Search for Charged Lepton Flavour Violation at LHCb experiment

Marcin Chrząszcz

Henryk Niewodniczanski Institute of Nuclear Physics Polish Academy of Sciences, Cracow, Poland

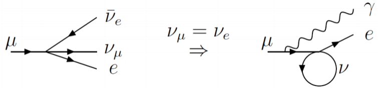
February 8, 2015

Outline

Lepton Flavour/Number Violation

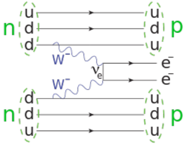
1. Lepton Flavour Violation (LFV) found in neutrino sector - the first phenomena outside the Standard Model.
2. The search for charged lepton flavour violation (CLFV) commenced with muon discovery (1936) and its identification as a separate particle.

- Expected: $B(\mu \to e\gamma) \approx 10^{-4}$
- Unless there is another ν.

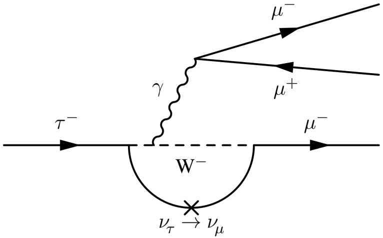


3. The observation of CLFV would be a clear signature of New Physics (NP) - paramount importance for flavour physics and the enigma of generations.
4. LFV vs LNV (Lepton Number Violation)

- Even with LFV, lepton number can be a conserved quantity.
- Many NP models predict LNV (Majorana neutrinos)
- LNV searched in so-called neutrinoless double β decays.



Status of searches for $\tau \to \mu\mu\mu$



- Charged Lepton Flavour Violation process.
- The Standard Model contribution: penguin diagram with neutrino oscillation.
- Negligible SM branching fraction.
- Large enhancement from NP models like: SUSY, Little Higgs, Fourth generation, etc.

Predictions

SM $O(10^{-40})$

var. SUSY 10^{-10}

non universal Z' 10^{-8}

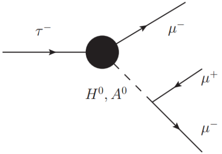
mSUGRA+seesaw 10^{-9}

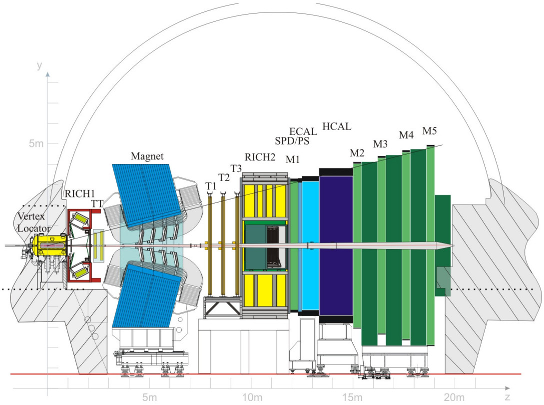
and many more...

Current limits (90 % CL)

BaBar 3.3×10^{-8}

Belle 2.1×10^{-8}





LHCb is a forward spectrometer:

- Excellent vertex resolution.
- Efficient trigger.
- High acceptance for τ and B.
- Superb particle identification (PID).

Strategy

1. Data sample: 1fb^{-1} 7 TeV and 2fb^{-1} 8TeV.
2. Normalization (control) decay channel: $\mathrm{D_s} \to \phi(\mu\mu)\pi$.
3. Blind analysis in the region of $|m_{\mu\mu\mu} - m_\tau| < 20$ MeV$/c^2$.
4. Event selection:
   - Preselection of three tracks that combine to give a mass close to m_τ, with displaced vertex.
   - Selection based on three classifiers:
     - Geometry and topology ($\mathcal{M}_{3body}$) - multivariate classifier
     - PID ($\mathcal{M}_{PID}$) - multivariate classifier
     - Three muon invariant mass ($m_{\mu\mu\mu}$)
5. Major background contributions: $\mathrm{D_s} \to \eta(\mu\mu\gamma)\mu\nu$ and $\mathrm{D} \to \mathrm{K}\pi\pi$ decays.
6. Evaluation of the upper limit on $\mathcal{B}(\tau \to \mu\mu\mu)$ using CL_s method.

