Measurement of the Ratio of Branching Fractions $B (\bar{B}^0 \to D^{*+} \to \bar{B} (\bar{B}^0 \to D^{*+} \to \bar{B}]) / B (\bar{B}^0 \to D^{*+} \to \bar{B})$

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Measurement of the Ratio of Branching Fractions $R(D^+) = \frac{B(\bar{B}^0 \to D^{+}\tau^-\bar{\nu}_\tau)}{B(\bar{B}^0 \to D^{+}\mu^-\bar{\nu}_\mu)}$

R. Aaij et al. (LHCb Collaboration)

(Received 30 June 2015; published 9 September 2015; corrected 14 September 2015)

The branching fraction ratio $R(D^+) = \frac{B(\bar{B}^0 \to D^{+}\tau^-\bar{\nu}_\tau)}{B(\bar{B}^0 \to D^{+}\mu^-\bar{\nu}_\mu)}$ is measured using a sample of proton-proton collision data corresponding to 3.0 fb$^{-1}$ of integrated luminosity recorded by the LHCb experiment during 2011 and 2012. The tau lepton is identified in the decay mode $\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$. The semitauonic decay is sensitive to contributions from non-standard-model particles that preferentially couple to the third generation of fermions, in particular, Higgs-like charged scalars. A multidimensional fit to kinematic distributions of the candidate $\bar{B}^0$ decays gives $R(D^+) = 0.336 \pm 0.027$ (stat) $\pm 0.030$ (syst). This result, which is the first measurement of this quantity at a hadron collider, is 2.1 standard deviations larger than the value expected from lepton universality in the standard model.

Lepton universality, enshrined within the standard model (SM), requires equality of couplings between the gauge bosons and the three families of leptons. Hints of lepton nonuniversality in $B^+ \to K^+e^+\mu^-$ and $B^+ \to K^+\mu^+\mu^-$ decays [1] have been seen, but no definitive observation of a deviation has yet been made. However, a large class of models that extend the SM contain additional interactions involving enhanced couplings to the third generation that would violate this principle. Semileptonic decays of $b$ hadrons (particles containing a $b$ quark) to third generation leptons provide a sensitive probe for such effects. In particular, the presence of additional charged Higgs bosons, which are often required in these models, can have a significant effect on the rate of the semitauonic decay $\bar{B}^0 \to D^{+}\tau^-\bar{\nu}_\tau$, [2]. The use of charge-conjugate modes is implied throughout this Letter.

Semitauonic $B$ meson decays have been observed by the BABAR and Belle collaborations [3–7]. Recently BABAR reported updated measurements [6,7] of the ratios of branching fractions, $R(D^+) = B(\bar{B}^0 \to D^{+}\tau^-\bar{\nu}_\tau)$/$B(\bar{B}^0 \to D^{+}\mu^-\bar{\nu}_\mu)$ and $R(D) = B(\bar{B}^0 \to D^{+}\tau^-\bar{\nu}_\tau)$/$B(\bar{B}^0 \to D^{+}\mu^-\bar{\nu}_\mu)$, which show deviations of $2.7 \sigma$ and $2.0 \sigma$, respectively, from the SM predictions [8,9]. These ratios have been calculated to high precision, owing to the cancellation of most of the uncertainties associated with the strong interaction in the $B$ to $D^{(*)}$ transition. Within the SM they differ from unity mainly because of phase-space effects due to the differing charged lepton masses.

This Letter presents the first measurement of $R(D^+)$ in hadron collisions using the data recorded by the LHCb detector at the Large Hadron Collider in 2011–2012. The data correspond to integrated luminosities of 1.0 fb$^{-1}$ and 2.0 fb$^{-1}$, collected at proton-proton ($pp$) center-of-mass energies of 7 TeV and 8 TeV, respectively. The $\bar{B}^0 \to D^{+}\tau^-\bar{\nu}_\tau$ decay with $\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$ (the signal channel) and the $\bar{B}^0 \to D^{+}\mu^-\bar{\nu}_\mu$ decay (the normalization channel) produce identical visible final-state topologies; consequently both are selected by a common reconstruction procedure. The selection identifies semileptonic $\bar{B}^0$ decay candidates containing a muon candidate and a $D^{*+}$ candidate reconstructed through the decay chain $D^{*+} \to D^0 (\to K^-\pi^+)\pi^+$. The selected sample contains contributions from the signal and the normalization channel, as well as several background processes, which include partially reconstructed $B$ decays and candidates from combinations of unrelated particles from different $b$ hadron decays. The kinematic and topological properties of the various components are exploited to suppress the background contributions. Finally, the signal, the normalization component and the residual background are statistically disentangled with a multidimensional fit to the data using template distributions derived from control samples or from simulation validated against data.

