MatrixNet, TMVA, Isolation

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

November 6, 2013

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

University of Zurich[UZH]

Iso optimisation

Until now every analysis that used track isolation parameter used the ones develeloped and optimised for $B_s \to \mu\mu$. This is based on an abstract definitions of isolating and non-isolating tracks:

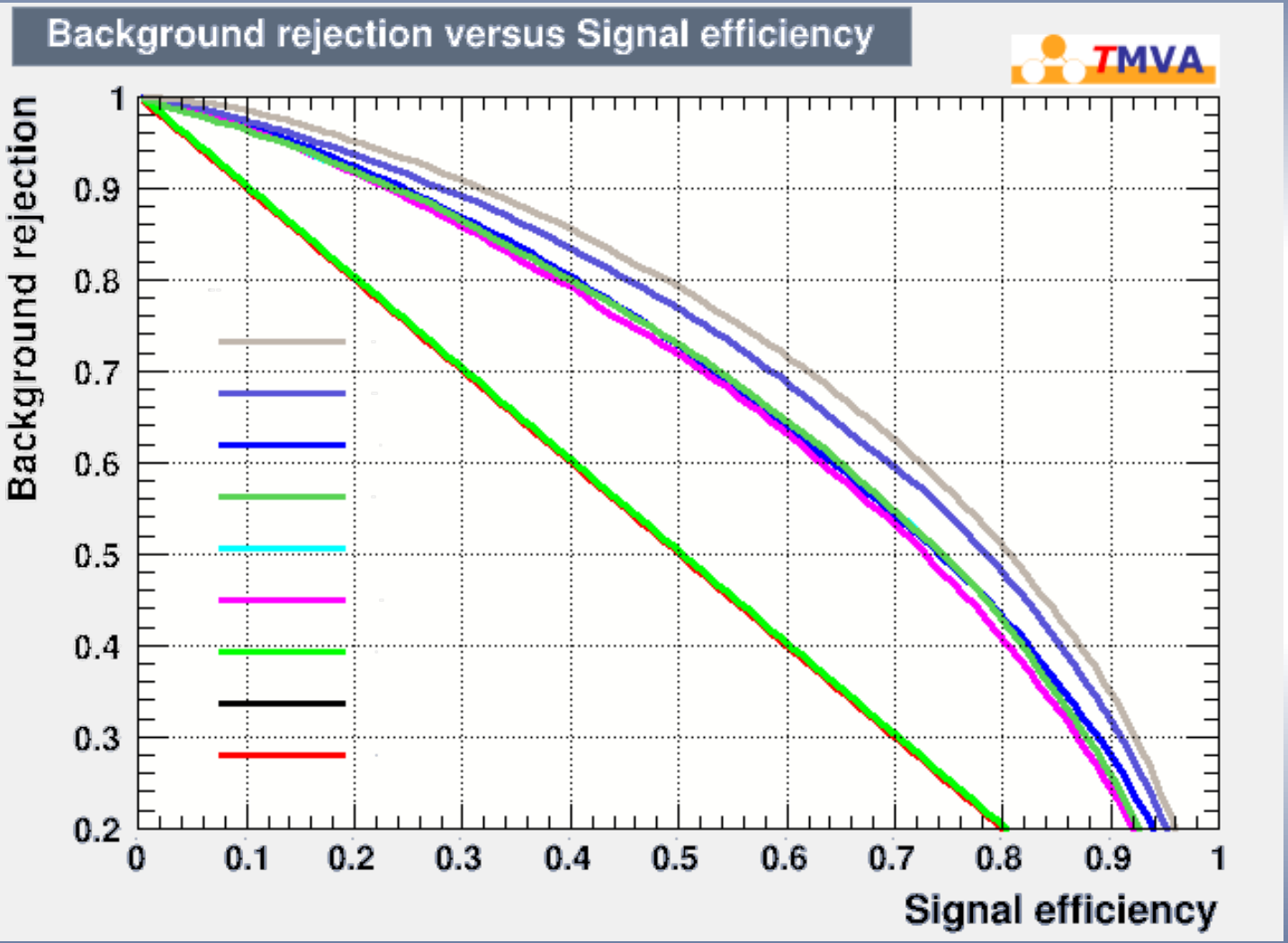
- Non-isolating track to a given track(μ from $B_s \to \mu\mu$ for example) will be a track that has the same primary mother as muon.
- Isolating is the negation of non-isolating.