τ production at LHCb

- τ's in LHCb come from five main sources:

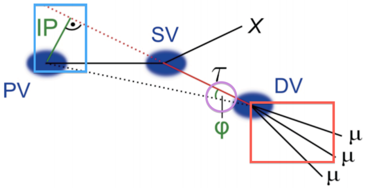
Mode	7 TeV	8 TeV
Prompt $D_s \to \tau$	$71.1 \pm 3.0\,\%$	$72.4 \pm 2.7\,\%$
Prompt $D^+ \to \tau$	$4.1 \pm 0.8\,\%$	$4.2 \pm 0.7\,\%$
Non-prompt $D_s \to \tau$	$9.0 \pm 2.0\,\%$	$8.5 \pm 1.7\,\%$
Non-prompt $D^+ \to \tau$	$0.18 \pm 0.04\,\%$	$0.17 \pm 0.04\,\%$
$X_b \to \tau$	$15.5 \pm 2.7\,\%$	$14.7 \pm 2.3\,\%$

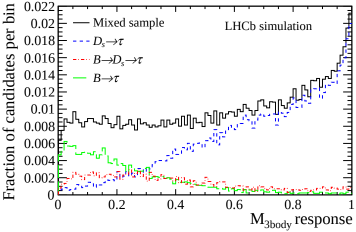
- Pythia produces them in wrong proportions
- The above decay channels were produced separately and added in the given proportions.
- Taken into account different trigger efficiencies for different channels.

Signal and background discrimination

- Two multivariate classifiers, $\mathcal{M}_{3body}$ and $\mathcal{M}_{\mathcal{PID}}$.

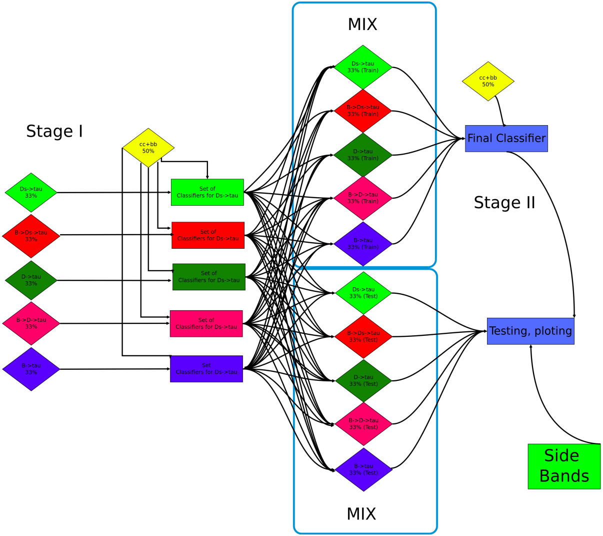
- $\mathcal{M}_{3body}$ trained using vertex and track fit quality, vertex displacement, vertex pointing, vertex isolation and τ p_T.
- Used Blending Technique (see the next slide).





- Trained on signal and background MC.
- Calibrated on $D_s \to \phi(\mu\mu)\pi$ sample.

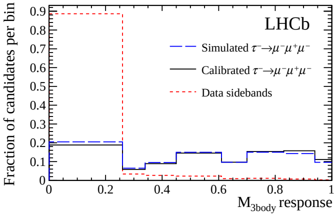
Blending technique



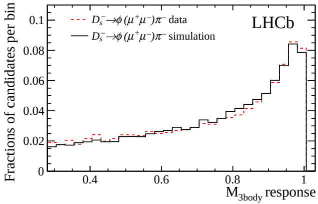
- Each of the τ lepton production channel have a different signature in terms of kinematic distributions.
- Signal blending technique improved the discriminating power by 6 %

Calibration

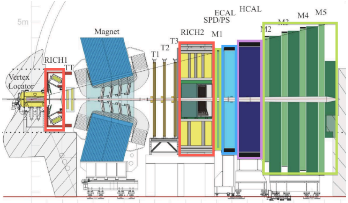
- Assume all differences between $\tau \to \mu\mu\mu$ and $D_s \to \phi\pi$ come from kinematics (mass, resonance, decay time), which is correct in MC.
- Get correction $D_s \Longrightarrow \tau$ from MC.
- Apply corrections to $D_s \to \phi\pi$ on data.
- Publication in preparation.



