"Test of Lepton Flavor Universality by the measurement of the $B_0 \to D^{*-}\tau^+\nu_\tau$ branching fraction using three-prong τ decays" by LHCb collaboration

Pedro M. F. Pereira

Instituto Superior Técnico

pedromanuelpereira@tecnico.ulisboa.pt

June 8, 2020

Overview

Motivation

- Measurement of the ratio $\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)/\mathcal{B}(B^0 \to D^{*-}\pi^+\pi^-\pi^+)$
- Using world average of $\mathcal{B}(B^0 \to D^{*-}\pi^+\pi^-\pi^+)$, derivation of $\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)$
- Using well-measured $\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)$ to test Lepton Flavour Universality (LFU) by computing $\mathcal{R}(D^{*-}) \equiv \mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)/\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)$

Who are these?

1. B^0 meson composed of a bottom antiquark, $\overline{b}$, and a down quark, d.
2. D^{*-} excited D^- meson composed of a down quark, d, and a charm antiquark, $\overline{c}$. D mesons are the lightest mesons that contain a charm quark, c.
3. $\pi^\pm$, π^0 are the lightest mesons, composed of: u$\overline{d}$ (π^+), d$\overline{u}$ (π^-) and u$\overline{u}$ or d$\overline{d}$ (π^0).

Introduction I: LFU

What is Lepton Flavour Universality (LFU)?

- Accidental Symmetry of the SM before SSB and without Yukawa interactions
- Broken by Yukawa Interactions, masses.
- Differences between given branching fractions in the SM are well established and can only originate from the different masses of leptons.
- Further deviations are a sign of BSM Physics.

Introduction II: What is new in this measurement?

- Previous measurements of tests of LFU reported by BaBar and Belle
- Previous measurement of $\mathcal{R}(D^*)$ by LHCb with results compatible with BaBar and Belle
- The average of all these $\mathcal{R}(D^*)$ measurements is in tension with the SM expectation at 3.3 standard deviations. All these $\mathcal{R}(D^{(*)-,0})$ measurements yield values that are above the SM predictions with a combined significance of 3.9 standard deviations.

Not the first time this is measured but,

- First time using the τ decay with three charged particles (three-prong) in the final state.
- **Signal:** $B^0 \to D^{*-}\tau^+\nu_\tau$
  Decays used to reconstruct τ: $\tau^+ \to \pi^+\pi^-\pi^+\overline{\nu}_\tau$ and $\tau^+ \to \pi^+\pi^-\pi^+\pi^0\overline{\nu}_\tau$
  D^{*-} reconstructed using $D^{*-} \to \overline{D}^0\pi^- \to (K^+\pi^-)\pi^-$[1].

[1] K^+ is a meson composed of $u\bar{s}$

Introduction III: Why is this method good?

- The three-prong τ decay modes lead to measurements with a better signal-to-background ratio and statistical significance.
- The absence of charged leptons in the final state avoids backgrounds originating from semileptonic decays of b or c hadrons.
- The three-prong topology enables the precise reconstruction of a τ decay vertex detached from the B_0 decay vertex due to the non zero τ lifetime, thereby allowing the discrimination between signal decays and the most abundant background due to $B \to D^{*-}\pi^+\pi^-\pi^+X$.
- The requirement of a 3π decay vertex detached from the B vertex suppresses the $D^{*-} \to 3\pi X$ background by three orders of magnitude, while retaining about 40% of the signal.
- Only one neutrino is produced in the τ decay, the measurements of the B_0 and τ lines of flight allow the determination of the complete kinematics of the decay, up to two quadratic ambiguities, leading to four solutions.

Introduction III: Dominant Background

Double-Charm: B decays with a D^{*-} and another charm hadron in the final state.

- Largest component: $B \to D^{*-} D_s^+(X)$. Same topology, D_s^+ decay vertex detached from B vertex.
- Suppressed by: Applying vetoes on the presence of additional particles around the direction of the τ and B candidates, and exploiting the different resonant structure of the 3π system in τ^+ and D_s^+ decays.

