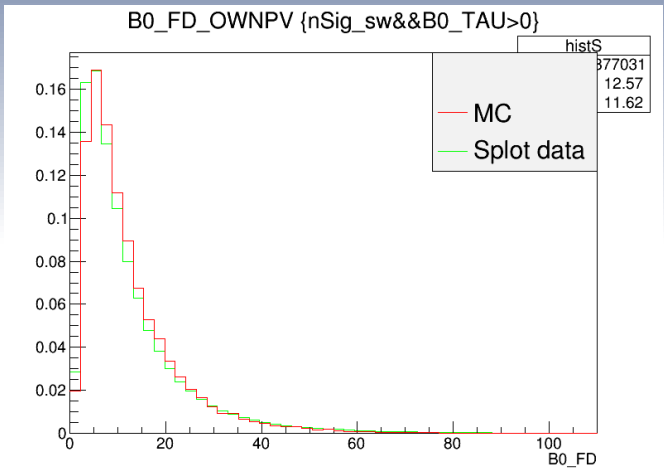
- Gain in performance with the chopping technique even more evident with more variables (as expected).
- Almost no extra complication/work needed in the data analysis.
- In the BACKUPS you have chopping for different configurations of variables.

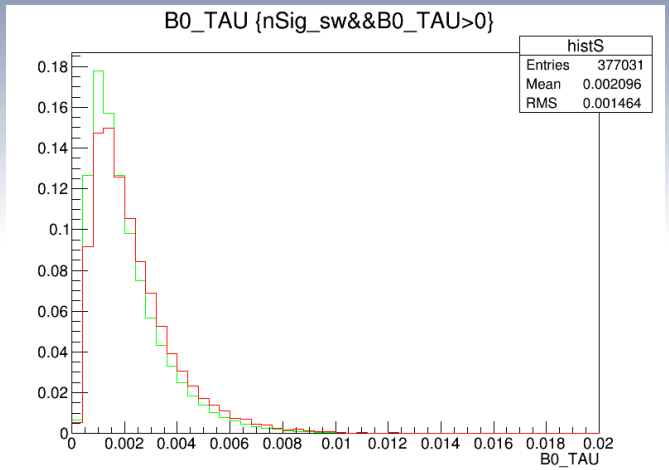
Proposal of new variables for the MVA

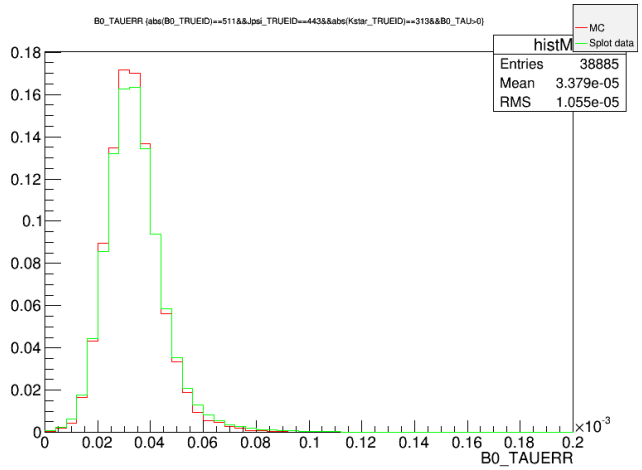
Definitions:

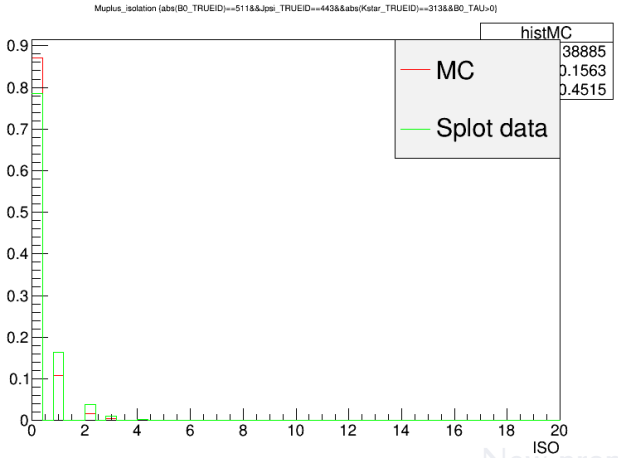
- Last BDT:
  BDT presented on 21^{th} of August 2013. Includes:Var5+probNN(π, K, μ) and isolation.
- Baseline:
  Var5+PiPIDK+KPDK
- New variables proposed for MVA:
  - TAUERR
  - FD
  - MUPID
  - ISO