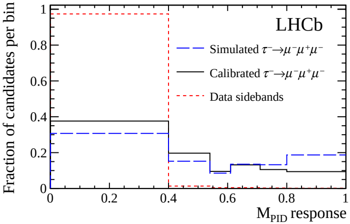
- $D_s \to \phi\pi$ decay well modelled in MC.



Signal and background discrimination - $\mathcal{M}_{PID}$



- $\mathcal{M}_{PID}$ trained using RICH, ECAL and muon chambers.



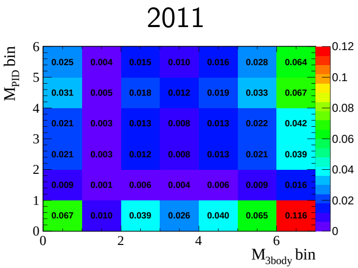
- Trained on signal and background MC.
- Calibrated on $B \to J/\psi K$ and $D_s \to \phi(\mu\mu)\pi$ decays.

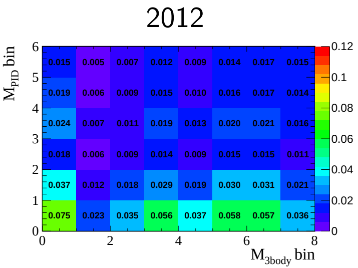
Binning optimisation

- Events are distributed among $\mathcal{M}_{3body}, \mathcal{M}_{PID}$ plane.
- In 2D we collect the events in groups(bins)
- Bins are optimised using CL_s method:

$$CL_s = \frac{\prod_{i=1}^{N_{\text{chan}}} \sum_{n=0}^{n_i} \frac{e^{-(s_i+b_i)}(s_i+b_i)^n}{n!}}{\prod_{i=1}^{n_{\text{chan}}} \sum_{n=0}^{n_i} \frac{e^{-b_i} b_i^n}{n!}},$$

- The lowest bins are rejected, because they do not contribute to the limit sensitivity.



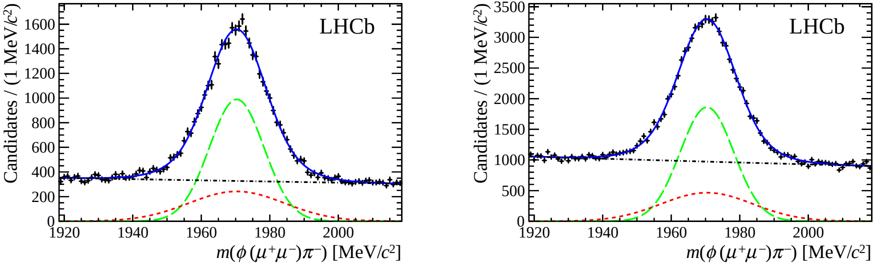


Relative normalisation

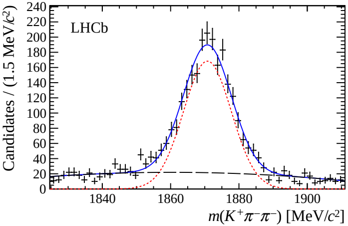
$$\mathcal{B}(\tau \to \mu\mu\mu) = \frac{\mathcal{B}(\mathrm{D_s} \to \phi\pi)}{\mathcal{B}(\mathrm{D_s} \to \tau\nu_\tau)} \times f^{\tau}_{\mathrm{D_s}} \times \frac{\varepsilon_{\mathrm{norm}}}{\varepsilon_{\mathrm{sig}}} \times \frac{N_{\mathrm{sig}}}{N_{\mathrm{norm}}} = \alpha \times N_{\mathrm{sig}}$$

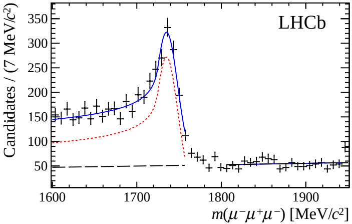
- where ε stands for trigger, reconstruction, selection efficiency.
- $f^{\tau}_{\mathrm{D_s}}$ is the fraction of τ coming from $\mathrm{D_s}$.
- norm = normalisation channel $\mathrm{D_s} \to \phi\pi$
  i.e. $(83 \pm 3)\,\%$ for 2012 data.