Iso optimisation

This definition may lead to false implications

- In detector we don't know reconstruct the full event.
- So we don't have the same information that we have on MC
- The idea is to train Signal vs Background.
- Idea developed for $\tau \to 3\mu$

Iso optimisation



Iso optimisation

For each event you build instead:

- Isolation is defined for tracks in the event.
- You have to move from track bases to even bases.
- Usually ideas: take mean, min, etc.
- See which performs better in final MVA

Folding data

There is one basic truth when it comes to MC:

Folding data

There is one basic truth when it comes to MC:

Folding data

There is one basic truth when it comes to MC:

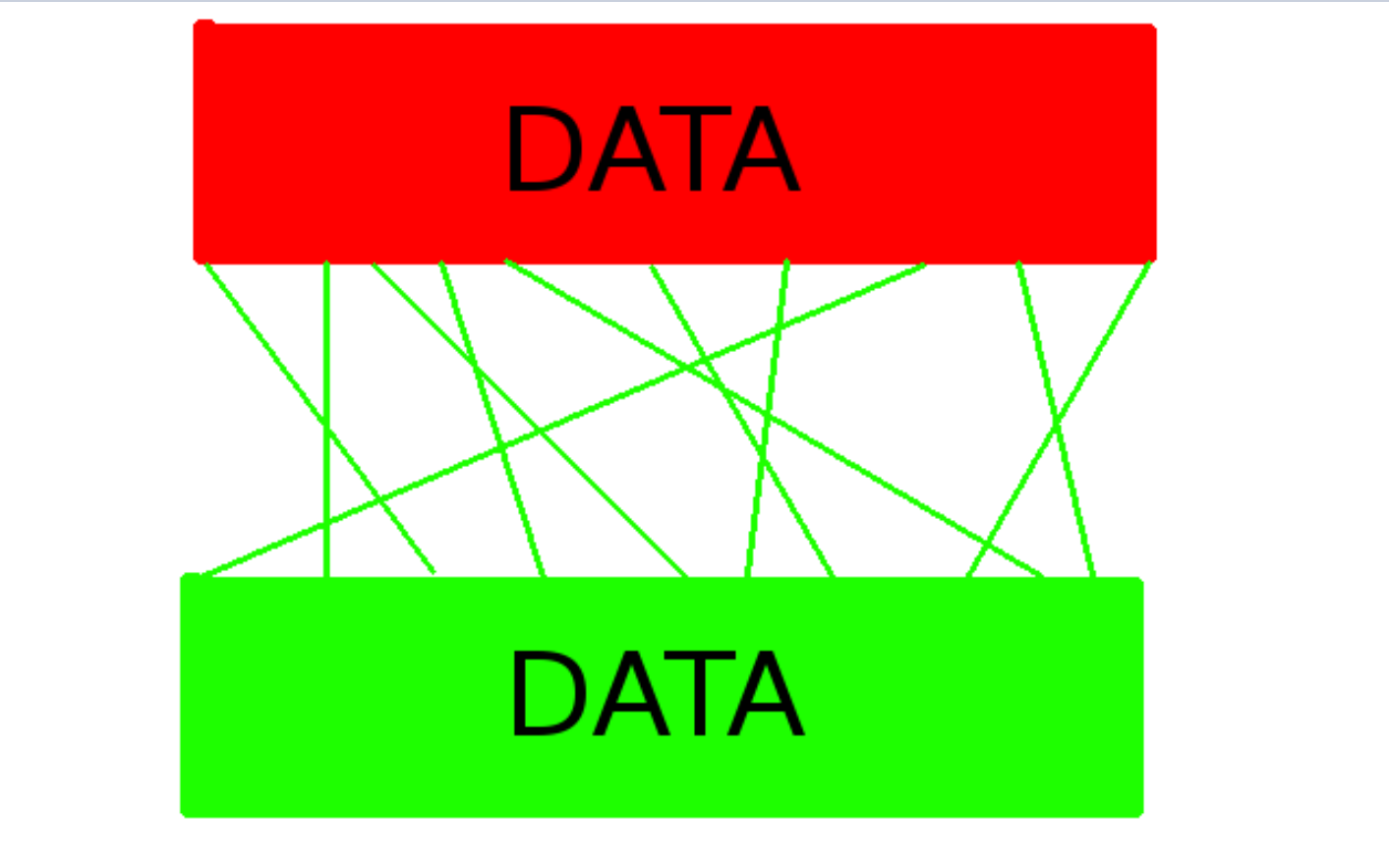
How to get more training set?