Measured quantity:

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

The absolute branching fraction is obtained as $\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau) = \mathcal{K}(D^{*-}) \times \mathcal{B}(B^0 \to D^{*-}3\pi)$, where the branching fraction of the $\mathcal{B}(B^0 \to D^{*-}3\pi)$ decay is taken by averaging the previous available measurements. A value for $\mathcal{R}(D^{*-})$ is then derived by using the branching fraction of the $\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)$ decay from previous works.

Detector and Simulation

- The LHCb detector is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing b or c quarks. Photons and Different types of charged particles can be distinguished.
- Simulations performed using *Pythia, EVTGEN, PHOTOS, TAUOLA, GEANT4*
- Trigger: Hardware stage and software stage.

At the hardware trigger stage, candidates are required to have a muon with high p_T or a hadron, photon or electron with high transverse energy. The software trigger requires a two-, three-, or four-track secondary vertex with significant displacement from any primary vertex (PV) consistent with the decay of a b hadron, or a two-track vertex with a significant displacement from any PV consistent with a $\overline{D}^0 \to K^+\pi^-$ decay.

Selection Criteria and Multivariate Analysis

The signal selection proceeds in two main steps:

1. The dominant background, consisting of candidates where the 3π system originates from the B_0 vertex, called prompt hereafter, is suppressed by applying a π detached-vertex requirement.
2. The double-charm background is suppressed using a multivariate analysis (MVA). This is the only background with the same vertex topology as the signal.

Because of the significant τ lifetime and boost along the forward direction, the 3π system is detached from the B_0 vertex, as shown in Fig. 1. The requirement for the detached vertex is that the distance between the 3π and the B_0 vertices along the beam direction,$\Delta z = z(3\pi) - z(B_0)$, is greater than four times its uncertainty, $\sigma\Delta z$. This leads to an improvement in the signal to noise ratio by a factor 160, c.f., Fig 2.

The detached-vertex topology

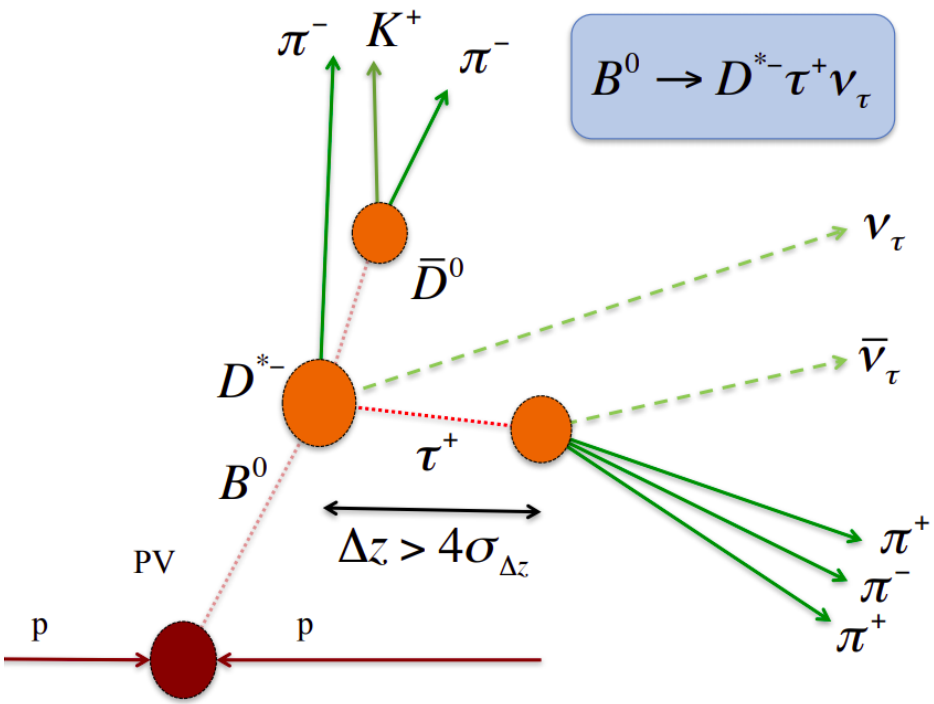

Figure 1: Topology of the signal decay.

