$B \to K^* \mu\mu$ 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

Reminder

On last $K^*\mu\mu$ meeting we agreed on:

- Use 10 Folds.
- Use DLL instead of ProbNN.
- Use isolation inside MVA.
- Use DLL for μ.

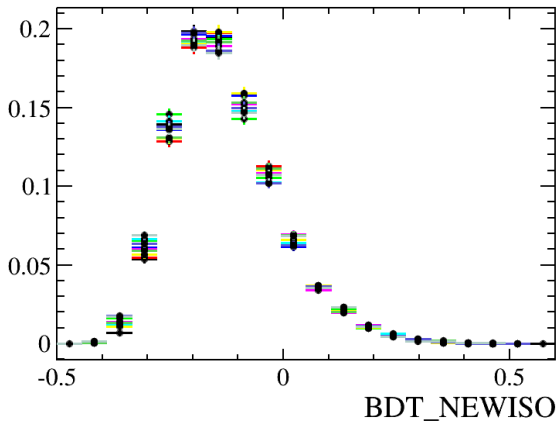
Remaining issues:

- Data agreement
- Use new isolation or the old one
- From me: TMVA vs MatrixNet

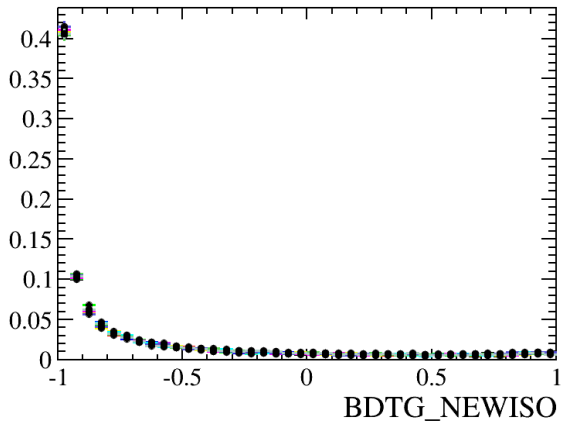
1. Comparison of the performance of each sub-sample from the chopping technique
2. TMVA vs MN performance
3. Conclusion

