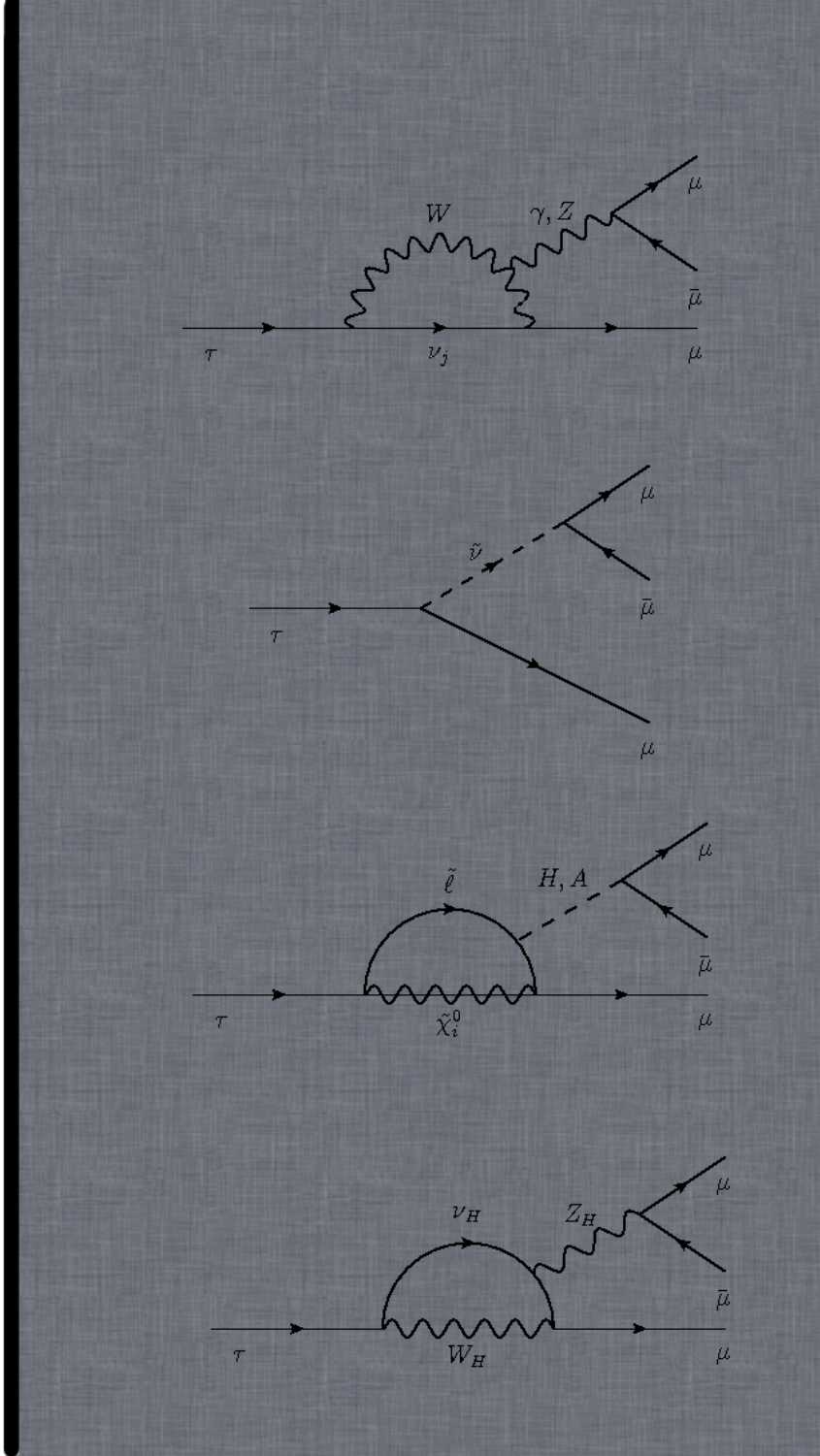
MC, η, TMVA

Marcin Chrząszcz[1,2] , Nicola Serra[1]

[1] University of Zurich , [2] Institute of Nuclear Physics, Krakow,

25th July 2013

University of Zurich[UZH]

