Test of Lepton Flavor Universality using $B_d \to D^{*-}\tau\nu$ decays at LHCb

JRJC 2022

Gaya Mogette   Supervised by Olivier Leroy

Aix Marseille Univ, CNRS/IN2P3, CPPM, Marseille, France

Monday october 24th 2022

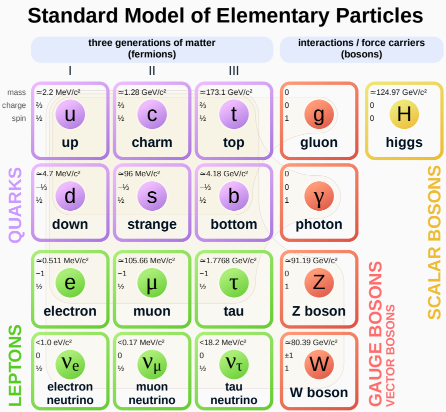
Overview

Introduction

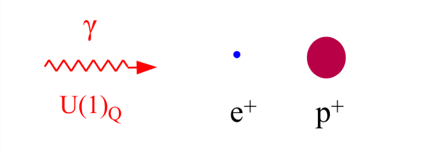
Standard Model of Particle Physics

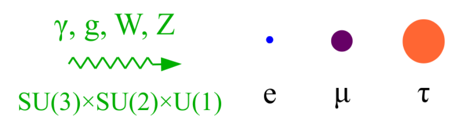
The Standard model of particle physics (SM) classifies elemetary particles into:

- 5 Bosons: The Higgs + 4 interaction mediators γ, $W^{\pm}$, Z^{0}, and gluons g
- 3 generations of fermions : quarks and leptons
- Their masses is determined by their coupling to the Higgs



Lepton Flavor Universality - a simple analogy

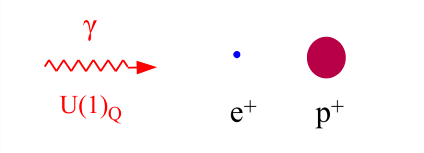




If we take a p and e^+

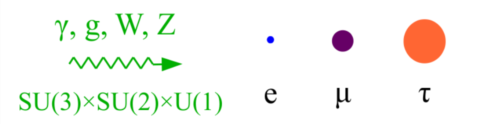
- Because they have the same charge under $U(1)_{\text{em}}$ the particles seems identical
- But they have different masses
- In the SM they are different:
    - p are not fundamental particles
    - e^+ is fundamental particle and 'do not feel' QCD

Lepton Flavor Universality - a simple analogy



If we take a p and e^+

- Because they have the same charge under $U(1)_{\text{em}}$ the particles seems identical
- But they have different masses
- In the SM they are different:
  - p are not fundamental particles
  - e^+ is fundamental particle and 'do not feel' QCD



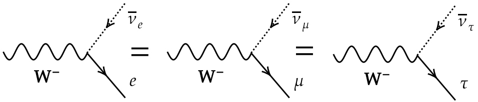
In the other hand, the SM classifies the different charged leptons : e, μ, and τ in three different family

- Under SM gauge group $U(1) \times SU(2) \times SU(3)$ the leptons seems identical because they have the same quantum numbers
- The only difference between them is their mass
- The different families may have a different behavior at high energies

Gino Isidori's Talk at LHCb Implication Workshop

Lepton Flavor Universality

The weak interaction in the Standard Model treats identically the three charged leptons: e, μ, and τ except for their different masses.



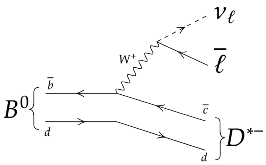
This property is referred as **Lepton Flavor Universality (LFU)**.

To test the LFU hypothesis we measure

$$R(D^*) \equiv \frac{\mathcal{B}(B^0 \to D^{*-} \tau^+ \nu_\tau)}{\mathcal{B}(B^0 \to D^{*-} \mu^+ \nu_\mu)}$$

LFU $\not\Longrightarrow$ $R(D^*) = 1$ because $m_\tau > m_\mu$

$R(D^{(*)})$ measurements $\longrightarrow$ charged flavor changing current $b \to c\ell\nu_\ell$:



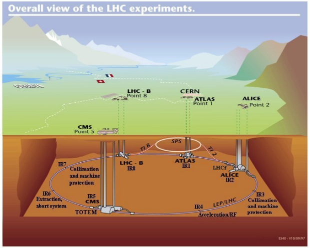
Previous measurements of $R(D^*)$:

- BaBar (2012)
  - $\frac{\mathcal{B}(B^0 \to D^{*-} \tau^+ \nu_\tau)}{\mathcal{B}(B^0 \to D^{*-} \ell \nu_\mu)}$ with $\ell = \mu, e$
- Belle (2015 and 2017)
  - Hadronic and leptonic τ
  - One-prong hadronic $\tau \to \pi\nu_\tau$ and $\tau \to \rho\nu_\tau$
- LHCb (2015 and 2018)
  - Muonic $\tau^- \to \mu^- \bar{\nu_\mu} \nu_\tau$
  - 3-prong hadronic τ (Run 1)