Fold comparison Background

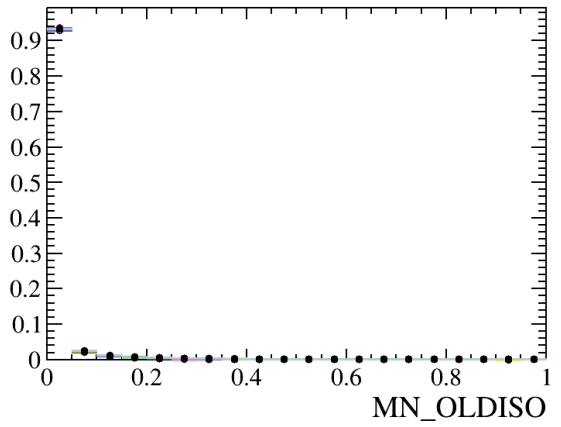
BDT



BDTG



MatrixNet

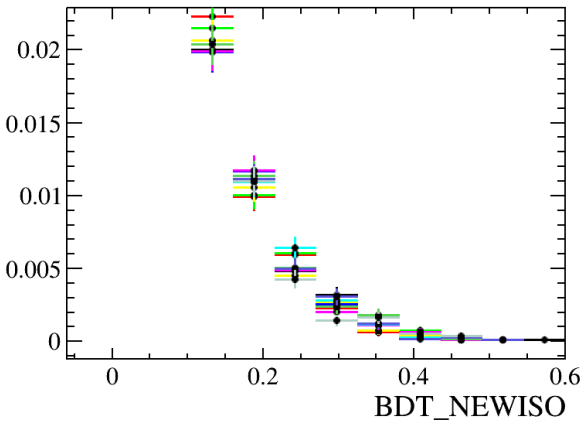


Conclusion

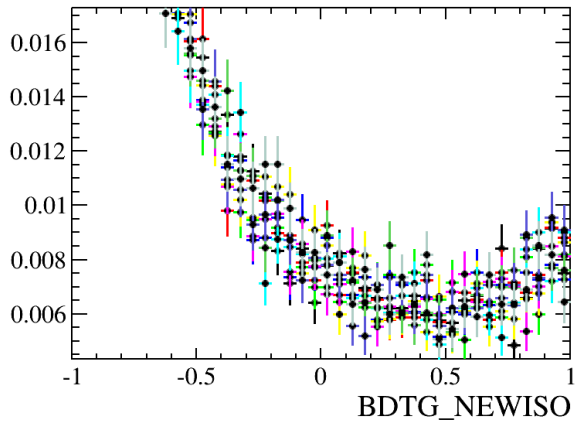
Everything is very consistent.

Fold comparison Background ZOOM

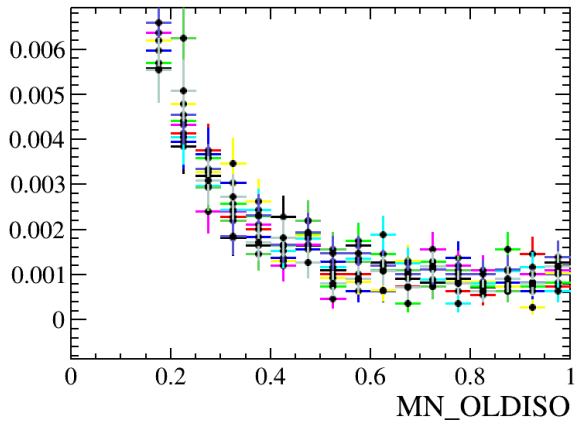
BDT



BDTG



MatrixNet

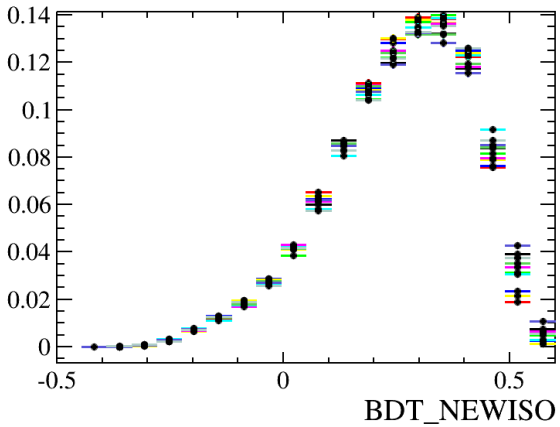


Conclusion

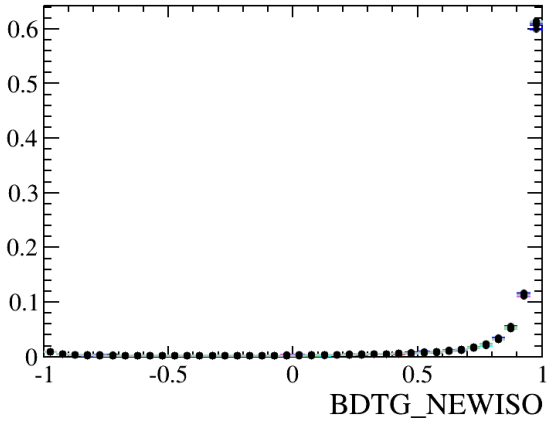
We only see statistical fluctuations within 2σ. What would one expect with 10 folds.