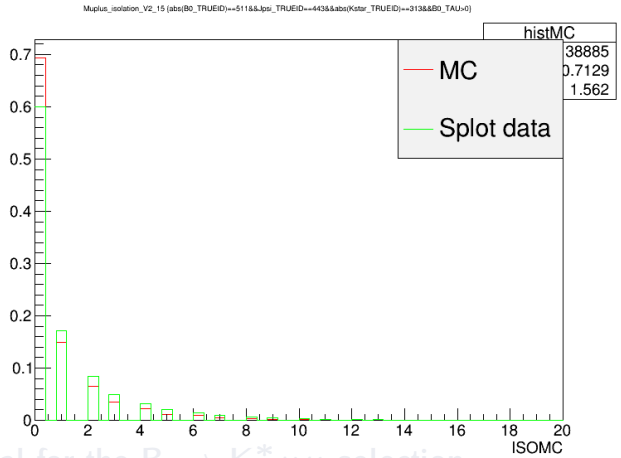
MC/DATA Comparison





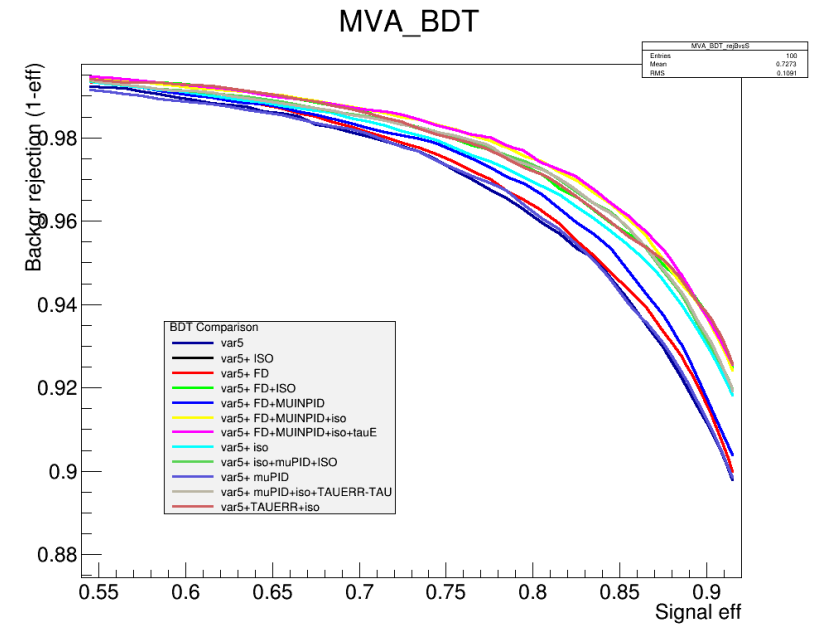




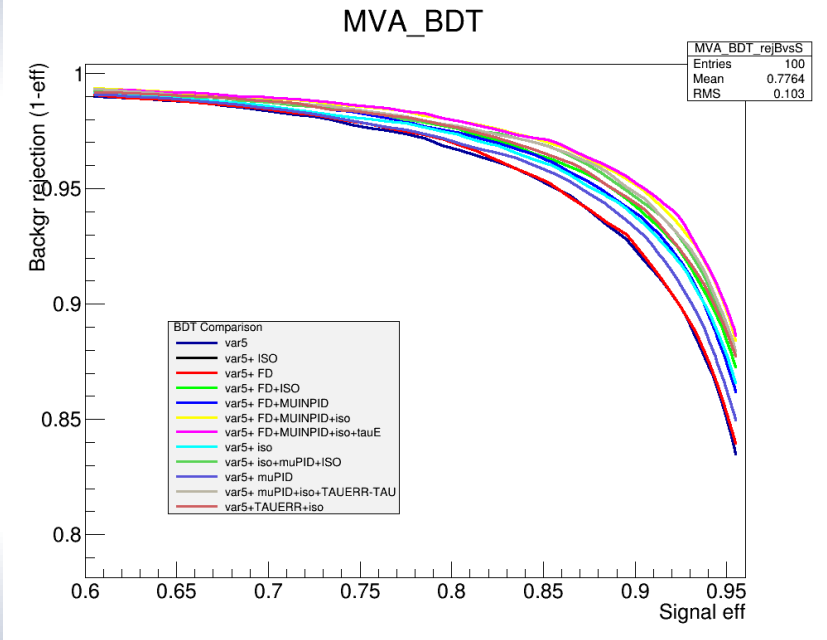


New variables performance(I)

