Lepton Flavor Universality test using $B \rightarrow D^*\tau\nu$ decays at LHCb

Thesis defense

Gaya Benane
Directed by: Olivier Leroy

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

29th October, 2024

Outline

1. Introduction
2. $\mathbf{R}(D^*)$ Measurement using hadronic τ decays
3. Selection
4. Results
5. Conclusions

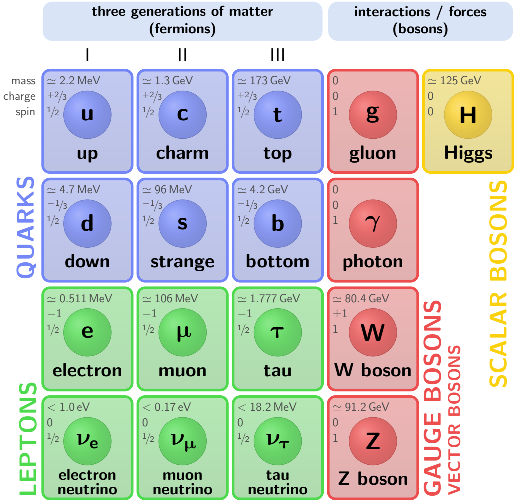
Introduction

Outline

Standard Model – Overview

Standard Model of Particle Physics (SM)
Theory that describes **three** of the four known **fundamental interactions** among elementary particles.

- **Successful Predictions:** The existence of the W and Z bosons, Higgs boson, and the top quark before their observational confirmation.
- **Precision Tests:** Confirmed through numerous experiments and precision measurements, making it one of the most tested theories in physics.



Open Questions

- **Hierarchy Problem**: Disparity between electroweak and Planck scales.
- **Strong *CP* Problem**: Lack of *CP* violation in strong interactions.
- **Fermion Mass Hierarchy**: Unclear pattern variation in quark and lepton masses.
- **Baryon Asymmetry**: Matter-antimatter asymmetry.
- **Dark Matter & Energy**: Unaccounted 95% of the universe's energy density.
- **Gravity**: Need for a theory of quantum gravity.
- **Neutrino Physics**: Neutrino masses and oscillations.
- ...

Two strategies to look for new physics

- **Direct searches**
  - Look for new particles or interactions at high-energy colliders.
- **Indirect searches**
  - Look for deviations from the SM in precision measurements.
  - Measurements of branching fractions (**Br**) and lifetimes.

Outline

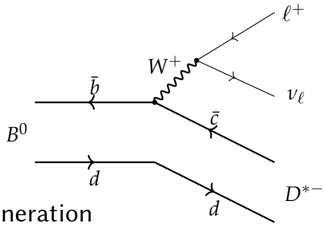
Lepton Flavor Universality

Lepton Flavor Universality (LFU) *All leptons should interact through the electroweak interaction with equal strength, independent of their flavor (i.e., electron, muon, tau).*

LFU can be tested via various processes, such as:

- ***B*-Meson Decays**
  - $\boldsymbol{R}(D)$ and $\boldsymbol{R}(D^*)$: Ratios of **Br** $B \to D^{(*)}\tau\nu$ versus $B \to D^{(*)}\ell\nu$ ($\ell = e, \mu$).
  - $\boldsymbol{R}(K)$ and $\boldsymbol{R}(K^*)$: Ratios of **Br** $B \to K^{(*)}\mu^+\mu^-$ versus $B \to K^{(*)}e^+e^-$.
- **Tau Lepton Decays**
  - Decay Rates Comparison: $\tau \to \mu\nu\nu$ versus $\tau \to e\nu\nu$.
- ***Z* and *W*-Boson Decays**
  - Equal Branching Fractions: $Z \to e^+e^-, \mu^+\mu^-, \tau^+\tau^-$; $W \to e\nu, \mu\nu, \tau\nu$.

To test the LFU, we measure the **ratio** of branching fractions **Br** in **semileptonic decays**:



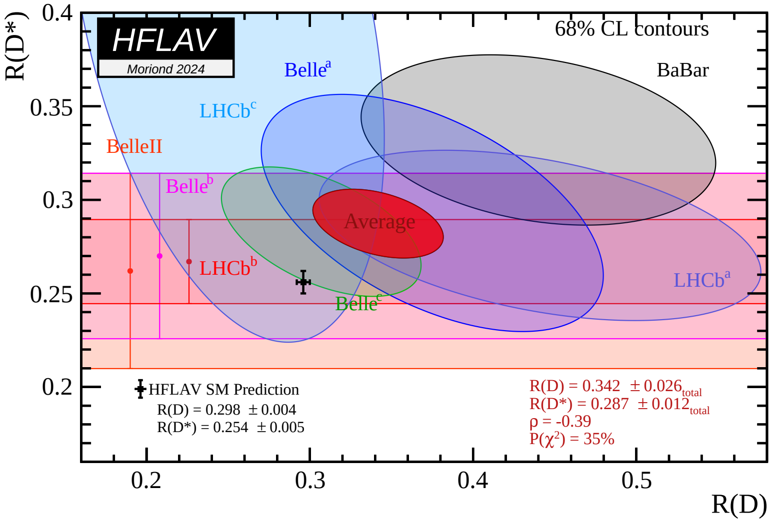
- Reduces theoretical uncertainties
- New physics may be more sensitive to the third generation

Charged current $b \to c\ell\nu_\ell$:

$$\boldsymbol{R}(X_c) \equiv \frac{\mathbf{Br}(X_b \to X_c\, \tau^+\, \nu_\tau)}{\mathbf{Br}(X_b \to X_c\, \ell^+\, \nu_\ell)}$$

where: $X_b = B^0, B^+_{(c)}, B^0_s, \Lambda_b, \ldots$ $X_c = D^{(*)}, J/\psi, D_s, \Lambda_c, \ldots$ and $\ell = \mu, e$

Main contribution: Tree-level digaram



$$R(D^{(*)}) = \frac{\mathbf{Br}(B \to D^{(*)}\tau^+\nu_\tau)}{\mathbf{Br}(B \to D^{(*)}\ell^+\nu_\ell)}$$

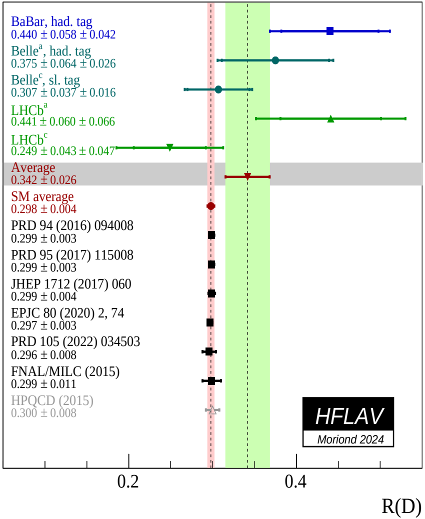
The world average

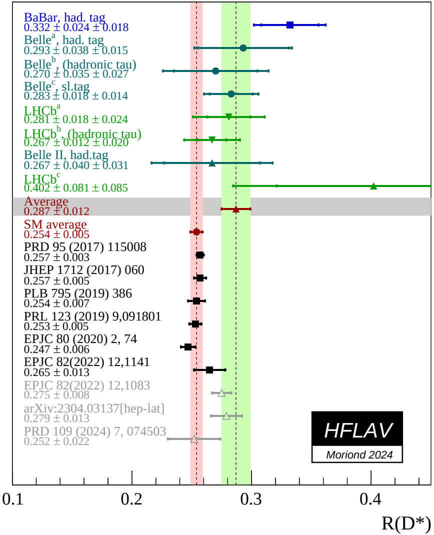
$$\begin{cases} R(D) & = 0.342 \pm 0.026 \\ R(D^*) & = 0.287 \pm 0.012 \end{cases}$$

The deviation w.r.t. the SM is at 3.3σ for the $R(D)$–$R(D^*)$ combination

[HFLAV]

$R(D)$–$R(D^*)$ Results

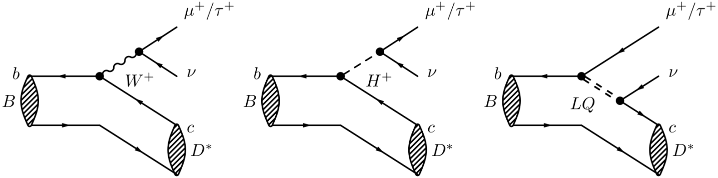




[HFLAV]

Possible Explanations for LFU Violation

SM | Charged Higgs | Leptoquark



Three typical candidates to account for the $R(D)$ and $R(D^*)$ anomalies:

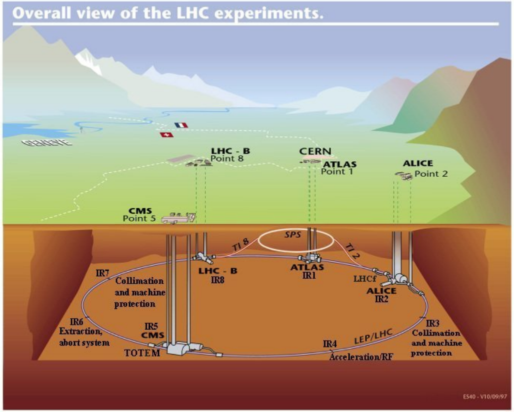
- Heavy vector bosons, *e.g.* W' [JHEP 07 (2015) 142 1506.01705, ...]
- Two-Higgs-doublet models H^+ [PRL 116, 081801, ...]
- Leptoquarks (LQ) [PRL 116, 081801, PRD 94, 115021, ...]

Outline

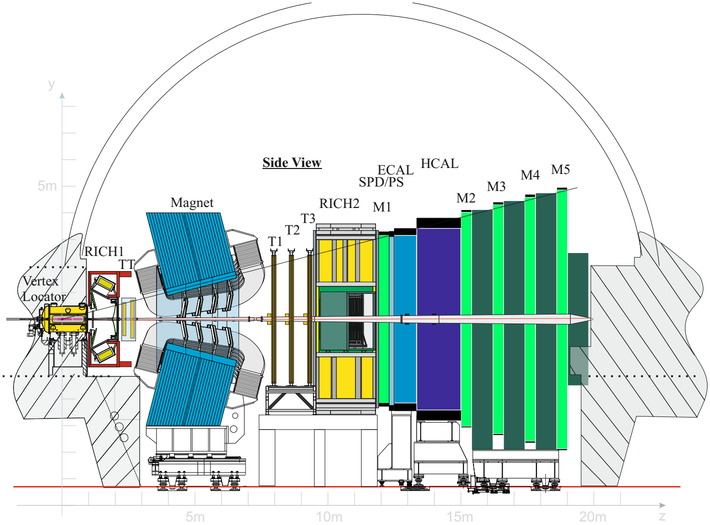
1. Introduction
   - Standard Model
   - Lepton Flavor Universality
   - LHCb Experiment
2. $\mathbf{R}(D^*)$ Measurement using hadronic τ decays
3. Selection
4. Results
5. Conclusions

The LHCb Experiment at the LHC

- LHCb is one of the experiments based at the Large Hadron Collider (LHC) at CERN, Geneva.
- Initially designed to perform *precision measurements* in the beauty quark sector
- It physics program has been extended:
  - Charm physics
  - Hadron spectroscopy
  - Top quark physics
  - Heavy ions
  - Electro-weak and Higgs physics



The LHCb Detector



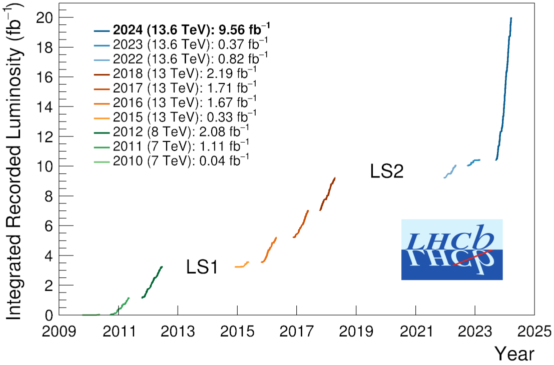
- A single-arm forward spectrometer.
- Pseudorapidity range $2 < \eta < 5$.
- Excellent vertex resolution
  - *xy*-plane: 10 – 40 μm
  - *z*-axis: 50 – 300 μm
  - Vertex reconstruction of *B* and τ^+.
- Particle identification efficiencies:
  - $\sim$ 97% for μ, *e*
  - $\sim$ 3% π misidentification
  - Good separation between π, *K*, *p*

Run 2 of LHCb

Period	$\int \mathcal{L}$ (fb^{-1})	$\sqrt{s}$ (TeV)	Number of $b\bar{b}$ events $\times 10^{11}$
Run 1: 2011–2012	3.2	$7-8$	2.5
Run 2: 2015–2016	2.0	13	2.9
Run 2: 2017–2018	3.9	13	5.7

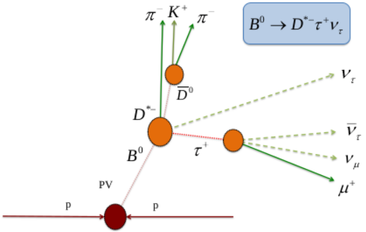
W.r.t Run 1, Run 2 has:

- Higher luminosity $\mathcal{L}$ and energy at the center of mass $\sqrt{s}$.
  - Higher $b\bar{b}$ cross-section.
- 15% higher trigger efficiency.



$R(D^*)$ measurements at LHCb

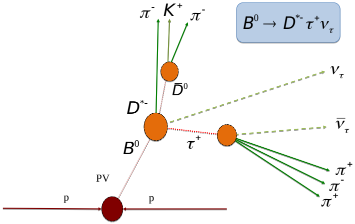
Muonic τ^+ decay



$\tau^+ \to \mu^+\nu_\mu\bar{\nu}_\tau$

- **Direct measurement of $R(D^*)$**
- **High statistics**
- Backgrounds from D^+ must be controlled well
- Sensitive to $D^{**}\mu^-\nu_\mu$

Hadronic τ^+ decay



- τ^+ decay **position** (suppress dominant backgrounds)
- **High purity sample**
- **Specific dynamics of $\tau^+ \to 3\pi^\pm\bar{\nu}_\tau$**
- $R(D^*)$ requires external inputs
- Lower statistics

Overview of semileptonic LFU tests in LHCb

- **Muonic** τ^- decays: $\tau^+ \to \mu^+ \bar{\nu}_\mu \nu_\tau$
  - $\boldsymbol{R}(D^*)$ (Run 1) [PRL **115**, 111803 (2015)]
  - $\boldsymbol{R}(J/\psi)$ (Run 1) [PRL **120**, 121801 (2018)]
  - $\boldsymbol{R}(D^*)$–$\boldsymbol{R}(D^0)$ (Run 1) [PRL **131**, 111802 (2023)]
  - $\boldsymbol{R}(D^+)$–$\boldsymbol{R}(D^{*+})$ (partial Run 2) [arxiv:2406.03387]
- **Hadronic** τ^+ decays: $\tau^+ \to \pi^+\pi^-\pi^+(\pi^0)\bar{\nu}_\tau$
  - $\boldsymbol{R}(D^*)$ (Run 1 and partial Run 2) [PRL **115**, 111803 (2015), PRD **108**, 012018 (2023)]
  - $\boldsymbol{R}(\Lambda_c)$ (Run 1) [PRL **128**, 191803 (2022)]
  - $\boldsymbol{R}(D^{**})$ (Run 1 and Run 2) [LHCb-PAPER-2024-037]

This thesis: $\boldsymbol{R}(D^*)$ with **hadronic** τ^+ decays using the **full Run 2** datasets

Aim: Enhance the precision of the $\boldsymbol{R}(D^*)$ measurement using a larger dataset.

R(D^*) Measurement using hadronic τ decays

Methodology

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

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

Methodology

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

$$\mathcal{K}(\mathcal{D}^*) = \frac{\mathcal{N}_{\mathtt{sig}}}{\mathcal{N}_{\mathtt{norm}}} \times \frac{\varepsilon_{\mathtt{norm}}}{\varepsilon_{\mathtt{sig}}} \times \frac{1}{\mathcal{B}(\tau^+ \to 3\pi^\pm\bar{\nu}_\tau) + \mathcal{B}(\tau^+ \to 3\pi^\pm\pi^0\bar{\nu}_\tau)}$$

Methodology

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

$$\mathcal{K}(\mathcal{D}^*) = \frac{\mathcal{N}_{\mathtt{sig}}}{\mathcal{N}_{\mathtt{norm}}} \times \frac{\varepsilon_{\mathtt{norm}}}{\varepsilon_{\mathtt{sig}}} \times \frac{1}{\mathcal{B}(\tau^+ \to 3\pi^\pm\bar{\nu}_\tau) + \mathcal{B}(\tau^+ \to 3\pi^\pm\pi^0\bar{\nu}_\tau)}$$

- $\mathcal{N}_{\mathtt{sig}}$ from a 3D binned Maximum Likelihood (ML) template fit.
  - No narrow peaks in the signal region to fit to.
  - $q^2 = (p_B - p_{D^*})^2$, τ^+ lifetime, and the output of a BDT classifier

Methodology

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

$$\mathcal{K}(\mathcal{D}^*) = \frac{\mathcal{N}_{\texttt{sig}}}{\mathcal{N}_{\texttt{norm}}} \times \frac{\varepsilon_{\texttt{norm}}}{\varepsilon_{\texttt{sig}}} \times \frac{1}{\mathcal{B}(\tau^+ \to 3\pi^\pm\bar{\nu}_\tau) + \mathcal{B}(\tau^+ \to 3\pi^\pm\pi^0\bar{\nu}_\tau)}$$

- $\mathcal{N}_{\texttt{sig}}$ from a 3D binned Maximum Likelihood (ML) template fit.
  - No narrow peaks in the signal region to fit to.
  - $q^2 = (p_B - p_{D^*})^2$, τ^+ lifetime, and the output of a BDT classifier
- $\mathcal{N}_{\texttt{norm}}$ from an unbinned ML fit to $m(D^* 3\pi^\pm)$.

Methodology

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

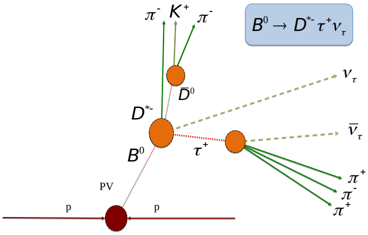
$$\mathcal{K}(\mathcal{D}^*) = \frac{\mathcal{N}_{\mathtt{sig}}}{\mathcal{N}_{\mathtt{norm}}} \times \frac{\varepsilon_{\mathtt{norm}}}{\varepsilon_{\mathtt{sig}}} \times \frac{1}{\mathcal{B}(\tau^+ \to 3\pi^\pm\bar{\nu}_\tau) + \mathcal{B}(\tau^+ \to 3\pi^\pm\pi^0\bar{\nu}_\tau)}$$

- $\mathcal{N}_{\mathtt{sig}}$ from a 3D binned Maximum Likelihood (ML) template fit.
  - No narrow peaks in the signal region to fit to.
  - $q^2 = (p_B - p_{D^*})^2$, τ^+ lifetime, and the output of a BDT classifier
- $\mathcal{N}_{\mathtt{norm}}$ from an unbinned ML fit to $m(D^* 3\pi^\pm)$.
- Efficiences $\varepsilon_{\mathtt{sig}}$ and $\varepsilon_{\mathtt{norm}}$ extracted from simulation samples.

Signal & Normalization mode

Signal mode

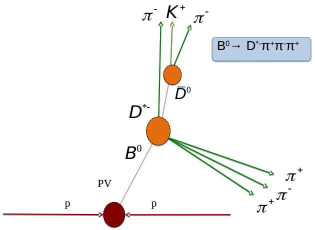
$B^0 \to D^{*-}\tau^+\nu_\tau$ where $\tau^+ \to 3\pi^\pm(\pi^0)\bar{\nu}_\tau$



- Partially reconstructed decay
  - Due to missing neutrinos

Normalization mode

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



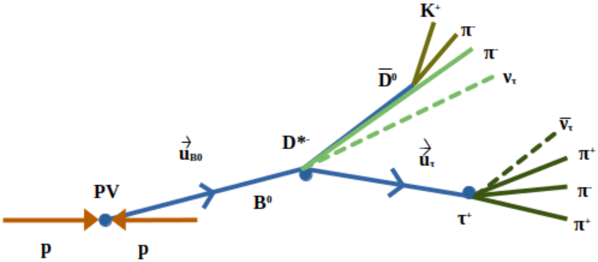
- Fully reconstructed decay

Identical visible final states in both modes $\implies$ systematic uncertainties cancel out

Signal Decay Kinematics

- In the signal mode, there are 2 (missing) neutrinos in the final state.
  $\rightarrow$ Approximations are needed for τ and B reconstruction.