Folding Data Set, How to

Take your data Set:

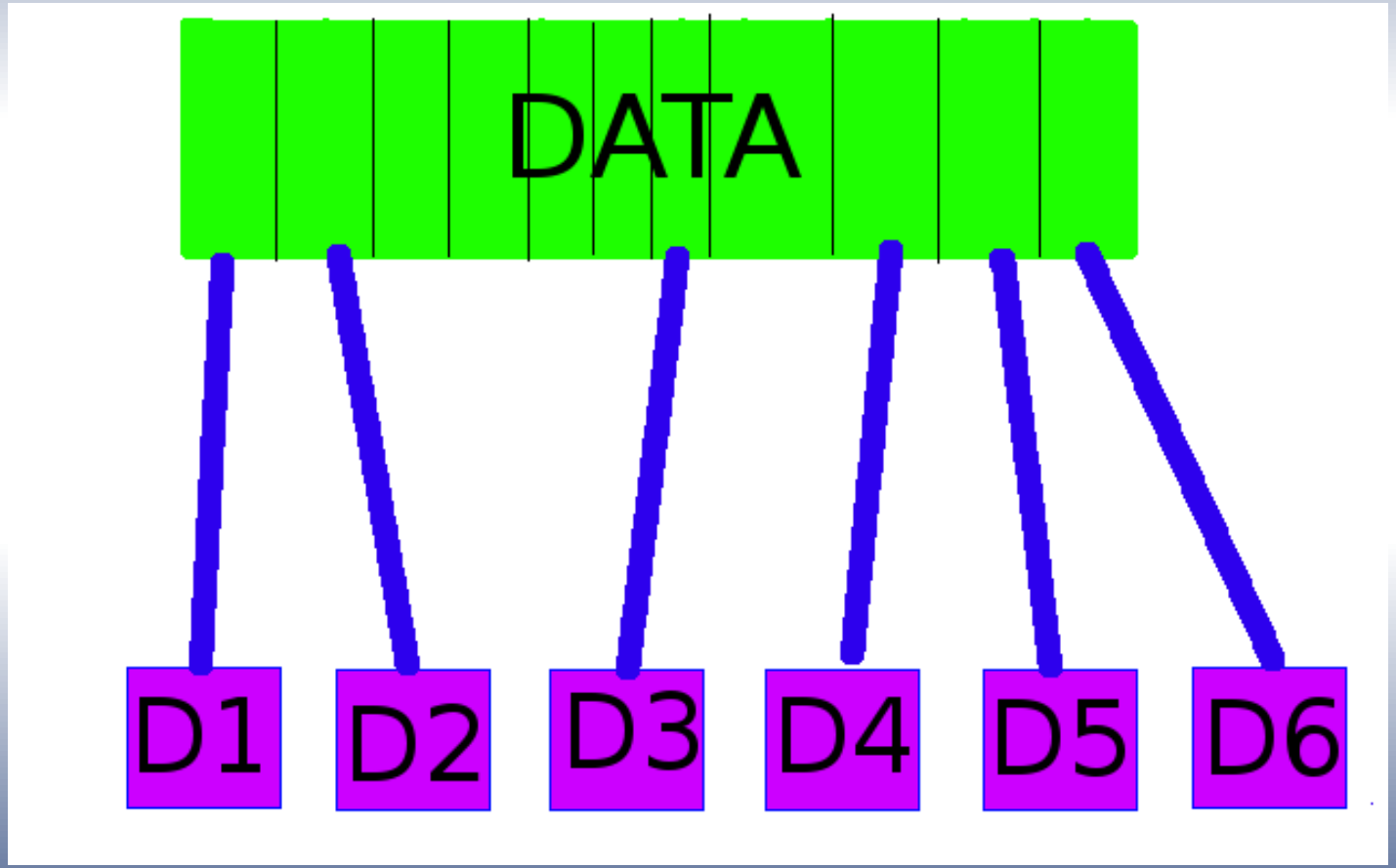
Folding Data Set, How to

Reshuffle it



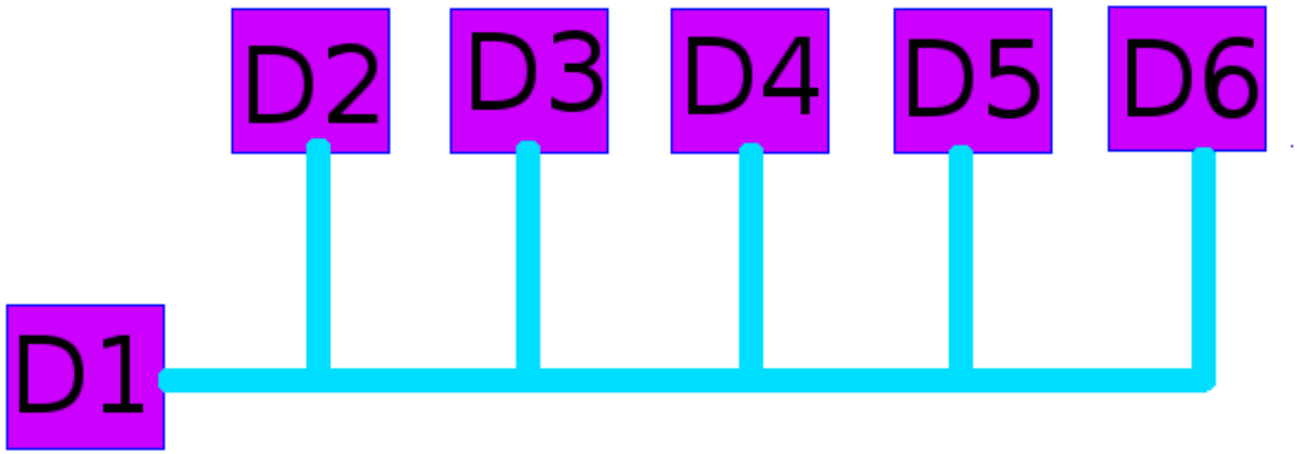
Folding Data Set, How to

Divide it into folds:



Folding Data Set, How to

Train i^{th} fold using $n-1$ folds



Folding Data Set, How to

- Using this simple technique we increase our training sample size :)
- One can better tune your TMVa.
- Procedure is bias free.

Folding results

- For $B \to K^* \mu\mu$ we used 10 folds.
- Train TMVA and MatrixNet.
- See the gain in the yields.

