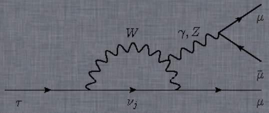
$\tau \to 3\mu$ Status Update

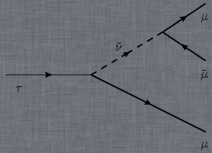
Johannes Albrecht[1], Marta Calvi[2], Marcin Chrząszcz[3,4], Jon Harrison[5], Basem Khanji[2], George Lafferty[5], Nicola Serra[3], Paul Seyfert[6]

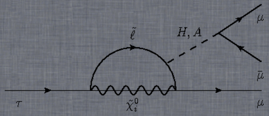
[1] Dortmund, [2] Milano, [3] Zurich, [4] Krakow, [5] Manchester, [6] Heidelberg

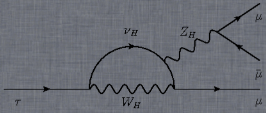
28th August 2013

University of Zurich[UZH]









Status

MC Samples

Normalization channel

Peaking backgrounds

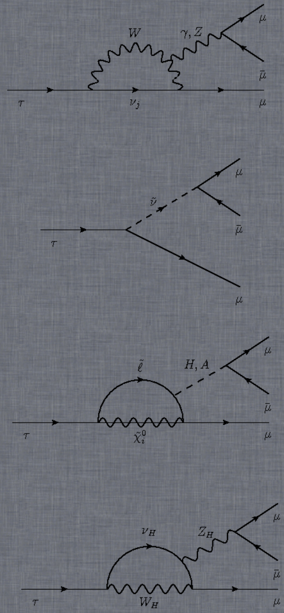
Normalization

Isolating parameter

Ensemble Selection

Binning optimisation

Conclusions



Status

1fb^{-1} analysis of $\tau \to \mu\mu\mu$ and $\tau \to p\mu\mu$ appeared in PLB.

Contents lists available at SciVerse ScienceDirect

Physics Letters B

www.elsevier.com/locate/physletb

Searches for violation of lepton flavour and baryon number in tau lepton decays at LHCb ☆

LHCb Collaboration

ARTICLE INFO

Article history:
Received 17 April 2013
Received in revised form 27 May 2013
Accepted 29 May 2013
Available online xxxx
Editor: L. Rolandi

ABSTRACT

Searches for the lepton flavour violating decay $\tau^- \to \mu^-\mu^+\mu^-$ and the lepton flavour and baryon number violating decays $\tau^- \to \bar{p}\mu^+\mu^-$ and $\tau^- \to p\mu^-\mu^-$ have been carried out using proton–proton collision data, corresponding to an integrated luminosity of 1.0 fb^{-1}, taken by the LHCb experiment at $\sqrt{s} = 7$ TeV. No evidence has been found for any signal, and limits have been set at 90% confidence level on the branching fractions: $\mathcal{B}(\tau^- \to \mu^-\mu^+\mu^-) < 8.0 \times 10^{-8}$, $\mathcal{B}(\tau^- \to \bar{p}\mu^+\mu^-) < 3.3 \times 10^{-7}$ and $\mathcal{B}(\tau^- \to p\mu^-\mu^-) < 4.4 \times 10^{-7}$. The results for the $\tau^- \to \bar{p}\mu^+\mu^-$ and $\tau^- \to p\mu^-\mu^-$ decay modes represent the first direct experimental limits on these channels.

2011 results:

1. Obtained limit for $\tau \to \mu\mu\mu$: 8.0×10^{-8}.
2. BaBar and Belle: $2.1(3.2) \times 10^{-8}$ at 90% CL.
3. For 2012 + 2011 planned to implement several improvements.