  Signal reconstruction is based on:
- The well-measured B^0 and τ^+ vertices allow for the reconstruction of flight directions.
- Momentum and Energy conservation



$\Longrightarrow$ Estimate of B^0 and τ^+ momentum.

Backgrounds

The background sources for the signal mode are:

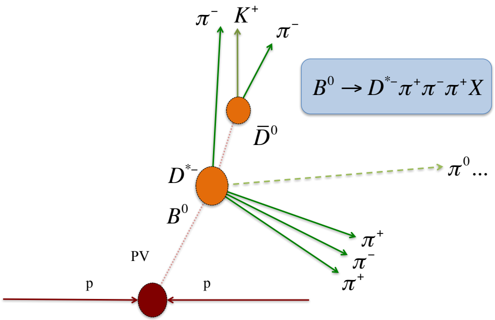
- **Prompt background**
  - $B \rightarrow D^{*-} 3\pi^{\pm} X$
  - The largest background contribution
- **Double charm decays**
  - $B \rightarrow D^{*-} D X$
  - The second largest contribution
- Other backgrounds:
  - **Feed–down** from excited D^{*} states $B \rightarrow D^{**} \tau \nu$
  - **Combinatorial backgrounds**

Backgrounds

Prompt background $B \to D^{*-}3\pi^{\pm}X$

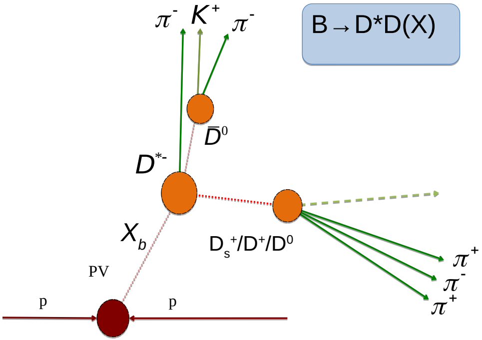
- $3\pi^{\pm}$ directly from B meson
- $\sim 100\times$ signal decays

Strategy: Apply cut on the τ^+ flight distance



Backgrounds

- **Double–charm decays $B \to D^{*-} DX$**
  - $D = D_s^+, D^+, D^0$
  - **Signal like topology** with a detached vertex due to non-negligible lifetime
  - $B \to D^{*-} D_s^+ X \sim 10\times$ signal decays
  - $B \to D^{*-} D^+ X \sim 1\times$ signal decays
  - $B \to D^{*-} D^0 X \sim 0.2\times$ signal decays
- **Strategy:** Use BDT output (trained to distinguish signal from double charm background) as a:
  - Cut variable.
  - **Variable in the signal extraction fit.**



Outline

Collision Data & Simulation Samples

- **Collision dataset**
  - $5.8\,\text{fb}^{-1}$ of LHCb Run 2 (2015–2018) pp collisions
  - Center of mass energy $\sqrt{s} = 13\,\text{TeV}$.
- **Simulated samples**
  - Two signal modes: $\tau^+ \to 3\pi^\pm \overline{\nu}_\tau$ and $\tau^+ \to 3\pi^\pm \pi^0 \overline{\nu}_\tau$
  - Normalization $B^0 \to D^{*-} 3\pi^\pm$
  - Various exclusive/inclusive background samples:
    - **Inclusive MC sample** $b\bar{b} \to D^{*-} 3\pi^\pm X$ (dominant background)
    - **Cocktail MC sample** $B^0 \to D^{*-} D^{0,+} X$

Decay	Generated Events [M]
$B^0 \to D^{*-}\tau^+\nu_\tau$	455.0
$B^0 \to D^{*-}3\pi^\pm$	651.0
$B^0 \to D^{**}\tau^+\nu_\tau$	298.0
$B^0_s \to D^{*-}D^+_s X$	251.0
$B^+ \to D^{*-}D^+_s X$	302.0
$B^0 \to D^{*-}D^+_s X$	972.0
$b\bar{b} \to D^{*-}3\pi^\pm X$	18132.0
$b\bar{b} \to D^{*-}D^{\{0,+\}} X$	822.0
Total	21884.0

Collision Data & Simulation Samples

- Simulation samples play an important role:
  - Selection optimization (MVA trainings)
  - Efficiency calculation
  - Templates building for the signal extraction fit
- The limited size of the simulation sample was the **dominant systematic uncertainty** in Run 1 analysis.

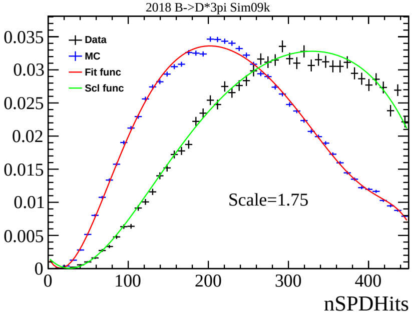
$\rightarrow$ To address this, over 21 **billion events were generated** (using the **fast simulation (ReDecay)** technique).

Decay	Generated Events [M]
$B^0 \rightarrow D^{*-}\tau^+\nu_\tau$	455.0
$B^0 \rightarrow D^{*-}3\pi^\pm$	651.0
$B^0 \rightarrow D^{**}\tau^+\nu_\tau$	298.0
$B^0_s \rightarrow D^{*-}D^+_s X$	251.0
$B^+ \rightarrow D^{*-}D^+_s X$	302.0
$B^0 \rightarrow D^{*-}D^+_s X$	972.0
$b\bar{b} \rightarrow D^{*-}3\pi^\pm X$	18132.0
$b\bar{b} \rightarrow D^{*-}D^{\{0,+\}}X$	822.0
Total	21884.0

Corrections to Simulation

To ensure that the simulation accurately represents the properties of the real data, corrections are applied on:

- Form-Factors of the simulated $B^0 \to D^{*-}\tau^+\nu_\tau$ decays.
- The kinematic properties of B candidates.
- number of Tracks and SPD detector hit counts
- The D^* and D_s^+ longitudinal polarizations.
- The uncertainty on the z-component of the $3\pi^\pm$ vertex position.



- Other corrections are applied to the background simulation samples (discussed later)

Selection

- Online selection

- Offline selection

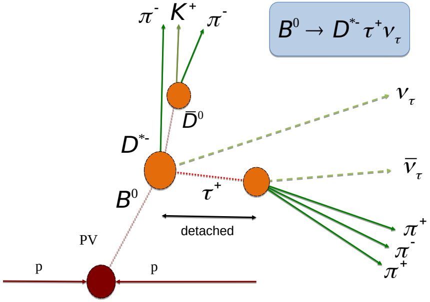
  - MultiVariate Analysis (MVA) training using Boosted Decision Trees (BDTs):
    - Anti-combinatorial background BDT
    - Charged isolation BDT
    - $3\pi^{\pm}$ vertex detachment BDT
    - Anti-D_s^+ BDT (**used in signal fit**)

  - Remaining cuts:
    - Mode specific : Control sample, signal and normalization modes

Outline

Preselection

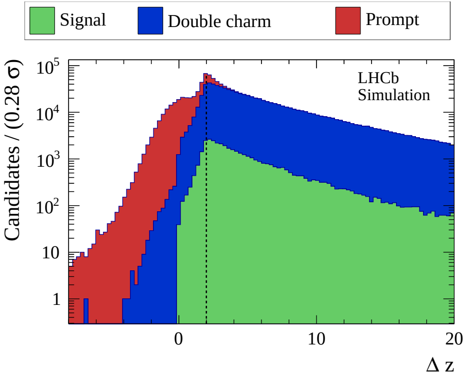
- The signal candidates are built based on the **6 final-state charged tracks**:
  - $3\pi^{\pm}$ from τ^{+}
  - K^{+} and π^{-} from D^{0} and slow π^{-} from D^{*-}.
- Common selection criteria:
  - Track and vertex quality
  - PID requirements
  - Mass constraints
  - 3π vertex detachment / uncertainty (Δz)



$$\Delta z(\tau^{+}, B^{0}) \equiv [\text{vtx}_z(\tau^{+}) - \text{vtx}_z(B^{0})]/\text{error}$$

Preselection

- The signal candidates are built based on the **6 final-state charged tracks**:
  - $3\pi^{\pm}$ from τ^{+}
  - K^{+} and π^{-} from D^{0} and slow π^{-} from D^{*-}.
- Common selection criteria:
  - Track and vertex quality
  - PID requirements
  - Mass constraints
  - 3π vertex detachment / uncertainty (Δz)

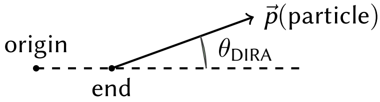


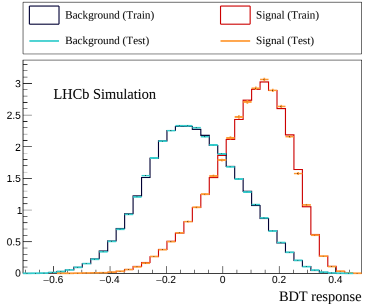
$$\Delta z(\tau^{+}, B^{0}) \equiv [\mathrm{vtx}_z(\tau^{+}) - \mathrm{vtx}_z(B^{0})]/\mathrm{error}$$

Outline

Anti-Combinatorial Background BDT

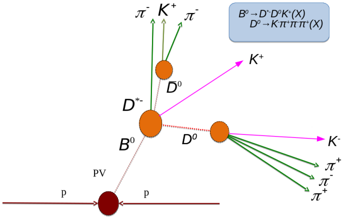
- **Objective:** Remove D^{*-} and $3\pi^{\pm}$ from different hadrons
- **Training samples:**
  - Signal: $B^0 \to D^{*-}\tau^+\nu_\tau$ MC
  - Background: wrong-sign Data (D^{*-} and $3\pi^{\mp}$)
- **Input variables:**
  - D^{*-} and τ^+ kinematic variables (p_T and η)
  - Direction angle of B^0, D^{*-} and τ^+ momentum w.r.t their *origin–end vertex* line (θ_{DIRA})
  - Radial distance between the τ decay vertex and the PV

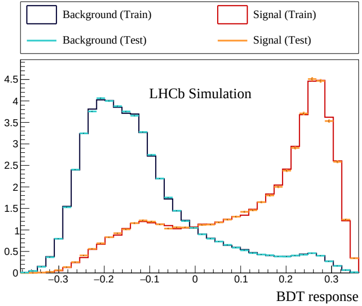




Charged Isolation BDT

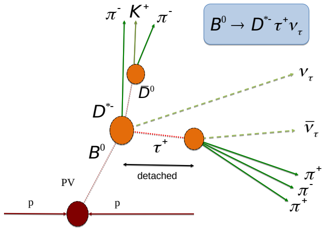
- **Objective:** remove partially reconstructed decays with additional charged tracks.
- **Training samples:** $b\bar{b} \to D^{*-}\ 3\pi^{\pm}$ MC
  - Signal: **without** extra tracks
  - Background: **with** extra tracks
- **Input variables:**
  - **Track**-based: one per each final states charged track
  - **Vertex**-based: number of charged tracks forming 'good' vertex with the τ^+ and B^0

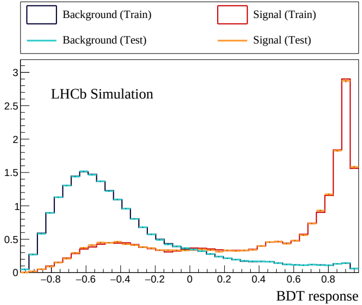




$3\pi^{\pm}$ Vertex Detachment BDT

- **Objective:** remove the 'prompt' background
- **Training samples:**
  - Signal: $B^0 \to D^{*-}\tau^+\nu_\tau$ MC sample
  - Background: $b\bar{b} \to D^{*-}3\pi^{\pm}$ MC ($3\pi^{\pm}$ don't come from the τ^+ or any D-meson)
- **Input variables:**
  - Distance between vertices:
    - τ^+, B^0
    - D^0, τ^+
  - Direction angle θ_{DIRA} of τ^+
  - Origin- and end-vertex quality of the τ^+

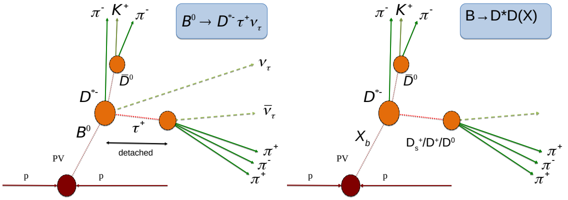


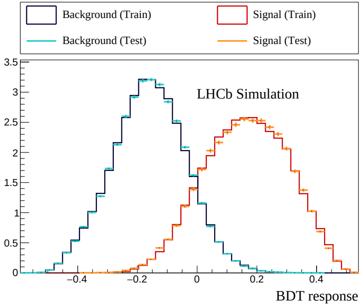


Anti-$B^0 \to D^{*-} D_s^+ X$ BDT

- **Objective:** reduce double-charm background $B^0 \to D^{*-} D_s^+ X$ where $D_s^+ \to 3\pi^\pm X$
- **Training samples:**
  - Signal: $B^0 \to D^{*-}\tau^+\nu_\tau$ MC
  - Background: $B^0 \to D^{*-} D_s^+ X$ MC
- **Input variables:**
  - Partial reconstruction variables (signal/double-charm background hypothesis)
  - Neutral and charged isolation in $3\pi^\pm$
  - Pion dynamics: $\min[m(\pi^+\pi^-)]$, $\max[m(\pi^+\pi^-)]$, $m(\pi^+\pi^+)$
  - PV–B^0 radial distance

This BDT is one of three variables used in the signal fit





Outline

Remaining Cuts

- Apply **similar cuts** wherever possible (reduce systematics related to tracking)
  - Tighter mass constraints: $m(D^0)$ and $\Delta m = m(D^{*-}) - m(D^0)$
  - PID on the pions
  - Combinatorial and isolation BDTs
- Signal mode specific cuts:
  - τ^+ and B^0 mass constraint
  - $q^2 \equiv (p_{B^0} - p_{D^{*-}})^2$
  - Δz cut between τ^+ and B^0
  - Detachment and anti-D_s^+ BDTs

Remaining Cuts

- Apply **similar cuts** wherever possible (reduce systematics related to tracking)
  - Tighter mass constraints: $m(D^0)$ and $\Delta m = m(D^{*-}) - m(D^0)$
  - PID on the pions
  - Combinatorial and isolation BDTs
- Normalization mode specific cuts:
  - Select $m(D^{*-}3\pi^{\pm})$ peak around the B^0 mass
  - Δz cut between τ^+ and D^0 (ensures the $3\pi^{\pm}$ system originates directly from B^0).
  - Detachment BDT not applied (prompt decay required)
  - Anti-D_s^+ BDT skipped (no intermediate resonances in $3\pi^{\pm}$)
  - q^2 skipped (no leptons)

Outline

Efficiency Calculation

The first input to obtain $\mathcal{K}(D^{*-})$:

$$\mathcal{K}(\mathcal{D}^*) = \frac{\mathcal{N}_{\mathtt{sig}}}{\mathcal{N}_{\mathtt{norm}}} \times \frac{\varepsilon_{\mathtt{norm}}}{\varepsilon_{\mathtt{sig}}} \times \frac{1}{\mathcal{B}(\tau^+ \to 3\pi^\pm \overline{\nu}_\tau) + \mathcal{B}(\tau^+ \to 3\pi^\pm \pi^0 \overline{\nu}_\tau)}$$

- Efficiencies are estimated using **simulated samples**.
- They are calculated at every stage of the selection process.
- Two signal modes ($\tau^+ \to 3\pi\overline{\nu}_\tau$ and $\tau^+ \to 3\pi\pi^0\overline{\nu}_\tau$) efficiencies are combined using their **Br**.

Efficiencies for the signal and normalization modes:

Efficiency	**Value**
$\varepsilon_{\mathtt{norm}}$	$3.466 \pm 0.019 \times 10^{-4}$
$\varepsilon_{\mathtt{sig}}$	$1.190 \pm 0.005 \times 10^{-4}$
$\varepsilon_{\mathtt{norm}}/\varepsilon_{\mathtt{sig}}$	2.952 ± 0.020

Outline

The Need for Control Samples

The signal yield $\mathcal{N}_{\text{sig}}$ is extracted using **templates** primarily derived from **simulation**

$\longrightarrow$ Important to apply **data-driven corrections** to the simulation

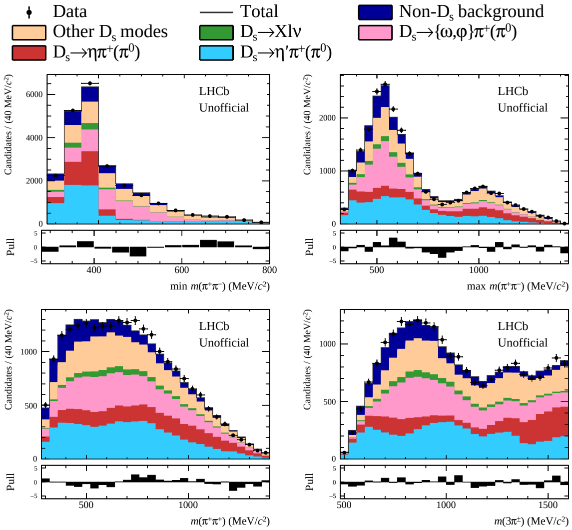
- D_s^+ **Decay Model:**
  - Some of the $D_s^+ \to 3\pi^\pm X$ branching fractions are not well known and/or correctly simulated.
- $B^0 \to D^{*-} D_s^+ X$ **Spectrum:**
  - Relative contributions of $B^0 \to D^{*-} D_s^+ X$ components are obtained from a fit to the relevant control sample.
- $B^0 \to D^{*-} D^+ X$ **and** $B^0 \to D^{*-} D^0 X$ **Control Samples**

D_s^+ Decay Model – Strategy

- Data sample selected with anti-D_s^+ BDT < -0.2.
- Simultaneous fit to: $\max[m(\pi^+\pi^-)]$, $\min[m(\pi^+\pi^-)]$, $m(\pi^+\pi^+)$ and $m(3\pi^\pm)$
- PDF constructed as a sum of PDFs for different $D_s^+ \to 3\pi$ decays
  - At least one pion originates directly from an **intermediate resonant state** :
    - π from η: $D_s^+ \to \eta\pi^+(\pi^0)$
    - π from η': $D_s^+ \to \eta'\pi^+(\pi^0)$
    - π from ω or ϕ: $D_s^+ \to \omega\pi^+(\pi^0)$ or $D_s^+ \to \phi\pi^+(\pi^0)$.
  - All pions originate directly from the **initial state** :
    - $\eta 3\pi$, ηa_1, $\eta' 3\pi$, $\eta' a_1$, $\omega 3\pi$, ωa_1,$\phi 3\pi$, ϕa_1, $K^0 3\pi$, $K^0 a_1$
  - $D_s^+ \to \tau\nu$ and non-resonant 3π

D_s^+ Decay Model – Fit Results

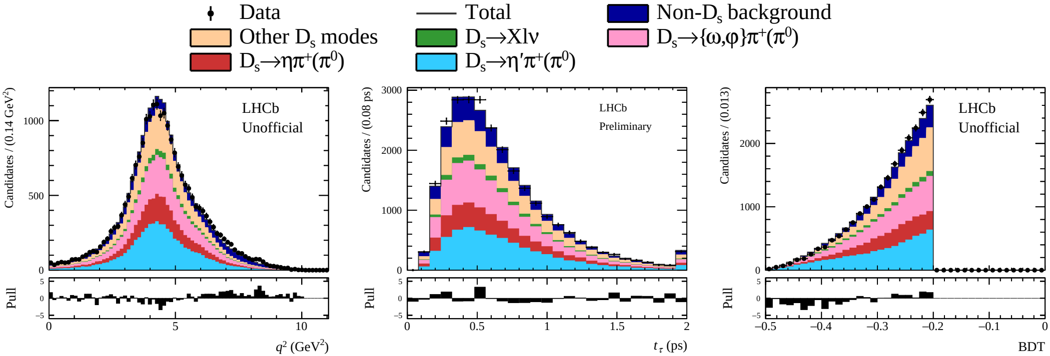
- Use cocktail MC to build templates



Index i	Template T_i	Fraction f_i	Correction Factor w_i
1	$\eta 3\pi^\pm$	0.007 ± 0.0	0.82 ± 0.059
2	$\eta' 3\pi^\pm$	0.007 ± 0.0	0.82 ± 0.059
3	$\omega 3\pi^\pm$	0.011 ± 0.001	0.82 ± 0.059
4	$\phi 3\pi^\pm$	0.024 ± 0.002	0.82 ± 0.059
5	$K^0 3\pi^\pm$	0.006 ± 0.0	0.82 ± 0.059
6	$D_s^+ \to \tau^+ \nu_\tau$	0.013 ± 0.001	0.82 ± 0.059
7	NR $3\pi^\pm$	0.101 ± 0.014	2.24 ± 0.315
8	$a_1 \eta$	0.026 ± 0.002	0.82 ± 0.059
9	$a_1 \eta'$	0.004 ± 0.0	0.82 ± 0.059
10	$a_1 \omega$	0.008 ± 0.001	0.82 ± 0.059
11	$a_1 \phi$	0.017 ± 0.001	0.82 ± 0.059
12	$a_1 K^0$	0.003 ± 0.0	0.82 ± 0.059
13	$5\pi^\pm$	0.026 ± 0.002	0.82 ± 0.059
14	$\omega\rho^+$ or $\phi\rho^+$	0.239 ± 0.015	0.99 ± 0.061
15	$\omega\pi^+$ or $\phi\pi^+$	0.022 ± 0.01	0.65 ± 0.305
16	$\eta\rho^+$	0.145 ± 0.008	0.67 ± 0.036
17	$\eta\pi^+$	0.026 ± 0.007	1.28 ± 0.355
18	$\eta'\rho^+$	0.232 ± 0.009	1.49 ± 0.059
19	$\eta'\pi^+$	0.084 ± 0.006	0.82 ± 0.059

Some correction factors are the same because their relative fractions are **fixed**.