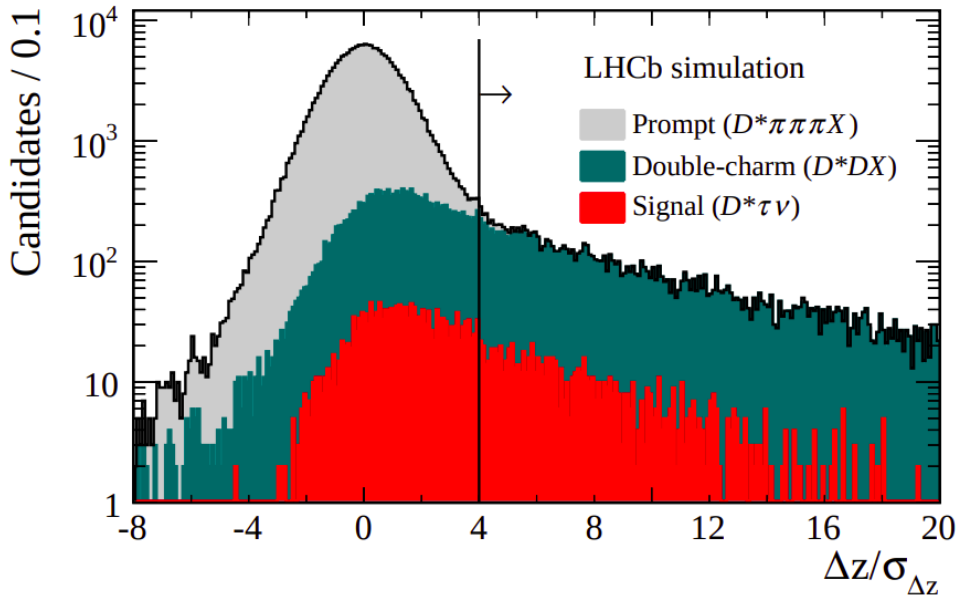

Figure 2: Distribution of the distance between the B_0 vertex and the 3π vertex along the beam direction, divided by its uncertainty, obtained using simulation.

The vertical line in Fig. 2 shows the $4\sigma\Delta z$ requirement used in the analysis to reject the prompt background component.

Isolation and Particle Identification Requirements

- A Isolation algorithm ensures that no extra tracks are compatible with either the B^0 or 3π decay vertices. Decays with additional neutral particles, for instance $B_0 \to D^{*-} D^0(X)$, are suppressed by requiring that an energy of at most 8 GeV is deposited in the electromagnetic calorimeter around the 3π direction.[2]
- In order to ensure that the tracks forming the 3π candidate are real pions, a positive pion identification is required and optimized taking into account the efficiency and rejection performance of particle identification (PID) algorithms, and the observed kaon to pion ratio in the 3π candidates, as measured through the D^- peak when giving a kaon mass to the negatively charged pion.

All of these PID requirements are chosen in order to get the best discrimination between signal and background. They form, together with the topology selection and the isolation requirement defined above, the final selection.

[2] This has an impact on signal, c.f., $\tau^+ \to 3\pi\pi^0\overline{\nu}_\tau$

How to reject double-charm?