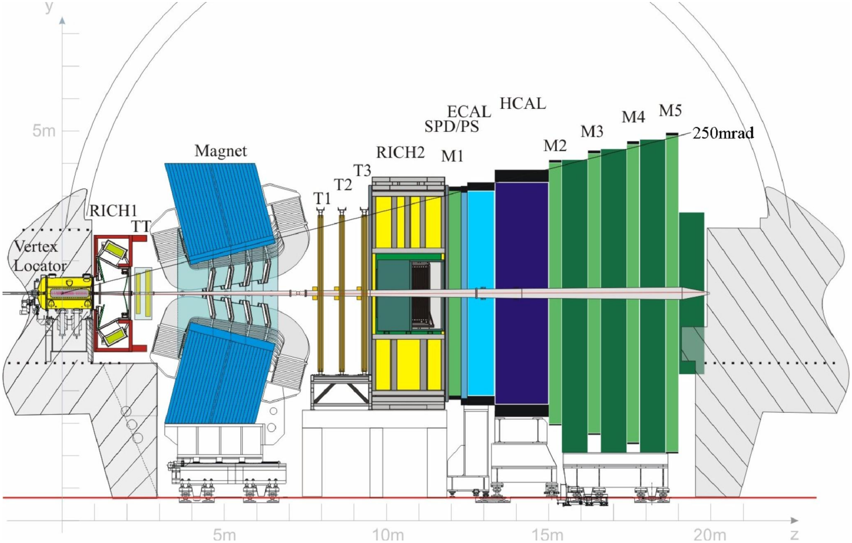
LHCb detector

The LHCb experiment at the Large Hadron Collider

- The Large Hadron Collider (LHC) is a proton-proton accelerator
- LHCb is one of experiments based at the LHC at CERN, Geneva
- Forward spectrometer initially designed to search for New Physics in the beauty quark sector
- Now very broad programme: charm and top quark, heavy ions, electro-weak physics, Higgs physics, ...
- Excellent vertex resolution (PV resolution: $10 - 40\,\mu\mathrm{m}$ in xy-plane and $50 - 300\,\mu\mathrm{m}$ in z-axis)
- Impact parameter (IP) resolution around $12\,\mu\mathrm{m}$ for high-momentum particles
- Momentum relative resolution of 0.5% below $20\,\mathrm{GeV}/c$ and 0.8% around $100\,\mathrm{GeV}/c$
- Typical PID efficiencies: 80% – 95% correct kaon ID and 3% – 10% misidentification of pion as kaon



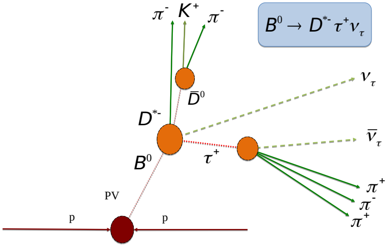
The LHCb detector - Run 2



$R(D^*)$ hadronic analysis (Run 2)

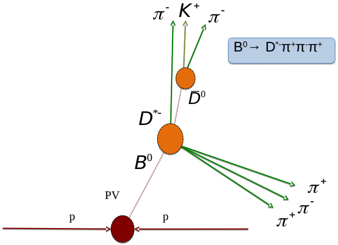
Signal & Normalisation mode

$B^0 \rightarrow D^{*-}\tau^+\nu_\tau$ and $\tau^+ \rightarrow 3\pi^\pm(\pi^0)\bar{\nu}_\tau$



Signal mode

$B^0 \rightarrow D^{*-}3\pi^\pm$



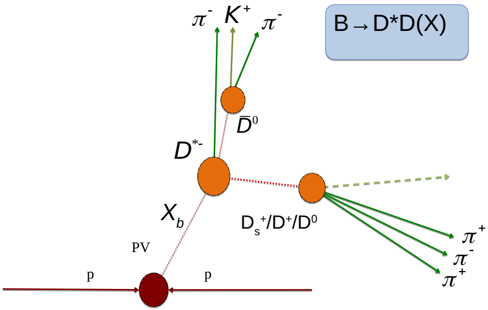
Normalisation mode

- Same final states in signal and normalisation modes
- Signal mode partially reconstructed (missing neutrino $\bar{\nu}_\tau$)
- Normalisation mode fully reconstructed
- Helps to cancel out systematic uncertainties

Backgrounds

Most relevant backgrounds are:

- Prompt decay: $B^0 \to D^* 3\pi^\pm$
- Double-charm decay: $B^0 \to D^* DX$ where $D \to 3\pi^\pm$
- Combinatorial backgrounds



$R(D^*)$ measurement

$$R(D^*) = \frac{\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)}{\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)} = \underbrace{\frac{\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)}{\mathcal{B}(B^0 \to D^{*-}\pi^+\pi^-\pi^+)}}_{\mathcal{K}(D^*)} \times \underbrace{\frac{\mathcal{B}(B^0 \to D^{*-}\pi^+\pi^-\pi^+)}{\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)}}_{\text{External branching fractions}}$$

We measure:

$$\mathcal{K}(\mathcal{D}^*) \equiv \frac{\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)}{\mathcal{B}(B^0 \to D^{*-}3\pi^\pm)} = \frac{N_{\text{sig}}}{N_{\text{norm}}} \frac{\varepsilon_{\text{norm}}}{\varepsilon_{\text{sig}}} \frac{1}{\mathcal{B}(\tau^+ \to 3\pi^\pm\bar{\nu}_\tau) + \mathcal{B}(\tau^+ \to 3\pi^\pm(\pi^0)\bar{\nu}_\tau)}$$

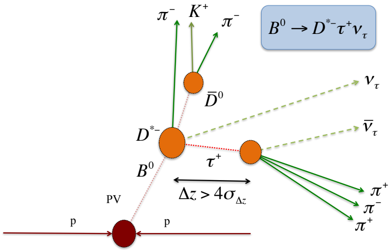
- N_{sig} from a 3D binned template fit:
  - $q^2 = (p_B - p_{D^*})^2$ momemtum transfered to the leptonic system (8 bins),
  - τ^+ lifetime t_τ (8 bins),
  - Anti-D_s^+ BDT (6 bins).
- N_{norm} from an unbinned fit to $m(D^* 3\pi^\pm)$
- Efficiencies ε_{sig} and $\varepsilon_{\text{norm}}$ extracted from MC samples

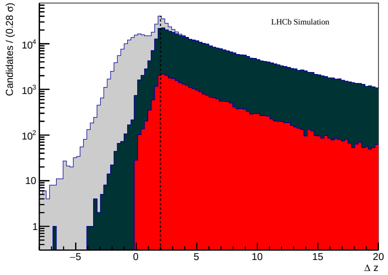
Selection

- Initial cuts
- Apply four BDTs (next slides):
  - $3\pi^{\pm}$ vertex detachment BDT
  - Anti-combinatorial background BDT
  - Charged isolation BDT
  - Anti-$B^0 \to D^{*-} D_s^+ X$ BDT (used in final fit)
- Remaining cuts: Signal and normalisation cuts

$3\pi^{\pm}$ vertex detachment BDT

- Remove 'prompt' background

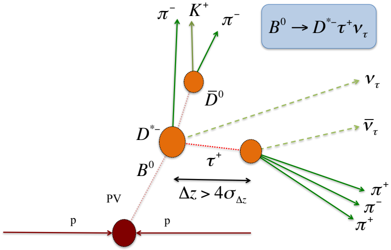


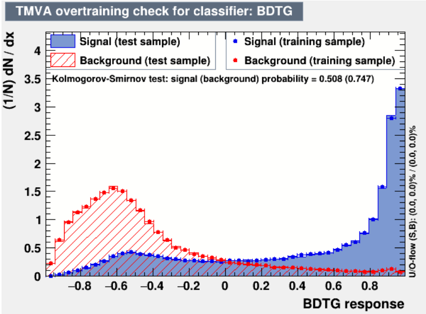


Δz/uncertainty distribution of the simulated signal (red), double charm background (black) and prompt background (grey), after the initial cuts. A cut at 2σ is shown.