Projections of fitting variables for the signal fit:

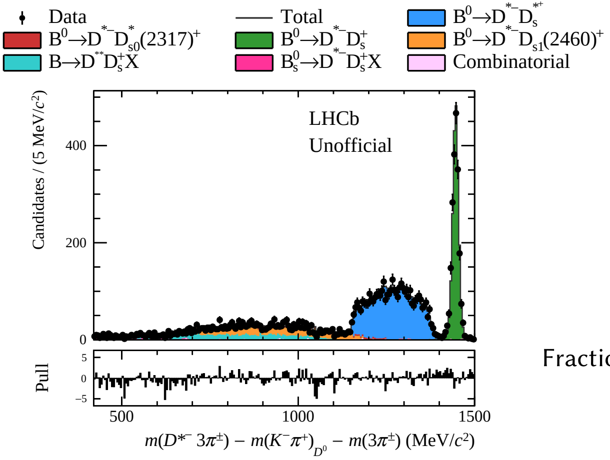


$B \to D^{*-} D_s^+ X$ Control – Strategy

- $B \to D^{*-} D_s^+ X$ branching fractions not all well known and/or correctly simulated
- Constrain fractions using fit to $B \to D^{*-} D_s^+ X$ data with D_s^+ fully reconstructed
  - $m(3\pi^\pm) \in m(D_s^+) \pm 20\,\mathrm{MeV}$

- Fit to $m(D^{*-} 3\pi^\pm) - m(D^0) - m(3\pi^\pm)$
  - PDF constructed as a sum of PDFs for different $B \to D^{*-} D_s^+ X$ decays:
    - $B^0 \to D^{*-} D_s^+$
    - $B^0 \to D^{*-} D_{s1}^+$
    - $B^0 \to D^{*-} D_s^{*+}$
    - $B^{(0,+)} \to D^{**(0,+)} D_s^+ X$
    - $B^0 \to D^{*-} D_{s0}^{*+}$
    - $B_s^0 \to D^{*-} D_s^+ X$
  - Combinatorial background

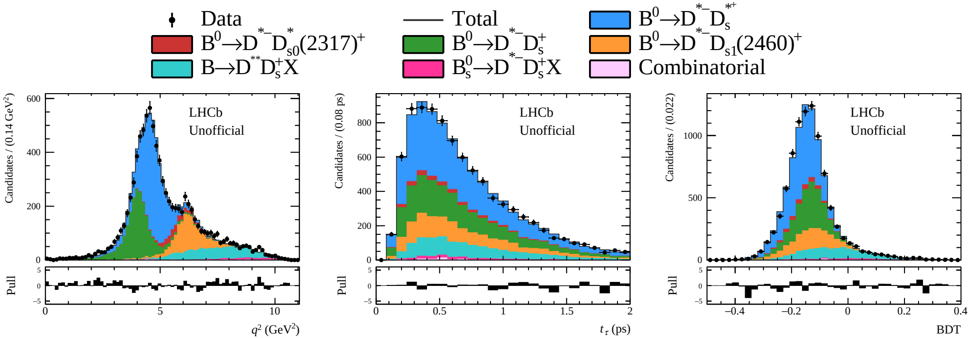
$B \to D^{*-} D_s^+ X$ Control – Fit Results



Parameter	Value
N_{D_s}	8674 ± 96
N_{bkg}	76 ± 26
$f_{B_s \to D^* D_s X}$	0.067 ± 0.017
$f_{D^{**} D_s X}$	0.188 ± 0.034
$f_{D'^+_{s1}}$	0.438 ± 0.033
f_{D_s}	0.556 ± 0.016
$f_{D^{*+}_{s0}}$	0.050 ± 0.021

Fractions of each component determined and used as constraints in the signal extraction fit

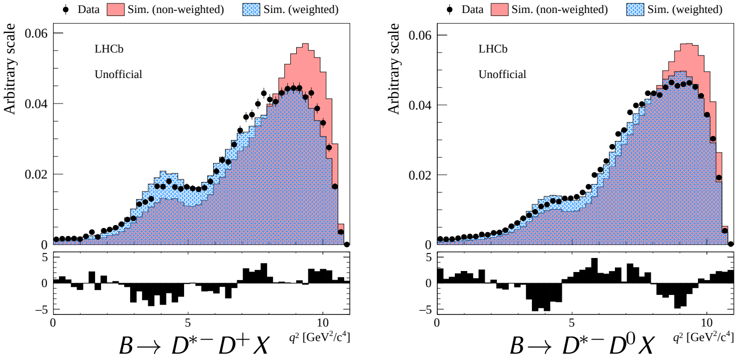
$B \to D^{*-} D_s^+ X$ Control – Fit Projections



Good agreement between model and data for the signal fit variables

$B \to D^{*-}(D^+, D^0)X$ Decays

- Data control samples for $B \to D^{*-}(D^+, D^0)X$ decays compared with simulation
- q^2 distribution in simulation is corrected for the observed differences



Outline

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

The total probability density function is:

$$
\begin{aligned}
\mathcal{P}_{\text{total}}(q^2, t_\tau, \text{BDT}) = 1/N_{\text{total}} \times \{ & N_{\text{sig}} \left[f_{\tau^+\to\pi^+\pi^-\pi^+\overline{\nu}_\tau} \; \mathcal{P}_{\tau^+\to\pi^+\pi^-\pi^+\overline{\nu}_\tau} + (1 - f_{\tau^+\to\pi^+\pi^-\pi^+\overline{\nu}_\tau}) \, \mathcal{P}_{\tau^+\to\pi^+\pi^-\pi^+\pi^0\overline{\nu}_\tau} \right. \\
& \left. + f_{D^{**}\tau\nu} \; \mathcal{P}_{B\to D^{**}\tau^+\nu_\tau} \right] + N_{D^0}^{\text{same}} \left[\mathcal{P}_{B\to D^{*-}D^0X \text{ SV}} + f_{D^0}^{v_1-v_2} \; \mathcal{P}_{B\to D^{*-}D^0X \text{ DV}} \right] \\
& + N_{D_s^+} / k \times \left[\mathcal{P}_{B^0\to D^{*-}D_s^{*+}} + f_{D_s^+} \; \mathcal{P}_{B^0\to D^{*-}D_s^+} + f_{D_{s0}^{*+}} \; \mathcal{P}_{B^0\to D^{*-}D_{s0}^{*+}} \right. \\
& \left. + f_{D_{s1}^+} \; \mathcal{P}_{B^0\to D^{*-}D_{s1}^+} + f_{D^{**}D_sX} \; \mathcal{P}_{B\to D^{**-}D_s^+X} + f_{B_s\to D^*D_s^+X} \; \mathcal{P}_{B_s^0\to D^{*-}D_s^+X} \right] \\
& + N_{D_s^+} \; f_{D^+} \; \mathcal{P}_{B\to D^{*-}D^+X} + N_{B\to D^{*-}3\pi^\pm X} \; \mathcal{P}_{B\to D^{*-}3\pi^\pm X} \\
& + N_{B_1-B_2} \; \mathcal{P}_{\text{combinatoric } B} + N_{\text{fake } D^0} \; \mathcal{P}_{\text{combinatoric } D^0} + N_{\text{fake } D^*} \; \mathcal{P}_{\text{combinatoric } D^{*-}} \}
\end{aligned}
$$

- 16 templates: 13 templates from MC , 3 templates from data
- 4 free parameters , 6 gaussian constrained parameters and 6 fixed parameters

Signal Fit – PDF

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

The total probability density function is:

$$\begin{aligned}\mathcal{P}_{\text{total}}(q^2, t_\tau, \text{BDT}) = 1/N_{\text{total}} \times \{\, & N_{\text{sig}} \, [\, f_{\tau^+\to\pi^+\pi^-\pi^+\overline{\nu}_\tau} \mathcal{P}_{\tau^+\to\pi^+\pi^-\pi^+\overline{\nu}_\tau} + (1 - f_{\tau^+\to\pi^+\pi^-\pi^+\overline{\nu}_\tau}) \mathcal{P}_{\tau^+\to\pi^+\pi^-\pi^+\pi^0\overline{\nu}_\tau} \\ & + f_{D^{**}\tau\nu} \mathcal{P}_{B\to D^{**}\tau^+\nu_\tau} \,] + N_{D^0}^{\text{same}} \, [\, \mathcal{P}_{B\to D^{*-}D^0X \text{ SV}} + f_{D^0}^{v_1-v_2} \mathcal{P}_{B\to D^{*-}D^0X \text{ DV}} \,] \\ & + N_{D_s^+}/k \times [\, \mathcal{P}_{B^0\to D^{*-}D_s^{*+}} + f_{D_s^+} \mathcal{P}_{B^0\to D^{*-}D_s^+} + f_{D_{s0}^{*+}} \mathcal{P}_{B^0\to D^{*-}D_{s0}^{*+}} \\ & + f_{D_{s1}^+} \mathcal{P}_{B^0\to D^{*-}D_{s1}^+} + f_{D^{**}D_sX} \mathcal{P}_{B\to D^{**-}D_s^+X} + f_{B_s\to D^*D_s^+X} \mathcal{P}_{B_s^0\to D^{*-}D_s^+X} \,] \\ & + N_{D_s^+} \, f_{D^+} \mathcal{P}_{B\to D^{*-}D^+X} + N_{B\to D^{*-}3\pi^\pm X} \mathcal{P}_{B\to D^{*-}3\pi^\pm X} \\ & + N_{B_1-B_2} \mathcal{P}_{\text{combinatoric } B} + N_{\text{fake } D^0} \mathcal{P}_{\text{combinatoric } D^0} + N_{\text{fake } D^*} \mathcal{P}_{\text{combinatoric } D^{*-}} \,\}\end{aligned}$$

- 16 templates: 13 templates from MC , 3 templates from data
- 4 free parameters , 6 gaussian constrained parameters and 6 fixed parameters

Signal Fit – PDF

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

The total probability density function is:

$$
\begin{aligned}
\mathcal{P}_{\rm total}(q^2, t_\tau, {\rm BDT}) = 1/N_{\rm total} \times \{ & N_{\rm sig} [f_{\tau^+\to\pi^+\pi^-\pi^+\overline{\nu}_\tau} \mathcal{P}_{\tau^+\to\pi^+\pi^-\pi^+\overline{\nu}_\tau} + (1 - f_{\tau^+\to\pi^+\pi^-\pi^+\overline{\nu}_\tau}) \mathcal{P}_{\tau^+\to\pi^+\pi^-\pi^+\pi^0\overline{\nu}_\tau} \\
& + f_{D^{**}\tau\nu} \mathcal{P}_{B\to D^{**}\tau^+\nu_\tau}] + N_{D^0}^{\rm same} [\mathcal{P}_{B\to D^{*-}D^0 X\ {\rm SV}} + f_{D^0}^{v_1-v_2} \mathcal{P}_{B\to D^{*-}D^0 X\ {\rm DV}}] \\
& + N_{D_s^+}/k \times [\mathcal{P}_{B^0\to D^{*-}D_s^{*+}} + f_{D_s^+} \mathcal{P}_{B^0\to D^{*-}D_s^+} + f_{D_{s0}^{*+}} \mathcal{P}_{B^0\to D^{*-}D_{s0}^{*+}} \\
& + f_{D_{s1}^+} \mathcal{P}_{B^0\to D^{*-}D_{s1}^+} + f_{D^{**}D_s X} \mathcal{P}_{B\to D^{**-}D_s^+ X} + f_{B_s\to D^* D_s^+ X} \mathcal{P}_{B_s^0\to D^{*-}D_s^+ X}] \\
& + N_{D_s^+} f_{D^+} \mathcal{P}_{B\to D^{*-}D^+ X} + N_{B\to D^{*-}3\pi^\pm X} \mathcal{P}_{B\to D^{*-}3\pi^\pm X} \\
& + N_{B_1-B_2} \mathcal{P}_{{\rm combinatoric}\ B} + N_{{\rm fake}\ D^0} \mathcal{P}_{{\rm combinatoric}\ D^0} + N_{{\rm fake}\ D^*} \mathcal{P}_{{\rm combinatoric}\ D^{*-}} \}
\end{aligned}
$$

- 16 templates: 13 templates from MC , 3 templates from data
- 4 free parameters, 6 gaussian constrained parameters and 6 fixed parameters

Signal Fit – PDF

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

The total probability density function is:

$$\begin{aligned}\mathcal{P}_{\text{total}}(q^2, t_\tau, \text{BDT}) = 1/N_{\text{total}} \times \{ & N_{\text{sig}} [f_{\tau^+\to\pi^+\pi^-\pi^+\overline{\nu}_\tau} \mathcal{P}_{\tau^+\to\pi^+\pi^-\pi^+\overline{\nu}_\tau} + (1 - f_{\tau^+\to\pi^+\pi^-\pi^+\overline{\nu}_\tau}) \mathcal{P}_{\tau^+\to\pi^+\pi^-\pi^+\pi^0\overline{\nu}_\tau} \\ & + f_{D^{**}\tau\nu} \mathcal{P}_{B\to D^{**}\tau^+\nu_\tau}] + N_{D^0}^{\text{same}} [\mathcal{P}_{B\to D^{*-}D^0 X \text{ SV}} + f_{D^0}^{v_1-v_2} \mathcal{P}_{B\to D^{*-}D^0 X \text{ DV}}] \\ & + N_{D_s^+}/k \times [\mathcal{P}_{B^0\to D^{*-}D_s^{*+}} + f_{D_s^+} \mathcal{P}_{B^0\to D^{*-}D_s^+} + f_{D_{s0}^{*+}} \mathcal{P}_{B^0\to D^{*-}D_{s0}^{*+}} \\ & + f_{D_{s1}^+} \mathcal{P}_{B^0\to D^{*-}D_{s1}^+} + f_{D^{**}D_s X} \mathcal{P}_{B\to D^{**-}D_s^+ X} + f_{B_s\to D^* D_s^+ X} \mathcal{P}_{B_s^0\to D^{*-}D_s^+ X}] \\ & + N_{D_s^+} f_{D^+} \mathcal{P}_{B\to D^{*-}D^+ X} + N_{B\to D^{*-}3\pi^\pm X} \mathcal{P}_{B\to D^{*-}3\pi^\pm X} \\ & + N_{B_1-B_2} \mathcal{P}_{\text{combinatoric } B} + N_{\text{fake } D^0} \mathcal{P}_{\text{combinatoric } D^0} + N_{\text{fake } D^*} \mathcal{P}_{\text{combinatoric } D^{*-}} \}\end{aligned}$$

- 16 templates: 13 templates from MC, 3 templates from data
- 4 free parameters, 6 gaussian constrained parameters and 6 fixed parameters

Before the Signal Fit

- D_s^+ decay model fit weights are applied
- Reweight q^2 distribution according to $D^{(0,+)}$ control sample corrections

Before the Signal Fit

- D_s^+ decay model fit weights are applied
- Reweight q^2 distribution according to $D^{(0,+)}$ control sample corrections

- Constrained parameters :
  - The no–Charm yield $N_{B \to D^{*-} 3\pi^{\pm} X}$ is estimated using a **data-driven method**
  - The fractions f_i are derived from the $B^0 \to D^* D_s^+ X$ control mode

Before the Signal Fit

- D_s^+ decay model fit weights are applied
- Reweight q^2 distribution according to $D^{(0,+)}$ control sample corrections

- Fixed parameters:
  - Fraction of τ^+ decays with/without π^0 **$\rightarrow$ efficiencies**
  - Fake D^* and D^0 yields **$\rightarrow$ a 2D fit** to $m(D^0)$ and $\Delta m = m(D^*) - m(D^0)$.
  - Feed-down fraction $f_{D^{**}}$ **$\rightarrow$ simulation** corrected according to theoretical predictions.
  - ...

Before the Signal Fit

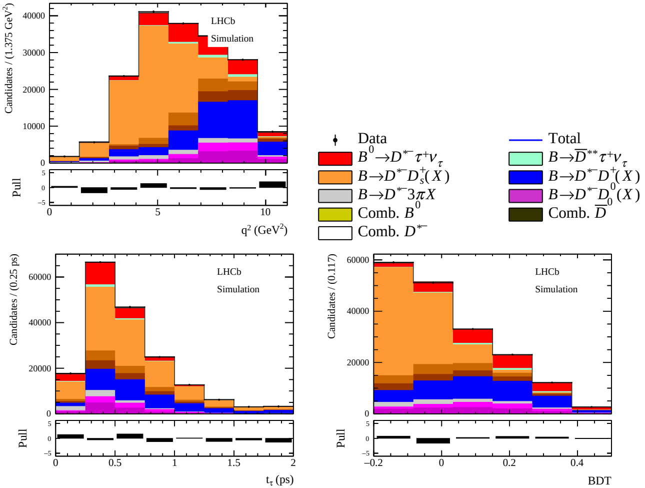
- D_s^+ decay model fit weights are applied
- Reweight q^2 distribution according to $D^{(0,+)}$ control sample corrections

- Fixed parameters:
  - Fraction of τ^+ decays with/without π^0 **$\rightarrow$ efficiencies**
  - Fake D^* and D^0 yields **$\rightarrow$ a 2D fit** to $m(D^0)$ and $\Delta m = m(D^*) - m(D^0)$.
  - Feed-down fraction $f_{D^{**}}$ **$\rightarrow$ simulation** corrected according to theoretical predictions.
  - ...

- Blinding strategy:
  - The signal yield in data is blinded: $\mathcal{N}_{\text{sig}} \rightarrow \mathcal{N}_{\text{sig}} \times x$
  - The relative uncertainty on the signal yield remains **unchanged** by the blinding.

Signal Fit – Inclusive MC (New Results)

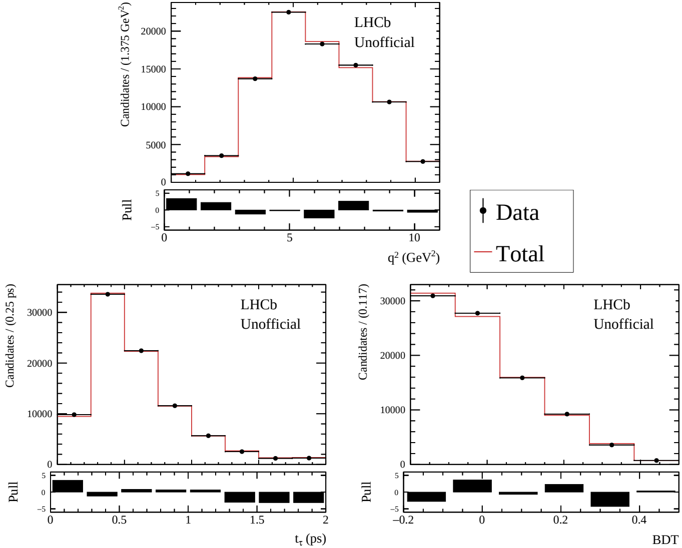
- Split the inclusive MC sample in two: one for **template building**, the other for **fitting**.
- Use fixed/constrained parameters from inclusive MC instead of data.