PID=DLL



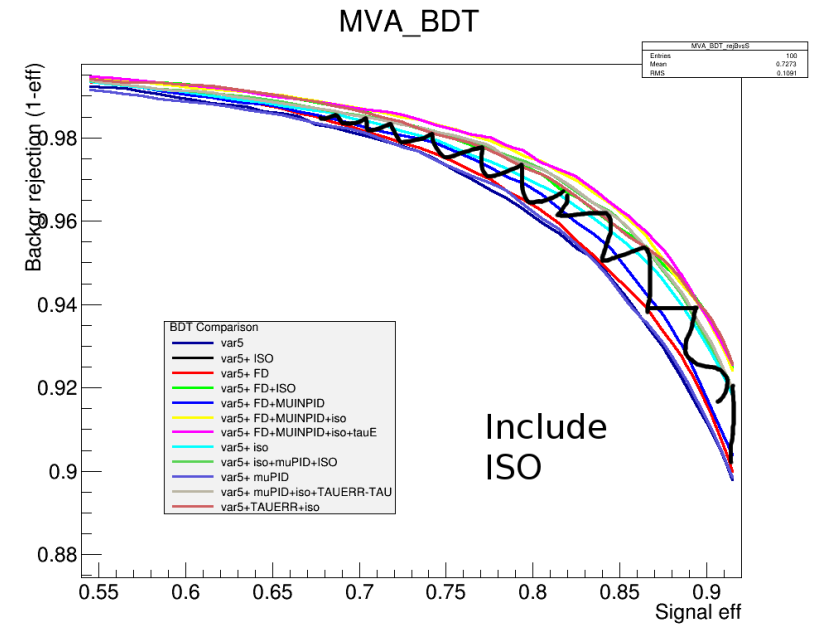
PID=ProbNN



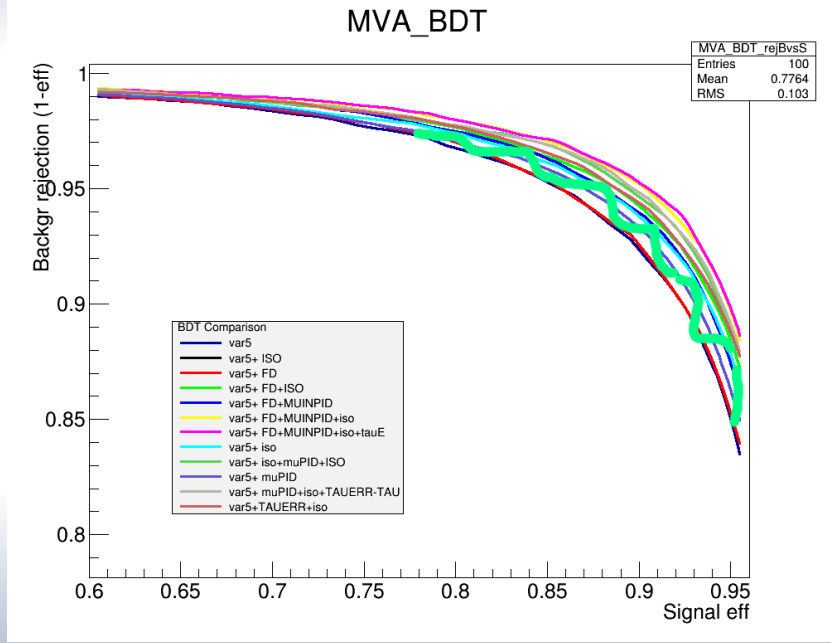
- ROC curve comparison of several variables configurations.
- Best performance adding FD, MUPID, TAUERR and ISO (pink line)
- Best performance of the ProbNN

New variables performance(II), ISO GAIN

PID=DLL



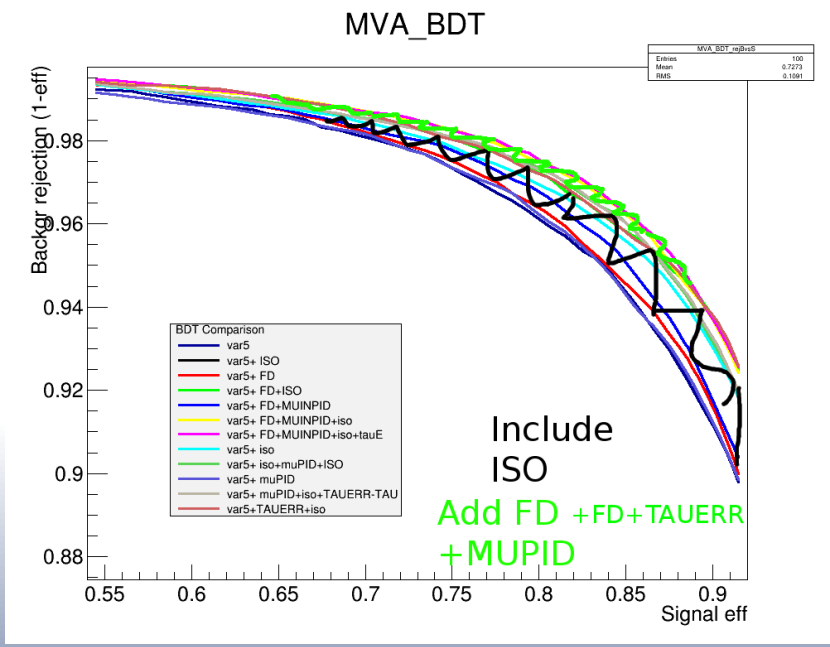
PID=ProbNN



- Include ISO

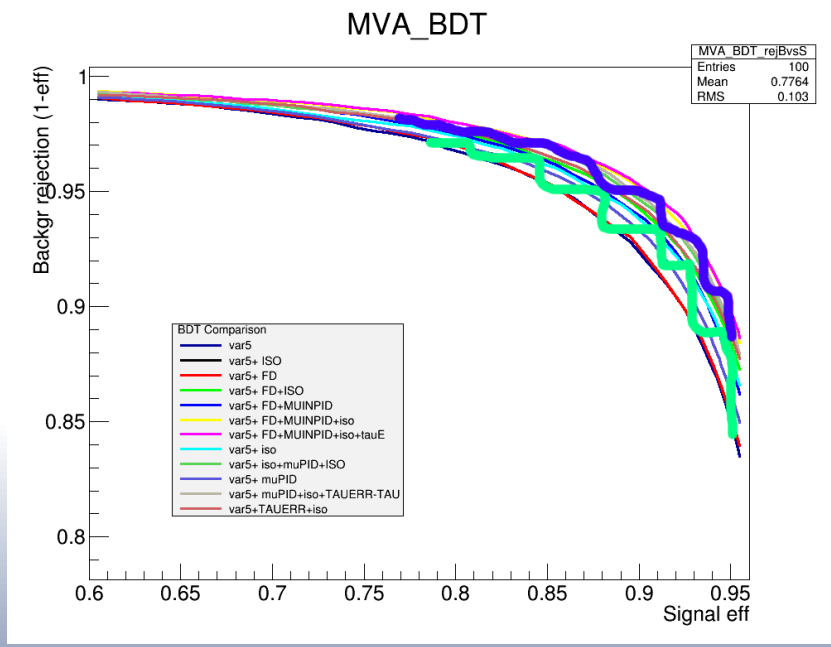
New variables performance(III), ISO+MUPID+TAUERR+FD

PID=DLL



- Include ISO and MUPID

PID=ProbNN



- Include ISO and MUPID

New variables performance(IV)

- Good data/MC agreement (reasonable for the ISO, and not worse than the old one)
- Best performance obtained adding FD, MUPID, TAUERR and ISO
- Best performance of the ProbNN
- MUPID will be also useful against peaking misidentified background.