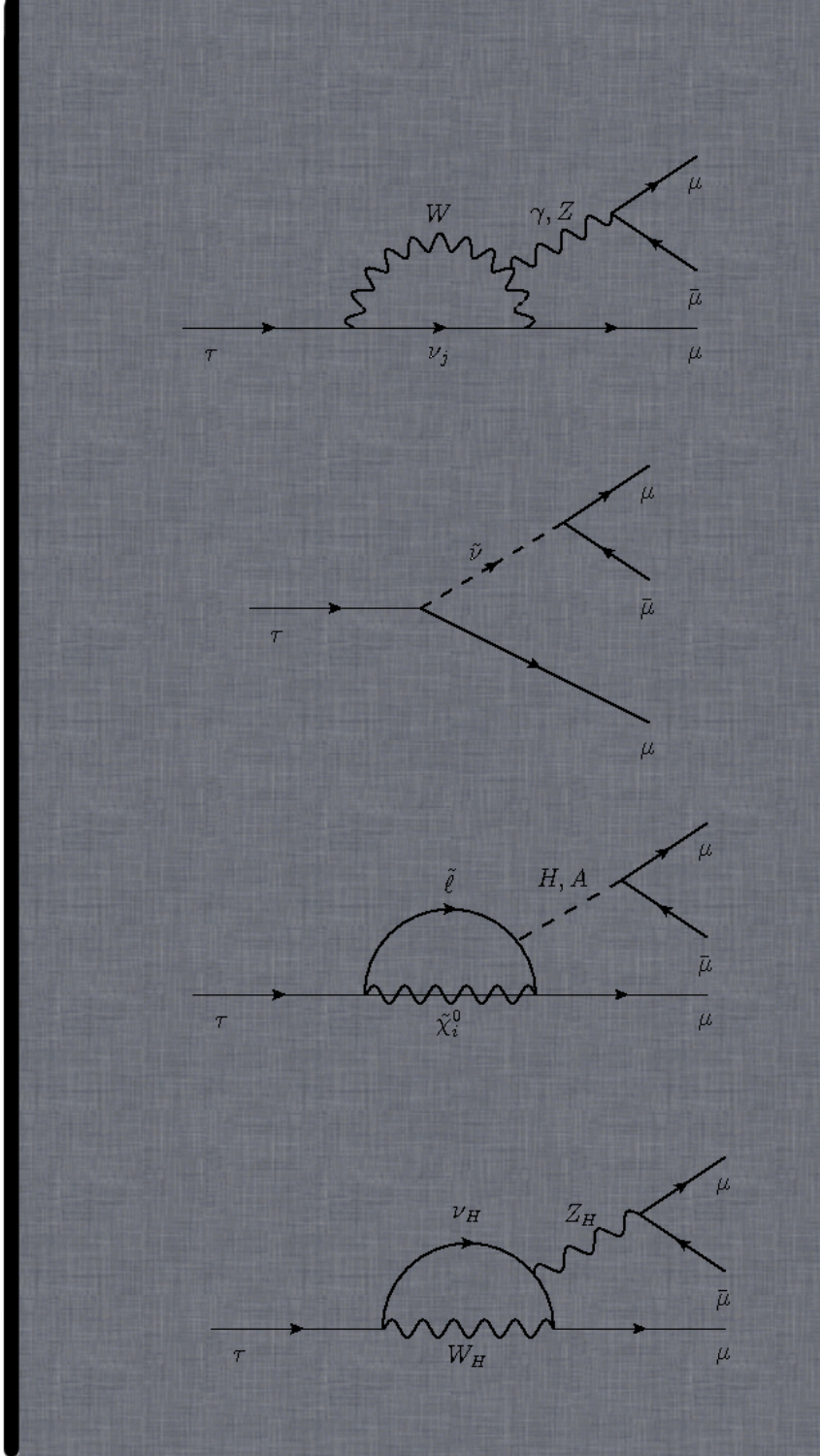


MC studies

η fits

TMVA

Plans for next week



MC Signal

Reminder:

- In 2011 we simulated a mixture of $\tau \rightarrow 3\mu$.
- We found out that the cross section is wrong in MC.
- We reweighed all this distributions to match the correct cross section.
- But what with DPC? This can't be reweighed!
- Let's check how ϵ_{DPC} depends on signal channel.

M.Chrząszcz, N.Serra 2013

Cross check procedure

Let's run Pythia6 with 8 TeV CM energy. With old decfile(aka the wrong mixture of $c\bar{c}$ and $b\bar{b}$. We get:

- $\epsilon_{DPC} = 17.9\%$
- For $7 TeV\%$ we had:17.7%
- This part looks reasonable. We would expected a small gain.

M.Chrząszcz, N.Serra 2013

Cross check procedure

We then simulate two samples for each of 5 sources of τ.

- 1st Sample with Geometry+Daughter[1] Cuts. $\epsilon_{DPC+DAU}$
- 2nd Sample with Daughter Cut. ϵ_{DAU}

[1] Daugher cuts forces τ to come from a specific mother. Ex. B.

M.Chrząszcz, N.Serra 2013

MC Signal

τ source	$\epsilon_{DPC+DAU}[\%]$	$\epsilon_{DAU}[\%]$	$\epsilon_{DPC}[\%]$
$D \to \tau$	12.12 ± 0.07	32.71 ± 0.13	18.5 ± 0.1
$B \to D \to \tau$	1.36 ± 0.01	3.99 ± 0.03	17.0 ± 0.1
$D_s \to \tau$	11.79 ± 0.07	31.53 ± 0.13	18.6 ± 0.1
$B \to D_s \to \tau$	1.75 ± 0.01	5.04 ± 0.03	17.4 ± 0.1
$B \to \tau$	5.16 ± 0.05	14.85 ± 0.13	17.4 ± 0.2

Let's take wrong weights from MC and calculate the ϵ_{DPC}:
$\epsilon_{DPC,WRONG} = 17.86$, with agriment with simulating the wrong mixture from beginning!

M.Chrząszcz, N.Serra 2013

MC Signal

τ source	$\epsilon_{DPC+DAU}[\%]$	$\epsilon_{DAU}[\%]$	$\epsilon_{DPC}[\%]$
$D \to \tau$	12.12 ± 0.07	32.71 ± 0.13	18.5 ± 0.1
$B \to D \to \tau$	1.36 ± 0.01	3.99 ± 0.03	17.0 ± 0.1
$D_s \to \tau$	11.79 ± 0.07	31.53 ± 0.13	18.6 ± 0.1
$B \to D_s \to \tau$	1.75 ± 0.01	5.04 ± 0.03	17.4 ± 0.1
$B \to \tau$	5.16 ± 0.05	14.85 ± 0.13	17.4 ± 0.2

Let's take wrong weights from MC and calculate the ϵ_{DPC}:
$\epsilon_{DPC,WRONG} = 17.86\%$, with agriment with simulating the wrong mixture from beginning!
If we take the correct weights we obtain:
$\epsilon_{DPC,CORRECT} = 18.60\%$. We underestimated our efficiency!

M.Chrząszcz, N.Serra 2013

MC Signal

τ source	$\epsilon_{DPC+DAU}[\%]$	$\epsilon_{DAU}[\%]$	$\epsilon_{DPC}[\%]$
$D \to \tau$	12.12 ± 0.07	32.71 ± 0.13	18.5 ± 0.1
$B \to D \to \tau$	1.36 ± 0.01	3.99 ± 0.03	17.0 ± 0.1
$D_s \to \tau$	11.79 ± 0.07	31.53 ± 0.13	18.6 ± 0.1
$B \to D_s \to \tau$	1.75 ± 0.01	5.04 ± 0.03	17.4 ± 0.1
$B \to \tau$	5.16 ± 0.05	14.85 ± 0.13	17.4 ± 0.2

How ever the overall effect will be smaller cuz the same thing will happen for the normalization channel.

M.Chrząszcz, N.Serra 2013

Pythia Wars

I have found an other disturbing thing. Lets compare pythia 6 with pythia8:

	$\epsilon_{DPC}[\%]$
Pythia 6	17.9
Pythia 8	19.1

This looks worse than it is. Jon checked and this happens not only to $\tau \to 3\mu$. Turn out this is common. $B_s \to \mu\mu$ aslo has the same problem. However thanks to normalization this the ratio of efficiencies changes by 0.1%.
We are safe anyway.