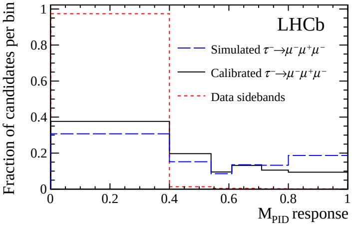
2011 2012



- Dominant: $D^+ \to K\pi\pi$.
- Also seen $D^+ \to \pi\pi\pi$ and $D_s \to \pi\pi\pi$.
- All contained in the lowest $\mathcal{M}_{PID}$ bin.

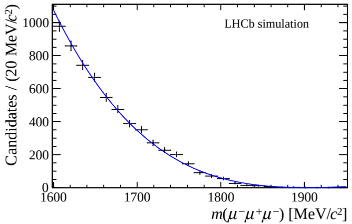


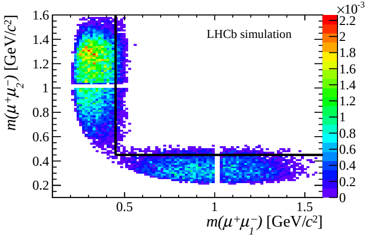




Other backgrounds

- $\phi \to \mu\mu + X$; narrow veto on dimuon mass.
- $D_s \to \eta(\mu\mu\gamma)\mu\nu_\mu$; not so easy:
  - Model it
  - <u>Remove it</u> with dimuon mass cut:
    - Fits better understood.
    - Sensitivity unchanged when removing veto.
    - Smaller uncertainty on expected background.

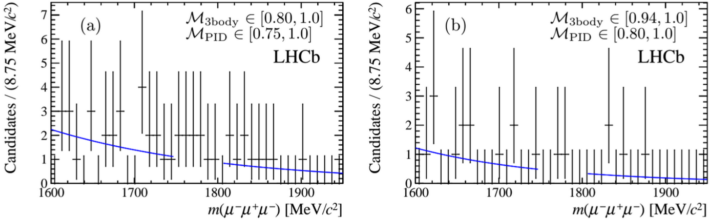




Remaining backgrounds

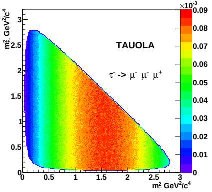
- Fit exponential to invariant mass spectrum in each likelihood bin.
- Don't use the ± 30 MeV region.

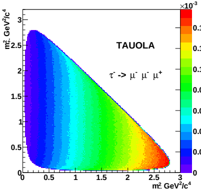
Example of most sensitive regions in 2011 and 2012

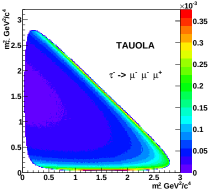


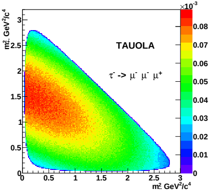
Model dependence

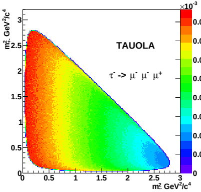
- η veto $\Rightarrow$ our limit not constraining to New Physics with small $m_{\mu^+\mu^-}$.
- Model description in arXiv:0707.0988 by S.Turczyk using Effective Field Theory approach.
- 5 relevant Dalitz distributions: 2 four-point operators, 1 radiative operator, 2 interference terms.







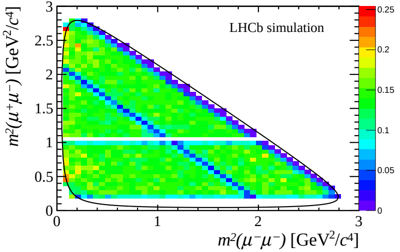




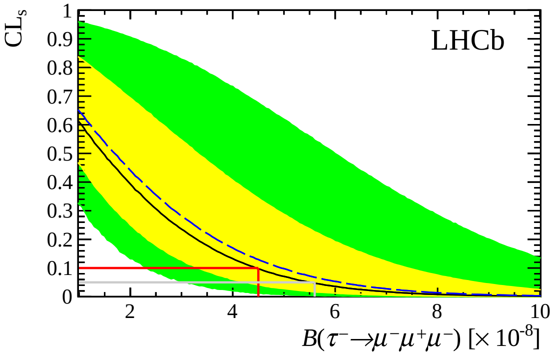
- All five cases implemented in TAUOLA.
- Publication in preparation.