Selections comparison

- MatrixNet outperformed TMVA.
- All fits made in the same way as Sam did(fixing the same parameters to the same values)
- 1:1 comparison with the last BDT and the baseline.
- Same data for both.

1:1 comparison with Last BDT (I)

q^2	Last BDT		MatrixNet	
$[GeV^2]$	Signal	Bck	Signal	Bck
$0.1, 2$	407 ± 25	58 ± 7	412 ± 22	39 ± 5
$2, 4.3$	202 ± 19	95 ± 7	220 ± 17	54 ± 5
$4.3, 8.68$	573 ± 32	170 ± 10	591 ± 28	131 ± 8
$10.09, 12.86$	508 ± 26	93 ± 7	508 ± 25	88 ± 7
$14.18, 16$	310 ± 20	49 ± 5	324 ± 20	$43. \pm 6$
$16, 19$	359 ± 29	34 ± 8	373 ± 21	35 ± 5
$0.1, 19$	2355.2 ± 63	510 ± 19	2365 ± 55	403 ± 15

Comparison between different configuration (II)

q^2	MN Baseline[1]		MN FULL[2]	
$[GeV^2]$	Signal	Bck	Signal	Bck
$0.1, 2$	384 ± 22	66 ± 8	419 ± 21	37 ± 5
$2, 4.3$	249 ± 21	120 ± 9	225 ± 18	50 ± 8
$4.3, 8.68$	641 ± 32	255 ± 12	591 ± 28	130 ± 8
$10.09, 12.86$	534 ± 27	140 ± 9	510 ± 25	86 ± 7
$14.18, 16$	328 ± 21	73 ± 6	328 ± 20	$46. \pm 5$
$16, 19$	386 ± 21	65 ± 8	361 ± 20	36 ± 5
$0.1, 19$	2501 ± 60	741 ± 22	2369 ± 55	396 ± 15

[1] Var+KPIDLL

[2] Var5+ProbNN+Iso+FD+TAUERR

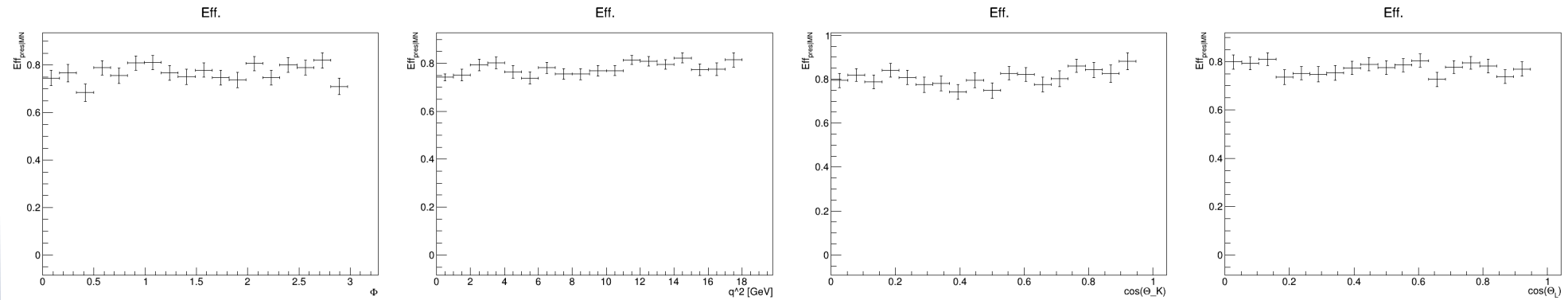
Comparison between different configuration (III)

q^2	MN FULL DLL[3]		MN FULL[4]	
$[GeV^2]$	Signal	Bck	Signal	Bck
$0.1, 2$	365 ± 23	43 ± 6	419 ± 21	37 ± 5
$2, 4.3$	227 ± 21	65 ± 6	225 ± 18	50 ± 8
$4.3, 8.68$	599 ± 29	154 ± 9	591 ± 28	130 ± 8
$10.09, 12.86$	511 ± 24	164 ± 7	510 ± 25	86 ± 7
$14.18, 16$	321 ± 20	47 ± 6	328 ± 20	$46. \pm 5$
$16, 19$	364 ± 21	39 ± 6	361 ± 20	36 ± 5
$0.1, 19$	2373 ± 56	468 ± 16	2369 ± 55	396 ± 15

[3] Var5+DLL+Iso+FD+TAUERR
[4] Var5+ProbNN+Iso+FD+TAUERR

MatrixNet efficiency

- Sim08 PHSP
- Efficiency defined as $\epsilon = n_{evts}(afterMN)/n_{evts}(afterpresel.)$