M.Chrząszcz 2013

MC Samples

1. In 2011 analysis one of the biggest contributions to the systematic error from MC was the reweighting the MC signal for the correct cross section.
2. For 2012 we solved this problem by simulating signal in 5 parts. One for each production channel(normalization to 1M events):

$$\tau \to \mu\mu\mu = \begin{cases} \mathrm{B} \to \tau \to \mu\mu\mu & 116,600 \\ \mathrm{B} \to \mathrm{D_s} \to \tau \to \mu\mu\mu & 87,200 \\ \mathrm{B} \to \mathrm{D} \to \tau \to \mu\mu\mu & 1,800 \\ \mathrm{D_s} \to \tau \to \mu\mu\mu & 750,600 \\ \mathrm{D} \to \tau \to \mu\mu\mu & 43,800 \end{cases}$$

M.Chrząszcz 2013

MC Generator Cuts

In order to reduce the number of unwanted events we introduced generator level cuts.

Signal sample[1]		Background sample(Dimuon)[2]	
$p_{t\mu}$	$> 250 MeV$	$p_{t\mu}$	$> 280 MeV$
p_{μ}	$> 2.5 GeV$	p_{μ}	$> 2.9 GeV$
		$m(\mu\mu)$	$< 4.5 GeV$
		DOCA$(\mu\mu)$	$< 0.35 mm$

Gain a factor of ~ 8 in statistics compared to 2011.

[1] $X \to \tau \to 3\mu$, $\mathrm{D_s} \to \eta(\mu\mu\gamma)\mu\nu$, $\mathrm{D_s} \to \phi(\mu\mu)\pi$
[2] $c\bar{c}$, $b\bar{b}$

Normalization channel

As last year we will use $\mathrm{D_s} \to \phi(\mu\mu)\pi$. Events are split into 2 categories:

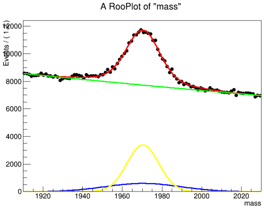
1. *cc* $\to \mathrm{D_s} \to \phi(\mu\mu)\pi$ 897,000
2. *bb* $\to \mathrm{D_s} \to \phi(\mu\mu)\pi$ 103,000

We avoid reweighting of the samples as in 2011.

M.Chrząszcz 2013

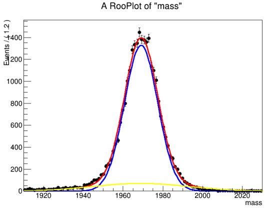
Mass correction

$D_s \to \phi(\mu\mu)\pi$ in data.



- mean = $1970.3 \pm 0.9 MeV$

$D_s \to \phi(\mu\mu)\pi$ in MC.

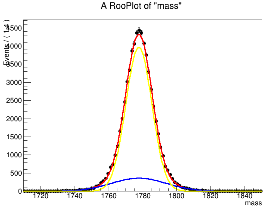


- mean = $1969.1 \pm 0.60 MeV$

- $m_{\tau\to3\mu} = \frac{1970.3}{1969.1} \times 1777.7 = 1778.8 \pm 1.1 MeV$

  In agreement with 2011.

Fit $\tau \to \mu\mu\mu$ in MC.

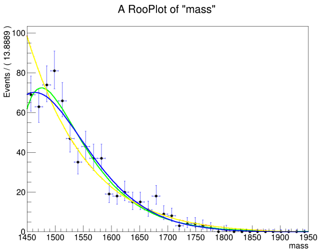


- mean = $1777.7 \pm 0.4 MeV$

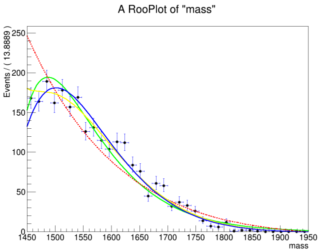
M.Chrząszcz 2013

Peaking backgrounds

1. The dominant background source of peaking background in this analysis is $\mathrm{D_s} \to \eta(\mu\mu\gamma)\mu\nu$
2. In 2011 we suffered from lack of MC statistics.
3. Thanks to generator cuts our pdfs became more stable.