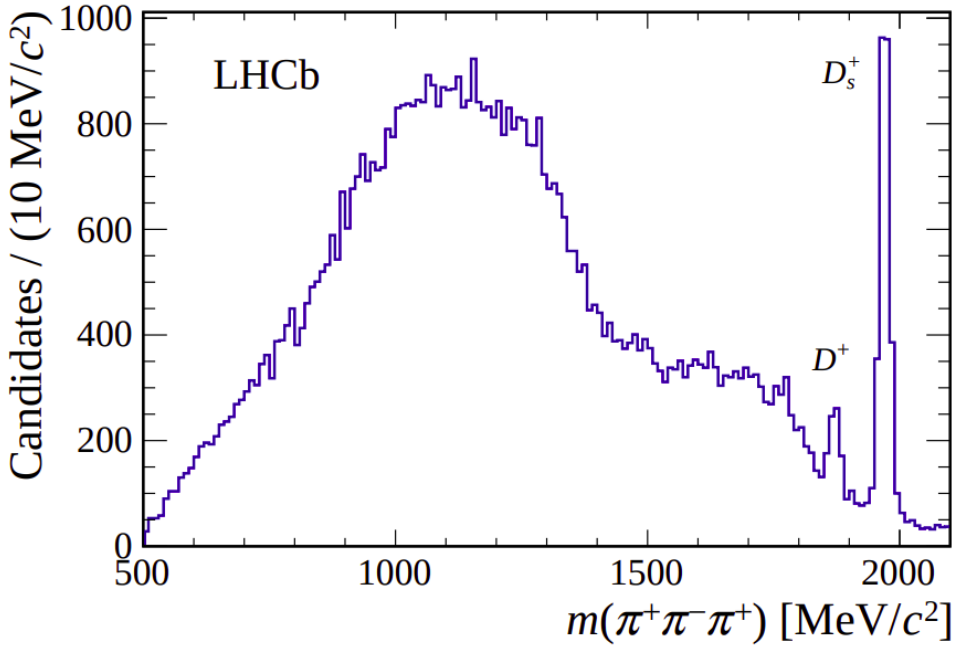

Figure 3: Distribution of the 3π mass for candidates after the detached-vertex requirement. The D^+ and D_s^+ mass peaks are indicated.

The double-charm $B \to D^{*-}D(X)$ decays are the only other B decays with the same vertex topology as the signal.

$q^2 \equiv (p_{B_0} - p_{D^{*-}})^2 = (p_\tau + p_{\nu_\tau})^2$

τ decay time: t_τ.

Rejecting double-charm: Multivariate Analysis

To suppress double-charm background, a set of 18 variables is used as input to a MVA based upon a Boosted Decision Tree (BDT).
This set is as follows: the output variables of the neutral isolation algorithm; momenta, masses and quality of the reconstruction of the decay chain under the signal and background hypotheses; the masses of oppositely charged pion pairs, the energy and the flight distance in the transverse plane of the 3 system; the mass of the six-charged-tracks system.
The BDT is trained using simulated samples of signal and double-charm background decays. The $B \to D^{*-} D_s^+(X)$, $B \to D^{*-} D_0(X)$ and $B \to D^{*-} D^+(X)$ control samples are used to validate the BDT.
The signal yield is determined from candidates in the region where the BDT output is greater than 0.075. According to simulation, this value gives the best statistical power in the determination of the signal yield.

Rejecting double-charm II

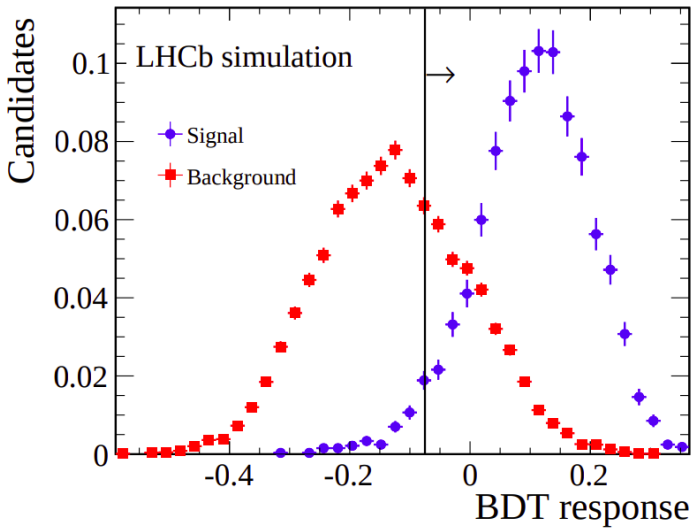

Figure 4: Distribution of the BDT response on the signal and background simulated samples

Candidates with the BDT output less than -0.075 are highly enriched in D_s^+ decays and contain very little signal, as shown in Fig. 4, and represent about half of the total data sample.