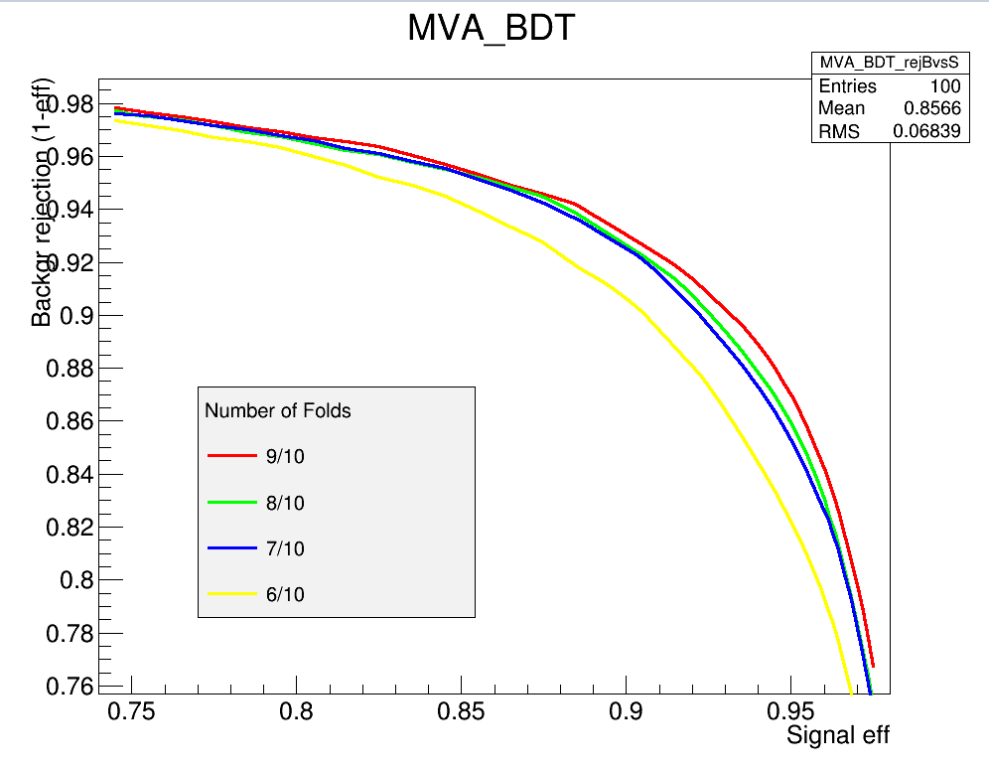
- Flat response in the angles and in q^2

Conclusions

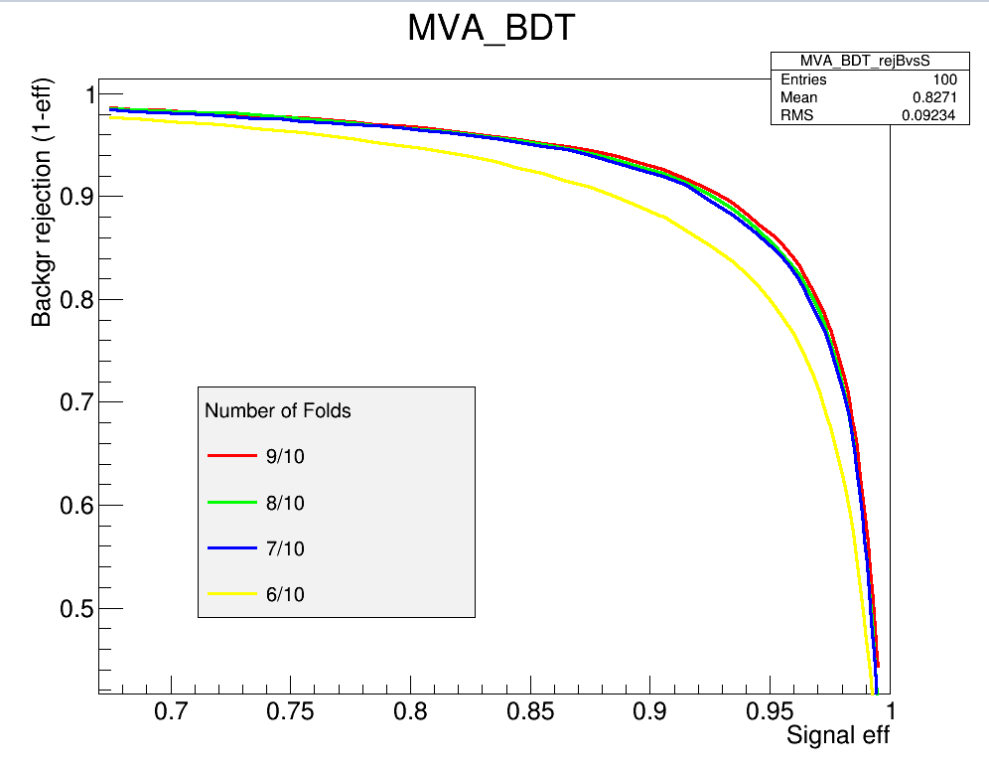
- Gain using the chopping technique, without extra complications in the analysis procedure
- Gain in performance using new variables which showed good agreement with MC
- Gain in performance using MatrixNet (respect to previous BDTs)
- Reduced background events keeping same signal efficiency.