Model dependence

- η veto $\Rightarrow$ our limit not constraining to New Physics with small $m_{\mu^+\mu^-}$.
- Model description in arXiv:0707.0988 by S.Turczyk using Effective Field Theory approach.
- 5 relevant Dalitz distributions: 2 four-point operators, 1 radiative operator, 2 interference terms.
- With radiative distribution limit gets worse by a factor of 1.5 (dominantly from the η veto).
- The other four Dalitz distributions remain stable (within 7 %).



Results

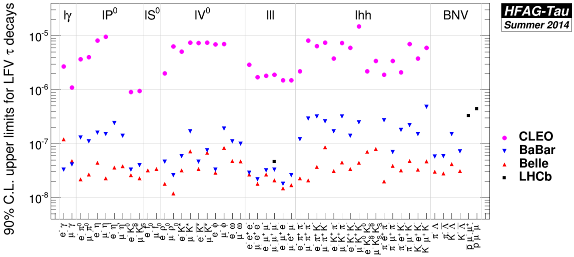


Limits(PHSP):
Observed(Expected)
$4.6\ (5.0) \times 10^{-8}$ at 90% CL
$5.6\ (6.1) \times 10^{-8}$ at 95% CL

Dalitz distribution	$\times 10^{-8}$
$\varrho_V^{(LL)(LL)}$	4.2 (4.7)
$\varrho_V^{(LL)(RR)}$	4.1 (4.6)
$\varrho_{rad}^{(LR)}$	6.8 (7.6)
$\varrho_{mix}^{(LL)(LL)}$	4.4 (5.1)
$\varrho_{mix}^{(LL)(RR)}$	4.6 (5.0)

Combination of LFV UL and Summary (1)

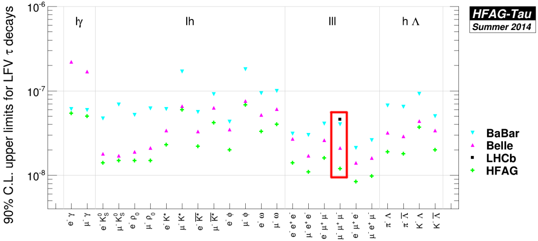
- Searches for LFV in τ sector is a domain of B factories.
- Over last years both BaBar and Belle set very strong limits on branching fractions of several rare τ decays.



  - First result from hadron collider comparable with B factories.
  - Since those limits are used to constraint NP models, their "official" combination is of paramount importance.
  - Various methods of limit computation used in Belle and BaBar's studies.
  - The HFAG group recomputed consistently all estimates using the CL_s method and the the same approach was involved in the average evaluation.

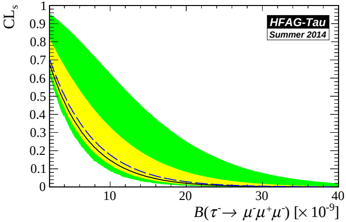
$\tau \to \mu\mu\mu$ limits (90 % CL)

BaBar(FC)	3.3×10^{-8}
Belle(FC)	2.1×10^{-8}
LHCb(CLs)	4.6×10^{-8}
HFAG(CLs)	1.2×10^{-8}



To conclude:

- LHCb is reaching B-factories limits.
- Many new techniques developed to perform this analysis.
- Combination of UL within HFAG gave the best sensitivity for $\mathcal{B}(\tau \to \mu\mu\mu) < 1.2 \times 10^{-8}$ at 90% CL.
- Erratum to bibliography:
  - [120] HFAG report published:arXiv:1412.7515 (previous cited preliminary web report).
  - [76] accepted for publication in JHEP.



Prof. J.Ciborowski comments

1. Podrozdział (4.10) ten jest de facto zapowiedzią większej pracy. Szkoda, że materiał tu przedstawiony potraktowany został bardzo skrótowo, co wymusiło na mnie konieczność kilkakrotnego przeczytania tego podrozdziału i utrudniło docenienie wyniku otrzymanego przez autora w konfrontacji z przewidywaniami teoretycznymi.