$3\pi^{\pm}$ vertex detachment BDT

- Remove ‘prompt’ background

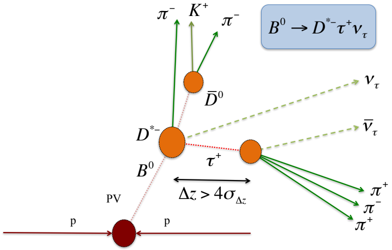


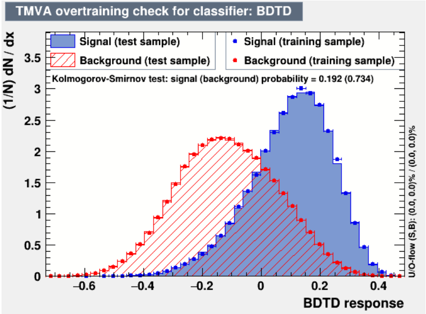


Signal efficiency: 70%; background rejection: 90% at BDT > 0.2

Anti-combinatorial background BDT

- Remove D^{*-} and $3\pi^{\pm}$ from different hadrons

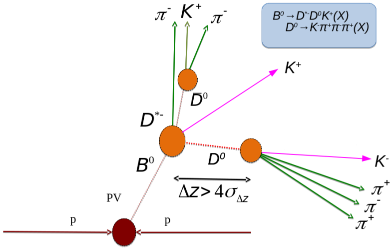


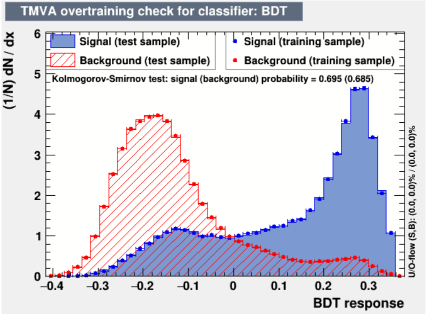


Signal efficiency: 85% & background rejection: 70% at BDT > 0

Charged isolation BDT

- Remove events with extra charged tracks associated with signal ones

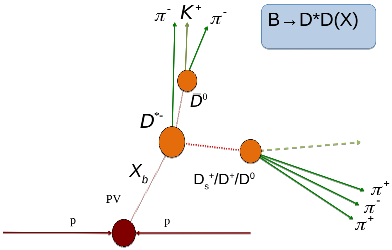


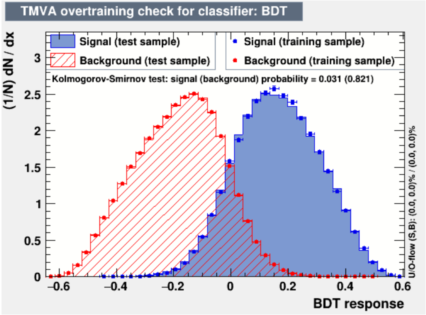


Signal efficiency: 80% and background rejection 77% at BDT > 0

Anti-$B^0 \to D^{*-} D_s^+ X$ BDT

- Distinguish $\tau^+ \to 3\pi^-$ from signal vs. $D_s^+ \to 3\pi^\pm X$ from $B^0 \to D^{*-} D_s^+ X$



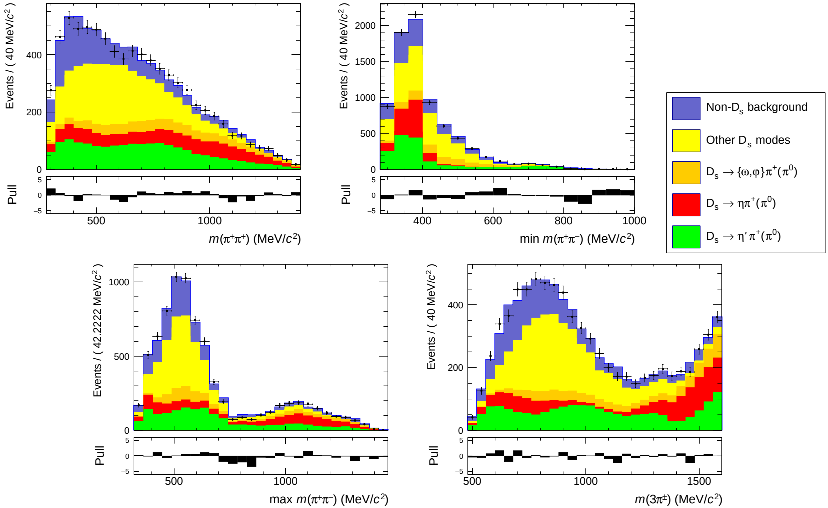


Signal efficiency: 99.7%; background rejection: 31% at BDT > -0.2

$\longrightarrow$ This BDT is used in final fit

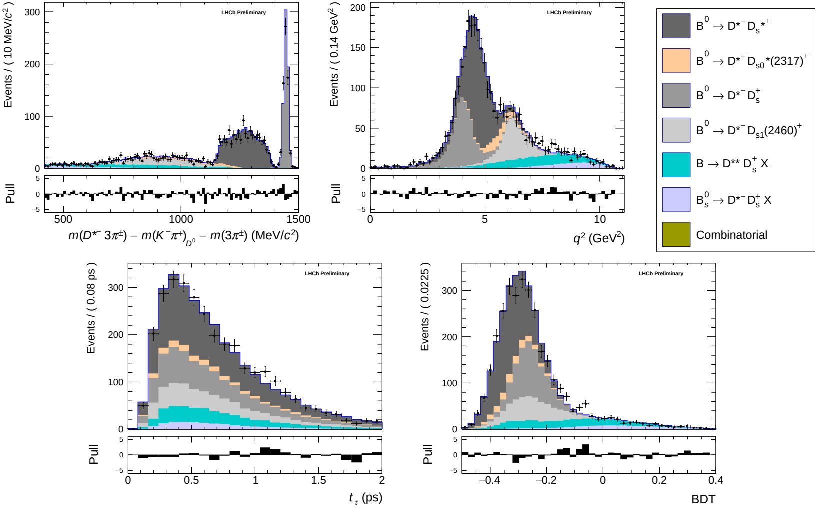
Control Samples

D_s decay model



Control Samples

$B \to D D_s^+(X)$ control mode



Results and implications

Signal fit

The total probability density function is:

$$\mathcal{P}_{\text{total}}(q^2, t_\tau, \text{BDT}) = 1/N_{\text{total}} \times \{N_{\text{sig}}[f\mathcal{P} + (1-f)\mathcal{P} + f_{\tau\nu}\mathcal{P}]$$
$$+ N^{\text{same}}_{D^0}[\mathcal{P}_{B\to D^{*-}D^0X\ \text{SV}} + f^{v_1-v_2}_{D^0}\mathcal{P}_{B\to D^{*-}D^0X\ \text{DV}}] + N_{D_s^+}/k \times [\mathcal{P}_{B^0\to D^{*-}D_s^{*+}} + f_{D_s^+}\mathcal{P}_{B^0\to D^{*-}D_s^+}$$
$$+ f\mathcal{P}_{B^0\to D^{*-}} + f\mathcal{P}_{B^0\to D^{*-}} + f_{D^{**}D_sX}\mathcal{P}_{B\to D_s^+X} + f_{B_s\to D^*D_s^+X}\mathcal{P}_{B_s^0\to D^{*-}D_s^+X}]$$
$$+ N_{D_s^+}f_{D^+}\mathcal{P}_{B\to D^{*-}D^+X} + N_{B\to D^{*-}3\pi^\pm X}\mathcal{P}_{B\to D^{*-}3\pi^\pm X} + N_{B_1-B_2}\mathcal{P}_{\text{combinatoric }B} + N_{\text{fake }D^0}\mathcal{P}_{\text{combinatoric }D^0}$$
$$+ N_{\text{fake }D^*}\mathcal{P}_{\text{combinatoric }D^{*-}}\},$$

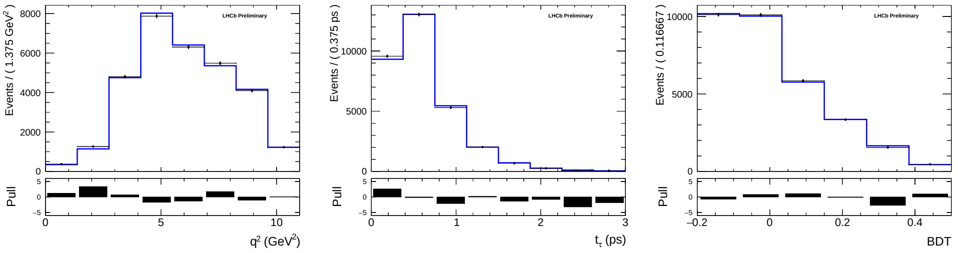
Where N_{total} is the total number of events and $k = 1 + f_{D_s^+} + f + f + f_{D^{**}D_sX} + f_{B_s\to D^*D_s^+X}$.

The 16 fit parameters can be grouped into three categories depending whether they are free, Gaussian-constrained or fixed in the fit:

- **free parameters (5)**: N_{sig}, $N_{D_s^+}$, f_{D^+} and $f^{v_1-v_2}_{D^0}$
- **Gaussian-constrained parameters (5)**: $N_{B\to D^{*-}3\pi^\pm X}$,$f_{D_s^+}$, f, f, $f_{D^{**}D_sX}$, $f_{B_s\to D^*D_s^+X}$
- **fixed parameters (6)**: $N_{B_1-B_2}$, $N_{\text{fake }D^0}$, $N_{\text{fake }D^*}$, $N^{\text{same}}_{D^0}$, $f_{\tau\nu}$ and f

Signal fit results

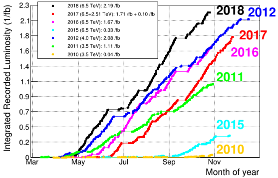
The signal yield is determined from a 3-dimensional maximum likelihood binned fit to q^2 (8 bins), decay time of the τ^+-candidate (8 bins), t_τ, and the anti-D_s^+ BDT (6 bins).



Signal fit projection of q2 (left), τ lifetime and the Anti-D_s BDT

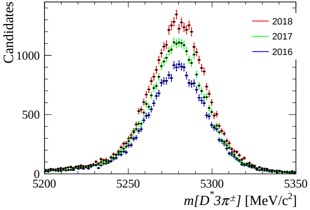
Preliminary results for $R(D^*)$ for Run 2 part 1 (2015-2016) (Under internal review)
$R(D^*)$ for Run 2 part 2 (2017-2018) (ongoing analysis)

Run 1 and Run 2



LHCb cumulative integrated luminosity from 2010 till 2018

Year	$\mathcal{L}$ (fb^{-1})	E (TeV)	Trigger efficiency
2011	1.11	3.5	
2012	2.08	4	
2015	0.33	6.5	increase
2016	1.67	6.5	by a
2017	1.71	6.5	factor of
2018	2.19	6.5	15%



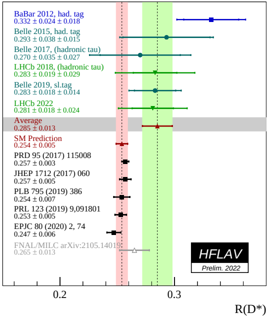
Mass distribution comparison of $D^* 3\pi^\pm$ between different datasets and after applying normalisation cuts

- **Run1** (2011+2012): $R(D^*) = 0.283 \pm \underbrace{0.019}_{6.7\%\ \text{stat.}} \pm \underbrace{0.029}_{10.3\%\ \text{syst.}}$ (PRL.120.171802)
- **Run2p1** (2015+2016): Blinded $R(D^*)$ (Resmi (CPPM) et al. LHCb internal review)
- **Run2p2** (2017+2018): Statistics increased by a factor > 2 w.r.t. 2016 (stat. uncertainty $\sim 3\%$)