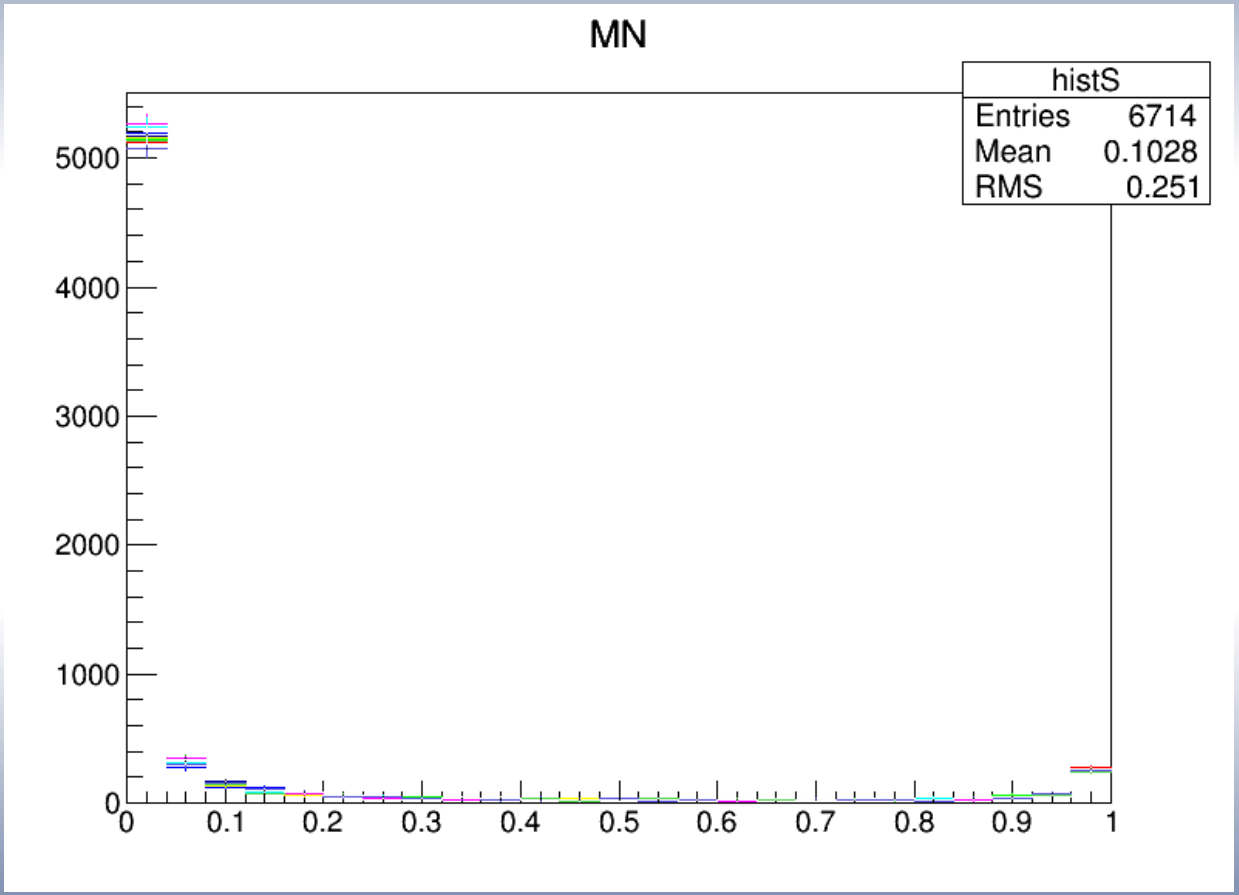
BACKUP

Base+MCISO


MVA_BDT
MVA_BDT_rejBvsS
Entries 100
Mean 0.8566
RMS 0.06839
Backgr rejection (1-eff)
Signal eff
Number of Folds
9/10
8/10
7/10
6/10
0.98
0.96
0.94
0.92
0.9
0.88
0.86
0.84
0.82
0.8
0.78
0.76
0.75
0.8
0.85
0.9
0.95