   - The theory part of this was presented in detail in 2.3.4. This chapter is just a showing how to reweight the distributions to a given NP model, that is why I tried to keep it short, but I agree I overdid it.

Prof. J.Ciborowski comments

2. Jedyne rzucające się w oczy uchybienie redakcyjne to pomyłki w numerach rozdziałów, których zawartość wymieniona jest pod koniec Wstępu.

- Mea Culpa. Completely missed that.

Prof. B.Muryn comments

1. Jest to kwestią subiektywną, ale myślę, że poziom rozprawy wyklucza nieobytego z tymi zagadnieniami czytelnika, a w związku z tym skróciłbym opisywane podstawy poświęcając więcej miejsca na trochę szerszą dyskusję modeli prowadzących do niezachowania liczby leptonowej L. Zgodziłbym się z ewentualną ripostą Autora, że przedstawia modele mające bardziej bezpośredni związek z badaną wartością współczynnika rozgałęzienia dla procesu (będącego w rzeczywistości miarą amplitudy tego procesu), niemniej ogólny niedostatek tych teorii wykazujący silną zależność od nieznanych parametrów modeli a prowadzący do niemożności wyboru żadnego z nich, został trochę pominięty.

- I said in Sec. 2.3: "The violation of charged lepton flavour is predicted in many extensions of the SM (generically named as BSM theories). The inclusion of the CLFV is usually straightforward and follows directly from the models assumptions".

  I however agree that I should said explicitly that there is no way which is correct and that is why we use a "Model independent" EFT approach.

Prof. B.Muryn comments

2. Prezentowane tabele i rysunki dowodzą dbałości o szczegóły i wskazują na samodzielność analizy przeprowadzonej przez mgr Marcina Chrząszcza. Z drugiej strony uważam, że jest ich za wiele i cytowanie niektórych tabel mija się z celem (np. tabele 4.13- 4.14, 4.16-4.18 lub 4.34-4.35 oraz 4 strony podobnych w wymowie rozkładów mas mionów). Zapełniają one miejsca, na których chętnie bym widział szerszy komentarz dotyczący Blending Technique oraz, co ważne, dyskusji niepewności systematycznych.

- I probably could make Table 4.14-18 shorter and only list most important trigger lines making them substantially shorter. The same goes for Tables 4.34-4.35. I agree that I have been to meticulous in those places.

Prof. B.Muryn comments continuation

- With the limit settings it's not as straightforward to rank the systematics as it is with a measurements.
- We can group systematics in three categories:
  - Normalization systematics.
  - Signal distribution among $\mathcal{M}_{3body}, \mathcal{M}_{PID}, m_{\mu\mu\mu}$.
  - Background systematics.
- Each of them influences the computation of limit in different way.
- The main systematic is the normalization, in which the dominant uncertainty comes from $\mathcal{B}(\mathrm{D_s} \to \phi(\mu\mu)\pi$ (7.5%).

Prof. B.Muryn comments

3. "... nie szkodziłoby, gdyby Autor niniejszej rozprawy w krótkim podsumowaniu wyszczególnił te etapy analizy, w których grał główną rolę."

- Apologies about that. I have focued on those things in the talk.

Triggers at LHCb

- LHCb uses complex trigger, $\mathcal{O}(100)$ trigger lines.
- Lines change with data taking.
- Optimized choice of triggers based on $\frac{s}{\sqrt{b}}$ FOM,

$$\varepsilon(\beta)'_{\mathrm{evt,line}} = \frac{N(\tau\ \text{MC(BKG) events triggered line, but not by any better line})}{N(\tau\ \text{MC(BKG) events triggered by any line})},$$

$$\mathrm{CTFM} = \frac{\sqrt{\sum_{\text{trigger lines}} \beta'_{\mathrm{evt,line}}}}{\sum_{\text{trigger lines}} \varepsilon'_{\mathrm{evt,line}}}$$

- Evaluated different triggers used in 2012 data taking.
- Found negligible differences in trigger efficiencies.