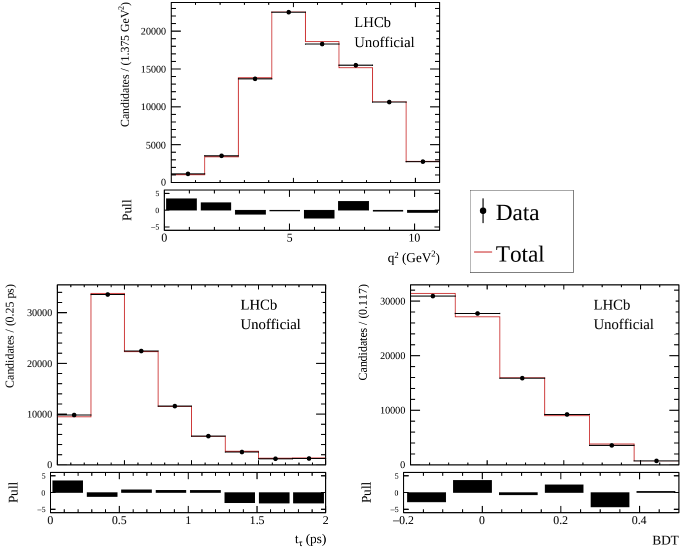
Parameter	Values	True Value	σ
Free Parameters			
N_{sig}	54689 ± 616	53439	2.0
N_{D_s}	272838 ± 1395	274799	-1.4
f_{D^+}	0.340 ± 0.005	0.333	1.7
$f_{D^0}^{\text{DV}}$	1.39 ± 0.06	1.48	-0.5
Constrained Parameters			
$N_{B \to D^{*-}3\pi X}$	15927 ± 646	16735	-1.3
f_{D_s}	0.491 ± 0.006	0.494	-0.5
$f_{D_{s0}^{*+}}$	0.092 ± 0.009	0.082	1.2
$f_{D_{s1}'^{+}}$	0.422 ± 0.009	0.464	-4.8
$f_{D^{**}D_sX}$	0.300 ± 0.016	0.267	
$f_{B_s \to D^* D_s X}$	0.257 ± 0.011	0.292	-3.3
Fixed Parameters			
$N_{D^0}^{\text{SV}}$	1912	1912	0
$f_{D^{**}\tau\nu}$	0.115	0.115	0
$f_{\tau \to 3\pi\overline{\nu}_\tau}$	0.738	0.738	0

(No combinatorial background in MC)

Signal Fit – Data (Blinded)



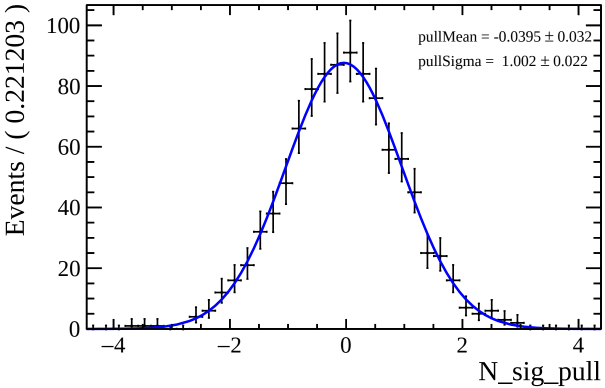
Parameter	Values	
Free Parameters		
N_{sig}	7815 ± 323	
N_{D_s}	53148 ± 752	
f_{D^+}	0.11 ± 0.01	
$f_{D^0}^{v_1-v_2}$	2.43 ± 0.20	
Constrained Parameters		
	Constrained Value	Fit Value
$N_{B\to D^{*-}3\pi^\pm X}$	6133 ± 400	6940 ± 347
f_{D_s}	0.546 ± 0.015	0.55 ± 0.01
$f_{D_{s0}^{*+}}$	0.052 ± 0.022	0.04 ± 0.02
$f_{D_{s1}^{\prime +}}$	0.489 ± 0.037	0.43 ± 0.02
$f_{D^{**}D_sX}$	0.220 ± 0.040	0.35 ± 0.03
$f_{B_s\to D^* D_s X}$	0.068 ± 0.017	0.08 ± 0.01
Fixed Parameters		
$N_{B_1 B_2}$	178 ± 13	
$N_{D^0}^{same}$	3233 ± 57	
$N_{fake\, D^0}$	802 ± 25	
$N_{fake\, D^*}$	1862 ± 49	
$f_{D^{**}\tau\nu}$	0.024 ± 0.000	
$f_{\tau^+\to 3\pi^\pm \overline{\nu}_\tau}$	0.768 ± 0.002	



- The fit quality: $\chi^2/\text{ndof} = 1.46$
- The relative uncertainty on the signal yield: $\sigma_{\mathcal{N}_{\text{sig}}}/\mathcal{N}_{\text{sig}} = 0.04$

Signal Fit – Pseudo-Experiments

Generate 1000 toy samples using the fit results from **the signal fit to data** .



$$\text{pull}(N_{\text{sig}}) = \frac{N^{toy}_{\texttt{sig}} - N_{\text{sig}}}{\sigma_{N_{\text{sig}}}}$$

The pull distributions for $\mathcal{N}_{\text{sig}}$ and other *free parameters* are centered at 0 with a width of 1.

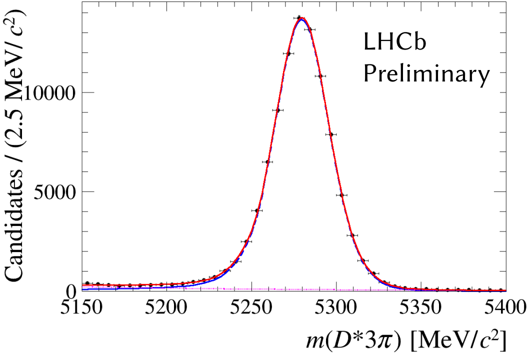
Outline

Normalization Mode – Fit Strategy

The normalization yield $\mathcal{N}_{\text{norm}}$ is extracted from an unbinned ML fit to the $D^*3\pi$ mass.

- **Total B^0 Yield** ($N_{\mathtt{tot}}$):
  - Fit to B^0 mass
- **Fake D^{*-}** combinatorial background can lead to B mass peak
  - Obtain the fake D^{*-} yield from Sidebands.
  - Estimate fake D^{*-} yield ($N_{\mathtt{fake}\ D^*}$).
- **Total D_s^+ Yield** ($N_{D_s^+}$):
  - Fit $3\pi^{\pm}$ mass around the D_s^+ mass.
  - Use same fake D^{*-} fraction to estimate true D_s^+ yield.
- **Normalization Yield** ($N_{\mathtt{norm}}$):
  - Subtract true D_s^+ yield: $N_{\mathtt{norm}} = N_{\mathtt{tot}} - N_{\mathtt{fake}\ D^*} - N_{D_s^+}$

Normalisation Fit – Result



Parameter	Value
N_{tot}	103590 ± 353
$N_{\mathrm{fake}D^*}$	9057 ± 315
N_B	94533 ± 473
$N_{D_s^+}$	1360 ± 60
N_{norm}	93173 ± 441

Outline

Recap.

At this stage, all the necessary components are in place to measure $\boldsymbol{R}(D^*)$:

$$\mathcal{K}(\mathcal{D}^*) = \frac{\mathcal{N}_{\mathtt{sig}}}{\mathcal{N}_{\mathtt{norm}}} \times \frac{\varepsilon_{\mathtt{norm}}}{\varepsilon_{\mathtt{sig}}} \times \frac{1}{\mathcal{B}(\tau^+ \to 3\pi^\pm \overline{\nu}_\tau) + \mathcal{B}(\tau^+ \to 3\pi^\pm \pi^0 \overline{\nu}_\tau)}$$

Recap.

At this stage, all the necessary components are in place to measure $\boldsymbol{R}(D^*)$:

$$\mathcal{K}(\mathcal{D}^*) = \frac{\mathcal{N}_{\texttt{sig}}}{\mathcal{N}_{\texttt{norm}}} \times \frac{\varepsilon_{\texttt{norm}}}{\varepsilon_{\texttt{sig}}} \times \frac{1}{\mathcal{B}(\tau^+ \to 3\pi^\pm \overline{\nu}_\tau) + \mathcal{B}(\tau^+ \to 3\pi^\pm \pi^0 \overline{\nu}_\tau)}$$

$\varepsilon_{\text{sig}} = 1.190 \pm 0.005 \times 10^{-4}$

$\varepsilon_{\text{norm}} = 3.466 \pm 0.019 \times 10^{-4}$

Recap.

At this stage, all the necessary components are in place to measure $\boldsymbol{R}(D^*)$:

$$\mathcal{K}(\mathcal{D}^*) = \frac{\mathcal{N}_{\texttt{sig}}}{\mathcal{N}_{\texttt{norm}}} \times \frac{\varepsilon_{\texttt{norm}}}{\varepsilon_{\texttt{sig}}} \times \frac{1}{\mathcal{B}(\tau^+ \to 3\pi^\pm \overline{\nu}_\tau) + \mathcal{B}(\tau^+ \to 3\pi^\pm \pi^0 \overline{\nu}_\tau)}$$

$\varepsilon_{\text{sig}} = 1.190 \pm 0.005 \times 10^{-4}$

$\varepsilon_{\text{norm}} = 3.466 \pm 0.019 \times 10^{-4}$

$\mathcal{N}_{\text{sig}} = 7\,815 \pm 323$

Recap.

At this stage, all the necessary components are in place to measure $\boldsymbol{R}(D^*)$:

$$\mathcal{K}(\mathcal{D}^*) = \frac{\mathcal{N}_{\texttt{sig}}}{\mathcal{N}_{\texttt{norm}}} \times \frac{\varepsilon_{\texttt{norm}}}{\varepsilon_{\texttt{sig}}} \times \frac{1}{\mathcal{B}(\tau^+ \to 3\pi^\pm \overline{\nu}_\tau) + \mathcal{B}(\tau^+ \to 3\pi^\pm \pi^0 \overline{\nu}_\tau)}$$

$\varepsilon_{\text{sig}} = 1.190 \pm 0.005 \times 10^{-4}$

$\mathcal{N}_{\text{sig}} = 7\,815 \pm 323$

$\varepsilon_{\text{norm}} = 3.466 \pm 0.019 \times 10^{-4}$

$\mathcal{N}_{\text{norm}} = 9\,3173 \pm 441$

Preliminary Results

$\mathcal{K}(D^{*-})$ is found to be:

$$\mathcal{K}(D^{*-}) = \frac{\mathcal{N}_{\mathtt{sig}}}{\mathcal{N}_{\mathtt{norm}}} \times \frac{\varepsilon_{\mathtt{norm}}}{\varepsilon_{\mathtt{sig}}} \times \frac{1}{\mathcal{B}(\tau^+ \to 3\pi^\pm \overline{\nu}_\tau) + \mathcal{B}(\tau^+ \to 3\pi^\pm \pi^0 \overline{\nu}_\tau)} = 1.829 \pm 0.067 \,.$$

The $\boldsymbol{R}(D^*)$ ratio is calculated as:

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

The blinded $\boldsymbol{R}(D^*)$ for the Run 2 dataset is obtained as:

$$\boldsymbol{R}(D^*)|_{Run\ 2} = 0.265 \pm 0.010\,(\mathtt{stat.}) \pm X(\mathtt{syst.}) \pm 0.012\,(\mathtt{ext.})\,.$$

The systematic uncertainty is yet **to be determined**.

Outline

Systematic Uncertainties

- The strategy will follow closely the previous analysis (partial Run 2).

Major sources are:

- Signal and background modelling.
- Selection criteria on signal $B^0 \to D^{*-}\tau^+\nu_\tau$ and normalisation $B^0 \to D^{*-}3\pi$ decay modes.
- Limited size of the simulation samples:
  - The largest contribution in Run 1: 4%.
  - Contributes to 2% in the partial Run 2 analysis.
- Empty bins in the templates.

Systematic Uncertainties – 2015–2016 Data

Source	Uncertainty (%)
PDF shapes uncertainty (size of simulation sample)	2.0
Fixing $B \to D^{*-} D_s^+(X)$ bkg model parameters	1.1
Fixing $B \to D^{*-} D^0(X)$ bkg model parameters	1.5
Fractions of signal τ^+ decays	0.3
Fixing the $\bar{D}^{**}\tau^+\nu_\tau$ and $D_s^{**+}\tau^+\nu_\tau$ fractions	$^{+1.8}_{-1.9}$
Knowledge of the $D_s^+ \to 3\pi X$ decay model	1.0
Specifically the $D_s^+ \to a_1 X$ fraction	1.5
Empty bins in templates	1.3
Signal decay template shape	1.8
Signal decay efficiency	0.9
Possible contributions from other τ^+ decays	1.0

Source	Uncertainty (%)
$B \to D^{*-} D^+(X)$ template shapes	$^{+2.2}_{-0.8}$
$B \to D^{*-} D^0(X)$ template shapes	1.2
$B \to D^{*-} D_s^+(X)$ template shapes	0.3
$B \to D^{*-} 3\pi X$ template shapes	1.2
Combinatorial background normalisation	$^{+0.5}_{-0.6}$
Preselection efficiency	2.0
Kinematic reweighting	0.7
Vertex error correction	0.9
PID efficiency	0.5
Signal efficiency (size of simulation sample)	1.1
Normalisation mode efficiency (modelling of $m(3\pi)$)	1.0
Normalisation efficiency (size of simulation sample)	1.1
Normalisation mode PDF choice	1.0

Total systematic uncertainty $^{+6.2}_{-5.9}\%$

Total statistical uncertainty 5.9%

Conclusions

- Analysis improvements w.r.t. the previous analysis:
  - Efficiency calculations.
  - Control samples, signal and normalization fits refinements.
- This study extends the dataset to include the full LHCb Run 2 data.
- The **blinded** value of $\mathcal{R}(D^*)$, using LHCb Run 2 data, is estimated to be:

$$\mathcal{R}(D^*)|_{\text{Run 2}} = 0.265 \pm 0.010\,(\texttt{stat.}) \pm X\,(\texttt{syst.}) \pm 0.012\,(\texttt{ext.})$$

The $\mathcal{R}(D^*)$ ratio was measured with a relative statistical uncertainty of:

- 6.7% with the **Run 1 dataset**.
- 5.8% using the **2015–2016 data samples only**.
- With the **Run 2 dataset**, the precision further improves to 3.8%.

Future Prospects

To-do list before publication:

- Improvements in the control samples and signal fits.
  - Improve $D \to 3\pi^{\pm} X$ model using BESIII results on inclusive $\mathbf{Br}(D^{(0,+)}_{(s)} \to 3\pi^{\pm} X)$.
  - Fits using inclusive and cocktail MC samples: control and signal fits.
- Systematic uncertainties estimation.
- Perform a fit to the combined Run 1 and Run 2 datasets.

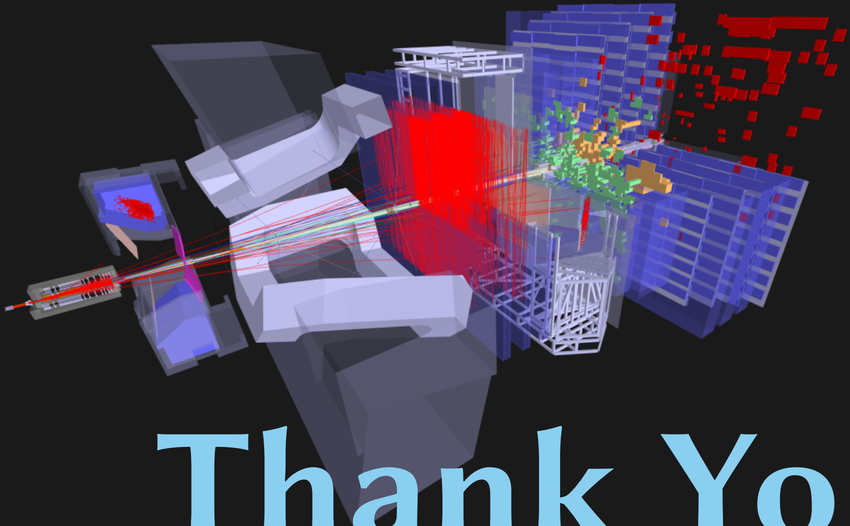
Future Prospects

To-do list before publication:

- Improvements in the control samples and signal fits.
  - Improve $D \to 3\pi^{\pm}X$ model using BESIII results on inclusive $\mathbf{Br}(D_{(s)}^{(0,+)} \to 3\pi^{\pm}X)$.
  - Fits using inclusive and cocktail MC samples: control and signal fits.
- Systematic uncertainties estimation.
- Perform a fit to the combined Run 1 and Run 2 datasets.

Exciting times ahead: Future investigations will explore LFU across:

- **Decay channels:** various leptons in the final states, ratios of other mesons and baryons.
- **Observable types:** angular observables, in q^2 bins, τ^+ and D^* polarizations.
- **Larger datasets:** currently Run 3 ($\sim 25\,\text{fb}^{-1}$) and future data collection periods.


LHCb Experiment at CERN
Run / Event: 255623 / 300064
Data recorded: 2022-11-25 09:40:16 GMT

Thank You!