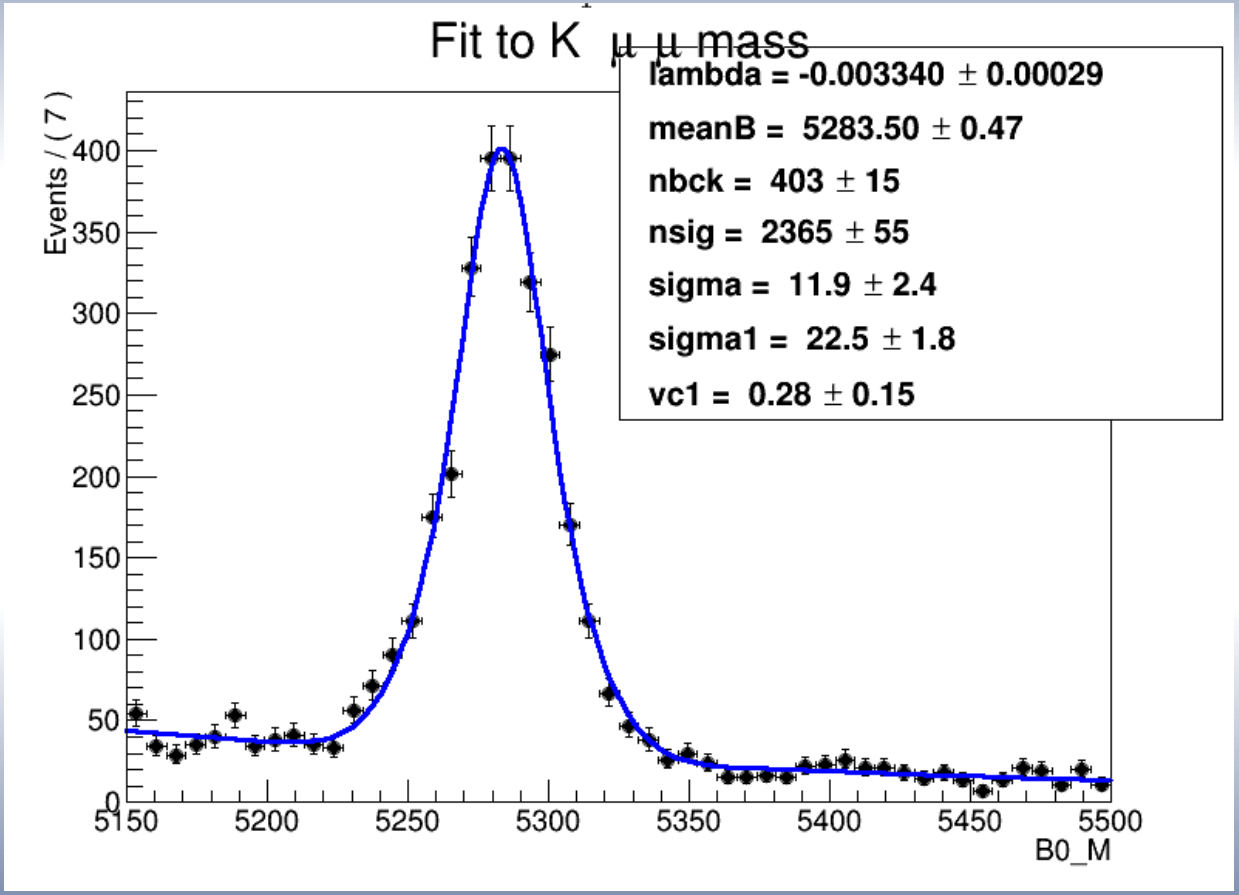
Base+ISO


MVA_BDT
MVA_BDT_rejBvsS
Entries 100
Mean 0.8271
RMS 0.09234
Backgr rejection (1-eff)
Signal eff
Number of Folds
9/10
8/10
7/10
6/10
1
0.9
0.8
0.7
0.6
0.5
0.7
0.75
0.8
0.85
0.9
0.95
1