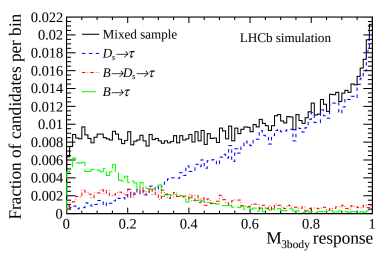
name	ε'	β'	CTFM
Hlt2TriMuonTauDecision	0.880708	0.736182	0.974228
Hlt2DiMuonDetachedDecision	0.0669841	0.173396	1.00636
Hlt2CharmSemilep3bodyD2KMuMuDecision	0.0206816	0.0182935	0.99472
Hlt2CharmHadD2HHHDecision	0.00554351	0.00666405	0.992604
Hlt2CharmSemilep3bodyD2KMuMuSSDecision	0.00195444	0.00470404	0.993106
Hlt2CharmSemilep3bodyD2PiMuMuDecision	0.00206105	0.00679472	0.994591
Hlt2TopoMu3BodyBBDTDecision	0.00394442	0.0121521	0.996937

Geometric likelihood

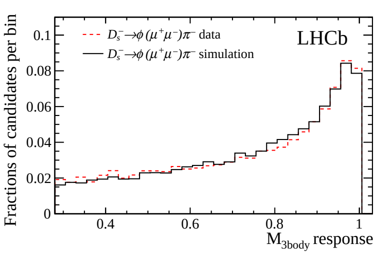
- As mentioned in LHC we have different production sources of τ's.
- Each source has different detector response signature.
- To maximise our performance we trained classifiers for each of the τ sources using:
  - Kinematic properties of τ candidate.
  - Geometric properties of τ candidate, like pointing angle, DOCA, Vertex χ^2, flight distance.
  - Isolations, for vertex and individual tracks.
- After training the individual classifiers one that combines all this information in a single classifier on mixed sample of τ's.
- This technique is known as Blending or Ensemble learning.
- Using this approach we gain 6% sensitivity!

- Classifier prefers τ's from prompt D_s, the dominant channel.

MC response for different τ production channels

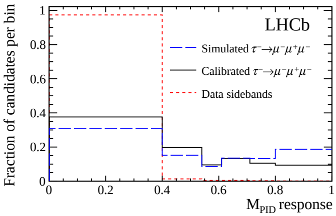


Response for $D_s \to \phi\pi$ data and MC



Particle Identification (PID)

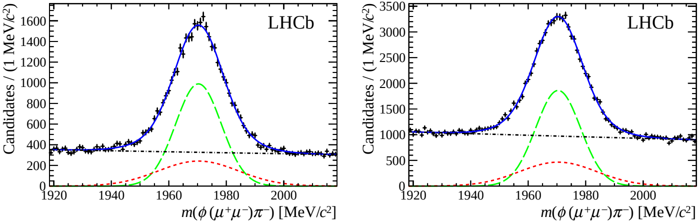
- Classifier trained on inclusive MC sample.
- Using information from: RICH, Calorimeters, Muon system and tracking.
- Correct for the MC efficiency using control channel: $\mathrm{D_s} \to \phi(\mu\mu)\pi$ and $\mathrm{B} \to \mathrm{J}/\psi(\mu\mu)\mathrm{K}$



Mass shape

- Double-Gaussian with fixed fraction (70 % inner Gaussian).
- Fix fraction to ease calibration.
- Correct mass by MC:

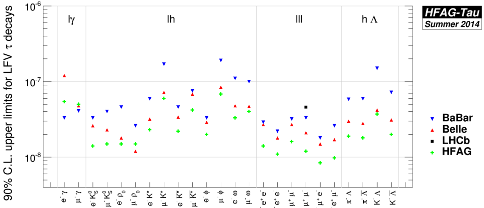
$$\sigma^{\tau}_{data} = \frac{\sigma^{\tau}_{MC}}{\sigma^{D_s}_{MC}} \times \sigma^{D_s}_{data}$$



Calibrated τ Mass shape	7 TeV	8 TeV
Mean (MeV)	1779.1 ± 0.1	1779.0 ± 0.1
σ_1 (MeV)	7.7 ± 0.1	7.6 ± 0.1
σ_2 (MeV)	12.0 ± 0.8	11.5 ± 0.5