Fold comparison Sploted $K^*\mu\mu$

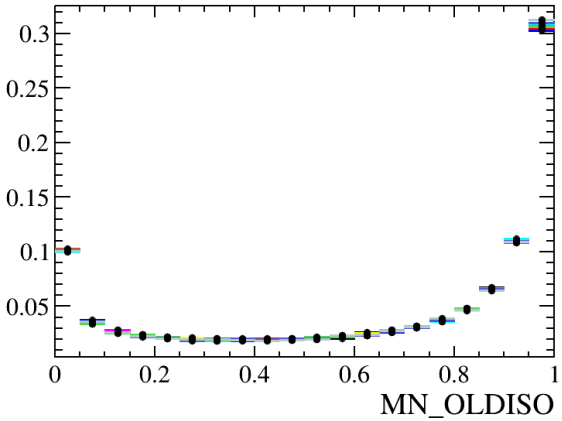
BDT



BDTG



MatrixNet

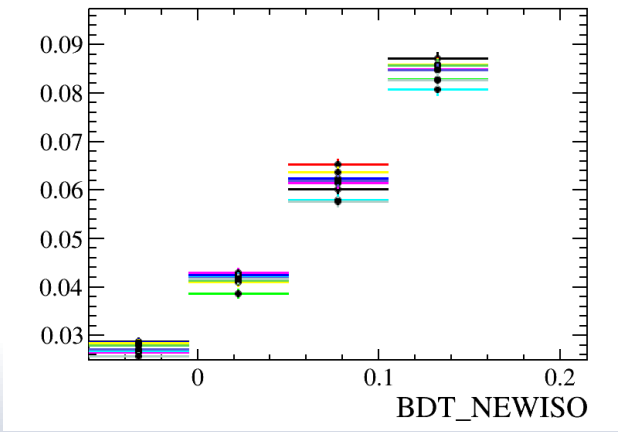


Conclusion

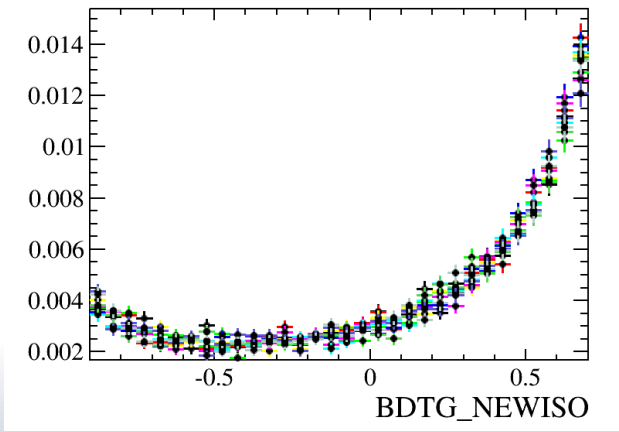
Again everything very consistent.