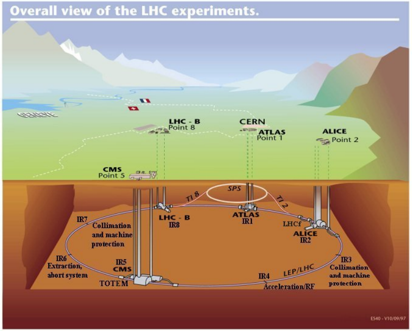
Part II

Backups

Outline of Backups

The LHCb experiment at the Large Hadron Collider

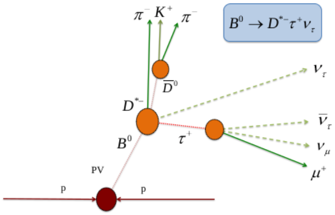
- The Large Hadron Collider (LHC) is a proton-proton accelerator
- LHCb is one of experiments based at the LHC at CERN, Geneva
- Forward spectrometer initially designed to search for New Physics in the beauty quark sector
- Now very broad programme: charm and top quark, heavy ions, electro-weak physics, 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

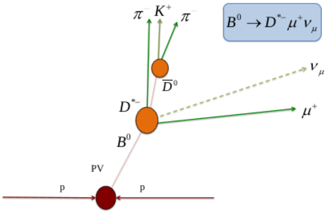


$R(D)$–$R(D^*)$ with muonic τ decays

[arXiv:2302.02886]

- Simultaneous measurement of $R(D)$ and $R(D^*)$ with **Run 1** data
    - **Muonic** channel $\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$
- No narrow peak to fit (3 neutrinos in final state)
- Backgrounds : partially reconstructed B decays
    - $B \to D^* \mu\nu$, $B \to D^{**} \mu\nu$, $B \to D^* DX$ with $D \to \mu X$, ...
- Select $D^0\mu^-$ and $D^{*+}\mu^-$ candidates where
    - $D^0 \to K^-\pi^+$, $D^{*+} \to D^0\pi^+$
    - Reconstructed $D^{*+} \to D^0\pi^+$ is vetoed in $D^0\mu^+$ sample
- Trigger on D^0 - preserve acceptance for soft muons
- Custom muon ID classifier, flatter in kinematic acceptance
    - Reduces misID background



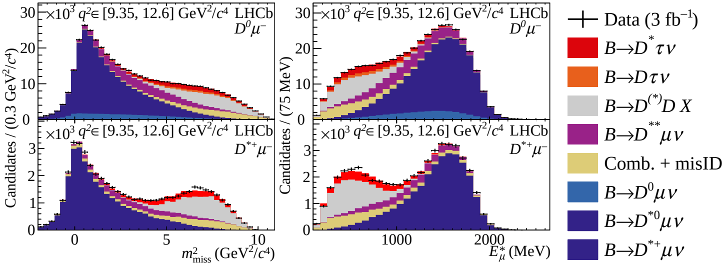


[arXiv:2302.02886]

- Simultaneous measurement of $R(D)$ and $R(D^*)$ with Run 1 data using **muonic** $\tau^+ \to \mu^+ \nu_\mu \bar{\nu}_\tau$

3D template fit to

- $q^2 \equiv (p_B - p_{D^*})^2$
- $m^2_{\text{miss}} \equiv (p_B - p_{D^*} - p_\mu)^2$
- E^*_μ energy of μ



$$\begin{cases} R(D) & = 0.441 \pm 0.060(\text{stat}) \pm 0.066(\text{syst}) \\ R(D^*) & = 0.281 \pm 0.018(\text{stat}) \pm 0.023(\text{syst}) \end{cases}$$

Agreement with SM within 1.9σ

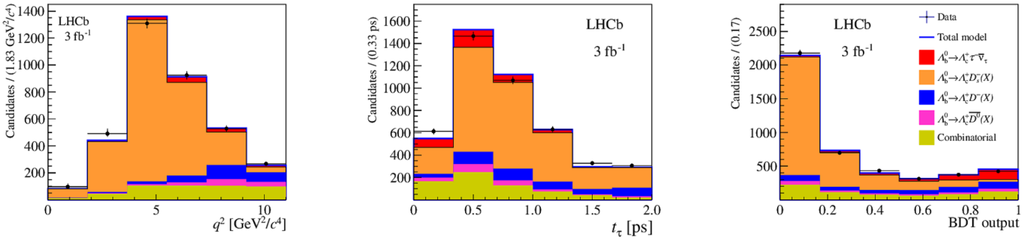
$R(\Lambda_c)$ with hadronic τ decays

- First LFU test in a **baryonic** $b \to c\ell\nu_\ell$ decay with Run 1 data using **hadronic** $\tau^+ \to 3\pi^\pm(\pi^0)$
- Normalisation channel $\Lambda_b^0 \to \Lambda_c^+ 3\pi^\pm$

$$\mathcal{K}(\Lambda_c^+) = \frac{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+\tau^-\overline{\nu}_\tau)}{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ 3\pi)}$$

$$R(\Lambda_c^+) = \mathcal{K}(\Lambda_c^+)\left\{\frac{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ 3\pi)}{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+\mu^-\overline{\nu}_\mu)}\right\}_{\text{ext. input}}$$

- 3D template fit to extract signal yield



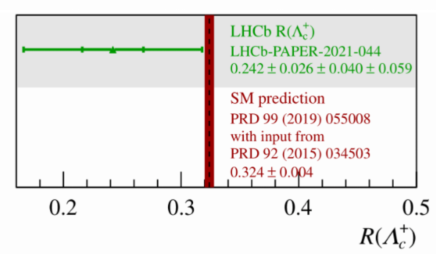
$R(\Lambda_c)$ with hadronic τ decays

- First LFU test in a **baryonic** $b \to c\ell\nu_\ell$ decay with Run 1 data using **hadronic** $\tau^+ \to 3\pi^\pm(\pi^0)$
- Normalisation channel $\Lambda_b^0 \to \Lambda_c^+ 3\pi^\pm$

$$\mathcal{K}(\Lambda_c^+) = \frac{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ \tau^- \overline{\nu}_\tau)}{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ 3\pi)}$$

$$R(\Lambda_c^+) = \mathcal{K}(\Lambda_c^+) \left\{ \frac{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ 3\pi)}{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ \mu^- \overline{\nu}_\mu)} \right\}_{\text{ext. input}}$$

- 3D template fit to extract signal yield

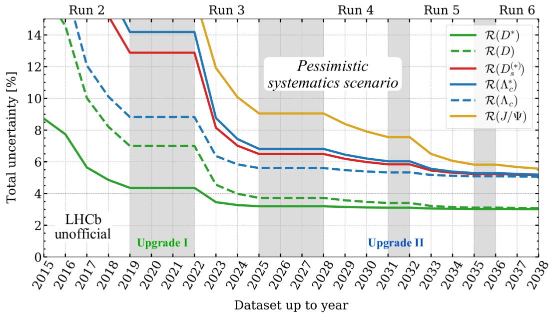


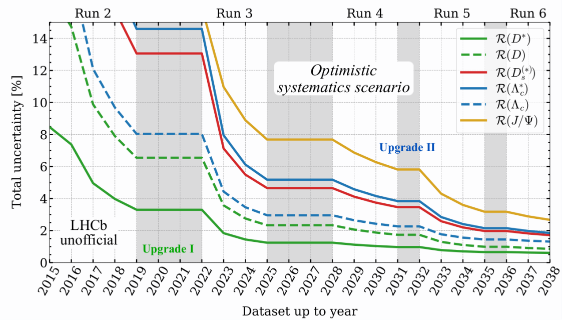
$$\begin{cases} \mathcal{K}(\Lambda_c^+) = 2.46 \pm 0.27(\text{stat}) \pm 0.40(\text{syst}) \\ R(\Lambda_c^+) = 0.242 \pm 0.026(\text{stat}) \pm 0.040(\text{syst}) \pm 0.059(\text{ext}) \end{cases}$$

Agreement within 1.0σ to SM

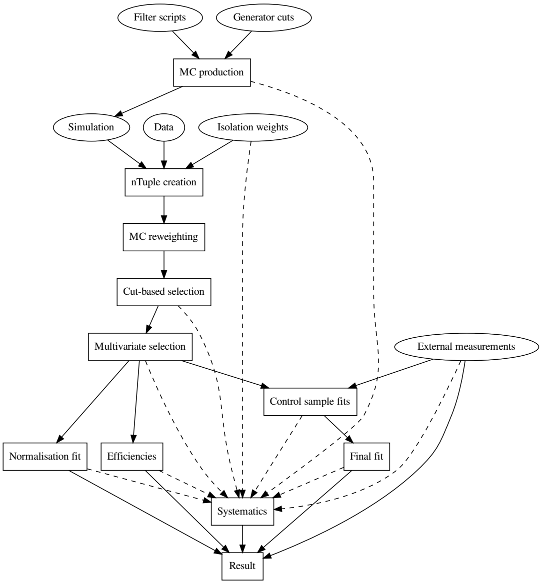
[PRL 128, 191803 (2022)]

Future of $R(X_c)$ at LHCb





Workflow



Filter scripts
Generator cuts
MC production
Simulation
Data
Isolation weights
nTuple creation
MC reweighting
Cut-based selection
Multivariate selection
External measurements
Control sample fits
Normalisation fit
Efficiencies
Final fit
Systematics
Result

Summary of Contributions – 1

- Performed thorough cross-check on the analysis using 2015–2016 datasets
  - Contributed to systematic studies, including effects from selection criteria such as BDT cuts, efficiency calculations, and PID efficiency.
- Extended dataset to full Run 2 (2017 and 2018 data) with several improvements:
  - Explored effects of changes in stripping and trigger configurations.
  - Developed a vertex correction scheme to reconcile differences between new data algorithms and existing MC simulations.
  - Improved efficiency calculations by eliminating multiple event candidates.

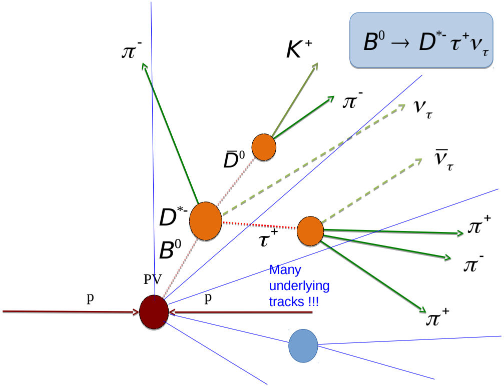
Summary of Contributions – 2

- Contributions to
  - New method for calculating the normalization yield and estimating bias via bootstrapping.
  - Parameterized FF using BGL parameterization.
  - Contributed to the production of D_s^+ MC samples, and their integration into the analysis: BDT training, control sample and signal fits.
- Enhanced analysis code
  - Code reproducibility, optimization, and parallelization
  - Separate results per dataset for cross-checks
  - Documentation

Outline

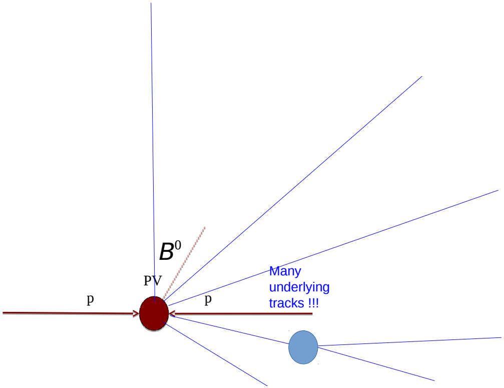
Simulated Samples

Decay Descriptor		EventType	Events [M] Generated	Filtered
$B^0 \to D^{*-}\tau^+\nu_\tau$	(with $\tau^+ \to 3\pi^\pm\overline{\nu}_\tau$)	11160001	257.0	30.0
	(with $\tau^+ \to 3\pi^\pm\pi^0\overline{\nu}_\tau$)	11563020	198.0	49.0
$B^0 \to D^{*-}3\pi^\pm$		11266018	651.0	12.0
$B^0 \to D^{**}\tau^+\nu_\tau$	(with $\tau^+ \to 3\pi^\pm\overline{\nu}_\tau$)	11566430	53.0	4.0
		11566431	245.0	16.0
$B_s^0 \to D^{*-}D_s^+X$	(with $D_s^+ \to 3\pi^\pm X$)	13996621	125.0	1.0
	(with $D_s^+ \to K\pi X$)	13996622	69.0	0.09
	(with $D_s^+ \to \ell\nu_\ell X$)	13996623	30.0	0.3
	(with $D_s^+ \to a_1(1260)^+X$ and $a_1(1260)^+ \to \pi^+2\pi^0$)	13996624	27.0	0.2
$B^+ \to D^{*-}D_s^+X$	(with $D_s^+ \to 3\pi^\pm X$)	12997623	148.0	2.0
	(with $D_s^+ \to K\pi X$)	12997624	78.0	0.2
	(with $D_s^+ \to \ell\nu_\ell X$)	12997625	45.0	0.5
	(with $D_s^+ \to a_1(1260)^+X$ and $a_1(1260)^+ \to \pi^+2\pi^0$)	12997626	31.0	0.3
$B^0 \to D^{*-}D_s^+X$	(with $D_s^+ \to 3\pi^\pm X$)	11896623	436.0	6.0
	(with $D_s^+ \to K\pi X$)	11896624	302.0	0.5
	(with $D_s^+ \to \ell\nu_\ell X$)	11896625	142.0	1.0
	(with $D_s^+ \to a_1^+X$ and $a_1(1260)^+ \to \pi^+2\pi^0$)	11896626	92.0	0.8
$b\overline{b} \to D^{*-}3\pi^\pm X$		27163970	18132.0	74.0
$b\overline{b} \to D^{*-}D^{\{0,+\}}X$		27163971	822.0	6.0
Total		–	21884.0	205.0



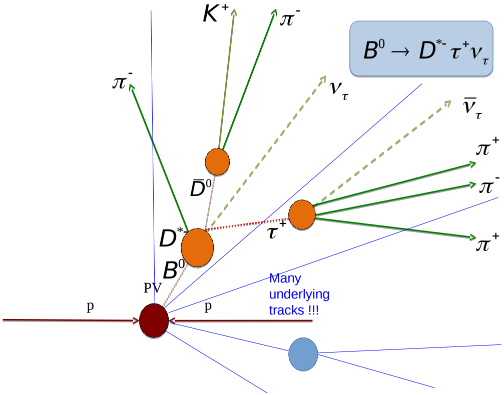
1. Generate 1 complete event: signal + underlying event

Fast simulation with ReDecay



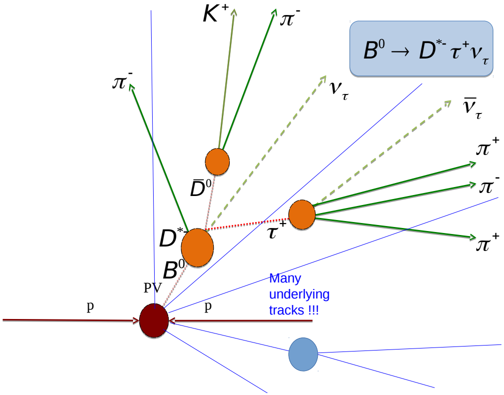
1. Generate 1 complete event: signal + underlying event
2. Re-generate the B decay 100 times and merge each with the underlying event

Fast simulation with ReDecay



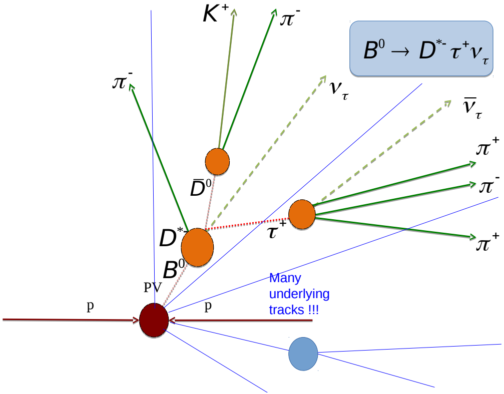
1. Generate 1 complete event: signal + underlying event
2. Re-generate the B decay 100 times and merge each with the underlying event

Fast simulation with ReDecay



1. Generate 1 complete event: signal + underlying event
2. Re-generate the B decay 100 times and merge each with the underlying event
3. Repeat 1 and 2 N times

Fast simulation with ReDecay



1 Generate 1 complete event: signal + underlying event

2 Re-generate the B decay 100 times and merge each with the underlying event

3 Repeat 1 and 2 N times

$\rightarrow$ Factor $\mathcal{O}(10)$ faster simulation

Vertex Error Parametrizatioon

Issue: Radiation damage in the detector led to the degradation of data quality in 2015-2016, which is not reflected in the simulation.

Improvement: An improvement of the vertex reconstruction algorithm was applied to data from 2017 onwards, **but not to simulation** , leading to disagreement between data and simulation samples.

Goal:

- Achieve the same resolution in data and MC.
- Width of the $\Delta z/\sigma_{\Delta z}$ distribution: Ensured to equal 1.

Solution: Parametrization of vertex position ($\Delta z/\sigma_{\Delta z}$) using control samples of $B \to J/\Psi K^+\pi^+\pi^-$ decays, providing two independent measurements of the B vertex position. then reweighting the distribution of modified uncertainties of vertex position.

Vertex correction strategy

- As is well known, for each data taking period we have:
  - Various Δz resolution
  - Various deviations of the pulls of the Δz resolution
  - Various effective luminosities

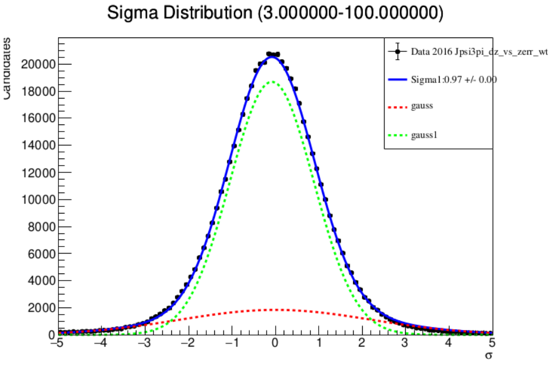
$$\Delta z = zvtx(\tau^+) - zvtx(B^0) \qquad \sigma_{\Delta z} = \sqrt{zvtx\_err(\tau^+)^2 + zvtx\_err(B^0)^2}$$

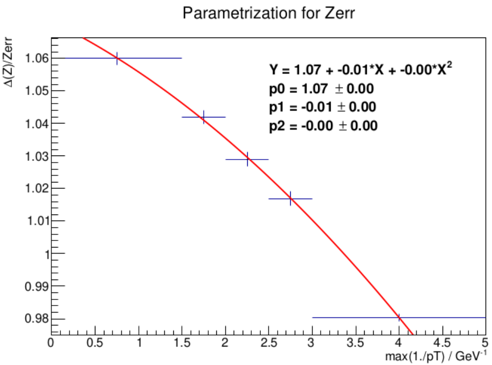
- The strategy is to pile up the data after the 3 following steps:
  - For each period, rescale the `tau_ENDVERTEX_ZERR` and `B0_ENDVERTEX_ZERR` so that the pulls have a unit gaussian width, both in data and MC samples
    - This is done using a $J/\Psi K\pi\pi$ sample for data and MC where Δz is **centered on** 0.
  - For the MC, reweight the corrected `tau_ENDVERTEX_ZERR` and `B0_ENDVERTEX_ZERR` to the corrected `tau_ENDVERTEX_ZERR` and `B0_ENDVERTEX_ZERR` data distributions
    - This is done using the sample used for the final fit.
  - Pile up the MC stats for the proper ratio of effective luminosity for each period
    - *i.e.* proportional to the number of $D^*3\pi$ events, corrected by the $\varepsilon_{\text{sig}}/\varepsilon_{\text{norm}}$ ratio.

Vertex Error Parametrizatioon

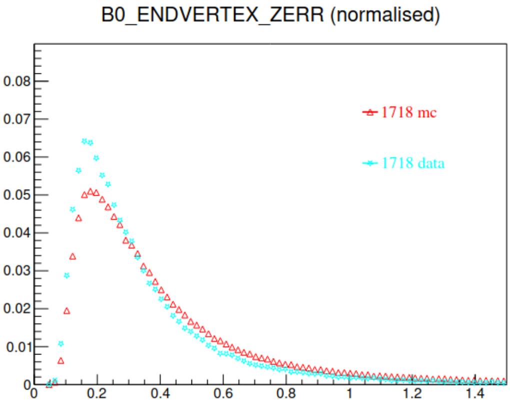
Using data and mc samples of $B^+ \to J/\Psi K^+ \pi^+ \pi^-$ decays as control samples:

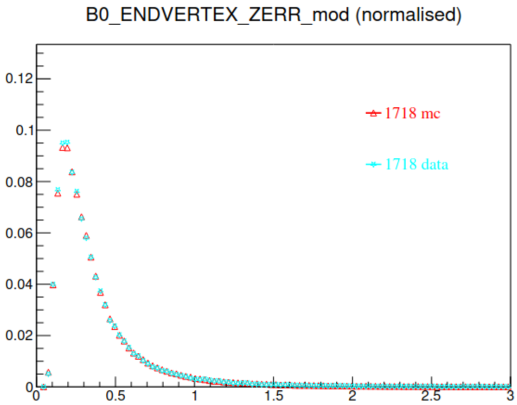
- Fit the Δz distribution in continuous $\max(1/p_T)$ bins with two Gaussian functions. Use the narrower Gaussian width (σ) as a proxy for Δz.
- Model σ as a function of $\max(1/p_T)$ with a polynomial of degree 2. Scale the error variable in each event by a factor derived from this polynomial.
- Reweight the modified error variable distribution in mc samples using GBReweighter to match the corresponding distribution in data.