M.Chrząszcz, N.Serra 2013

η fits

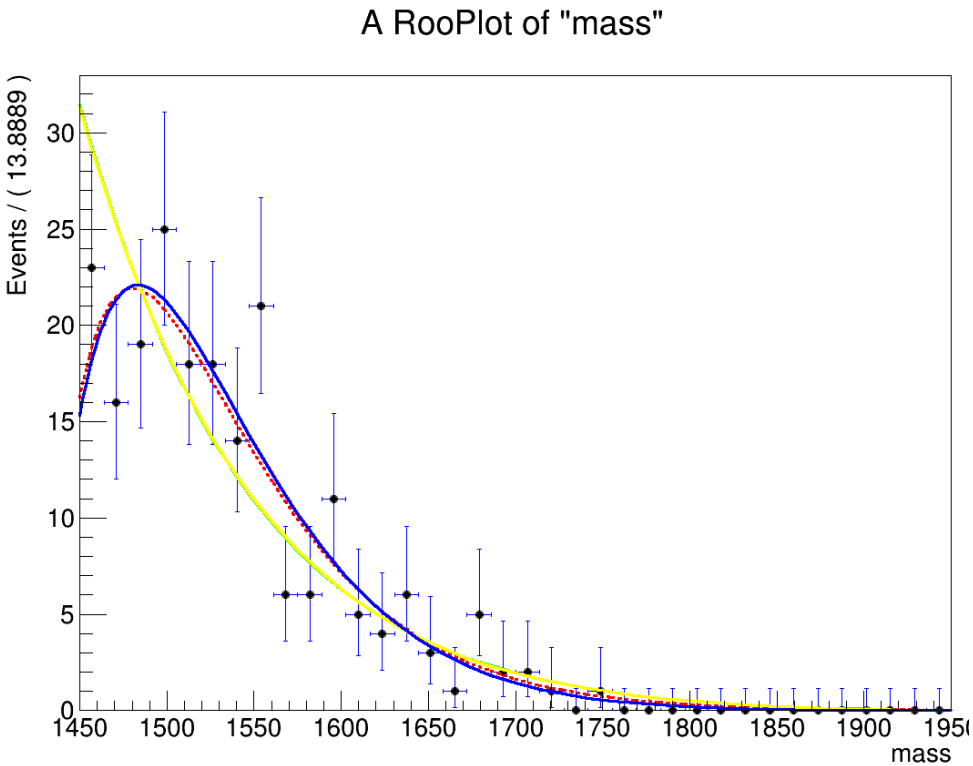
- Till yesterday we took η for fitting directly from MC.
- But how much eta is there?
- We might have combinatorial background with partially reconstructed η.
- Lots of thanks to Paul for speedy implementation of this idea!
- To increase the sensitivity I took left mass range larger! Make the fit more stable.

M.Chrząszcz, N.Serra 2013

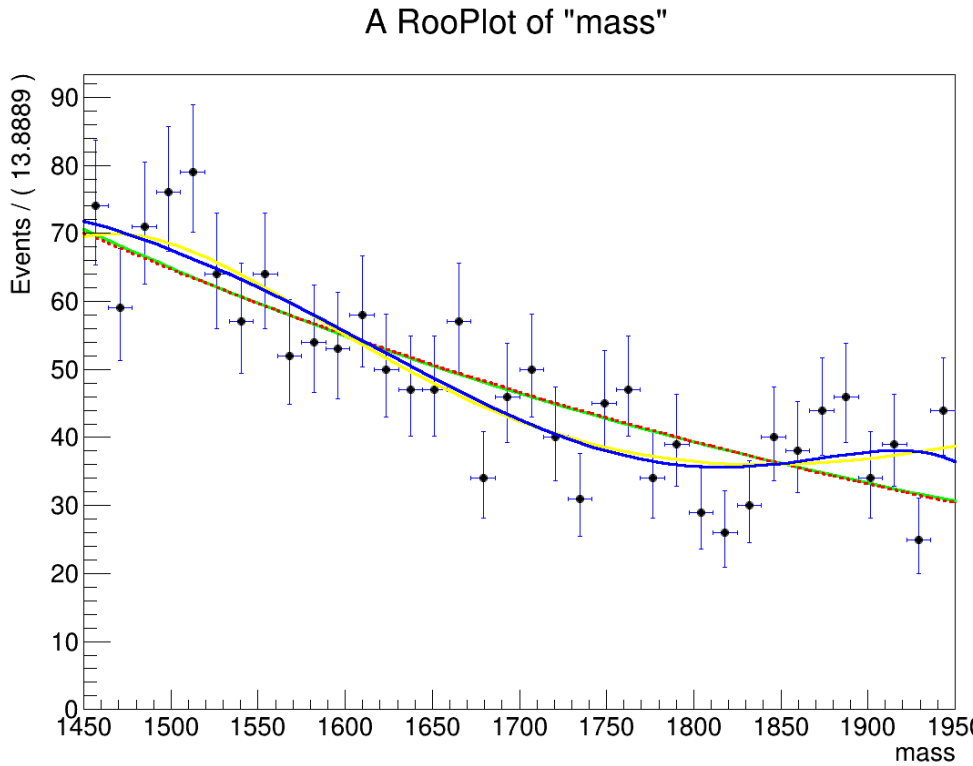
η fits

Extreme case: Trash bins

Only η



η with combinatorics.