Fold comparison Sploted K*$\mu\mu$ ZOOM

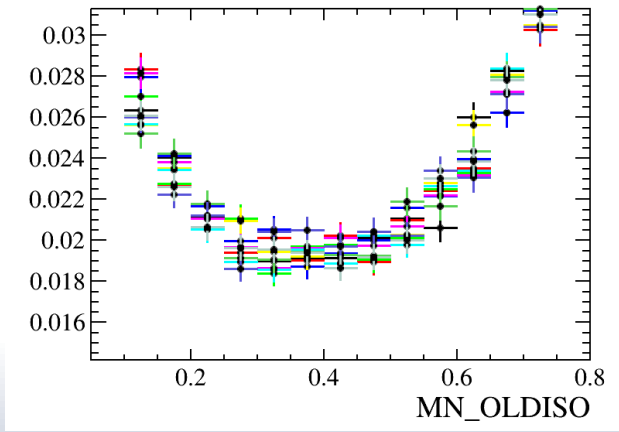
BDT



BDTG



MatrixNet

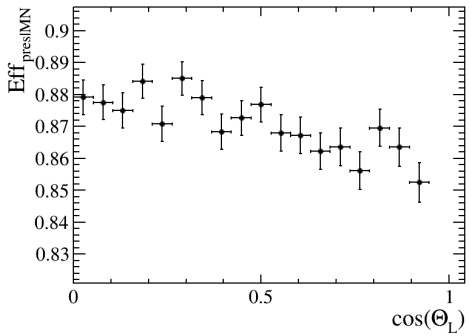


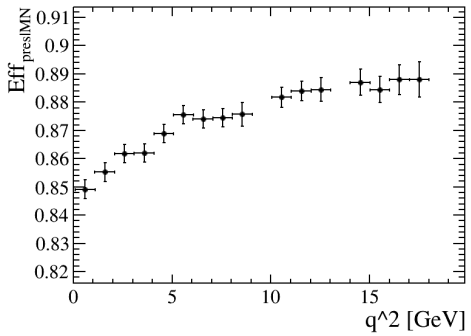
Conclusion

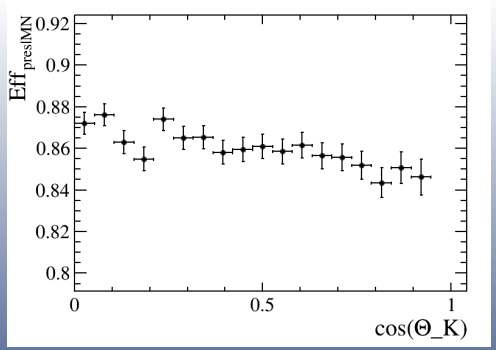
Again everything very consistent with statistical fluctuations.

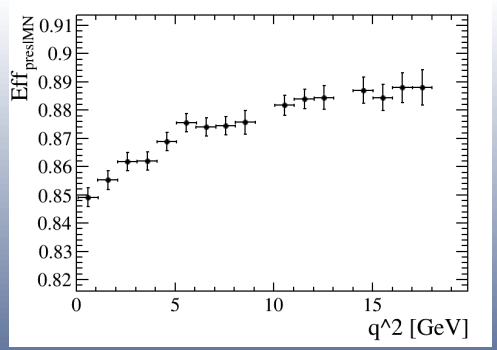
Acceptance

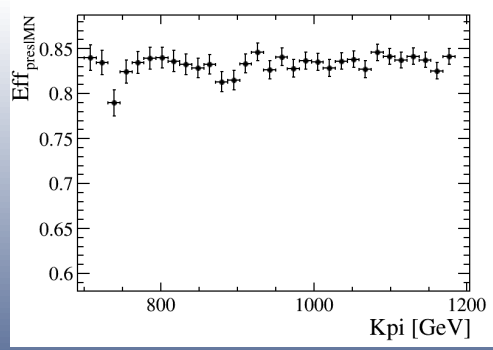
All the classifiers show the same behavior: only the BDT is shown here (others in the backup)











Chopping stability I

I have tested each formula on MC following the procedure:

- Add 10 formulas to MC ntuples.
- For each formula make a cut.
- Calculate the efficiency.

Chopping stability I, BDTG

Fold	Eff.[%]	Err.[%]
0	86.86	0.09
1	87.04	0.09
2	87.02	0.09
3	86.88	0.09
4	86.94	0.09
5	87.02	0.09
6	87.10	0.09
7	86.99	0.09
8	87.14	0.09
9	87.12	0.09

Conclusion

Everything consistent with statistical fluctuations!

Chopping stability II, MatrixNet

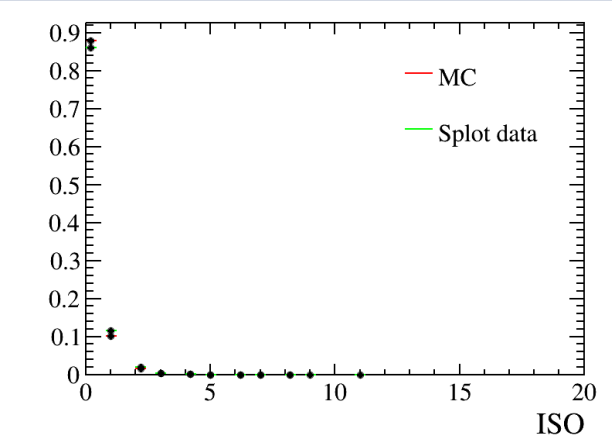
Fold	Eff.[%]	Err.[%]
0	89.56	0.08
1	89.59	0.08
2	89.59	0.08
3	89.60	0.08
4	89.56	0.08
5	89.56	0.08
6	89.60	0.08
7	89.63	0.08
8	89.65	0.08
9	89.55	0.08