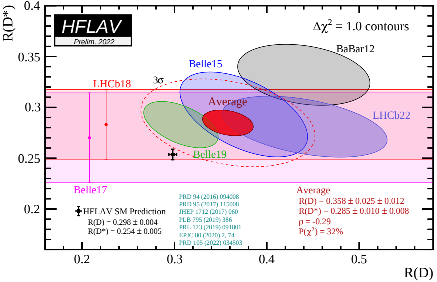
LFU Anomalies: $R(D)$ vs $R(D^*)$ plot

$$R(D) \equiv \frac{\mathcal{B}(B^0 \to D^- \tau^+ \nu_\tau)}{\mathcal{B}(B^0 \to D^- \mu^+ \nu_\mu)} \quad \text{and} \quad R(D^*) \equiv \frac{\mathcal{B}(B^0 \to D^{*-} \tau^+ \nu_\tau)}{\mathcal{B}(B^0 \to D^{*-} \mu^+ \nu_\mu)}$$

3 experiments, 6 measurements, different analysis techniques



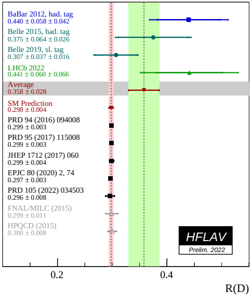
[Heavy Flavor Averaging Group, HFLAV]

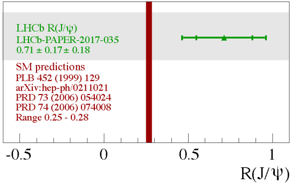


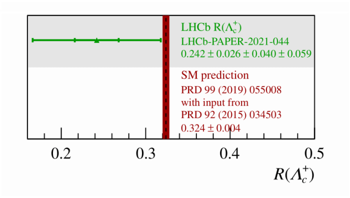
All the measurements lie above the SM expectation

The current world-average measured $R(D)$ and $R(D^*)$ are **3.2σ** away from the SM

Other R ratios





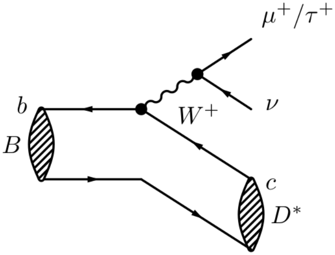


$$R(D) \equiv \frac{\mathcal{B}(B^0 \to D^- \tau^+ \nu_\tau)}{\mathcal{B}(B^0 \to D^- \mu^+ \nu_\mu)}$$

$$R(J/\Psi) = \frac{\mathcal{B}(B_c^+ \to J/\Psi \tau \nu_\tau)}{\mathcal{B}(B_c^+ \to J\Psi \mu^+ \nu_\mu)}$$

$$R(\Lambda_c^+) = \frac{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ \tau \nu_\tau)}{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ \mu \nu_\tau)}$$

New Physics behind LFU?



Contributions for B to D^* semileptonic decay: Left: SM, Middle: Charged Higgs, Right: Leptoquark.

Conclusions

Summary and Prospects

Summary

- $R(D^*)$ is an important analysis that may shed light on the intriguing Lepton Flavor Anomaly
- Complicated analysis because of many backgrounds due to proton-proton collision and partially reconstructed signal
- Analysis of $R(D^*)$ hadronic for Run 2 part 1 (2015-2016) in final step of internal review

Prospects

- Complete Run 2 part 2 (2017-2018) analysis
- Use the full dataset Run1 + Run2 (2011-2018) to determine $R(D^*)$
  - Expected statistical uncertainty of the order of 3%
    (while: 6.7% for run1 only and for 4.3% run1+run2p1)
- $R(D)$ and simultaneous measurements $R(D)$ and $R(D^*)$
- Angular analysis study of $B_d \to D^* \tau \nu_\tau$ decay
- Study other Beauty mesons

Thank you!

Backups

$R(D^*)$ hadronic ($\tau \to 3\pi\nu_\tau$) Systematics (run1) [PRL 120, 171802 2018]

$$R(D^*) = 0.280 \pm 0.018(\text{stat}) \pm 0.026(\text{syst}) \pm 0.013(\text{ext})$$

Breakdown of relative uncertainties:

Source	$\frac{\delta R(D^{*-})}{R(D^{*-})}$ [%]
orange Simulated sample size	4.7
Empty bins in templates	1.3
Signal decay model	1.8
D^{**} τ ν and D_s^{**} τ ν feed-downs	2.7
$D_s^+ \to 3\pi X$ decay model	2.5
lightblue $B \to D^{*-} D_s^+ X$, $D^{*-} D^+ X$, $D^{*-} D^0 X$ backgrounds	3.9
Combinatorial background	0.7
$B \to D^{*-} 3\pi X$ background	2.8
lightblue Efficiency ratio	3.9
Normalization channel efficiency (modeling of $B^0 \to D^{*-} 3\pi$)	2.0
Total systematic uncertainty	9.1
Statistical uncertainty	6.5
Total uncertainty	11.9

Will be improved in the next iteration of the analysis

$R(D^*)$ hadronic ($\tau \to 3\pi\nu_\tau$) Systematics [PRL 120, 171802 2018]

$$R(D^*) = 0.280 \pm 0.018(\text{stat}) \pm 0.026(\text{syst}) \pm 0.013(\text{ext})$$

Breakdown of relative uncertainties:

Source	$\frac{\delta R(D^{*-})}{R(D^{*-})}$ [%]	Future
orange Simulated sample size	4.7	Produce more MC !
Empty bins in templates	1.3	
Signal decay model	1.8	
D^{**} τ ν and D_s^{**} τ ν feed-downs	2.7	Measure $R(D_1(2420)^0)$
$D_s^+ \to 3\pi X$ decay model	2.5	BESIII
lightblue $B \to D^{*-} D_s^+ X,\ D^{*-} D^+ X,\ D^{*-} D^0 X$ bkgs	3.9	Improves with stat
Combinatorial background	0.7	
$B \to D^{*-} 3\pi X$ background	2.8	Kill with $\lvert z\tau - zD \rvert > 5\sigma$
lightblue Efficiency ratio	3.9	Improves with stat
Normalization channel efficiency (modeling of $B^0 \to D^{*-} 3\pi$)	2.0	
Total systematic uncertainty	9.1	
Statistical uncertainty	6.5	
Total uncertainty	11.9	