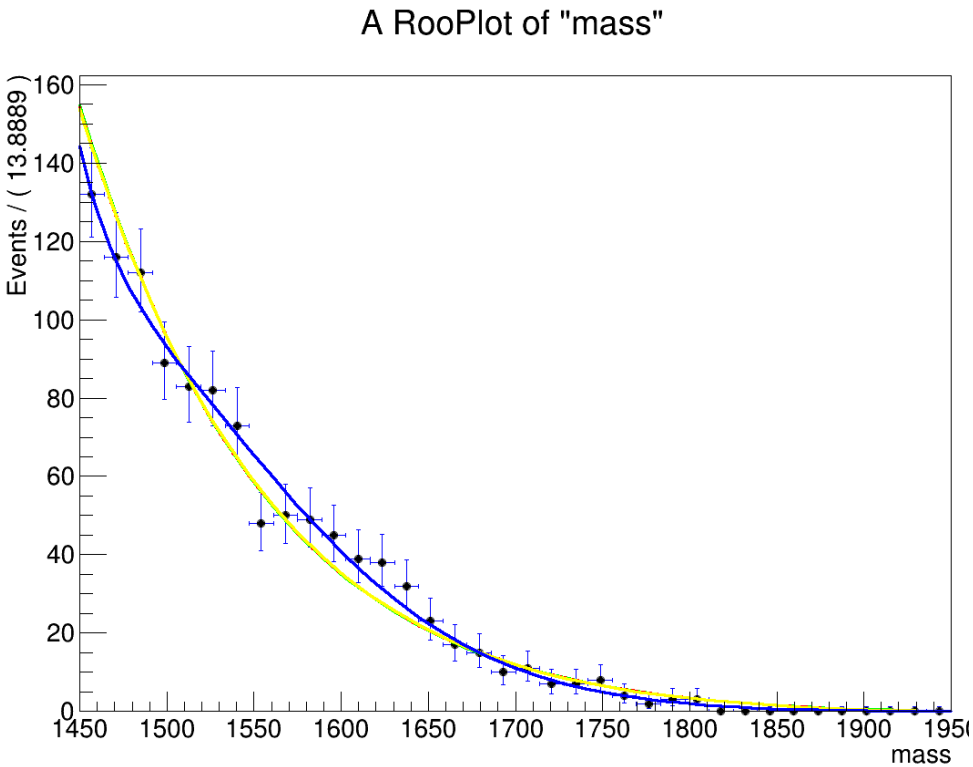


M.Chrząszcz, N.Serra 2013

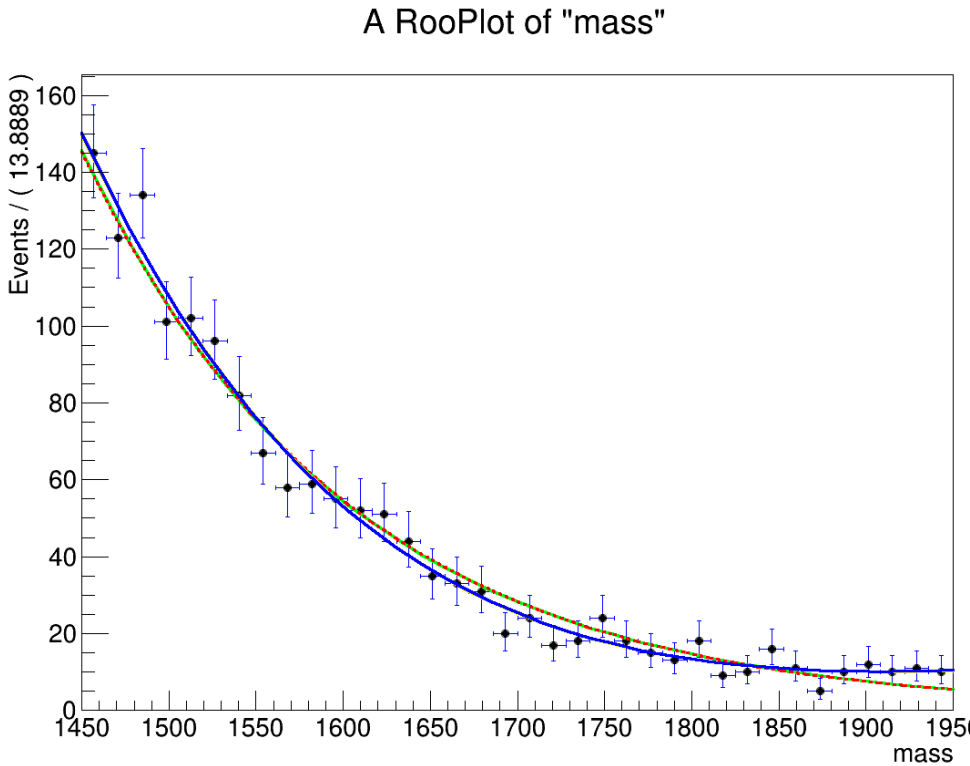
η fits

Not only the trash bin is affected: pid $0.725 - 0.86$
geo: $-0.48 - 0.05$

Only η



η with combinatorics.