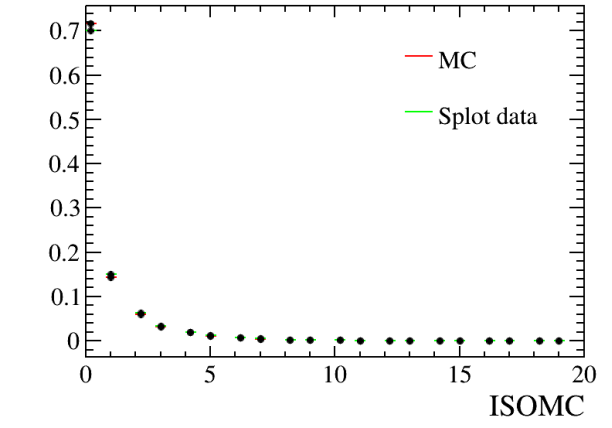
Conclusion

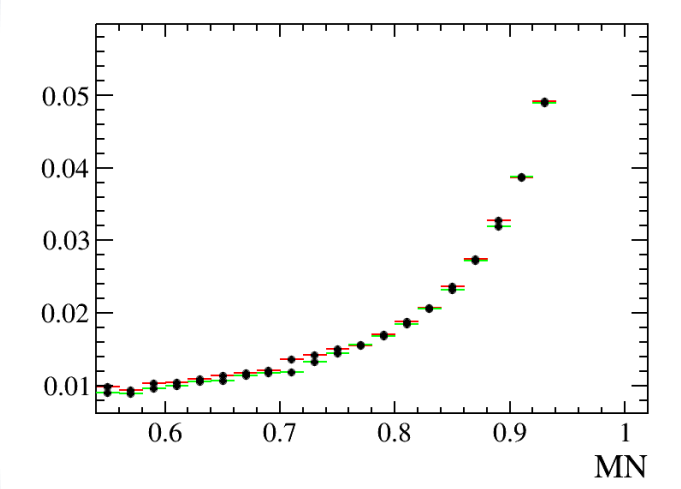
Everything consistent with statistical fluctuations!

MC Data comparison

We performed MC/Data comparison using weights provided by Sam.







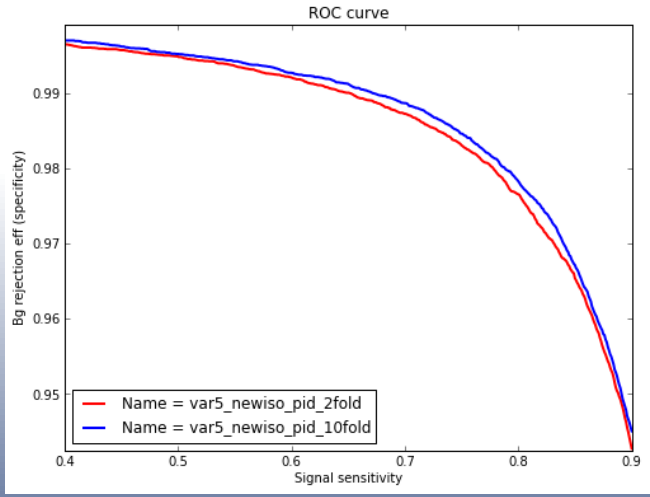
Conclusion

New Isolation in slightly better agreement.
MatrixNet in similar agreement as BDT.

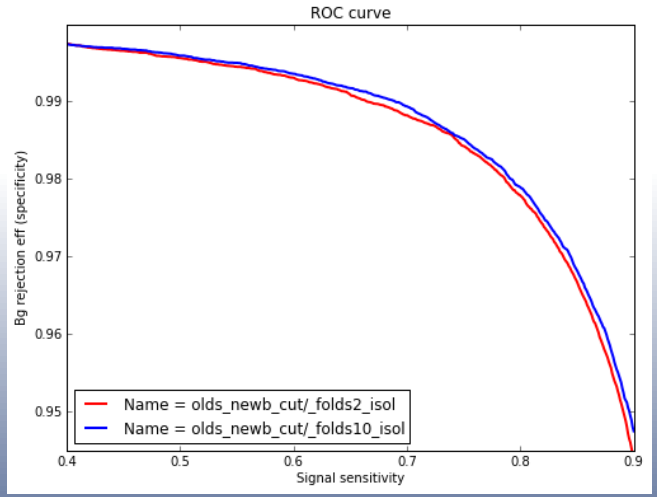
Effect of new K-pi mass window on the chopping I

Due to enlarged K-pi mass window the gain of chopping is reduced (but remember, the chopping helps also in keeping the results more homogeneous)

SMALL K* mass.