The LHCb detector [10,11] is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing $b$ or $c$ quarks. The detector includes a high-precision tracking system consisting of a silicon-strip vertex detector surrounding the $pp$ interaction region [12], a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about 4Tm, and three stations of silicon-strip detectors and straw drift tubes [13] placed downstream of the magnet. The tracking system provides a measurement of momentum, $p_T$, of charged particles with a relative uncertainty that varies from 0.5% at low momentum to 1.0% at 200 GeV/$c$. The minimum distance of a track to a primary vertex (PV), the impact parameter, is measured with a resolution of $(15 + 29/p_T)$ $\mu$m, where $p_T$ is the
component of the momentum transverse to the beam, in GeV/c. Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov detectors [14]. Photons, electrons, and hadrons are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic calorimeter, and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers [15]. The online event selection is performed by a trigger [16], which consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which applies a full event reconstruction.

A simulation of pp collisions is provided by Pythia [17,18] with a specific LHCb configuration [19]. Decays of hadronic particles are described by EvtGen [20], in which final-state radiation is generated using Photos [21]. The interaction of the generated particles with the detector, and its response, are implemented using the Geant4 toolkit [22,23] as described in Ref. [24].

The trigger requirements are chosen to avoid the imposition of any pT selection on the muon, or invariant mass requirements on the $D^+\mu^-$ system, crucial for preserving the distinct kinematic distributions of the $B^0 \rightarrow D^{*+}\tau^- (\rightarrow \mu^- \nu_{\mu} \bar{\nu}_{\mu})$ decay. Events are required to pass the hardware trigger either because the decay products of the $D^+$ candidate satisfy the hadron trigger requirements or because high-$p_T$ particles in the event, independent of the $D^+\mu^-$, satisfy one of the hardware trigger requirements. In the software trigger, the events are required to meet criteria designed to accept $D^0 \rightarrow K^-\pi^+$ candidates with $p_T > 2$ GeV/c. Quality requirements are applied to the tracks of the charged particles that originate from a candidate $D^0$ decay: their momenta must exceed 5 GeV/c and at least one must have $p_T > 1.5$ GeV/c. The momentum vector of the $D^0$ candidate must point back to one of the PVs in the event and the reconstructed mass must be consistent with the known $D^0$ mass [25].

In the offline reconstruction, the $D^0$ candidates satisfying the trigger are further required to have well-identified $K^-$ and $\pi^+$ daughters, and the decay vertex is required to be significantly separated from any PV. The invariant mass of the $D^0$ candidate is required to be within 23.5 MeV/c² of the peak value, corresponding to approximately three times the $D^0$ mass resolution. These candidates are combined with low-energy pions to form candidate $D^{*+} \rightarrow D^0\pi^+$ decays, which are subjected to a kinematic and vertex fit to the decay chain. Candidates are then required to have a mass difference $\Delta m \equiv m(D^0\pi^+) - m(D^0)$ within 2 MeV/c² of the known value, corresponding to approximately 2.5 times the observed resolution. The muon candidate is required to be consistent with a muon signature in the detector, to have momentum $3 < p < 100$ GeV/c, to be significantly separated from the primary vertex, and to form a good vertex with the $D^0$ candidate. The $D^{*+}\mu^-$ combinations are required to have an invariant mass less than 5280 MeV/c² and their momentum vector must point approximately to one of the reconstructed PV locations, which removes combinatoric candidates while preserving a large fraction of semileptonic decays. In addition to the signal candidates, two independent samples of “wrong sign” candidates, $D^{*+}\mu^+$ and $D^0\pi^-\mu^-$, are formed for estimating the combinatorial background. The former represents random combinations of $D^{*+}$ candidates with muons from unrelated decays, and the latter is used to model the contribution of misreconstructed $D^{*+}$ decays. Mass regions $5280 < m(D^{*+}\mu^-) < 10000$ MeV/c² and $139 < \Delta m < 160$ MeV/c² are included in all samples for study of the combinatorial backgrounds. Finally, a sample of candidates is selected where the track paired with the $D^{*+}$ fails all muon identification requirements. These $D^{*+}h^\pm$ candidates are used to model the background from hadrons misidentified as muons.