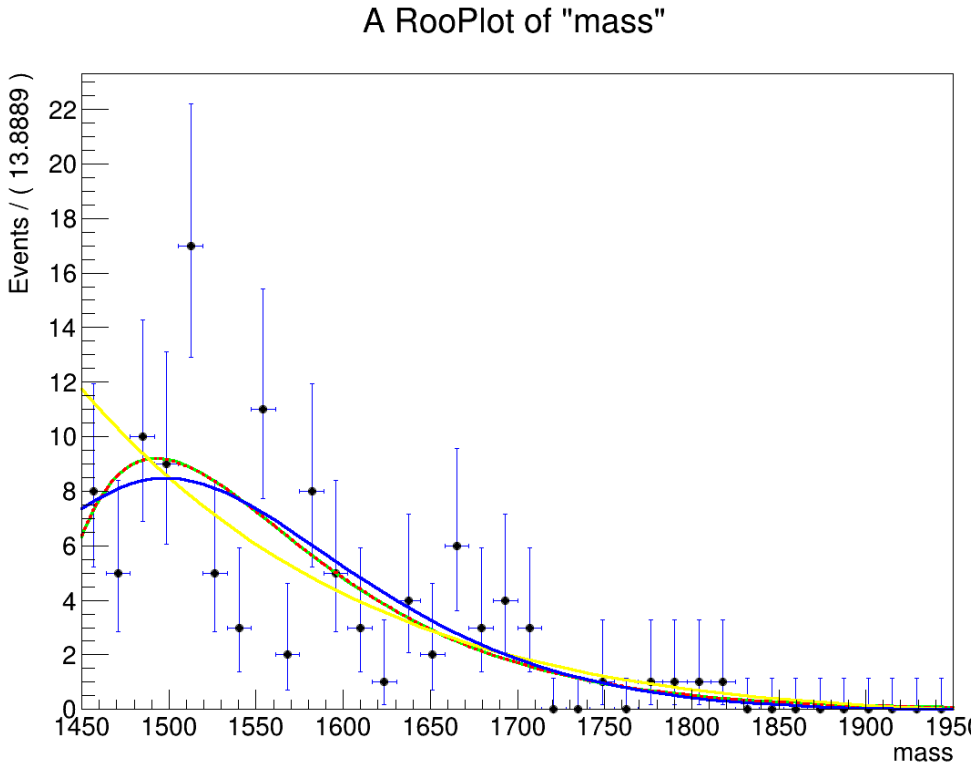


M.Chrząszcz, N.Serra 2013

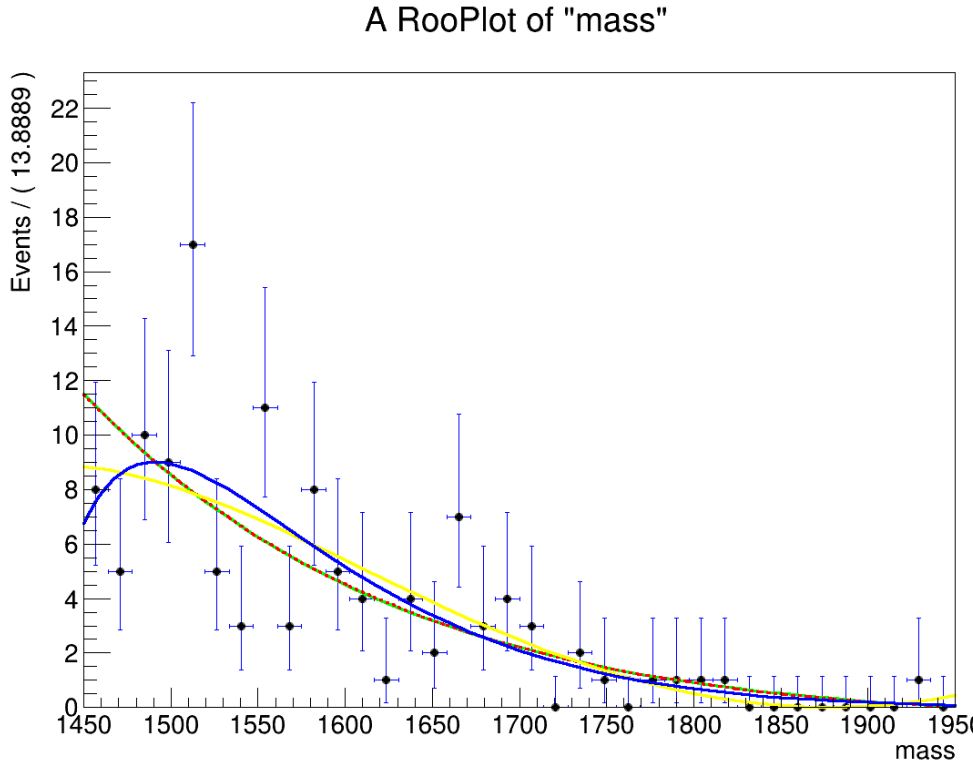
η fits

As old Chinese wisdom says: "One event can make a difference"
Not only the trash bin is affected: pid $0.6 - 0.65$
geo: $0.65 - 0.74$

Only η



η with combinatorics.



M.Chrząszcz, N.Serra 2013

Conclusions on η

- 23% of events in the ntuple are background.
- Much better shape of η.
- PDF similar in each bin!
- Much smaller linkage of η to mass window!
- PDFs are ready for fitting with 2012 data!

M.Chrząszcz, N.Serra 2013

Introduction

Kaggle (leading machine learning competition platform).

If you notice how people win this competition; you'll notice that sometimes people combine two or more algorithm into ensemble and get better results.
This is called blending.
Isn't $\tau \rightarrow 3\mu$ perfect environment to play?

M.Chrząszcz, N.Serra 2013

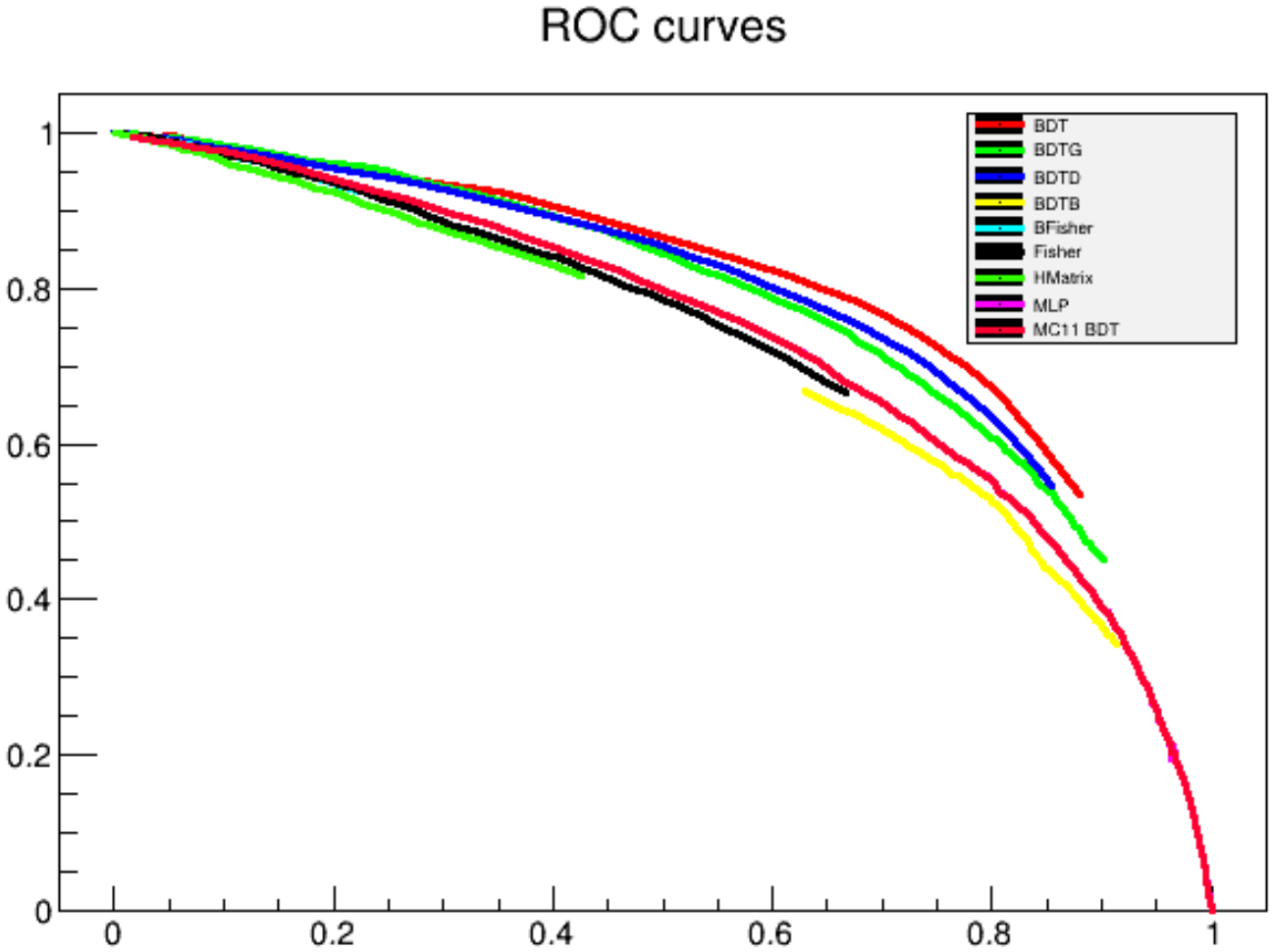
First attempts

- Let's take our background produced so far.
- Already a comparable sample to 2011! Generator cuts are doing their job.
- Let's train each signal on separate source of τ.