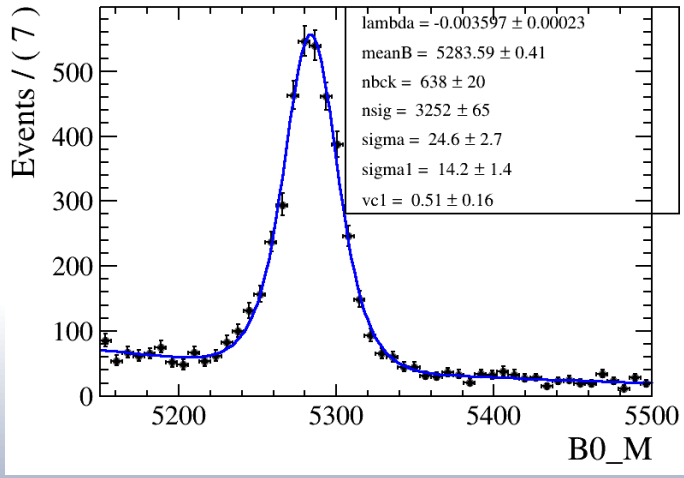
LARGE K* mass.



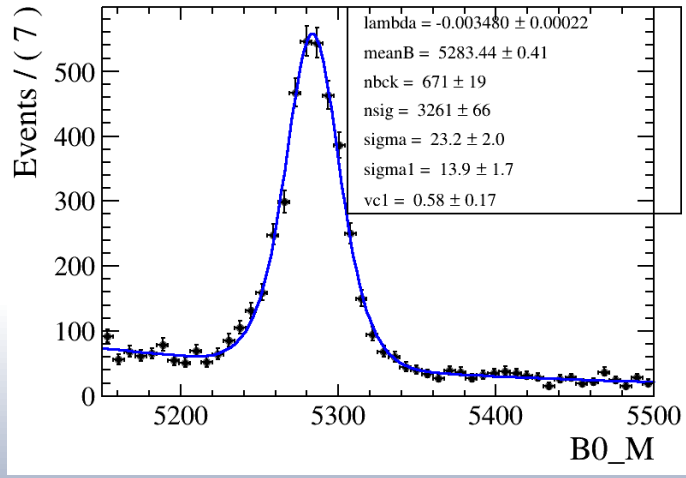
Effect of new K-pi mass window on the chopping II

We disagree with Sam on this issue.
Not easy to drive conclusion on ROC curve. Numerical results from 2 vs 10 Folds training.

10 Folds, Large K^* mass cut.



2 Folds, Large K^* mass cut.



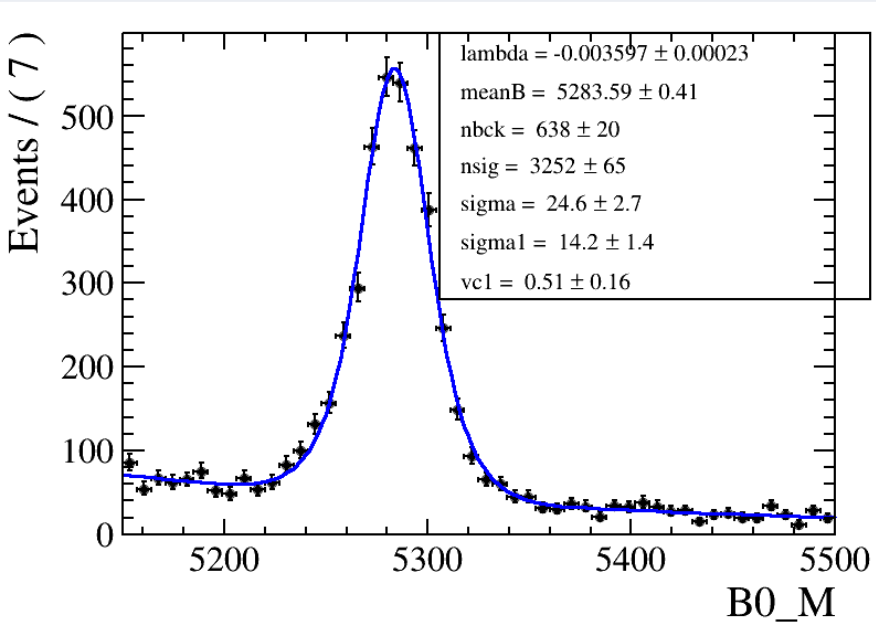
Conclusion

One gaisn 5% background rejection.