They are used to validate the simulation of the various components in $D_s^+ \to 3\pi X$ decays used in the parametrization of the templates entering in the fit that determines the signal yield.

Study of double-charm candidates

The fit that determines the signal yield uses templates that are taken from simulation. It is therefore of paramount importance to verify the agreement between data and simulation for the remaining background processes.
To study:

1. $B_0 \to D^{*-} D_s^+$ (X)
2. $B_0 \to D^{*-} D^0$ (X)
3. $B_0 \to D^{*-} D^+$ (X)

The relative contributions of double-charm backgrounds and their q^2 distributions from simulation are validated, and corrected where appropriate, by using data control samples enriched in such processes. Inclusive decays of D_0, D^+, D_s^+ mesons to 3π are also studied in this way.

Study of double-charm candidates: D_s^+

After a study of the possible decays of D_s^+, one can obtain a pure sample of $B \to D^- D_s^+(X)$ decays by using candidates where the D_s^+ meson decays exclusively to the 3π final state.

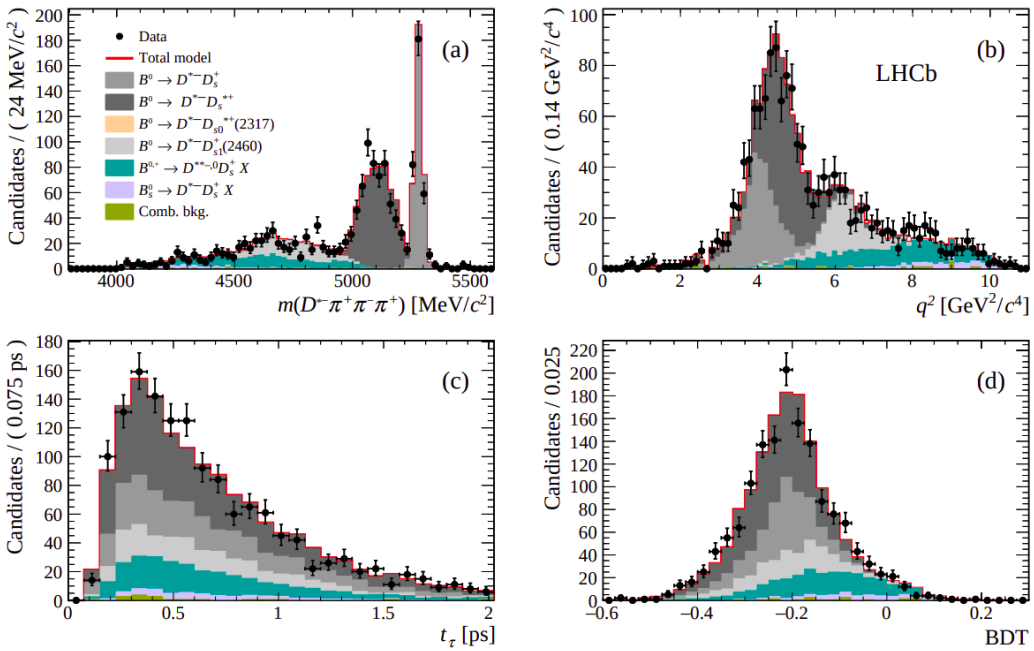

Figure 5: Results from the fit to data for candidates containing a $D D_s^+$ pair, where $D_s^+ \to 3\pi$.

Good agreement between simulation and data.

Study of double-charm candidates: D^0 and D^+

The decays of D^0 and D^+ mesons into final states with three pions are dominated by the $D^{0,+} \to K^{-,0} 3\pi(\pi^0)$ modes.

D^0: Isolation algorithm identifies a Kaon with charge opposite to the total charge of the 3π system, and compatible with originating from the 3π vertex. The mass of the $K - 3\pi$ system must be compatible with the known D_0 mass.
D^+: A pure sample of $B \to D^- D^+(X)$ decays is obtained by inverting the PID requirements on the negative pion of the 3π system, assigning to this particle the Kaon mass and selecting 3π candidates with mass compatible with the known D^+ mass.