Preselection

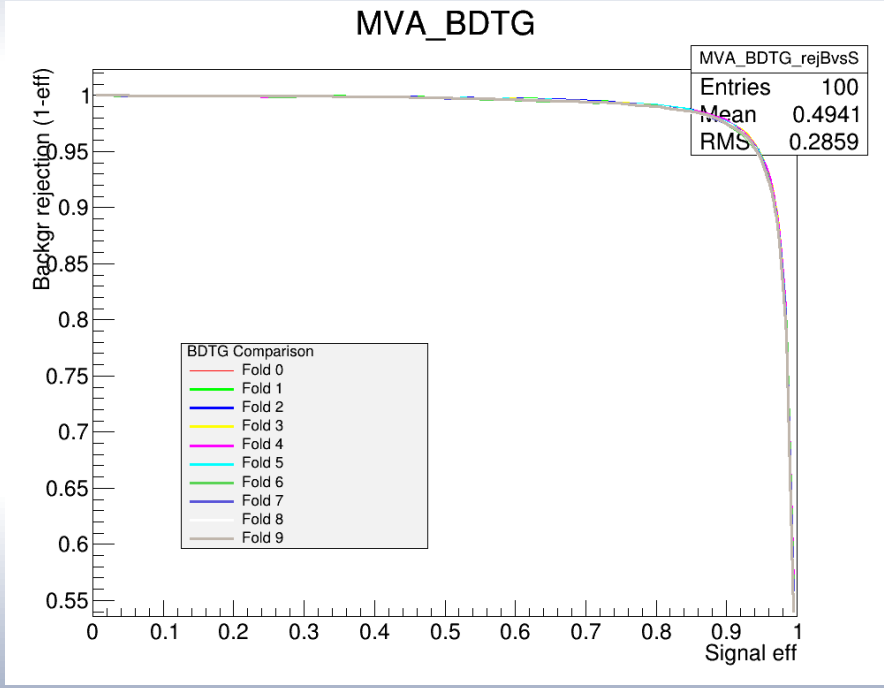
- Apply the following presection cuts:
  1. PID cuts for K, π.
  2. K^* mass cut.
  3. cc vetos.
  4. Trigger as in 2011
  5. Swap:$\pi \longleftrightarrow K$, $\pi \longleftrightarrow \mu$, $K \longleftrightarrow \mu$
  6. ϕ veto
  7. B_s^0 veto.
  8. Λ_b veto.

MatrixNet/TMVa training

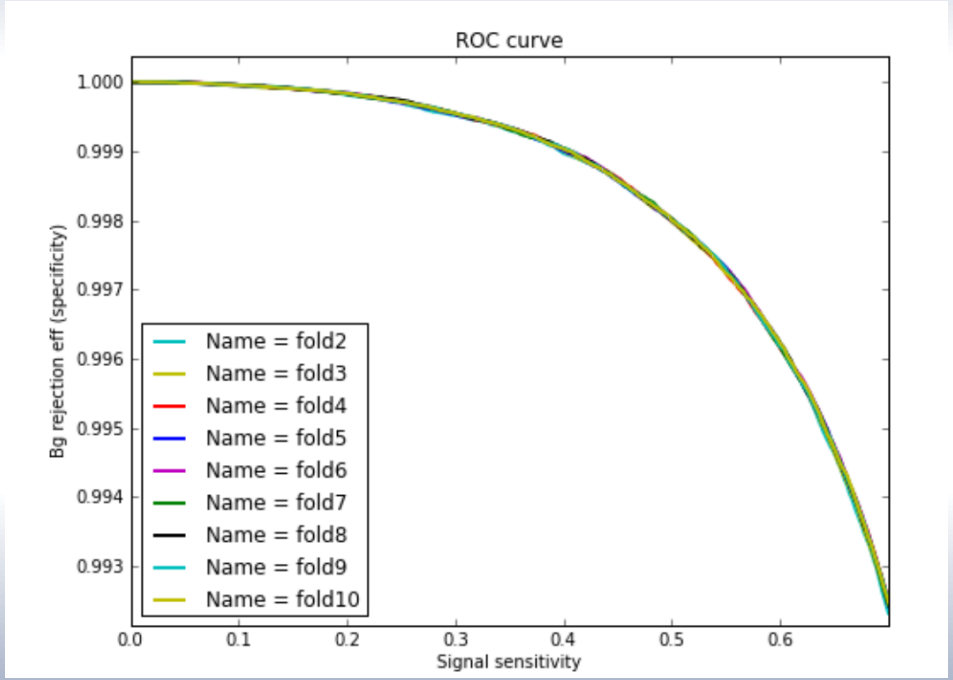
- Put the standard variables inside classifiers:
  1. Isolations
  2. B_DIRA_OWNPV
  3. P, PT
  4. VERTEX_CHI2
  5. LF, LFE, FD
  6. PID