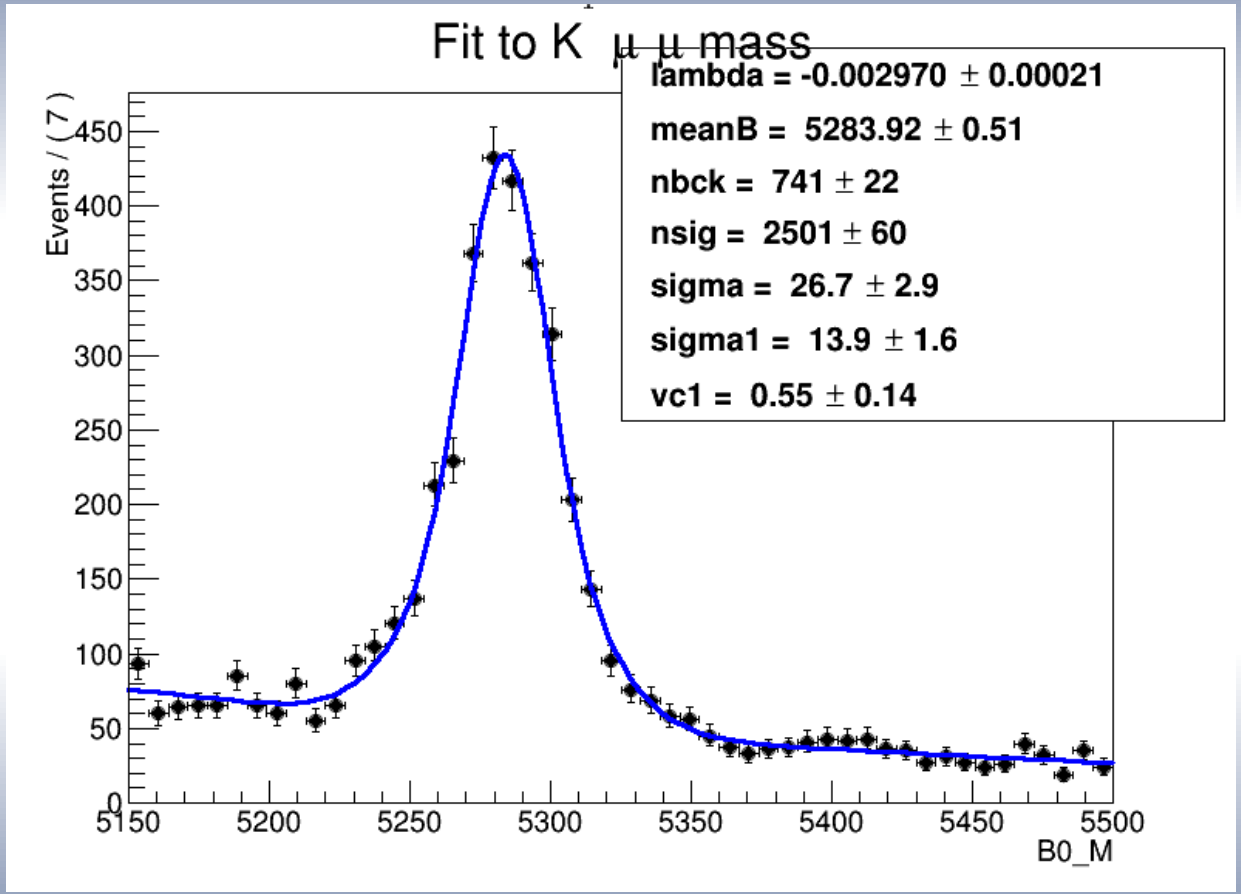

MN
histS
Entries 6714
Mean 0.1028
RMS 0.251
5000
4000
3000
2000
1000
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

MN:Iso+var+probNNx4

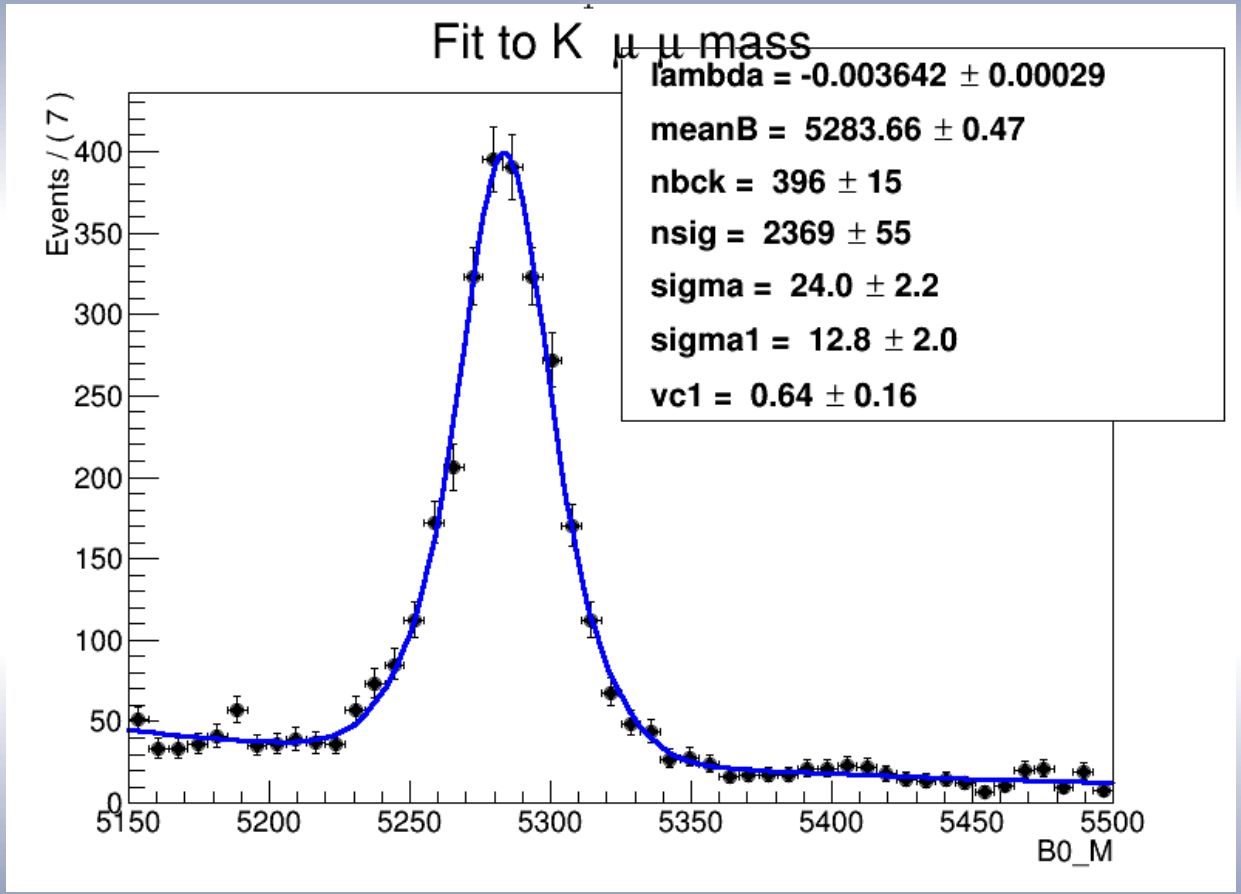
Base+ISO



BASE


Fit to K μ μ mass
lambda = -0.002970 ± 0.00021
meanB = 5283.92 ± 0.51
nbck = 741 ± 22
nsig = 2501 ± 60
sigma = 26.7 ± 2.9
sigma1 = 13.9 ± 1.6
vc1 = 0.55 ± 0.14
Events / (7)
B0_M

FULL


Fit to K μ μ mass
lambda = -0.003642 ± 0.00029
meanB = 5283.66 ± 0.47
nbck = 396 ± 15
nsig = 2369 ± 55
sigma = 24.0 ± 2.2
sigma1 = 12.8 ± 2.0
vc1 = 0.64 ± 0.16
Events / (7)
0
50
100
150
200
250
300
350
400
5150
5200
5250
5300
5350
5400
5450
5500
B0_M

matrix Net ROC