Disagreement between data and simulation is found and corrected for.

Determination of the Signal Yield

The yield of $\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)$ decays is determined from a three-dimensional binned maximum likelihood fit to the distributions of q^2, 3π decay time, and BDT output.

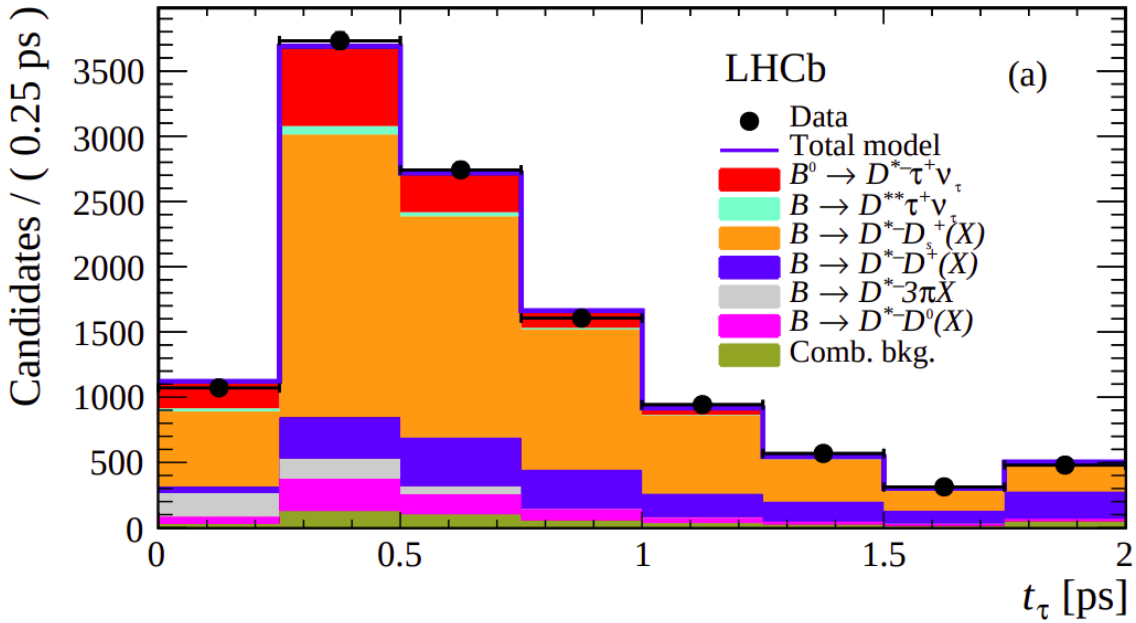

Figure 6: Projections of the three-dimensional fit on the 3π decay time

$\tau_{HL} \sim 0.29\ ps$, $D_s^+{}_{HL} \sim 0.5\ ps$.