$R(D^*)$ hadronic ($\tau \to 3\pi\nu_\tau$) [PRD 97,072013 2018]

List of the individual systematic uncertainties for $R(D^*)$:

Contribution	Value in %
$\mathcal{B}(\tau^+ \to 3\pi\overline{\nu}_\tau)/\mathcal{B}(\tau^+ \to 3\pi(\pi^0)\overline{\nu}_\tau)$	0.7
Form factors (template shapes)	0.7
Form factors (efficiency)	1.0
τ polarization effects	0.4
Other τ decays	1.0
$B \to D^{**}\tau^+\nu_\tau$	2.3
$B_s^0 \to D_s^{**}\tau^+\nu_\tau$ feed-down	1.5
$D_s^+ \to 3\pi X$ decay model	2.5
D_s^+, D^0 and D^+ template shape	2.9
$B \to D^{*-}D_s^+(X)$ and $B \to D^{*-}D^0(X)$ decay model	2.6
$D^{*-}3\pi X$ from B decays	2.8
Combinatorial background (shape + normalization)	0.7
Bias due to empty bins in templates	1.3
Size of simulation samples	4.1
Trigger acceptance	1.2
Trigger efficiency	1.0
Online selection	2.0
Offline selection	2.0
Charged-isolation algorithm	1.0
Particle identification	1.3
Normalization channel	1.0
Signal efficiencies (size of simulation samples)	1.7
Normalization channel efficiency (size of simulation samples)	1.6
Normalization channel efficiency (modeling of $B^0 \to D^{*-}3\pi$)	2.0
Total uncertainty	9.1

Generated MC samples

Decay descr.	EventType	Generated evts [M]	Filtered evts [M]
$B^0 \to D^{*-}\tau^+(\to 3\pi^\pm\overline{\nu}_\tau)\nu_\tau$	11160001	93.0	0.4
$B^0 \to D^{*-}\tau^+(\to 3\pi^\pm\pi^0\overline{\nu}_\tau)\nu_\tau$	11563020	90.0	0.3
$B^0 \to D^{*-}3\pi^\pm$	11266018	363.0	1.0
$B^0 \to D^{**}\tau^+(\to 3\pi^\pm\overline{\nu}_\tau)\nu_\tau$	11566431	7.0	0.09
$B_s^0 \to D^{*-}D_s^+X$	13996612	50.0	0.4
$B^+ \to D^{**0}D_s^+X$	12997613	354.0	4.0
$B^0 \to D^{*-}D_s^+X$	11896612	692.0	8.0
$B^0 \to D^{**-}D_s^+X$	11996413	42.0	0.4
$b\overline{b} \to D^{*-}3\pi^\pm X$	27163970	8202.0	16.0
$b\overline{b} \to D^{*-}D^{\{0,+\}}X$	27163971	427.0	2.0
Total	-	10320.0	32.0

$R(D^*)$ hadronic kinematics

Two-fold ambiguities in determining τ momentum:

$$|\vec{p}_\tau| = \frac{(m_{3\pi}^2 + m_\tau^2)|\vec{p}_{3\pi}|\cos\theta_{\tau,3\pi} \pm E_{3\pi}\sqrt{(m_\tau^2 - m_{3\pi}^2)^2 - 4m_\tau^2|\vec{p}_{3\pi}|^2\sin^2\theta_{\tau,3\pi}}}{2(E_{3\pi}^2 - |\vec{p}_{3\pi}|^2\cos^2\theta_{\tau,3\pi})},$$

where $\theta_{\tau,3\pi}$ is the angle between the 3π system three-momentum and the τ line of flight.

Approximation: take the maximum allowed angle

$$\theta_{\tau,3\pi} \approx \theta_{\tau,3\pi}^{\max} = \arcsin\left(\frac{m_\tau^2 - m_{3\pi}^2}{2m_\tau|\vec{p}_{3\pi}|}\right),$$

The B^0 momentum is obtained similarly:

$$|\vec{p}_{B^0}| = \frac{(m_Y^2 + m_{B^0}^2)|\vec{p}_Y|\cos\theta_{B^0,Y} \pm E_Y\sqrt{(m_{B^0}^2 - m_Y^2)^2 - 4m_{B^0}^2|\vec{p}_Y|^2\sin^2\theta_{B^0,Y}}}{2(E_Y^2 - |\vec{p}_Y|^2\cos^2\theta_{B^0,Y})}$$

with

$$\theta_{B^0,Y}^{\max} = \arcsin\left(\frac{m_{B^0}^2 - m_Y^2}{2m_{B^0}|\vec{p}_Y|}\right),$$

where Y represents the $D^{*-}\tau^+$ system.

Properties of charged leptons

Particle	Mass (MeV/c^2)	Lifetime	Main decay modes
e^-	0.5109989461(31)	$>6.6\times10^{26}$ years	None
μ^-	105.6583745(24)	2.1969811(22) μs	$e^-\bar{\nu}_e\nu_\mu$
τ^-	1776.86(12)	290.3(5) fs	$\pi^-\pi^0\nu_\tau$ (25.5%) $e^-\bar{\nu}_e\nu_\tau$ (17.8%) $\mu^-\bar{\nu}_\mu\nu_\tau$ (17.39%) $\pi^-\nu_\tau$ (10.8%) $\pi^-\pi^+\pi^-\nu_\tau$ (9.3%)

τ lepton Branching Ratios [PDG 2018]

Mode	$\mathcal{BR}$ (%)
$\tau^- \to \pi^-\pi^0\nu_\tau$	25.49 ± 0.09
$\tau^- \to e^-\bar{\nu}_e\nu_\tau$	17.82 ± 0.04
$\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$	17.39 ± 0.04
$\tau^- \to \pi^-\nu_\tau$	10.82 ± 0.05
$\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$	9.31 ± 0.05
$\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$	4.62 ± 0.05