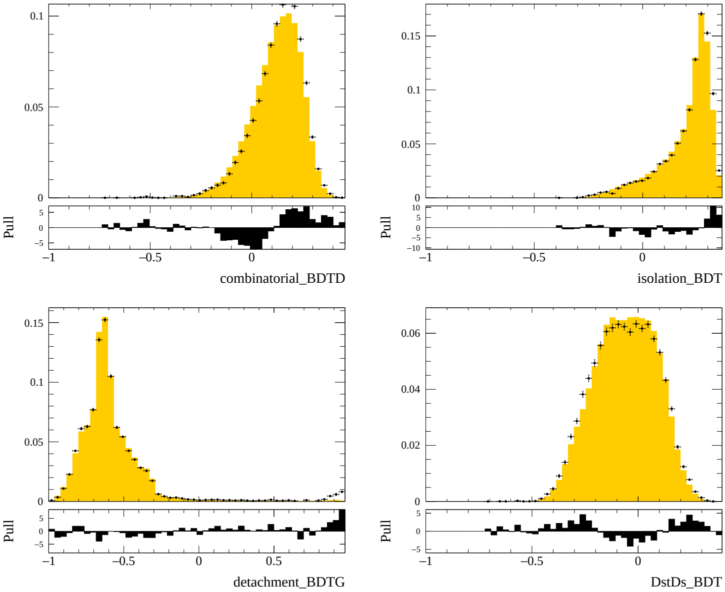


Vertex Error Parametrizatioon





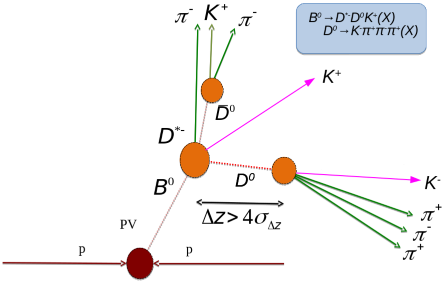
BDTs – Data and MC Comparison



Double-charm backgrounds

- $B \to D^{*-}(D^+, D^0)X$ decays are sub-leading contributors
- $D^+ \to K^-\pi^+\pi^+(\pi^0)$ contributes to the $B \to D^{*-}D^+X$ backgrounds
  - Significant when π^- is misidentified as K^-
  - Tight particle identification requirements

- $D^0 \to K^-3\pi^\pm$ contributes to the $B \to D^{*-}D^0X$ backgrounds
  - When there is an extra charged track
  - A BDT classifier is used to reject such events



Double-charm Backgrounds

- Double-charm decays being the largest fraction in the final sample, need to be modelled well in the final signal extraction fit
  - Templates used in the signal fit are derived from simulation and corrections need to be applied wherever necessary
- Specific control samples derived using the peculiarities of these decays for further studies

Decay	**Sample selection**
$B \to D^{*-} D_s^+ X$	reversing the anti-D_s^+ BDT selection
$B \to D^{*-} D_s^+ (\to 3\pi^\pm) X$	$m(3\pi^\pm)$ around D_s^+ mass
$B \to D^{*-} D^+ X$	kaon mass hypothesis given to π^- among the $3\pi^\pm$ candidates
$B \to D^{*-} D^0 X$	additional charged track (kaon) selected in an event

Signal Mode Selection

- Remaining cuts for the signal mode

Requirement		Targeted Background
$\Delta z(B^0, \tau^+)$	> 2	prompt
$m(K^-\pi^+)$	$\in [1840, 1890]\,\text{MeV}/c^2$	combinatorial D^0
$m(D^{*-}) - m(K^-\pi^+)$	$\in [143, 148]\,\text{MeV}/c^2$	combinatorial D^{*-}
$m(\tau^+)$	$< 1600\,\text{MeV}/c^2$	double-charm
$m(B^0)$	$< 5100\,\text{MeV}/c^2$	combinatorial
q^2	$\in [0, 11]\,\text{GeV}^2/c^4$	combinatorial
mva cuts		
DstDs_BDT	> -0.2	$D^{*-} D_s^+ X$
Isolation_BDT	> 0.1	double-charm
Combinatorial_BDTD	> 0.0	combinatorial
Detachment_BDTG	> 0.2	prompt
pid cuts		
$\text{ProbNN}_\pi(\pi^-)$	> 0.1 (π^- from D^{*-})	misidentification
$\text{ProbNN}_\pi(\pi^\pm)$	> 0.6 ($\pi^\pm$ from τ^+)	
$\text{ProbNN}_K(\pi^-)$	< 0.1 (π^- from τ^+)	

Normalization Mode Selection

- Remaining cuts for the normalization mode

Requirement		Targeted Background
$m(B^0)$	$\in [5150, 5400]$ MeV	combinatorial
$m(D^{*-}) - m(\bar{D}^0)$	$\in [143, 148]$ MeV$/c^2$	combinatorial D^{*-}
$m(K^-\pi^+)$	$\in [1840, 1890]$ MeV$/c^2$	combinatorial D^0
$\Delta z(\tau^+, \bar{D}^0)$	> 4	non-prompt
mva cuts		
Isolation_BDT	> 0.1	double-charm
Combinatorial_BDTD	> 0.0	combinatorial
pid cuts		
$\text{ProbNN}_\pi(\pi^-)$	> 0.1 (π^- from D^{*-})	misidentification
$\text{ProbNN}_\pi(\pi^\pm)$	> 0.6 ($\pi^\pm$ from τ^+)	
$\text{ProbNN}_K(\pi^-)$	< 0.1 (π^- from τ^+)	

Efficiency Calculation Overview

- Efficiencies for signal $\varepsilon_{\texttt{sig}}$ and normalization modes $\varepsilon_{\texttt{norm}}$ are calculated using simulated samples.
- Separate calculations are done for each dataset and polarity configuration $\varepsilon(p, d)$.
- Online efficiency is the product of geometrical acceptance and preselection cuts efficiency:

$$\varepsilon_{\text{online}} = \varepsilon_{\text{geometry}} \cdot \varepsilon_{\text{preselection}}$$

- Offline efficiency is the product of efficiencies obtained from different selection steps:

$$\varepsilon_{\text{offline}} = \prod_{i \in \text{steps}} \varepsilon_{\text{i}} = \prod_{i \in \text{steps}} \frac{N_i}{N_{i-1}}$$

where N_i is the number of events passing the step i, and i the offline selection step: common, mode-specific and PID cuts.

Efficiency Steps

Online and offline efficiency averages at different stages of the selection:

	Selection Step	Efficiency (%) sig 1 $\tau^+ \to 3\pi\overline{\nu}_\tau$	sig 2 $\tau^+ \to 3\pi\pi^0\overline{\nu}_\tau$	norm –
Online	Geometry	16.44 ± 1.98	15.76 ± 1.81	15.67 ± 1.70
	Filtering & Stripping	45.55 ± 0.09	45.52 ± 0.08	45.95 ± 0.12
	DaVinci	54.05 ± 0.47	54.04 ± 0.54	53.63 ± 0.51
Offline	Common & Trigger	63.26 ± 1.62	59.94 ± 1.57	41.40 ± 1.58
	Mode-Specific	39.67 ± 1.83	30.91 ± 1.73	61.43 ± 1.97
	PID	79.10 ± 1.37	81.59 ± 1.27	77.01 ± 1.40

Fraction Calculation for $\tau^+ \to 3\pi\overline{\nu}_\tau$ Decays

The fraction of $\tau^+ \to 3\pi^\pm$ decays is calculated according to the formulas:

$$f_{\tau^+\to 3\pi\overline{\nu}_\tau} = \frac{\langle\varepsilon_{\mathtt{sig\,1}}\rangle \cdot \mathrm{Br}(\mathtt{sig\,1})}{\langle\varepsilon_{\mathtt{sig\,1}}\rangle \cdot \mathrm{Br}(\mathtt{sig\,1}) + \langle\varepsilon_{\mathtt{sig\,2}}\rangle \cdot \mathrm{Br}(\mathtt{sig\,2})}$$

Decay		Branching Fraction (%)
		PDG
sig 1	$\tau \to 3\pi\nu$	9.02 ± 0.05
sig 2	$\tau \to 3\pi\pi^0\nu$	4.49 ± 0.05

using the branching fractions, the fraction is calculated as:

$$f_{\tau^+\to 3\pi\overline{\nu}_\tau} = 0.768 \pm 0.002$$

This fraction is used as a fixed parameter in the signal extraction fit.

D_s^+ decay model

New logic for Ds fit:

- Applying scaling factors for four types of MC ($3\pi X$, SL, $K\pi h$, $a_1\pi^0$) to align them since they were generated independently and events number didn't align with branching fraction.
- Extracting templates from these MC samples.
- Fixing $a_1X/(a_1X+3\pi X)$ fractions from anti-D_s^+ BDT efficiencies and branching fractions. ($X=\eta, \eta', \phi, \omega, K_s^0$)
- Fit with templates.

			Correction factor		
Sample	Decay descr.	EvtType	2015/16	2017	2018
old MC	$B^0 \to D^{*-} D_s^+ X$	11896612	1	1	1
new MC	$B^0 \to D^{*-} D_s^+ X$ (with $D_s^+ \to 3\pi^\pm X$)	11896623	1	1	1
	$B^0 \to D^{*-} D_s^+ X$ (with $D_s^+ \to K\pi h$)	11896624	0.948	0.976	0.943
	$B^0 \to D^{*-} D_s^+ X$ (with $D_s^+ \to \ell\nu_\ell X$)	11896625	0.419	0.531	0.345
	$B^0 \to D^{*-} D_s^+ X$ (with $D_s^+ \to a_1^+ X$ and $a_1^+ \to \pi^+\pi^0\pi^0$)	11896626	0.396	0.365	0.373

Table: D_s decay model studies

D_s^+ decay model

New logic for Ds fit:

- Applying scaling factors for four types of MC ($3\pi X$, SL, $K\pi h$, $a_1\pi^0$) to align them since they were generated independently and events number didn't align with branching fraction.
- Extracting templates from these MC samples.
- Fixing $a_1X/(a_1X+3\pi X)$ fractions from anti-D_s^+ BDT efficiencies and branching fractions. ($X=\eta, \eta', \phi, \omega, K_s^0$)
- Fit with templates.

			Correction factor		
Sample	Decay descr.	EvtType	2015/16	2017	2018
old MC	$B^0 \to D^{*-} D_s^+ X$	11896612	1	1	1
new MC	$B^0 \to D^{*-} D_s^+ X$ (with $D_s^+ \to 3\pi^\pm X$)	11896623	1	1	1
	$B^0 \to D^{*-} D_s^+ X$ (with $D_s^+ \to K\pi h$)	11896624	0.948	0.976	0.943
	$B^0 \to D^{*-} D_s^+ X$ (with $D_s^+ \to \ell\nu_\ell X$)	11896625	0.419	0.531	0.345
	$B^0 \to D^{*-} D_s^+ X$ (with $D_s^+ \to a_1^+ X$ and $a_1^+ \to \pi^+\pi^0\pi^0$)	11896626	0.396	0.365	0.373

Table: D_s decay model studies

Outline

Signal Fit Components

- Number of $B^0 \to D^{*-}\tau^+\nu_\tau$ events : $\mathcal{N}_{\rm sig}$
- Fraction of $\tau^+ \to 3\pi\overline{\nu}_\tau$ decays relative to the sum of $\tau^+ \to 3\pi\overline{\nu}_\tau$ and $\tau^+ \to 3\pi\pi^0\overline{\nu}_\tau$ decays: $f_{\tau^+\to 3\pi\overline{\nu}_\tau}$
- Fraction of $B \to \bar{D}^{**}\tau^+\nu_\tau$ decays relative to $\mathcal{N}_{\rm sig}$ decays: $f_{D^{**}\tau\nu}$
- Number of $B \to D^{*-} D^0 X$ events where all pions in the $3\pi^\pm$ system originate from the D^0 vertex: $N_{D^0}^{\rm same}$
- Ratio of $B \to D^{*-} D^0 X$ decays where at least one pion originates from the D^0 vertex and the other pions from a different vertex, normalized to $N_{D^0}^{\rm same}$: $f_{D^0}^{v_1-v_2}$
- Ratio of $B \to D^{*-} D^+ X$ decays relative to those containing a D_s^+: f_{D^+}
- Yield of events involving a D_s^+: $N_{D_s^+}$
- Yield of the inclusive prompt $B \to D^{*-}3\pi X$ events where the three pions come from the B vertex: $N_{B\to 3\pi^\pm X}$
- Combinatorial bkg events yield where the D^{*-} and the 3π system come from different B decays: $N_{B_1 B_2}$
- Yield of combinatorial bkg events with a fake D^0: $N_{{\rm fake}D^0}$
- Yield of combinatorial bkg events with a fake D^{*-}: $N_{{\rm fake}D^*}$

Component	Normalization
$B^0 \to D^{*-}\tau^+\nu_\tau$ ($\tau^+ \to 3\pi\overline{\nu}_\tau$)	$\mathcal{N}_{\rm sig} \times f_{\tau^+\to 3\pi\overline{\nu}_\tau}$
$B^0 \to D^{*-}\tau^+\nu_\tau$ ($\tau^+ \to 3\pi\pi^0\overline{\nu}_\tau$)	$\mathcal{N}_{\rm sig} \times (1 - f_{\tau^+\to 3\pi\overline{\nu}_\tau})$
$B \to \bar{D}^{**}\tau^+\nu_\tau$	$\mathcal{N}_{\rm sig} \times f_{D^{**}\tau\nu}$
$B \to D^{*-}3\pi^\pm X$	$N_{B\to D^{*-}3\pi^\pm X}$
$B \to D^{*-}D^0 X$ (SV)	$N_{D^0}^{\rm same}$
$B \to D^{*-}D^0 X$ (DV)	$N_{D^0}^{\rm same} \times f_{D^0}^{v_1-v_2}$
$B \to D^{*-}D^+ X$	$N_{D_s^+} \times f_{D^+}$
$B^0 \to D^{*-}D_s^+$	$N_{D_s^+} \times f_{D_s^+}/k$
$B^0 \to D^{*-}D_s^{*+}$	$N_{D_s^+} \times 1/k$
$B^0 \to D^{*-}D_{s0}^{*+}$	$N_{D_s^+} \times f_{D_{s0}^{*+}}/k$
$B^0 \to D^{*-}D_{s1}^+$	$N_{D_s^+} \times f_{D_{s1}^+}/k$
$B \to D^{**-}D_s^+ X$	$N_{D_s^+} \times f_{D^{**}D_s X}/k$
$B_s^0 \to D^{*-}D_s^+ X$	$N_{D_s^+} \times f_{B_s\to D^* D_s^+ X}/k$
Combinatorial B	$N_{B_1 B_2}$
Combinatorial D^*	$N_{{\rm fake}\ D^*}$
Combinatorial D^0	$N_{{\rm fake}\ D^0}$

Constrained Parameters

Parameter	Fit result	Efficiency correction	Corrected fraction
f_{D_s}	0.556 ± 0.016	0.981	0.546 ± 0.015
$f_{D_{s0}^{*+}}$	0.050 ± 0.021	1.039	0.052 ± 0.022
$f_{D_{s1}^{\prime +}}$	0.438 ± 0.033	1.116	0.489 ± 0.037
$f_{D^{**} D_s X}$	0.188 ± 0.034	1.170	0.220 ± 0.04
$f_{B_s \to D^* D_s X}$	0.067 ± 0.017	1.010	0.068 ± 0.017

Estimation of no–Charm Yield in Data

- Methodology involves selecting yield from peak and prompt regions.
  - Peak region: $m(D^{*-}3\pi^{\pm}) \in [5200, 5350]\,\text{MeV}/c^2$
  - Prompt region: $m(D^{*-}3\pi^{\pm}) \leq 5100\,\text{MeV}/c^2$
- Correction factor accounts for MC and data discrepancies:

$$N(B \to D^* 3\pi X)^{\text{data}} = N(B \to D^* 3\pi)^{\text{data}} \cdot \frac{N(B \to D^* 3\pi X)^{\text{MC}}}{N(B \to D^* 3\pi)^{\text{MC}}} \cdot f_{\text{Corr.}}$$

- The correction factor is determined from the ratio of the yields in MC and data.
- The correction factor $f_{\text{Corr.}}$, accounting for discrepancies between MC modeling and real data, is derived from a purely prompt sample:

$$f_{\text{Corr.}} = \frac{N(B \to D^* 3\pi X)^{\text{data}}_{\text{prompt}}}{N(B \to D^* 3\pi)^{\text{data}}_{\text{prompt}}} \Bigg/ \frac{N(B \to D^* 3\pi X)^{\text{MC}}_{\text{prompt}}}{N(B \to D^* 3\pi)^{\text{MC}}_{\text{prompt}}}$$

Prompt yield

$$N^{\text{data}}_{B\to D^* 3\pi X} = N^{\text{data}}_{B\to D^* 3\pi} \times \frac{N^{\text{MC}}_{B\to D^* 3\pi X}}{N^{\text{MC}}_{B\to D^* 3\pi}} \times R_{\text{Corr.}}$$

To select prompt we require: $m(B^0) < 5100$ MeV
In MC we have events with: $m(B^0) > 4500$ MeV
If we require $4500 < m(B^0) < 5100$ MeV

	2015	2016	2017	2018
Prompt data	460	73	439	565
Peak data	71	68	368	505
Prompt mc	779	872	886	1444
Peak mc	183	1698	1611	2706
Rcorr	1.52	2.09	2.17	2.1
Yield	123	1678	1720	1867

Table: If we require $m(B^0) > 4500$ MeV

2015	2016
545	573
71	68
900	6912
183	1698
1.56	2.07
147	1662

Table: wo $m(B^0) > 4.5$ GeV

2015	2016
149	133
79	80
103	693
87	771
1.59	1.85
233	1822

Table: Resmi

Correction ratio compatible between **with** and **without** $m(B^0) > 4500$ MeV requirement

Fake D^* Template Construction

Definition: Fake D^* & true D^0

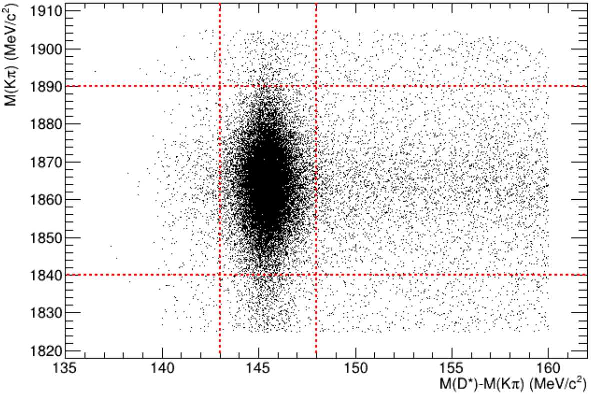
Components in data:

1. D^* sideband & D^0 mass window: fake D^* & (fake D^0 + **true** D^0)
   - True D^0 dominated
2. D^* sideband & D^0 sideband: fake D^* & (**fake** D^0 + true D^0)
   - Fake D^0 dominated
     (Still some true D^0 events because the D^0 sideband is just 25 MeV far away from the peak)

(quasi) Data-driven method to obtain fake D^* & true D^0 template:

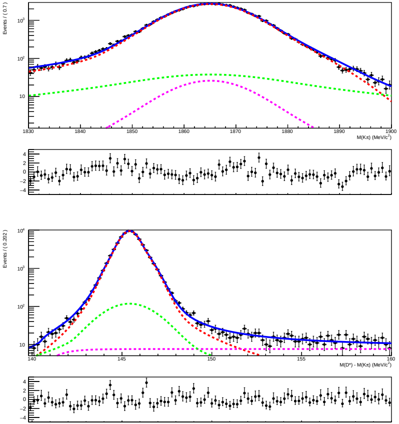
3. fake D^* & fake D^0 template: use bb_Dst3piX inclusive sample to subtract the fake D^* & true D^0 in item 2
4. fake D^* & true D^0 template: 1-3

Fake D^* and D^0 Events



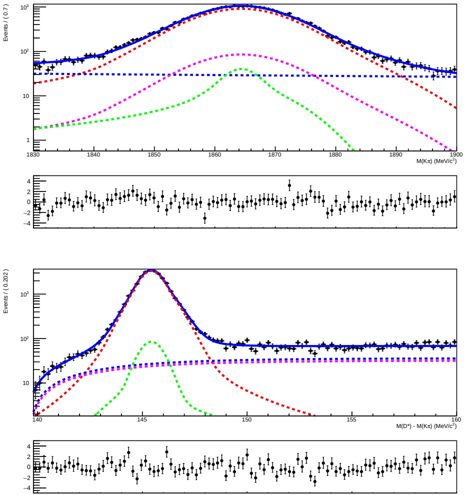
- We select events in the region $M(K\pi) \in (1840, 1890)$ MeV/c^2 and $M(D^*) - M(K\pi) \in (143, 148)$ MeV/c^2
- The sideband region is used to get the templates for fake D^* and D^0 events

Fake D^* and D^0 - Two dimensional fit in MC



- Signal, fake D^0, fake D^* but true D^0 and unmatched tracks
- Signal shape parameters except mean and sigma are fixed
- Unmatched tracks PDF parameters except sigma are fixed, same mean as signal
- The bkg 1 PDF for $M(K\pi)$ is exponential and for $M(D^*) - M(K\pi)$, it is RooDstD0Bkg PDF. They are fixed from 1D fits
- Bkg 2 is modeled by signal PDF in $M(K\pi)$ and bkg 1 PDF in $M(D^*) - M(K\pi)$.