Obtained: $N_{sig} = 1296 \pm 86$

Determination of Normalization Yield

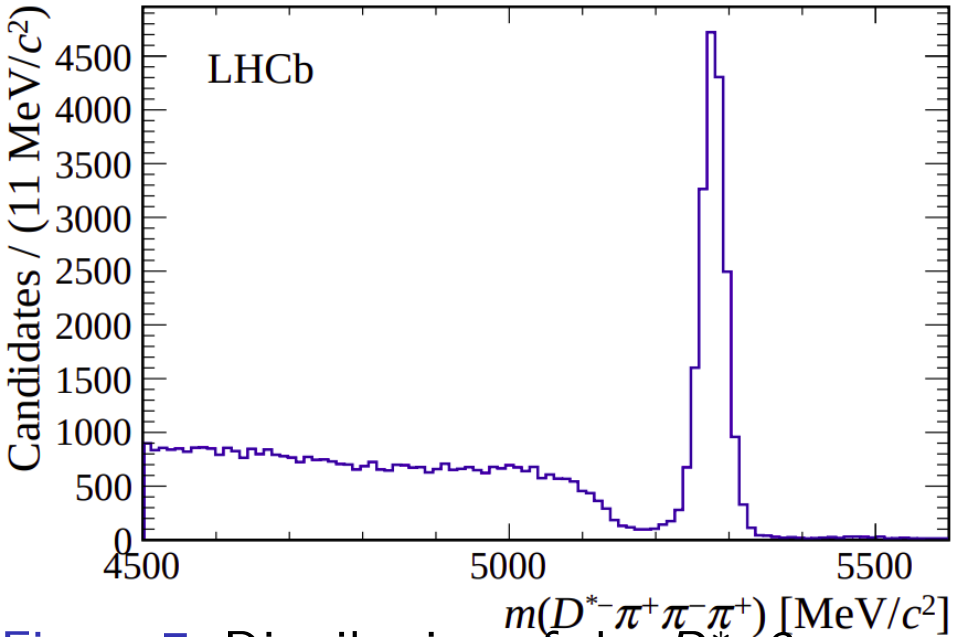

Figure 7: Distribution of the $D^{*-}3\pi$ mass for candidates passing the selection.

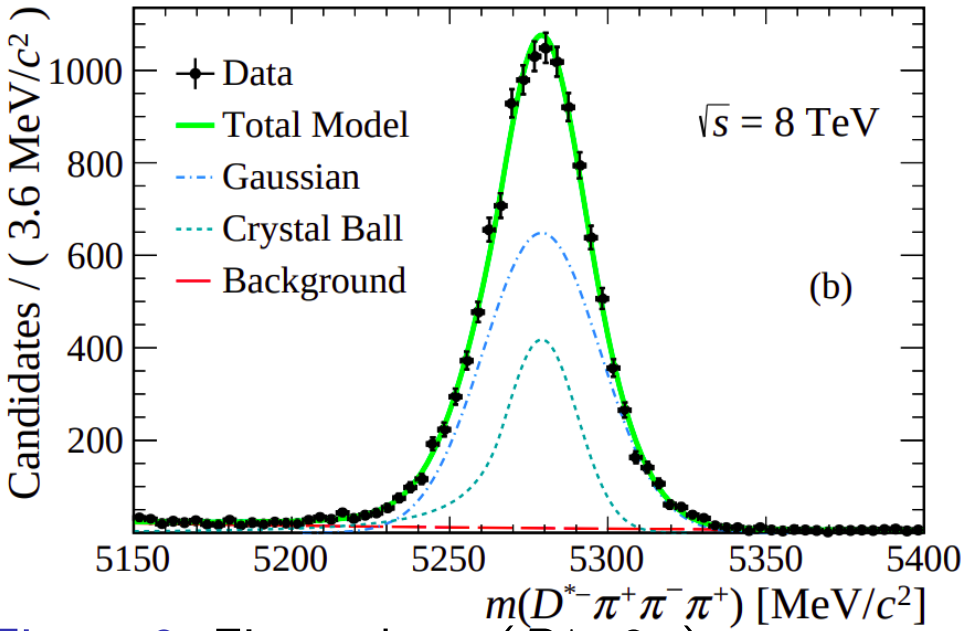

Figure 8: Fit to the $m(D^{*-}3\pi)$ distribution after the full selection in the $\sqrt{s} = 8\ TeV$ data samples.

In order to determine the normalization yield, a fit is performed in the region between 5150 and 5400 MeV/c^2, $m(B_0) = 5279.61 \pm 0.16\ MeV/c^2$.
The signal component is described by the sum of a Gaussian function and a Crystal Ball function. An exponential function is used to describe the background.
Obtained: $N_{norm} = 17660 \pm 143\ (stat)\ \pm 64\ (syst)\ \pm 22\ (sub)$, where the third uncertainty is due to the subtraction of the $B_c \to D^{*-}D^+$ component.