M.Chrząszcz, N.Serra 2013

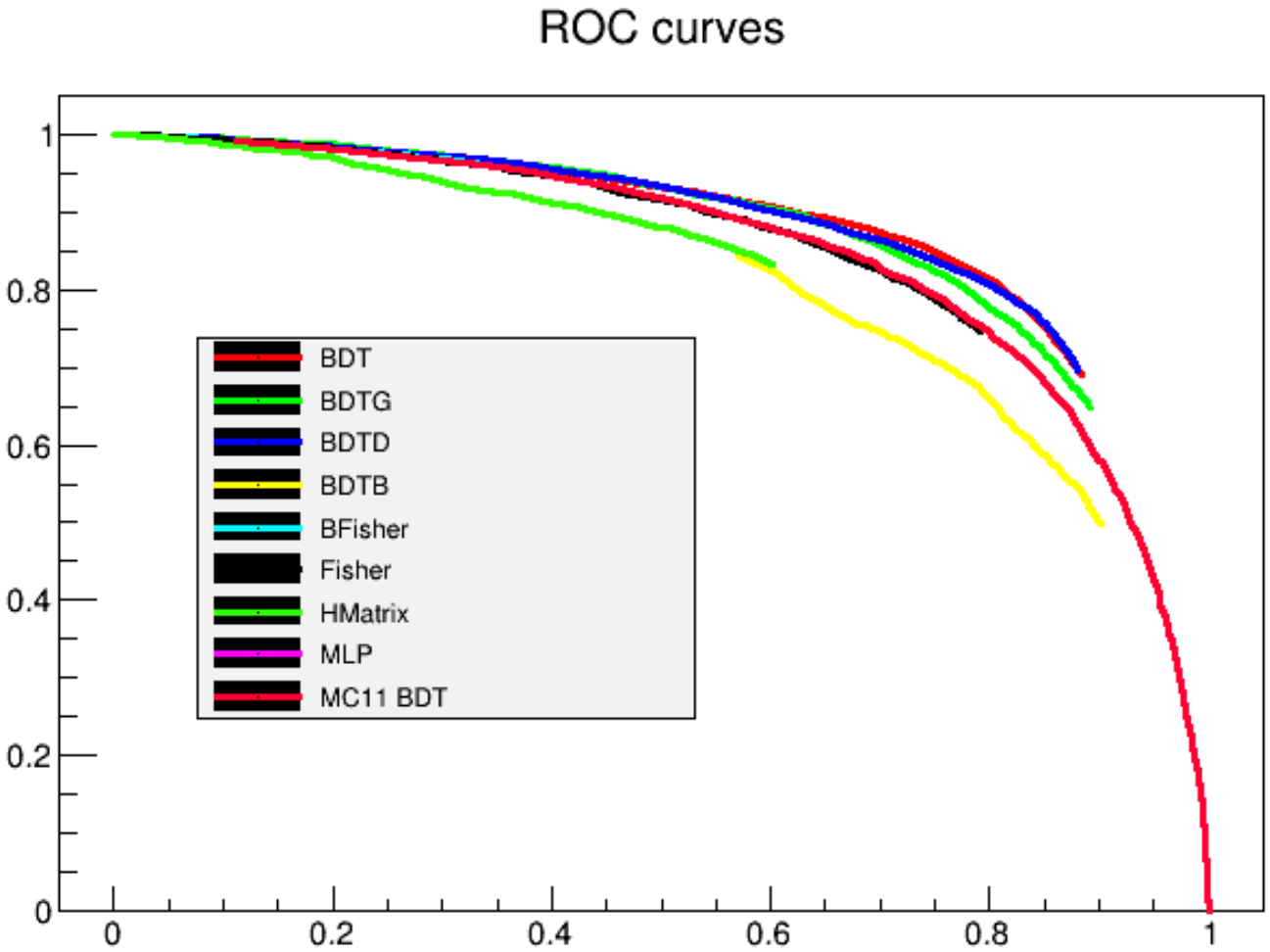
$B \to \tau$

We really suck in selecting this channel.



M.Chrząszcz, N.Serra 2013

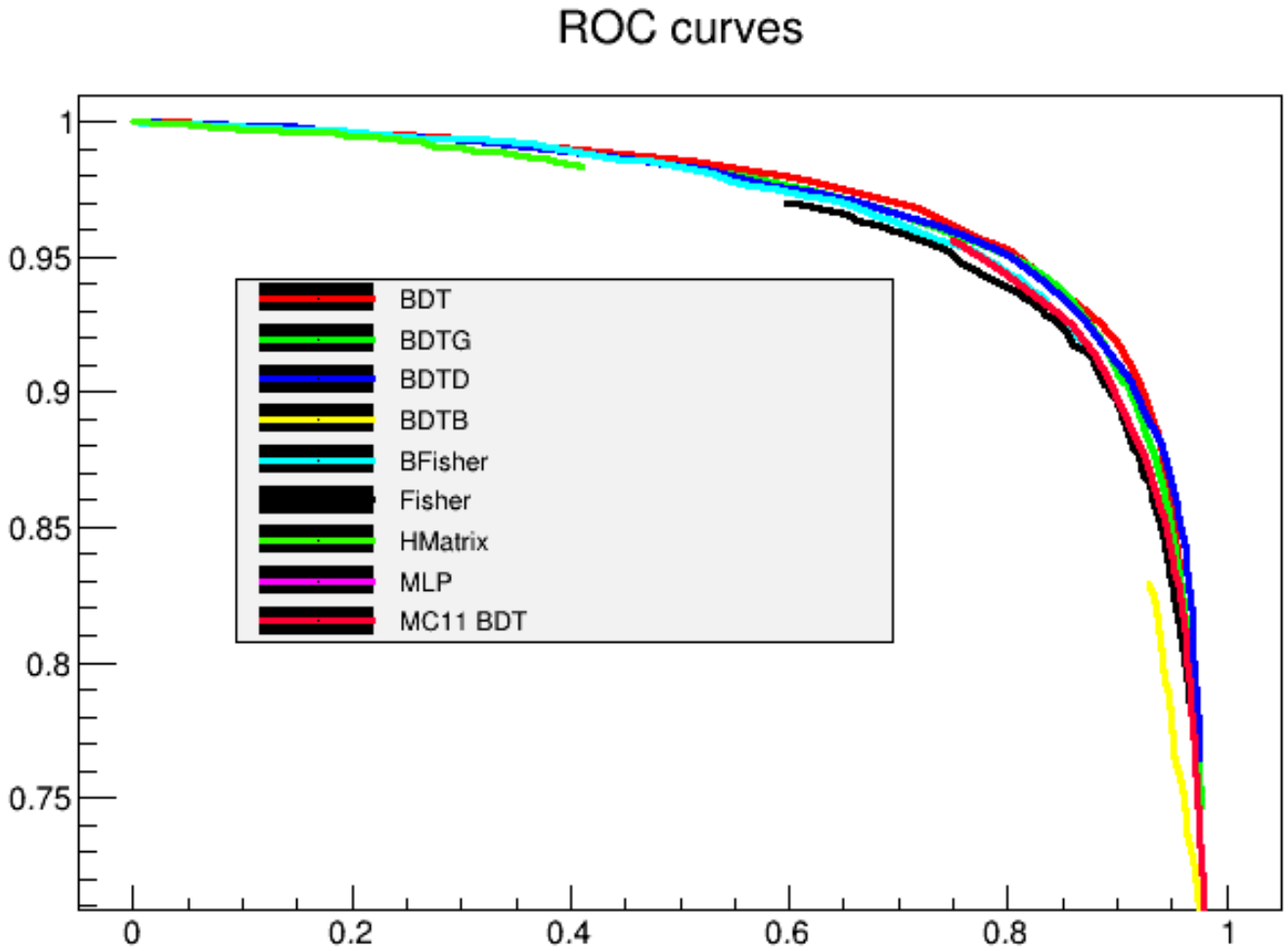
On the biggest contributing channel we are quite optimal.