Fake D^* and D^0 - Two dimensional fit in data

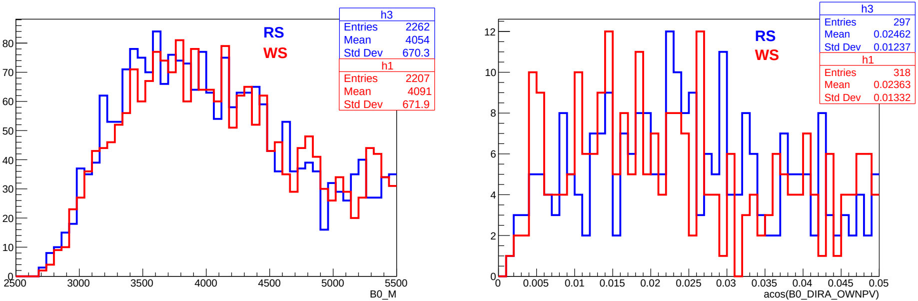


- Bkg shape parameters are floated.
- The fake D^0 is negligible in MC
- So the MC sideband is subtracted from data sideband to get rid of signal leakage while making fake D^0 template
- The fake D^0 events are subtracted from ΔM sideband to get fake D^* but true D^0 template

Fixed Yields in Combinatorial Events

- Combinatorial events' yields, denoted N_{B1B2}, are fixed based on expected sample values.
- Analysed using Wrong-Sign and Right-Sign data samples.
- Do not apply isolation and combinatorial suppression requirements.
- Characteristics of combinatorial B events:
  - Non-isolated
  - Flat distribution in $\cos(\mathrm{DIRA}(B0, PV))$
  - Large values in $m(D^{*}{-}3\pi\pm)$
- Yield of combinatorial B events assessed using the Right-Sign sample.

- These events are characterised by the following properties: non-isolated, flat in $\mathrm{acos}(\mathrm{DIRA}(B^0, \mathrm{PV}))$, large values of $m(D^{*-}3\pi^{\pm})$.



$m(D^{*-}3\pi^{\pm})$ (left) and $\mathrm{acos}(\mathrm{DIRA}(B^0, \mathrm{PV}))$ (right) distributions for combinatoric B events in the RS and WS data samples.

- The fraction of $B \to D^{**}\tau\nu$ events is overestimated in the simulated sample.
- To correct for this, a weight adjustment formula is used:

$$\mathcal{W} = \frac{R(D^{**}) \times \mathbf{Br}(B \to D^{**}\mu\nu)}{\mathbf{Br}(B^0 \to D^{**}\tau\nu)|_{\text{Sim.}} \times \mathbf{Br}(D^{**} \to D^{*})|_{\text{Sim.}}}$$

- These weights are calculated according to theoretical predictions.
- The feed-down fraction is obtained by summing over the D** decay modes, indexed by j:

$$f_{D^{**}\tau\nu} = \sum_{j=1}^{8} \mathcal{W}_j \times N_j/N_{tot} = 2.4\%$$

Where N_{tot} is the total D* yield and $W_j \times N_j$ is the weighted D** yield.

Fixing the Fraction of $B \to D^{**}\tau\nu$ Events

$B \to D^{**}\tau\nu$ events in $b\bar{b} \to D^{*}3\pi X$ inclusive mc sample and the corresponding weighted yields for different D^{**} states ($B \in \{B^0, B^+, B_s^0\}$).

Index j	State	$\mathcal{W}_j$	N_j	$\mathcal{W}_j \times N_j$
1	$D_1(2420)^0$	0.3497	1618	566
2	$D_2^*(2460)^0$	0.1667	736	123
3	$D_1(2430)^0$	0.1023	2366	242
4	$D_1(2420)^+$	0.3263	2321	757
5	$D_2^*(2460)^+$	0.1225	1164	143
6	$D_1(H)^+$	0.1174	3438	404
7	$D_{s1}(2536)^+$	0.1700	2525	429
8	$D_{s2}^*(2573)^+$	0.5	227	113
Total				2777

Recent BESIII results [PRD 107, 032002 (2023), arXiv:2212.13072]

$$\mathcal{B}(D_s^+ \to 3\pi^\pm X) = (32.81 \pm 0.35 \pm 0.63)\%$$

$$\mathcal{B}(D^+ \to 3\pi^\pm X) = (15.25 \pm 0.09 \pm 0.18)\%$$

$$\mathcal{B}(D^0 \to 3\pi^\pm X) = (17.60 \pm 0.11 \pm 0.22)\%$$

In our MC $\mathcal{B}(D_s^+ \to 3\pi^\pm X) = 28.653\%$ with various modes:

$$\mathcal{B}(D_s^+ \to K\pi h) = 18.894\% \qquad \mathcal{B}(D_s^+ \to \ell\nu) = 3.82\%$$

And separately, we produced:

$$\mathcal{B}(D_s^+ \to a_1\pi^0) = 2.27\%$$

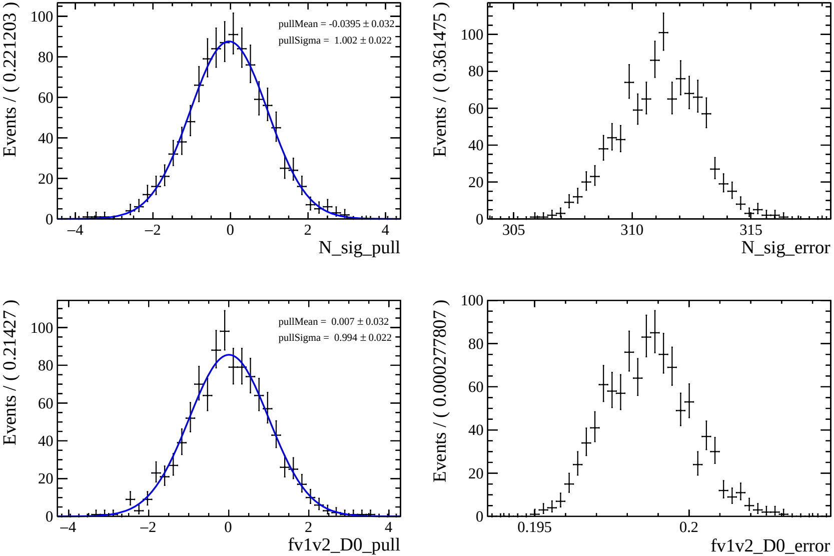
From free to constrained parameter

$$N_{D_s^+} = N(B^0 \to D^* D_s^+ (D_s^+ \to 3\pi^\pm X)) = \frac{\mathcal{B}(D_s^+ \to 3\pi^\pm X)}{\mathcal{B}(D_s^+ \to 3\pi^\pm)} \times N(B^0 \to D^* D_s^+ (D_s^+ \to 3\pi^\pm)) \times r_{\text{eff.}}$$

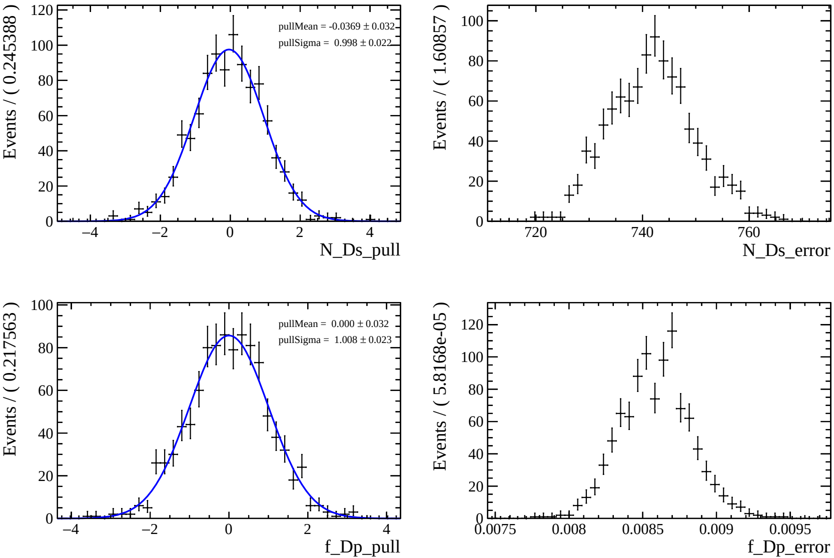
Blinding strategy

- Multiply signal yield by random factor drawn from Gaussian distribution $\mu = 1, \sigma = 0.3$
  - Multiplicative blinding preserves relative uncertainties $\rightarrow$ more useful for calculating correct relative systematic uncertainties and comparing with Run 1
- Multiply background yields by correction factor such that sum of yields $=$ number of datapoints
- When plotting, plot correct PDF and no components
  - Check fit quality
  - No 'visual unblinding'

Signal Fit – Pseudo-experiments

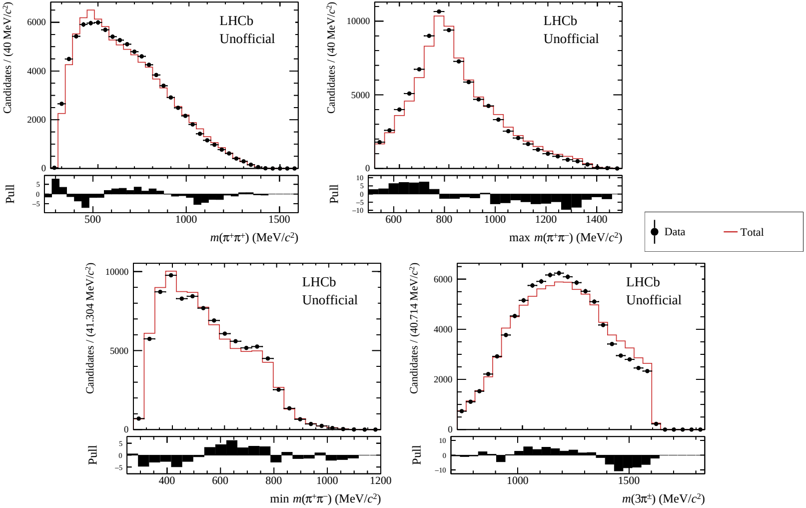


Signal Fit – Pseudo-experiments



Projections on the D_s^+ Variables

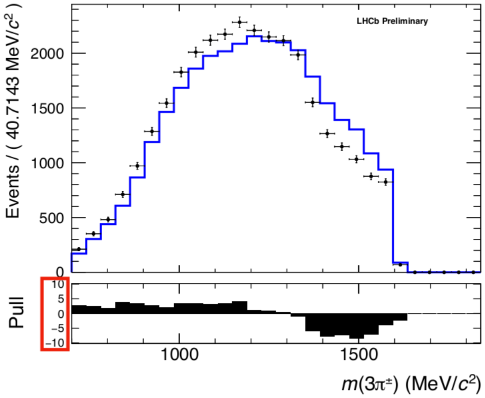
The signal fit projections on the D_s^+ control variables to the Run 2 data sample,



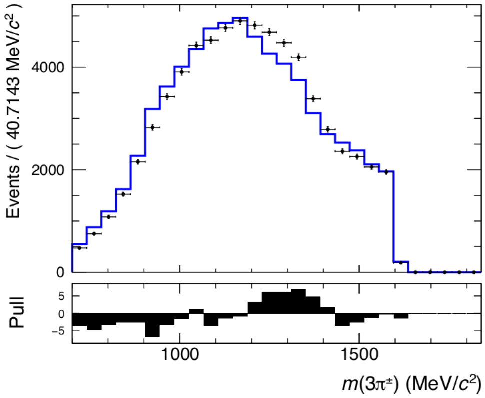
Projections on the D_s^+ Variables

Comparison between the $3\pi^{\pm}$ mass (blinded) projections from the signal fit.

Previous analysis (2015–2016)



Current analysis (2017–2018)



Signal Fit Results – Different Datasets

Parameters	Datasets: Run 2	2018	2017	2015–2016
Free Parameters				
N_{sig}	7815 ± 323	3177 ± 193	2557 ± 176	2201 ± 171
N_{D_s}	53148 ± 752	18622 ± 393	15790 ± 334	17902 ± 373
f_{D^+}	0.11 ± 0.01	0.1 ± 0.01	0.09 ± 0.01	0.16 ± 0.01
$f_{D^0}^{v_1-v_2}$	2.43 ± 0.2	3.09 ± 0.29	3.22 ± 0.31	1.5 ± 0.31
Constrained Parameters				
Constrained Value				
$N_{B\to D^{*-}3\pi^\pm X}$	6133 ± 400	2020 ± 191	1800 ± 192	2329 ± 303
f_{D_s}	0.546 ± 0.015	0.532 ± 0.025	0.554 ± 0.028	0.564 ± 0.028
$f_{D_{s0}^{*+}}$	0.052 ± 0.022	0.064 ± 0.032	0.063 ± 0.034	0.127 ± 0.037
$f_{D_{s1}^{\prime +}}$	0.489 ± 0.037	0.481 ± 0.049	0.416 ± 0.053	0.440 ± 0.055
$f_{D^{**}D_sX}$	0.220 ± 0.040	0.200 ± 0.052	0.242 ± 0.051	0.281 ± 0.059
$f_{B_s\to D^* D_s X}$	0.068 ± 0.017	0.078 ± 0.023	0.065 ± 0.024	0.080 ± 0.025
Fit Value				
$N_{B\to D^{*-}3\pi^\pm X}$	6940 ± 347	2467 ± 174	2184 ± 171	1616 ± 217
f_{D_s}	0.55 ± 0.01	0.52 ± 0.02	0.55 ± 0.02	0.59 ± 0.02
$f_{D_{s0}^{*+}}$	0.04 ± 0.02	0.07 ± 0.03	0.07 ± 0.03	0.09 ± 0.03
$f_{D_{s1}^{\prime +}}$	0.43 ± 0.02	0.44 ± 0.03	0.42 ± 0.03	0.38 ± 0.03
$f_{D^{**}D_sX}$	0.35 ± 0.03	0.28 ± 0.04	0.31 ± 0.04	0.38 ± 0.04
$f_{B_s\to D^* D_s X}$	0.08 ± 0.01	0.09 ± 0.02	0.05 ± 0.02	0.09 ± 0.02
Fixed Parameters				
$N_{B_1B_2}$	178 ± 13	80 ± 9	64 ± 8	34 ± 6
$N_{D^0}^{same}$	3233 ± 57	1231 ± 35	1016 ± 32	985 ± 31
$N_{fake D^0}$	802 ± 25	384 ± 13	301 ± 12	304 ± 12
$N_{fake D^*}$	1862 ± 49	762 ± 27	674 ± 25	627 ± 25
$f_{D^{**}\tau\nu}$	0.024 ± 0.00	0.024 ± 0.00	0.023 ± 0.00	0.025 ± 0.00
$f_{\tau^+\to 3\pi^\pm\overline{\nu}_\tau}$	0.768 ± 0.002	0.770 ± 0.003	0.768 ± 0.002	0.767 ± 0.004

	Datasets: Run 2	2018	2017	2015–2016
$\sigma_{N_{sig}}/N_{sig}$	4.1%	6.1%	6.9%	7.8%
$\chi^2/374$	1.46	1.08	1.43	1.05
Integrated Luminosity (fb^{-1})	5.8	2.9	1.7	1.9

Signal Fit Results – Fixing D^0 and D_s^+ Fractions

Parameter	Value
Free Parameters	
N_{sig}	7909 ± 280
N_{D_s}	50446 ± 530
f_{D^+}	0.12 ± 0.01
Constrained Parameters	
$N_{B\to D^{*-}3\pi^\pm X}$	6778 ± 340
$f_{B_s\to D^* D_s X}$	0.07 ± 0.01
$f_{D^{**} D_s X}$	0.35 ± 0.02
$f_{D'^+_{s1}}$	0.44 ± 0.02
$f_{D^{*+}_{s0}}$	0.05 ± 0.02
Fixed Parameters	
$N_{B_1-B_2}$	178 ± 13
$N^{\text{same}}_{D^0}$	3233 ± 57
$N_{\text{fake}\,D^0}$	802 ± 25
$N_{\text{fake}\,D^*}$	1862 ± 49
$f_{D^{**}\tau\nu}$	0.024 ± 0.00
$f_{\tau^+\to 3\pi^\pm\overline{\nu}_\tau}$	0.768 ± 0.002
$f^{v_1-v_2}_{D^0}$	2.43 ± 0.20
f_{D_s}	0.55 ± 0.01

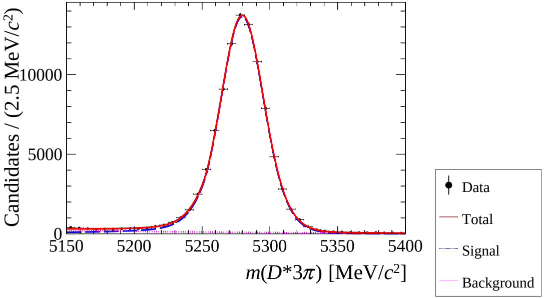
Outline

Normalization Fit – Strategy

The normalization fit is performed using an unbinned maximum likelihood fit to the $D^*3\pi$ mass.

- **Total B^0 Yield** (N_{tot}):
  - Fit to B^0 mass
- **B^0 Yield in Δm Sidebands**:
  - Fit B^0 yield in 2 MeV bins from 138 to 160 MeV.
  - Yield in each sideband region.
- **Extrapolation to the signal region**:
  - Estimate fake D^{*-} yield (N_{fake}).
- **Total D_s^+ Yield** ($N_{D_s^+}$):
  - Fit $3\pi^\pm$ mass
  - Use same fake D^{*-} fraction to estimate true D_s^+ yield.
- **Normalization Yield** (N_{norm}):
  - Subtract true D_s^+ yield: $N_{\mathrm{norm}} = N_{\mathrm{tot}} - N_{\mathrm{fake}} - N_{D_s^+}$

Normalisation Fit – B^0 Yield



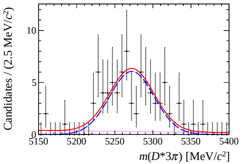
Fit to B^0 mass:

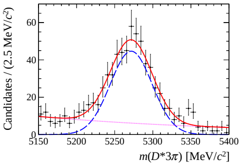
- Signal: Crystal Ball + 2 Gaussians
- Background: Exponential

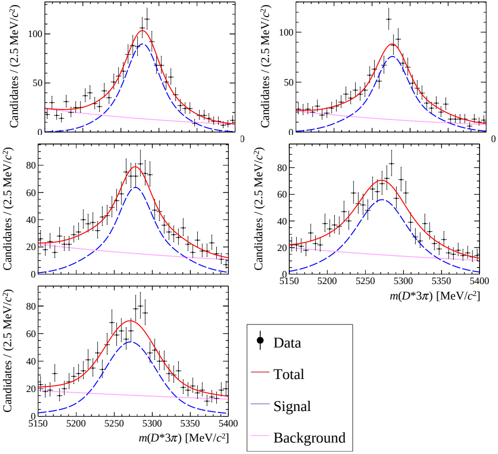
Total B^0 yield: $N_{\mathtt{tot}} = 103590 \pm 353$

Extrapolation method to estimate fake D^{*-} yield from data in Δm sidebands.

- Determine fake D^{*-} yields in Δm sidebands:
  - $[138, 140], [140, 142]$ MeV
  - $[150, 152], \ldots, [158, 160]$ MeV
  - Fit to B^0 mass



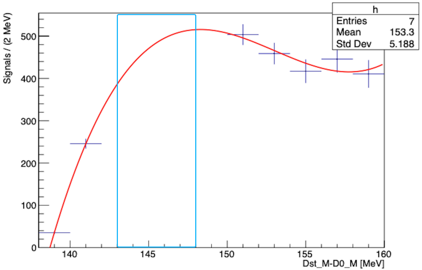




Normalization Fit – Fake D^{*-}

Extrapolation method to estimate fake D^{*-} yield from data in Δm sidebands.