Determination of $\mathcal{K}(D^{*-})$

Using the eq. from slide 7:

$$\mathcal{K}(D^{*-}) \equiv \frac{\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)}{\mathcal{B}(B^0 \to D^{*-}3\pi)} = 1.97 \pm 0.13\ (stat)\ \pm 0.18\ (syst) \quad (2)$$

The ratio of efficiencies between the signal and normalization modes differs from unity due to the softer momentum spectrum of the signal particles and the correspondingly lower trigger efficiency. The effective sum of the branching fractions for the $\tau^+ \to 3\pi\overline{\nu}_\tau$ and $\tau^+ \to 3\pi\pi^0\overline{\nu}_\tau$ decays is $(13.81 \pm 0.07)\%$ as from PDG.

Systematic Uncertainties

1. **Uncertainty in the signal model**
   The uncertainty in the relative proportion of signal events, i.e, discrepancies between $\tau^+ \to 3\pi\overline{\nu}_\tau$ and $\tau^+ \to 3\pi\pi^0\overline{\nu}_\tau$, etc.
2. **Uncertainty in Background**
   Systematic uncertainties due to the modeling of different background sources, i.e., D_s^+ decay model, etc.
3. **Fit-related uncertainties**
   Fit-bias and limited size of simulated samples. Repeating the study implementing different smoothing parameters or new samples based on the original, after a procedure based on random selection with replacement.
4. **Uncertainties in the data selection**
   Such uncertainties stem from the choice of the trigger strategy, the online and offline selection of the candidates, the normalization and external inputs, and the efficiency of the PID criteria.

Systematic Uncertainties

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

Figure 9: List of the individual systematic uncertainties for the measurement of the ratio $\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)/\mathcal{B}(B^0 \to D^{*-}\pi^+\pi^-\pi^+)$

The main systematic uncertainties are due to the limited size of the simulated samples and the knowledge of the doubly-charmed decays $B_0 \to D^{*-}D(X)$.

Conclusions: Results

$$\mathcal{K}(D^{*-}) \equiv \frac{\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)}{\mathcal{B}(B^0 \to D^{*-}3\pi)} = 1.97 \pm 0.13\ (stat)\ \pm 0.18\ (syst) \quad (3)$$

Using the branching fraction $\mathcal{B}(B_0 \to D^{*-}3\pi) = (7.214 \pm 0.28) \times 10^{-3}$ from the weighted average of the measurements by the LHCb, BaBar, and Belle collaborations, one can obtain:

$$\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau) = (1.42 \pm 0.094\ (stat)\ \pm 0.129\ (syst)\ \pm 0.054\ (ext)\ ) \times 10^{-2} \quad (4)$$

Precision comparable to current world average.
Using the well measured branching fraction of $\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu) = (4.88 \pm 0.10) \times 10^{-2}$:

$$\begin{aligned} \mathcal{R}(D^{*-}) \equiv & \mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)/\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu) \\ = & 0.291 \pm 0.019\ (stat)\ \pm 0.026\ (syst)\ \pm 0.013\ (ext) \end{aligned} \quad (5)$$

Conclusions: Physics

1. $\mathcal{R}(D^{*-})$ is one of the most precise single measurements performed so far. It is 1.1 standard deviations higher than the SM prediction (0.252 ± 0.003).
2. An average of this measurement with the LHCb result using $\tau^+ \to \mu^+ \nu_\mu \overline{\nu}_\tau$ decays gives a value of $\mathcal{R}(D^{*-}) = 0.310 \ \pm \ 0.0155 \ (stat) \ \pm 0.0219 \ (syst)$, consistent with the world average and 2.2 standard deviations above the SM prediction.
3. After inclusion of the above result, the combined discrepancy of $\mathcal{R}(D)$ and $\mathcal{R}(D^{*-})$ determinations with the SM prediction is 4.1σ.
4. The novel technique presented in this paper, allowing the reconstruction and selection of semitauonic decays with $\tau^+ \to 3\pi(\pi^0)\overline{\nu}_\tau$ transitions, can be applied to all the other semitauonic decays, such as those of B^+, B_s^0, B_c^+ and Λ_b^0.
5. The inclusion of further data collected by LHCb at $\sqrt{s} = 13 \ TeV$ will result in an overall uncertainty on $\mathcal{R}(D^{*-})$ using this technique comparable to that of the current world average.

The End

The End