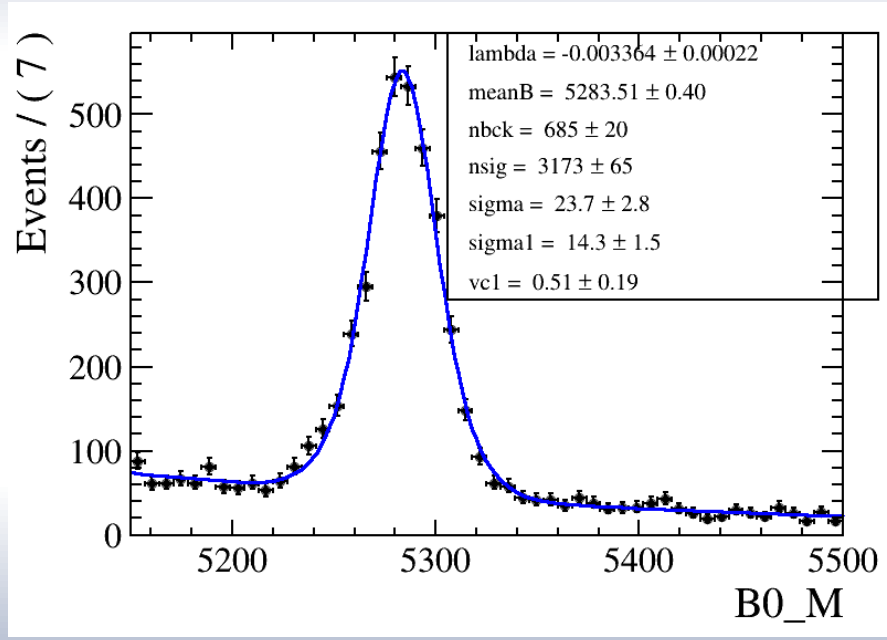
TMVA vs MN

Both classifiers show the same level of correlation in angles. Let's see their performance.

MatrixNet



BDTG



Conclusion

Matrix Net we gain 2.4% in signal and 6.9% of bck rejection.

MatrixNet numbers

q^2	MatrixNet		BDTG	
$[GeV^2]$	Signal	Bck	Signal	Bck
$0.1, 2$	484 ± 24	54 ± 7	465 ± 24	58 ± 7
$2, 4.3$	291 ± 21	84 ± 8	270 ± 20	98 ± 7
$4.3, 8.68$	823 ± 34	221 ± 11	807 ± 34	235 ± 11
$10.09, 12.86$	660 ± 28	138 ± 7	658 ± 28	142 ± 8
$14.18, 16$	481 ± 24	58 ± 5	467 ± 24	66 ± 6
$16, 19$	532 ± 25	60 ± 7	529 ± 25	61 ± 7
$0.1, 19$	3252 ± 65	638 ± 20	3173 ± 65	685 ± 20

Conclusion

1. Agreement in the MVA distribution for different sub-samples from the chopping: chopping helps in keeping results more homogeneous!
2. Effectively we can use one classifier; simplification of the analysis.
3. Slightly better performance of chopping in case of Larger K^* mass window.
4. New isolation is slightly better then the $B_s^0 \to \mu\mu$.
5. MatrixNet performed slightly better then BDT.

From last time (in agreement with present Sam's studies)

- ProbbNN performs better than DLL (from our studies 15% less background)
- NewIso slightly better than $B_s^0 \to \mu\mu$ (but with better Data/MC agreement).

BACKUP

Acceptance MatrixNet

MatrixNet