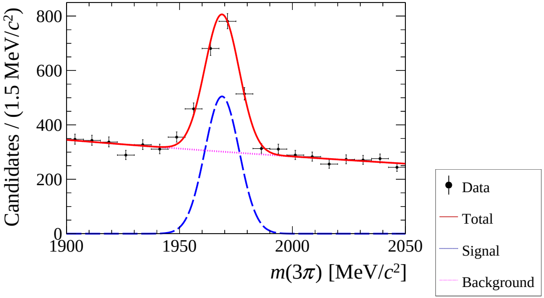
- Determine fake D^{*-} yields in Δm sidebands:
  - $[138, 140], [140, 142]$ MeV
  - $[150, 152], \ldots, [158, 160]$ MeV
  - Fit to B^0 mass



Fake D^{*-} yield: $N_{\mathtt{fake}} = 9057 \pm 315$

- Extrapolate to the signal region ($\Delta m \in [144, 148]$ MeV) to estimate the fake D^{*-} yield

Normalisation Fit – $B^0 \to D^{*-} D_s^+$ Subtraction



Fit to $m(3\pi)$ to obtain the $B^0 \to D^{*-} D_s^+$ yield

- Signal: Crystal Ball + 2 Gaussians
- Background: Exponential
- Use same fake D^{*-} fraction to estimate true D_s^+ yield.

$$N_{D_s^+} = 1360 \pm 60$$

Normalization Yield Extraction (Old Way)

The previous strategy to extract the normalization yield was as follows:

1. Perform a 2D fit on m_{D^0} and $\Delta m = \Delta m$
   $\rightarrow$ sWeights for **true** D^{*-} and **true** D^0 contributions.
2. Perform a B mass fit using the sWeights
   $\rightarrow$ determine the normalization yield.

Fake D^{*-} can also contribute to the B^0 peak due to the narrow Δm range.

- This occurs in $B \rightarrow D^* 3\pi$ decays where $D^{*0/-} \rightarrow D^0 \pi^{0/-}$, and the $\pi^{0/-}$ is replaced by a slow pion during reconstruction.

Issues on the (Old) Normalization Fit

- sWeights are obtained in a **wide** m_{D^0} and Δm range:
  - m_{D^0}: $(1825, 1905)$ MeV
  - Δm: $(138, 160)$ MeV
- In B^0 mass fit, **tighter cuts** on m_{D^0} and Δm are applied:
  - m_{D^0}: $(1840, 1890)$ MeV
  - Δm: $(143, 148)$ MeV

This results in some events with negative sWeights being excluded $\Longrightarrow$ improper background subtraction.

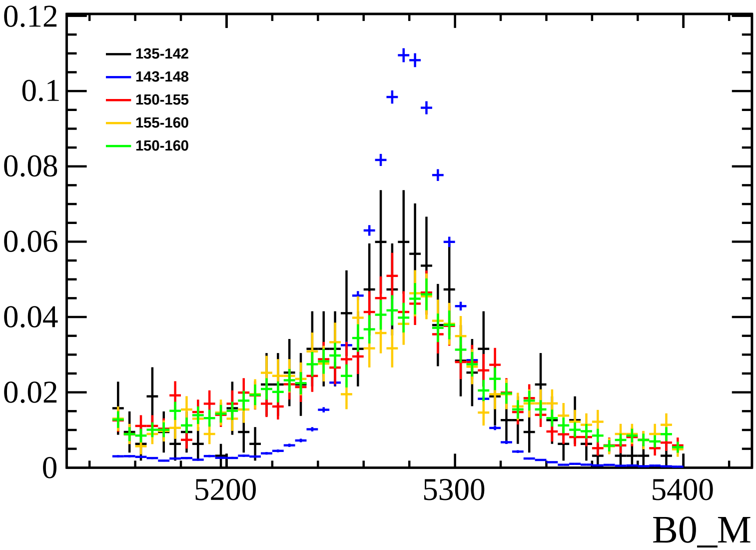
Illustration: As a consequence, the normalization yield is overestimated:

- $N_1 = 30990 \pm 184$ (fit with sWeight applied)
- $N_2 = 30282 \pm 183$ (fit without sWeight applied, **including** fake D^{*-} contribution)
- N_1 should be smaller than N_2

($B^0 \to D^{*-} D_s^+$ not subtracted in the above results)

Issues on the (Old) Normalization Fit

Correlation between discriminating variable (Δm) and fit variable (B^0 mass):



Issues on the (Old) Normalization Fit

Minor Issue:

- Fake D^{*-} in MC samples from replacement of slow pion by another pion in the $B^0 \to D^{*-}3\pi^{\pm}$, $D^{*-} \to \bar{D}^0\pi^-$ decay are not accounted for in efficiency computation.
- Overestimation of efficiency.
- Expected effect on $R(D^*)$: no more than 1% (estimated from MC).

$R(D^*)$ hadronic ($\tau \to 3\pi\nu_\tau$) Systematics

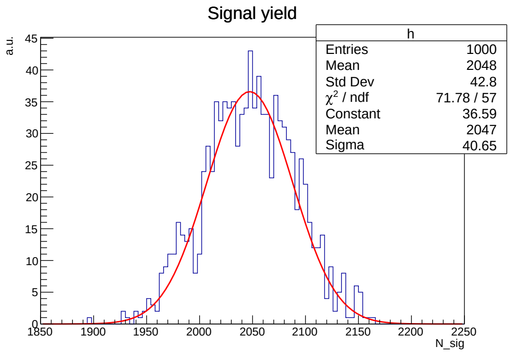
Breakdown of relative uncertainties:

Source	$R(D)/R(D^*)$ $\frac{\delta R(D^{*-})}{R(D^*)}$ [%]	Future
Simulated sample size	4.7	Produce more MC !
Empty bins in templates	1.3	
Signal decay model	1.8	
$D^{**}\tau\nu$ and $D_s^{**}\tau\nu$ feed-downs	2.7	Measure $R(D_1(2420)^0)$
$D_s^+ \to 3\pi X$ decay model	2.5	BESIII
$B \to D^{*-}D_s^+X$, $D^{*-}D^+X$, $D^{*-}D^0X$ bkgs	3.9	Improves with stat
Combinatorial background	0.7	
$B \to D^{*-}3\pi X$ background	2.8	Kill with $\vert z\tau - zD\vert > 5\sigma$
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	

[PRL 120, 171802 2018]

Systematic uncertainties

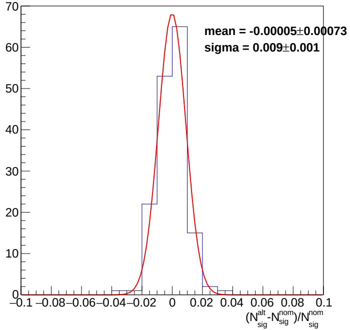
- The systematic uncertainty due to the MC statistics is estimated by from 1000 template bootstraps
- The contribution is 1.99%



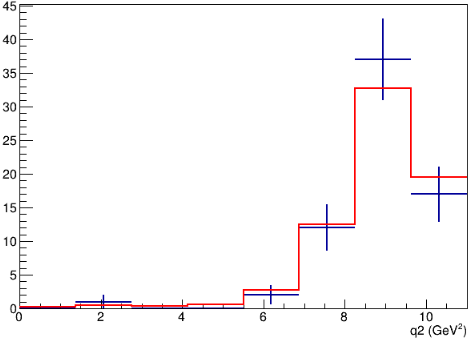
Systematic uncertainties

- Effects of different fixed parameters in the signal fit are estimated by $\pm 1\sigma$ variation
- Systematic deviation due to D^{**} feed-down is estimated by varying the fixed value by $\pm 50\%$; the contribution is $^{+1.80}_{-1.85}\%$
- The fraction of a_1 component in the D_s^+ decay model fit is varied by $\pm 30\%$ and the relative difference is $^{+1.46}_{-1.51}\%$

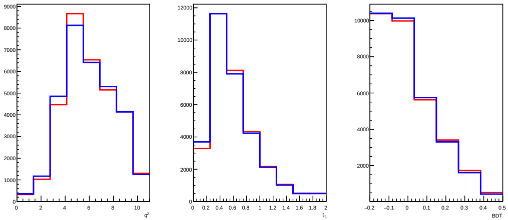
- The variation due to D_s^+ weights is estimated from fits using alternate templates for D_s^+ modes in the final signal fit, with Gaussian variation of the D_s^+ weights of each sub-component
- The resultant variation is 1%.



Empty Bins in Templates



q^2 projection of the nominal (blue) and KDE (red) templates for D^+ background in low BDT and t_τ regions.



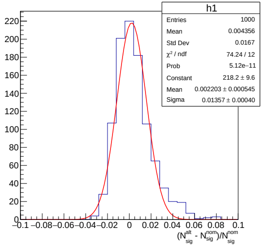
Fit projections from signal fit using of the nominal (blue) and KDE (red) templates.

Systematic uncertainties – Template Shape Variation

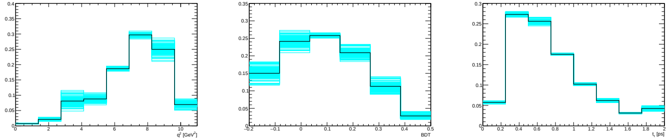
- The template shape of different background components contribute significantly to the systematic uncertainties
- The template shapes of D^+ and D^0 backgrounds are varied by weights estimated in relation to the kinematic variables $m(D^*3\pi)$, $m(3\pi^{\pm})$, $m(\pi^+\pi^-)_{\text{min}}$ and $m(\pi^+\pi^-)_{\text{max}}$
- Template shapes of D_s backgrounds are varied with weights related to $m(D^*3\pi)$
- Two limiting templates for the variation are obtained and a set of 1000 alternate templates are made with a linear interpolation between the two limits, for each of the background categories
- The interpolation parameter is drawn from a Gaussian random distribution

Systematic uncertainties

- The relative difference between the nominal signal yield and each of the 1000 alternate fits for D^+ background
- Its width, 1.4%, taken as systematic uncertainty

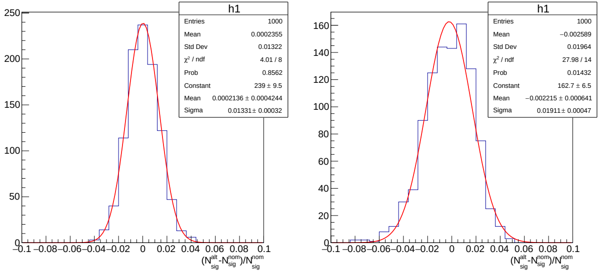


- The distribution of 1000 alternate templates (in cyan) along with the nominal one



Systematic uncertainties

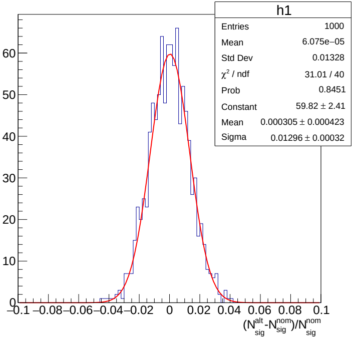
- Similarly, alternate templates are built for D^0 and D_s backgrounds
- For D_s backgrounds, only the variation w.r.t $m(D^* 3\pi)$ is considered; the effect of others included in D_s weight variation study
- The relative difference between the nominal signal yield and each of the alternate fits is 1.3% and 1.9% for D^0 (2 templates) and D_s (6 templates) backgrounds, respectively



Systematic uncertainties

- The systematic contribution from combinatorial background is estimated by changing the fixed normalisation by $\pm 1\sigma$ and noting the relative change in the signal yield
  - The resultant deviation is $^{+0.53}_{-0.63}\%$
  - The templates are built from data

- The prompt template variation w.r.t the kinematic variables $m(D^{*}3\pi)$, $m(3\pi^{\pm})$, $m(\pi^{+}\pi^{-})_{\text{min}}$ and $m(\pi^{+}\pi^{-})_{\text{max}}$ is studied in a similar manner as in the case of double-charm bkgs; the systematic deviation is 1.3%



Systematic Uncertainties

- Template shapes are varied with weights

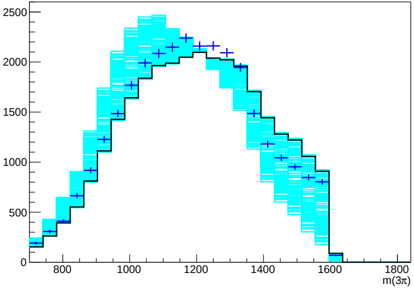
$$\omega_{m(D^*3\pi)} = 1 + 2\alpha_{m(D^*3\pi)} \left(\frac{m(D^*3\pi) - 2700}{5100 - 2700} - \frac{1}{2} \right),$$

$$\omega_{m(3\pi)} = 1 + 2\alpha_{m(3\pi)} \left(\frac{m(3\pi) - 3m_\pi}{1600 - 3m_\pi} - \frac{1}{2} \right),$$

$$\omega_{\min[m(\pi^+\pi^-)]} = 1 + 2\alpha_{\min[m(\pi^+\pi^-)]} \left(\frac{min[m(\pi^+\pi^-)] - 2m_\pi}{880 - 2m_\pi} - \frac{1}{2} \right),$$

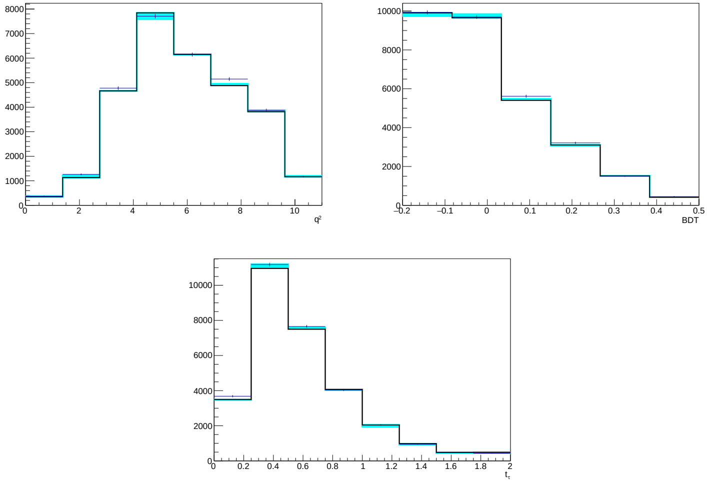
$$\omega_{\max[m(\pi^+\pi^-)]} = 1 + 2\alpha_{\max[m(\pi^+\pi^-)]} \left(\frac{max[m(\pi^+\pi^-)] - 2m_\pi}{1440 - 2m_\pi} - \frac{1}{2} \right);$$

Systematic Uncertainties – D_s^+ template variation



Total PDF variations (in cyan) along with the nominal one (in black) and data (blue) for $m(3\pi)$. The variations are applied on D_s^+ templates.

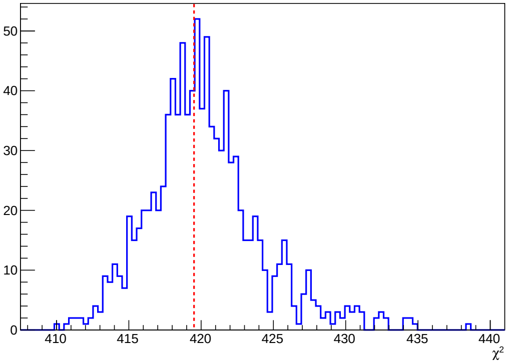
Systematic Uncertainties – D_s^+ template variation



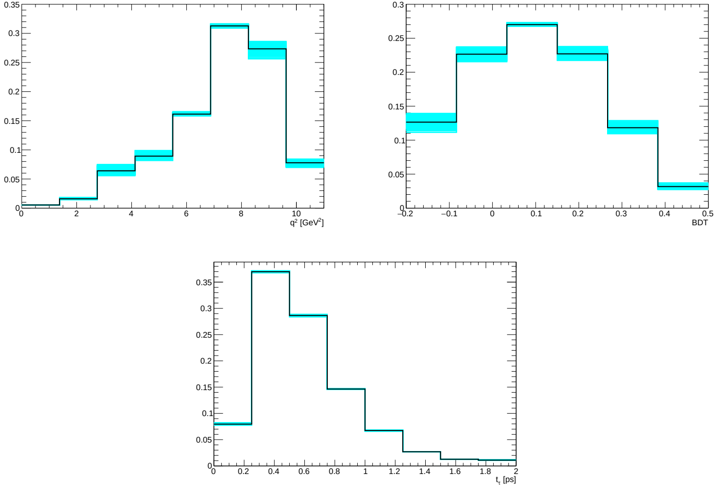
Total PDF variations (in cyan) along with the nominal one (black) and data (blue) for the fit variables q^2, BDT and t_τ. The variations are applied on D_s^+ templates.

D^{+} template variation

- Distribution of χ^2 of the 1000 fits; nominal χ^2 is marked in red line

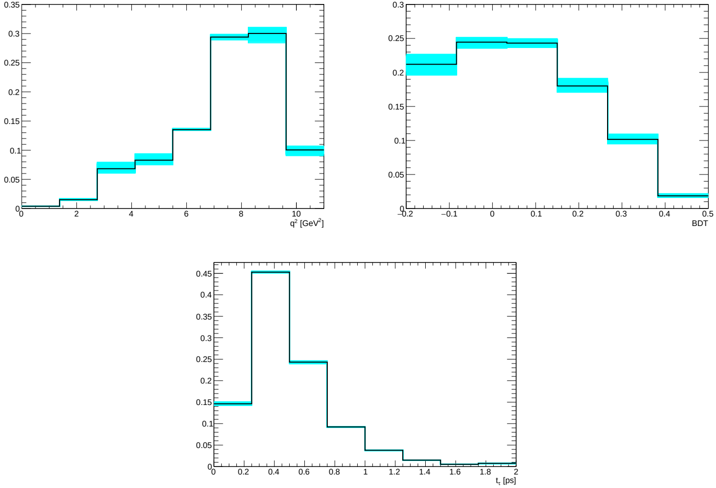


D^0 same vertex template variation



Nominal Variations

D^0 different vertex template variation

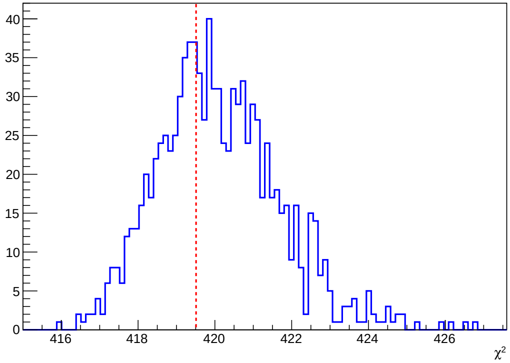


Nominal

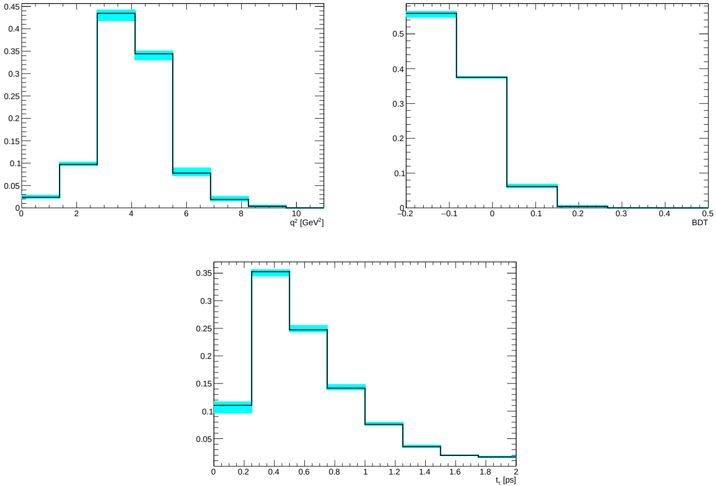
Variations

D^0 template variation

- Distribution of χ^2 of the 1000 fits; nominal χ^2 is marked in red line



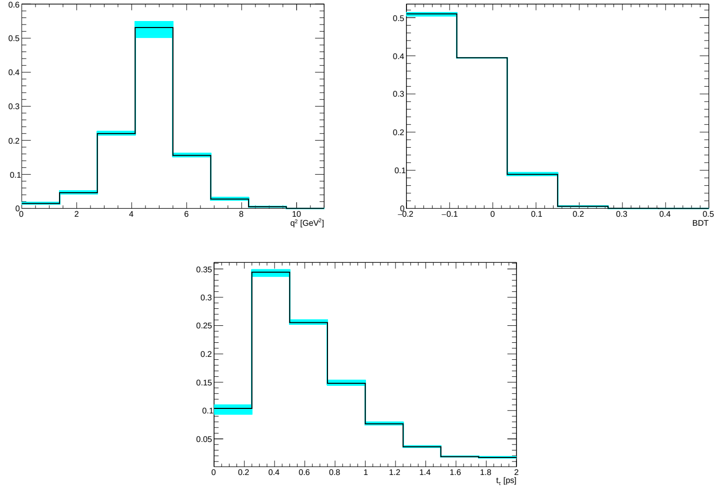
D_s^+ template variation



Nominal

Variations

D_s^* template variation



Nominal Variations