PID:0.65; 0.725,GEO:−0.48; 0.05



PID:0.725; 0.0.86,GEO:0.35; 0.65

M.Chrząszcz 2013

Normalization

For the normalization of background samples($c\bar{c}$ and $b\bar{b}$) we used generator cuts efficiencies and corrected the nominal cross section accordingly:

$$\mathcal{L} = \frac{N_{MC}}{\varepsilon_{acc} \times \varepsilon_{gen} \times \sigma LHCb}$$

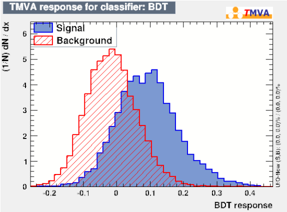
The obtained luminosities(per 1M events):

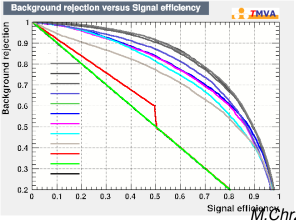
1. $\mathcal{L}_{cc} = 0.25 \pm 0.04 pb^{-1}$
2. $\mathcal{L}_{bb} = 1.20 \pm 0.15 pb^{-1}$

Dominant uncertainty from the cross section.

Isolating parameters

1. In 2011 we used the isolation parameter developed for $B^0_s \to \mu\mu$. For 2012 data we optimised the isolation parameter for our channel based on MVA(BDT).
2. Instead of training on isolating vs non-isoalting tracks we train on combinatorial background vs signal.
3. We see big improvement compared to old isolation.





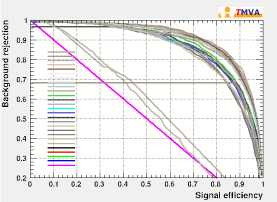
M.Chrząszcz 2013

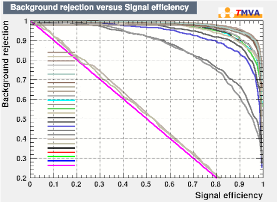
Ensemble Selection

1. In the last few years people winning leading machine learning contests started to combine their classifiers to squeeze the best out of them.
2. This technique/method is know as Ensemble Selection or Blending.
3. The plan for $\tau \to \mu\mu\mu$ is to take it to the next level.
4. Combine not only different channels, but also different τ sources(slide 4).

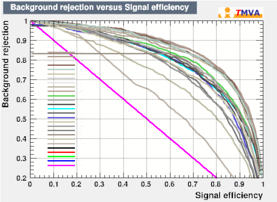
M.Chrząszcz, N.Serra 2013

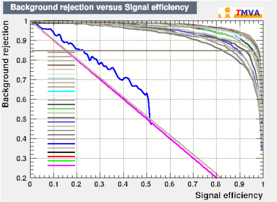
Ensemble Selection





- $\mathrm{B} \to \mathrm{D} \to \tau$
- $\mathrm{D} \to \tau$



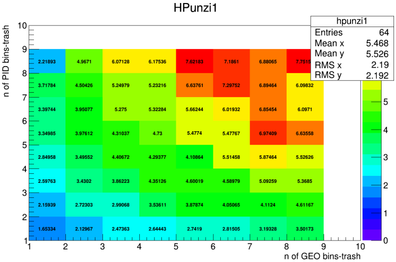


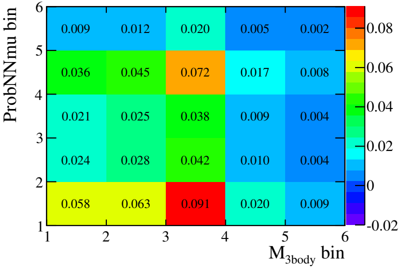
- $\mathrm{B} \to \mathrm{D_s} \to \tau$
- $\mathrm{D_s} \to \tau$

M.Chrząszcz, N.Serra 2013

Binning optimisation

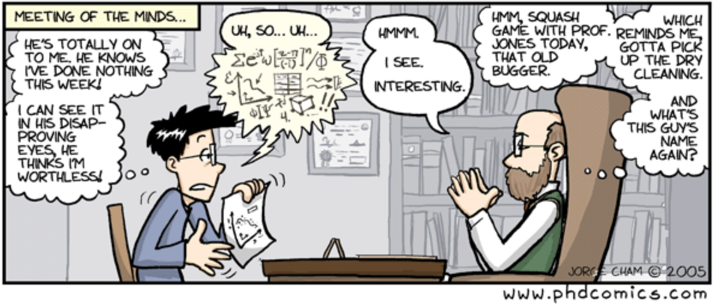
For the 2011 analysis we had two classifiers: *PIDNN* and M_{GEO}. Each of them we optimised separately. For the 2012 analysis we are performing a simultaneous 2D optimisation.