- For each measurement take integrated luminosity ($\mathcal{L}$), cross section ($\sigma_{\tau\tau}$), efficiencies (ϵ), background expected(b) and all systematics.
- Calculate number of signal: $s = \mathcal{L}\sigma_{\tau\tau}\epsilon^{tot}\mathcal{B}(\tau \to LFV)$.
- Scan the CL_s wrt. $\mathcal{B}(\tau \to LFV)$:

$$CL_s = \frac{\prod_{i=1}^{N_{\mathrm{chan}}} \sum_{n=0}^{n_i} \frac{e^{-(s_i+b_i)}(s_i+b_i)^n}{n!}}{\prod_{i=1}^{n_{\mathrm{chan}}} \sum_{n=0}^{n_i} \frac{e^{-b_i} b_i^n}{n!}},$$



Variables 1/2

The multi-variate classifiers were trained using the following variables:

- **DOCA:** the minimum of distances of the closest approach of two muons in each of three possible two muon pairings,
- $\boldsymbol{\tau}$**($\mathbf{D_s}$) Vertex** χ^2**:** the quality of the vertex parametrized as the chi square of the τ secondary vertex fit (as defined in,
- **c**τ**:** The measured decay length of the τ lepton, assuming its production at the primary vertex. To smooth out the distribution, the decay time is transformed according to the formula $T = \exp(-1000 \cdot \tau)$,
- **IP** χ^2 **(**$\boldsymbol{\tau}$**):** τ lepton impact parameter χ^2/ndf,
- **Min. IP** χ^2 **(**$\boldsymbol{\mu}$**):** the minimum value of the three μ impact parameter (χ^2/ndf)s,
- **Track** χ^2**/ndf:** maximum of track's (χ^2)s of the three muons,

Variables 2/2

The multi-variate classifiers were trained using the following variables:

- **Pointing angle α:** the angle between the direction of τ momentum and a straight line from the τ decay vertex to the primary vertex,
- $\mathbf{p_T}$: the τ transverse momentum,
- **Track isolation:** the sum of three track isolations variables, each parametrising how far in space is an individual muon candidate w.r.t. the rest of event.
- **BDT (Boosted Decision Tree) isolation:** the response of multivariate analysis (MVA) working at the charged track level and aimed at discriminating between isolated and non-isolated tracks.
- **Cone isolation:** the fraction of the τ candidate transverse momentum among the sum of all transverse momenta within a certain cone around the τ candidate.

Isolations 1/2

The track isolation (TI) variable is constructed on the basis of the respective studies performed by the LHCb collaboration for the needs of $B_s^0 \to \mu^+\mu^-$ analysis. The TI is defined as the number of extra tracks (i.e. excluding tracks that are attributed to the $\tau \to \mu\mu\mu$ candidate) that can form a vertex with a muon track. The assignment to the above SV is based on the selection criteria imposed on the following variables:

- minimum distance between the **track** and the PV (**pvdist**),
- minimum distance between the **track** and the $\tau \to \mu\mu\mu$ vertex (**svdist**),
- the distance of the closest approach between the **muon** and the **track** (DOCA),
- IP χ^2,
- angle between the **muon** and the **track** (β),
- the quantity

$$f_c = \frac{|\overrightarrow{p}_h + \overrightarrow{p}_{trk}|\alpha^{h+trk,PV}}{|\overrightarrow{p}_h + \overrightarrow{p}_{trk}|\alpha^{h+trk,PV} + p_{\mathrm{T},h} + p_{\mathrm{T},trk}}, \tag{1}$$

where $\alpha^{h+trk,PV}$ is the angle between the **muon** and the **track** candidate, $P_{\mathrm{T},h}$ and $P_{\mathrm{T},trk}$ are the transverse momentum with respect to the beam line.

Isolations 2/2

The track is considered as "isolated" if it satisfies the following requirements (imposed on the above mentioned variables):

- pvdist $\in [0.5, 40]$ mm,
- svdist $\in [-0.15, 30]$ mm,
- DOCA < 0.13 mm,
- Track IP significance > 3,
- $\beta < 0.27$ rad,
- $f_c < 0.6$.