M.Chrząszcz, N.Serra 2013

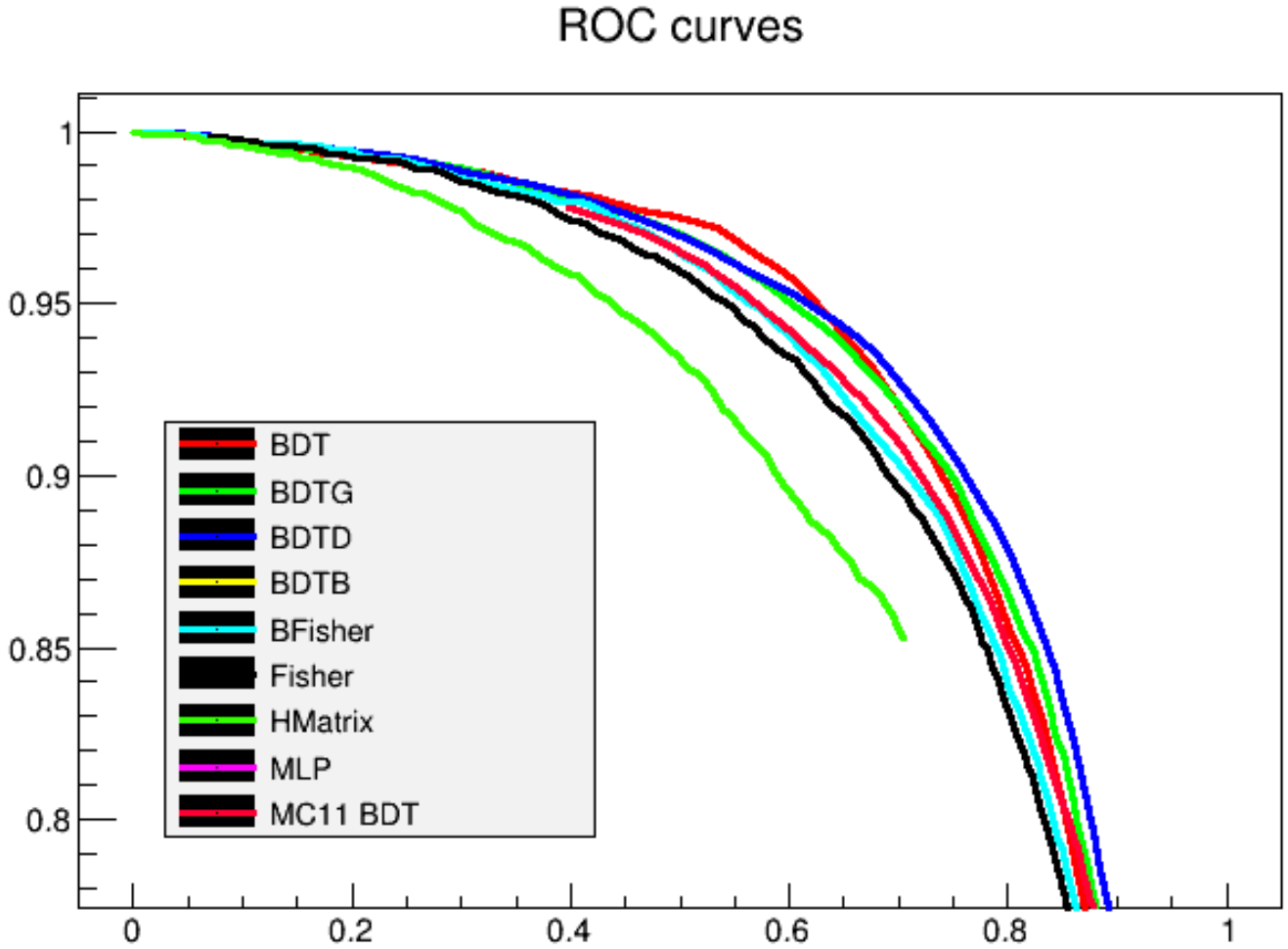
$D_s \to \tau$

On the biggest contributing channel we are quite optimal.



M.Chrząszcz, N.Serra 2013

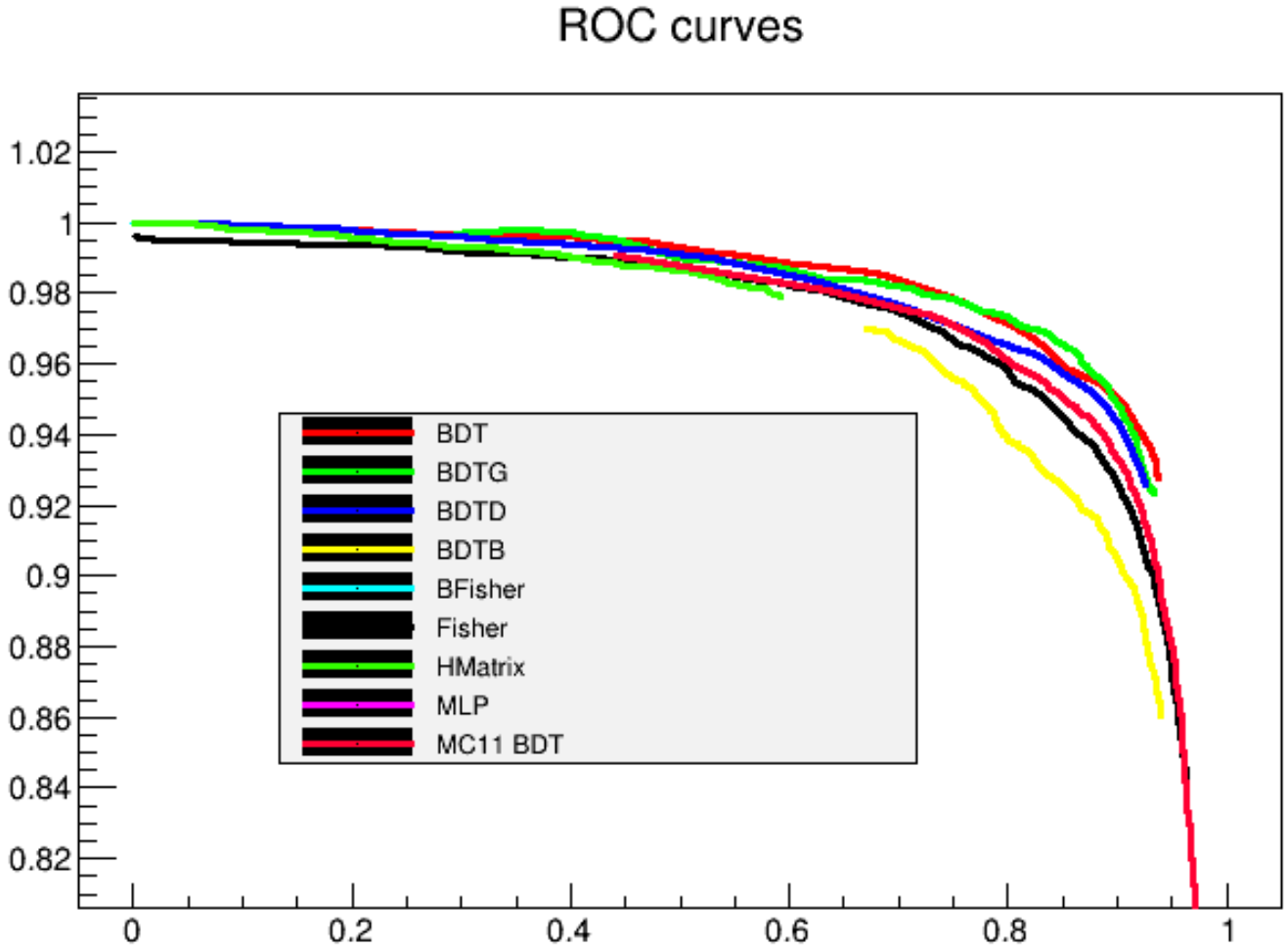
On the biggest contributing channel we are quite optimal.