Conclusions

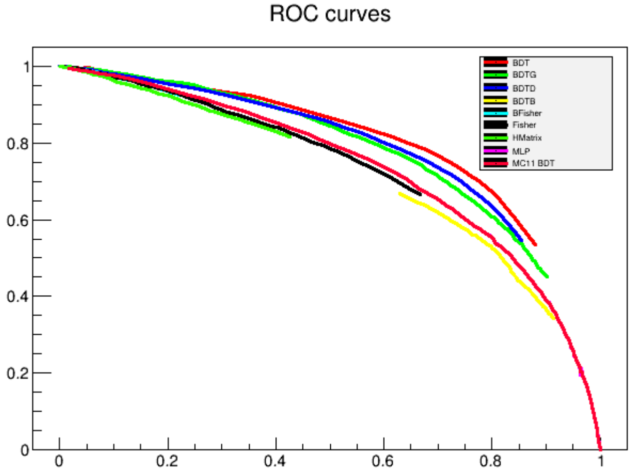
1. Analysis is well underway.
2. MC samples are almost there.
3. Hope to improve the selection.
4. $\tau \to p\mu\mu$ mode will be studied in parallel.



BACKUP

$B \to \tau$

We really suck in selecting this channel.

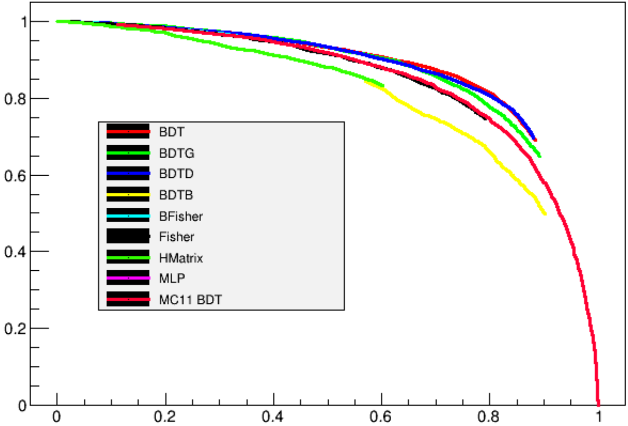


M.Chrząszcz 2013

$B \to D_s \to \tau$

On the biggest contributing channel we are quite optimal.

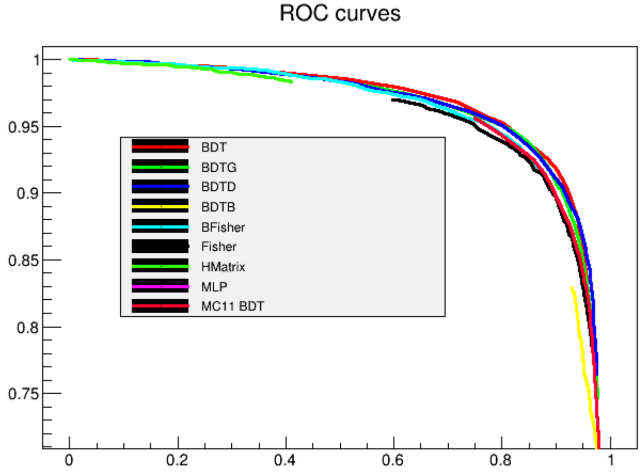
ROC curves



M.Chrząszcz 2013

$D_s \to \tau$

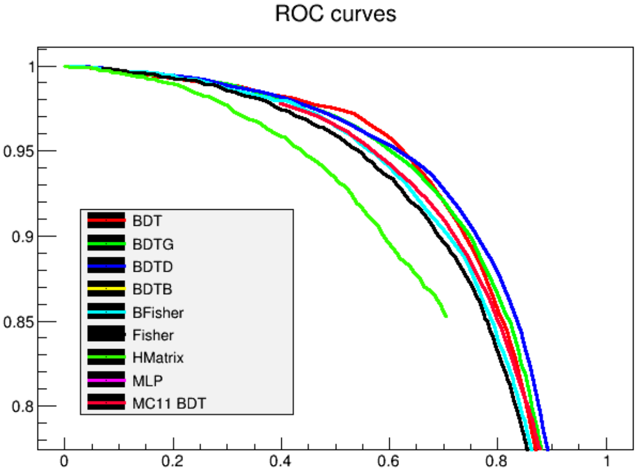
On the biggest contributing channel we are quite optimal.