D^* branching ratios

Mode	BR
$D^*(2007)^0 \to D^0\pi^0$	$(64.7 \pm 0.9)\%$
$D^*(2007)^0 \to D^0\gamma$	$(35.3 \pm 0.9)\%$
$D^*(2010)^+ \to D^0\pi^+$	$(67.7 \pm 0.5)\%$
$D^*(2010)^+ \to D^+\pi^0$	$(30.7 \pm 0.5)\%$
$D^*(2010)^+ \to D^+\gamma$	$(1.6 \pm 0.4)\%$

Particle	Mass (MeV/c^2)	Lifetime
D^+	1869.65 ± 0.05	(1.040 ± 0.007) ps
D^0	1864.83 ± 0.05	(0.4101 ± 0.0015) ps
D_s^+	1968.34 ± 0.07	(0.504 ± 0.004) ps
Λ_c^+	2286.46 ± 0.14	(0.200 ± 0.006) ps
$D^*(2007)^0$	2006.85 ± 0.05	-
$D^*(2010)^-$	2010.26 ± 26	-

Relevant branching ratios

Mode	$\mathcal{BR}$
$B^0 \to D^*(2010)^- D_s^+$	$(8.0 \pm 1.1) \times 10^{-3}$
$B^0 \to D^*(2010)^- D_s^{*+}$	$(1.77 \pm 0.14) \times 10^{-2}$
$B^0 \to D^*(2010)^- D^0 K^+$	$(2.47 \pm 0.10 \pm 0.18) \times 10^{-3}$
$B^0 \to D^*(2010)^- D^*(2007) K^+$	$(10.6 \pm 0.33 \pm 0.86) \times 10^{-3}$
$B^0 \to D^*(2010)^- \pi^+ \pi^+ \pi^- \pi^0$	$(1.67 \pm 0.27)\%$
$B^0 \to D^*(2010)^- 3\pi^+ \pi^+ 2\pi^-$	$(4.7 \pm 0.9)\%$
$B^0 \to D^*(2010)^- D_{s0}(2317)^+$	$(1.5 \pm 0.6)\%$
$B^0 \to D^*(2010)^- D_{sJ}(2457)^+$	$(9.3 \pm 2.2) \times 10^{-3}$
$B^0 \to D^*(2010)^- D_{s1}(2536)^+,\ D_{s1}^+ \to D^{*0} K^+ + D^{*+} K^0$	$(5.0 \pm 1.4) \times 10^{-3}$

Generated events 2017 and 2018

Decay descr.	EventType	Generated evts [M]	Filtered evts [M]
$B^0 \to D^{*-}\tau^+(\to 3\pi^\pm\overline{\nu}_\tau)\nu_\tau$	11160001	65.0	0.3
$B^0 \to D^{*-}\tau^+(\to 3\pi^\pm\pi^0\overline{\nu}_\tau)\nu_\tau$	11563020	60.0	0.2
$B^0 \to D^{*-}3\pi^\pm$	11266018	269.0	1.0
$B^0 \to D^{**}\tau^+(\to 3\pi^\pm\overline{\nu}_\tau)\nu_\tau$	11566431	6.0	0.1
$B_s^0 \to D^{*-}D_s^+X$	13996612	35.0	0.3
$B^+ \to D^{**0}D_s^+X$	12997613	242.0	3.0
$B^0 \to D^{*-}D_s^+X$	11896612	472.0	6.0
$B^0 \to D^{**-}D_s^+X$	11996413	29.0	0.3
$b\bar{b} \to D^{*-}3\pi^\pm X$	27163970	5688.0	13.0
$b\bar{b} \to D^{*-}D^{\{0,+\}}X$	27163971	302.0	1.0
Total	-	7170.0	26.0

Decay descr.	EventType	Generated evts [M]	Filtered evts [M]
$B^0 \to D^{*-}\tau^+(\to 3\pi^\pm\overline{\nu}_\tau)\nu_\tau$	11160001	93.0	0.4
$B^0 \to D^{*-}\tau^+(\to 3\pi^\pm\pi^0\overline{\nu}_\tau)\nu_\tau$	11563020	90.0	0.3
$B^0 \to D^{*-}3\pi^\pm$	11266018	363.0	1.0
$B^0 \to D^{**}\tau^+(\to 3\pi^\pm\overline{\nu}_\tau)\nu_\tau$	11566431	7.0	0.09
$B_s^0 \to D^{*-}D_s^+X$	13996612	50.0	0.4
$B^+ \to D^{**0}D_s^+X$	12997613	354.0	4.0
$B^0 \to D^{*-}D_s^+X$	11896612	692.0	8.0
$B^0 \to D^{**-}D_s^+X$	11996413	42.0	0.4
$b\bar{b} \to D^{*-}3\pi^\pm X$	27163970	8202.0	16.0
$b\bar{b} \to D^{*-}D^{\{0,+\}}X$	27163971	427.0	2.0
Total	-	10320.0	32.0

Selection efficiencies

Cut	Absolute efficiencies			Cumulative efficiencies		
	$3\pi\overline{\nu}_\tau$	$3\pi\pi^0\overline{\nu}_\tau$		$3\pi\overline{\nu}_\tau$	$3\pi\pi^0\overline{\nu}_\tau$	
Initial selection						
L0	89.51	86.60	89.08	89.51	86.60	89.08
Hlt1	89.76	87.32	90.92	87.14	83.88	88.02
Hlt2	79.90	77.31	90.33	73.25	69.02	85.10
$PV(\overline{D}^0) = PV(\tau^+)$	69.76	65.73	79.94	69.76	65.73	79.94
totCandidates $= 1$	60.89	52.22	71.97	58.06	49.87	67.75
$[\text{vtx}_z(\tau^+) - \text{vtx}_z(PV)]/\text{error} > 10$	71.66	66.59	78.60	57.01	48.29	62.64
nSPDHits < 450	72.24	67.78	83.97	56.37	47.56	61.99
Signal selection						
$m(D^{*-}) - m(K^-\pi^+) \in [143, 148]\ \text{MeV}/c^2$	94.63	93.98	-	94.63	93.98	-
$m(K^-\pi^+) \in [1840, 1890]\ \text{MeV}/c^2$	97.36	97.39	-	92.28	91.70	-
$m(3\pi) < 1825\ \text{MeV}/c^2$	98.24	98.77	-	90.73	90.68	-
$m(B^0) < 5100\ \text{MeV}/c^2$	99.29	99.03	-	90.46	90.27	-
$q^2 \in [0, 12]\ \text{GeV}^2/c^4$	97.52	97.22	-	88.74	88.53	-
combinatorial BDTD > 0	80.37	76.71	-	74.72	71.89	-
$[\text{vtx}_z(\tau^+) - \text{vtx}_z(B^0)]/\text{error} > 2$	99.81	99.78	-	74.72	71.89	-
isolation BDT > 0	87.85	83.86	-	67.42	62.41	-
anti D_s^+ BDT > -0.2	98.30	86.10	-	67.12	54.87	-
PID	76.23	78.86	-	-	-	-
Normalisation selection						
$[\text{vtx}_z(\overline{D}^0) - \text{vtx}_z(\tau^+)]/\text{error} > 4$	-	-	94.30	-	-	94.30
$m(D^*3\pi^\pm) \in [5150, 5400]\ \text{MeV}$	-	-	97.87	-	-	93.32
$m(D^{*-}) - m(\overline{D}^0) \in [143, 148]\ \text{MeV}$	-	-	94.97	-	-	89.04
combinatorial BDTD > 0	-	-	81.37	-	-	74.19
isolation BDT > 0	-	-	88.33	-	-	66.94
PID	-	-	73.96	-	-	-