M.Chrząszcz, N.Serra 2013

$D^+ \to \tau$

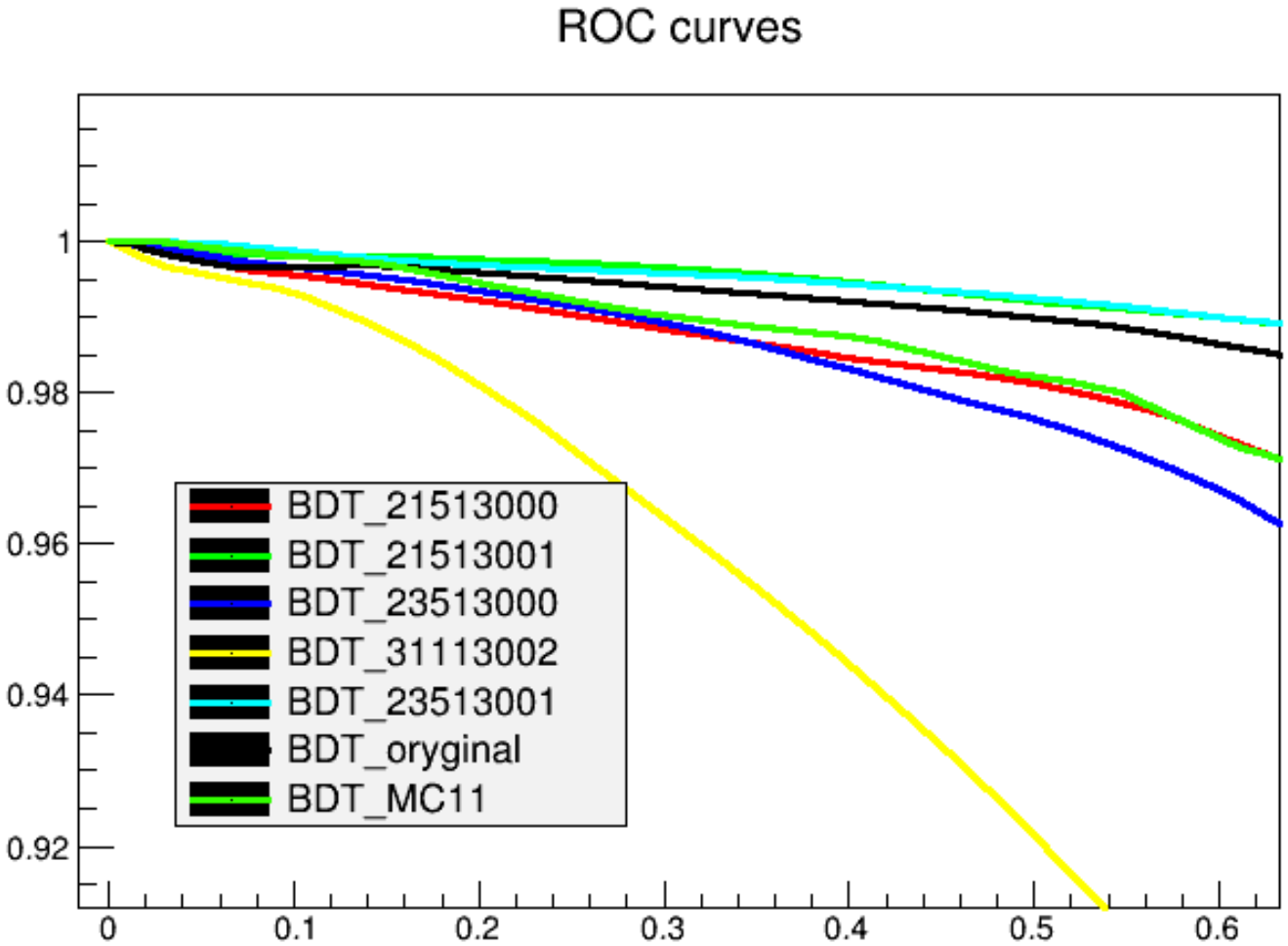
On the biggest contributing channel we are quite optimal.



M.Chrząszcz, N.Serra 2013

Comparison on mix sample

On the biggest contributing channel we are quite optimal.



M.Chrząszcz, N.Serra 2013

Conclusions on TMVA

- Each of the signal components is enormously larger than MVA trained on mix.
- Method looks very promising if we can find a nice blending method(work for next week).
- Mayby discusion on TMVA/MatrixNet/Neurobayes is next to leading order effect compared to this method?

M.Chrząszcz, N.Serra 2013

Conclusions on TMVA

- Finish producing cc bck
- Continue blending.
- Finish calculating new 2D binning optimisation(last night it was still calculating).
- Start Normalizing the η
- Produce Normalization channel MC.

M.Chrząszcz, N.Serra 2013