M.Chrząszcz 2013

$B \to D^+ \to \tau$

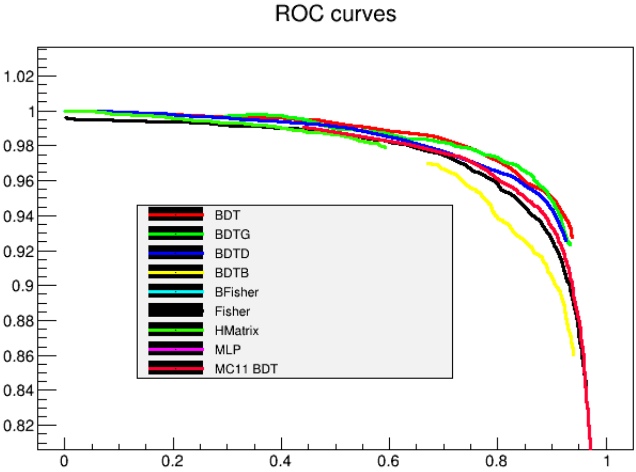
On the biggest contributing channel we are quite optimal.



M.Chrząszcz 2013

$D^+ \to \tau$

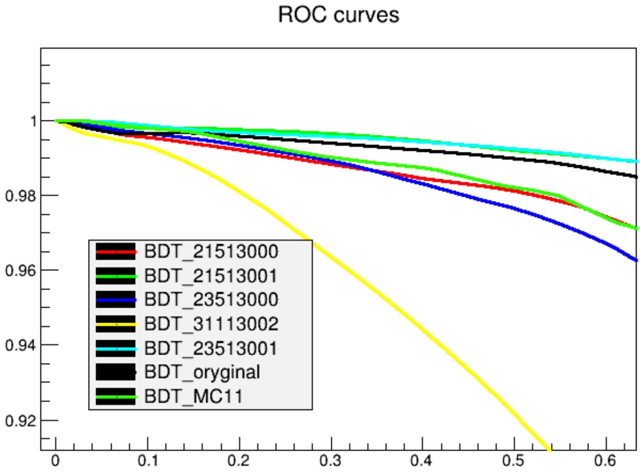
On the biggest contributing channel we are quite optimal.



M.Chrząszcz 2013

Comparison on mix sample

On the biggest contributing channel we are quite optimal.

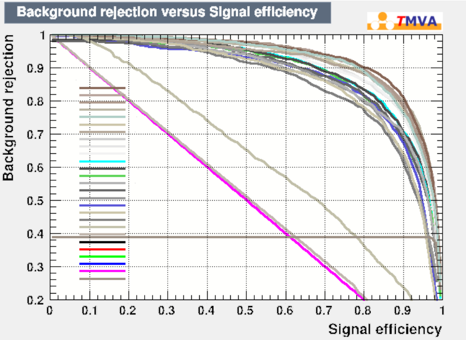


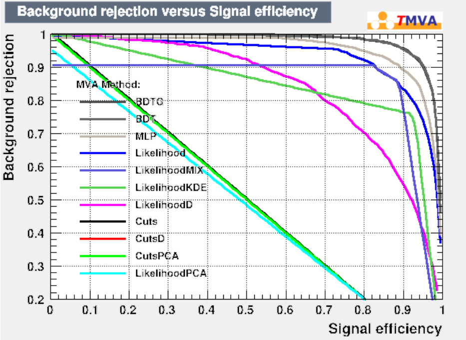
M.Chrząszcz 2013

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?

Comparison on mix sample





M.Chrząszcz 2013