MatrixNet/TMVa ROCs

TMVA



MATRIXNET



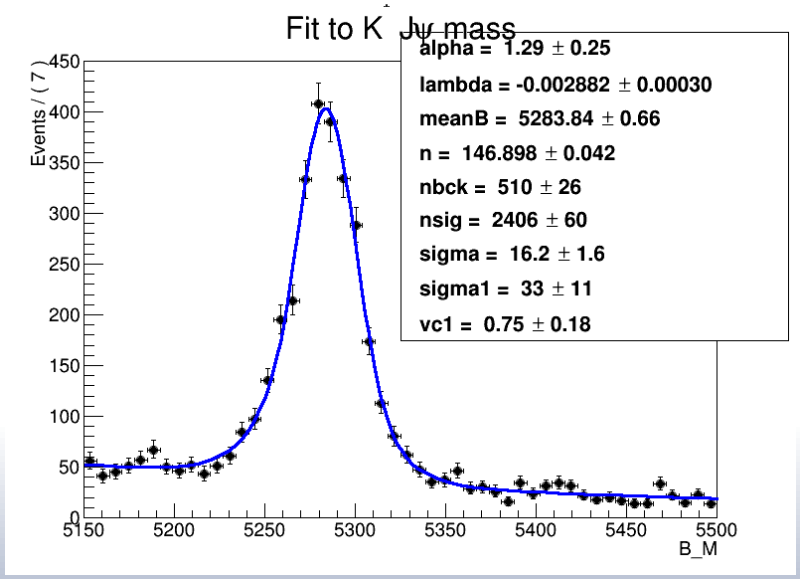
MatrixNet/TMVa ROCs

- All folds have the same ROC curve prediction.
- Good agreement between the folds.

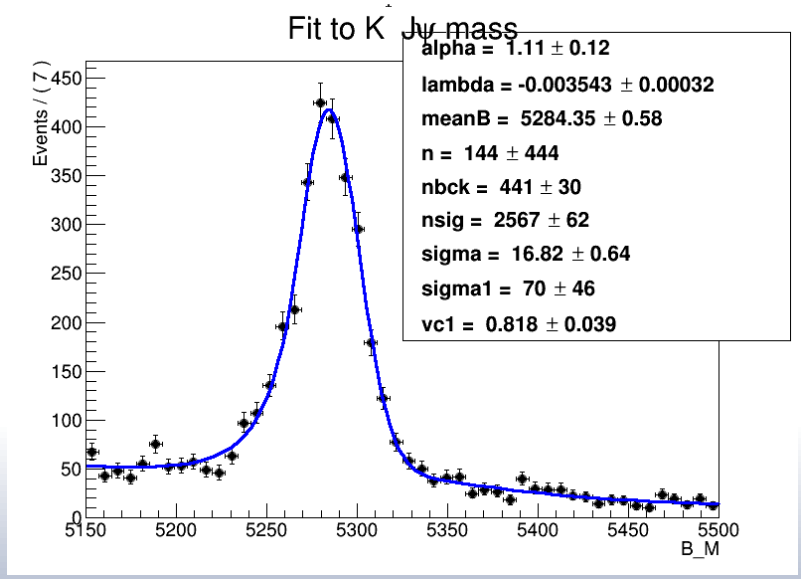
Results

After optimisation of FOM: $\frac{s}{s+b}$ Fit to full q^2 range

TMVA



MATRIXNET



Results

q^2	Last BDT		NEW BDT		MatrixNet	
$[GeV]$	Signal	Bck	Signal	Bck	Signal	Bck
$0.1, 2$	407	58	407	27	405	41
$2, 4.3$	202	95	232	75	233	66
$4.3, 8.68$	573	170	599	180	644	174
$10.09, 12.86$	508	93	515	109	516	115
$14.18, 16$	310	49	322	48	346	39
$16, 19$	359	34	374	32	385	34

Conclusions

1. New MVAs perform better then the old one.
2. Ready for $2fb^{-1}$ data.
3. Have one more trick we want to try :)
4. Once we freeze 2012 we will do the same to 2011 data.

Questions:

1. Is FOM $\frac{s}{s+b}$ really the right one to use?
2. Maybe some toy MC studies?
3. Investigate the angle efficiency on MC.