Signal fit templates

The signal yield is determined from a 3-dimensional maximum likelihood binned fit to q^2 (8 bins), decay time of the τ^+-candidate (8 bins), t_τ, and the anti-D_s^+ BDT (6 bins).
There are 16 templates, 13 of them come from MC and three from data. The latter ones are combinatorial B, D^0 and D^* events. The templates are grouped into the 12 following categories, due to similar shapes:

- $B^0 \to D^{*-}\tau^+\nu_\tau$ – the signal; includes $\tau^+ \to 3\pi^-\overline{\nu}_\tau$ and $\tau^+ \to 3\pi^\pm\pi^0\overline{\nu}_\tau$
- $B^0 \to D^{**-}\tau^+\nu_\tau$ – excited D^{*-} states
- $B^0 \to D^{*-}D_s^{+(*)}$ – includes $B^0 \to D^{*-}D_s^+$, $B^0 \to D^{*-}D_s^{+*}$, $B^0 \to D^{*-}$ and $B^0 \to D^{*-}$
- $B \to D_s^+X$
- $B_s^0 \to D^{*-}D_s^+X$
- $B \to D^{*-}D^+X$
- $B \to D^{*-}3\pi^\pm X$

Signal fit templates

The signal yield is determined from a 3-dimensional maximum likelihood binned fit to q^2 (8 bins), decay time of the τ^+-candidate (8 bins), t_τ, and the anti-D_s^+ BDT (6 bins).
There are 16 templates, 13 of them come from MC and three from data. The latter ones are combinatorial B, D^0 and D^* events. The templates are grouped into the 12 following categories, due to similar shapes:

- $B \to D^{*-} D^0 X$ SV, 'Same Vertex' where all 3 pions come from D^0
- $B \to D^{*-} D^0 X$ DV, 'Different Vertices' where at least 1 of the 3 pions comes from the D^0 vertex and the other(s) from a different vertex, *e.g.* the slow pion from D^{*-} is reconstructed as coming from the D^0
- combinatorial B^0 whose template is made from the collision data with the $D^{*\pm}$ of the same sign as the $3\pi^\pm$ system (*i.e.* wrong sign data w.r.t. the signal)
- combinatorial D^0
- combinatorial D^{*-} but genuine D^0

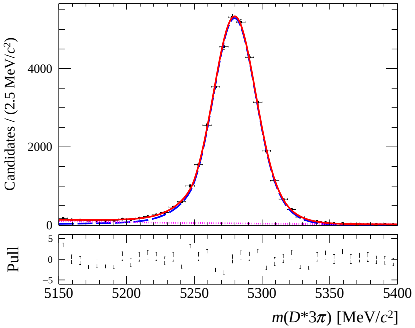
NTuples production

- 2016 (real data + MC) dataset (reproduced as a cross-check)
- 2017-2018 (real data + MC) dataset (new production)

Year	2016	2017	2018
# MC event types	14	13	13
# MC events [M]	21.8	39.8	49.4
# data events [M]	7	7.2	9.1

In total, we generate $\mathcal{O}(100M)$ events.

Normalisation fit result



Fit results

Sweight fit result at the $D^*3\pi^{\pm}$ mass peak using 2018 dataset

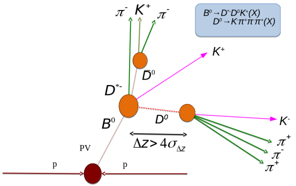
year	N_{norm}	N_{norm} expected	N_{bkg}	N_{bkg} expected	$N_{\text{norm}}/N_{\text{bkg}}$
2016	$26\,434 \pm 190$	-	$1\,446 \pm 107$	-	18.28
2017	$31\,200 \pm 207$	27 067	$2\,002 \pm 117$	1 481	15.58
2018	$37\,137 \pm 225$	34 664	$2\,170 \pm 126$	1 896	17.11

B^0 and τ vertex error reweighting

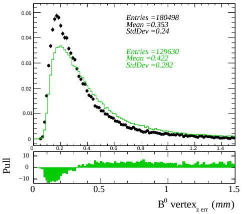
- **Main background** Prompt decay where $3\pi^\pm$ system comes directly from the B^0 vertex.
- To suppress this background we require a separation between the B^0 and τ vertex

$$\Delta z = vtx_z(\tau^+) - vtx_z(B^0) \qquad \sigma_{\Delta_z} = \sqrt{vtx_{err\ z}(\tau^+)^2 + vtx_{err\ z}(B^0)^2}$$

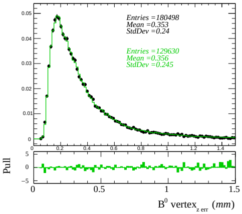
- Difference between 'run2p1' and 'run2p2' MC: B^0 vertex error in the beam direction
  - Due to reconstruction algorithms (applied for data but not for MC)
  - Need to apply simultaneous gradient-boosted weight for the B^0 and τ vertex error



Signal mode



(a) No reweighting



(b) Reweighted MC

B^0 vertex$_z$ error from sweighted data and $D^*3\pi^\pm$ MC 2018 sample