To suppress the contributions of partially reconstructed $B$ decays, including $B$ decays to pairs of charged hadrons, and semileptonic $\bar{B} \rightarrow D^{*+}(n\pi)\mu^-\bar{\nu}_{\mu}$ decays with $n \geq 1$ additional pions, the $D^{*+}\mu^-$ candidates are required to be isolated from additional tracks in the event. An algorithm is developed and trained to determine whether a given track is likely to have originated from the signal $B$ candidate or from the rest of the event based on a multivariate analysis (MVA) method. For each track in the event, the algorithm employs information on the track separation from the PV, the track separation from the decay vertex, the angle between the track and the candidate momentum vector, the decay length significance of the decay vertex under the hypothesis that the track does not originate from the candidate and the change in this significance under the hypothesis that it does. A signal sample, enriched in $B^0 \rightarrow D^{*+}\tau^-\bar{\nu}_{\tau}$ and $\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_{\mu}$ decays, is constructed by requiring that no tracks in the event reach a threshold in the MVA output. In addition, the output is used to select three control samples enriched in partially reconstructed $B$ decays of interest for background studies by requiring that only one or two tracks be selected by the MVA ($D^{*+}\mu^-\pi^-$ or $D^{*+}\mu^-\pi^+\pi^-$) or that at least one track selected by the MVA passes $K^\pm$ identification requirements ($D^{*+}\mu^-K^\pm$). These samples are depleted of $B^0 \rightarrow D^{*+}\mu^-\bar{\nu}_{\mu}$ and $\bar{B}^0 \rightarrow D^{*+}\tau^-\bar{\nu}_{\tau}$ decays and are used to study and constrain the shapes of remaining backgrounds in the signal sample.

The efficiencies $\varepsilon_s$ and $\varepsilon_n$ for the signal and the normalization channels, respectively, are determined in simulation. These include the effects of the trigger, event reconstruction, event selection, particle identification procedure, isolation method, and the detector acceptance. To account for the effect of differing detector occupancy distributions between simulation and data, the simulated samples are reweighted to match the occupancy observed in data. The overall efficiency ratio is $\varepsilon_s/\varepsilon_n = (77.6 \pm 1.4)\%$, with the deviation from unity primarily due to the particle
identification, which dominantly removes low-$p_T$ muon candidates, and vertex quality requirements.

The separation of the signal from the normalization channel, as well as from background processes, is achieved by exploiting the distinct kinematic distributions that characterize the various decay modes, resulting from the $\mu - \tau$ mass difference and the presence of extra neutrinos from the decay $\tau \rightarrow \mu \bar{\nu}_\mu \nu_\tau$. The most discriminating kinematic variables are the following quantities, computed in the $B$ rest frame: the muon energy, $E_\mu^\tau$; the missing mass squared, defined as $m^2_{\text{miss}} = (p_B^2 - p_D^2 - p_\mu^2)$; and the squared four-momentum transfer to the lepton system, $q^2 = (p^\mu_B - p^\mu_D)^2$, where $p^\mu_B$, $p^\mu_D$ and $p^\mu_\mu$ are the four-momenta of the $B$ meson, the $D^{*+}$ meson and the muon. The determination of the rest-frame variables requires knowledge of the $B$ candidate momentum vector in the laboratory frame, which is estimated from the measured parameters of the reconstructed final-state particles. The $B$ momentum direction is determined from the unit vector to the $B$ decay vertex from the associated PV. The component of the $B$ momentum along the beam axis is approximated using the relation $(p_B z) = (m_B/m_\text{rec})(p_\text{rec}) z$, where $m_B$ is the known $B$ mass, and $m_\text{rec}$ and $p_\text{rec}$ are the mass and momentum of the system of reconstructed particles. The rest-frame variables described above are then calculated using the resulting estimated $B$ four-momentum and the measured four-momenta of the $\mu$ and $D^{*+}$. The rest-frame variables are shown in simulation studies to have sufficient resolution ($\approx 15\% - 20\%$ full width at half maximum) to preserve the discriminating features of the original distributions.

Simulated events are used to derive kinematic distributions from signal and $B$ backgrounds that are used to fit the data. The hadronic transition-matrix elements for $B^0 \rightarrow D^{*+} \tau^+ \bar{\nu}_\tau$ and $B^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu$ decays are described using form factors derived from heavy quark effective theory [26]. Recent world averages for the corresponding parameters are taken from Ref. [27]. These values, along with their correlations and uncertainties, are included as external constraints on the respective fit parameters. The hadronic matrix elements describing $B^0 \rightarrow D^{*+} \tau^+ \bar{\nu}_\tau$ decays include a helicity-suppressed component, which is negligible in $B^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu$ decays [28]. This parameter is not well constrained by data; hence, the central value and uncertainty from the sum rule presented in Ref. [8] are used as a constraint. It is assumed that the kinematic properties of the $B^0 \rightarrow D^{*+} \tau^+ \bar{\nu}_\tau$ decay are not modified by any SM extensions.

For the background semileptonic decays $B \rightarrow [D_1(2420), D_2^0(2460), D'_1(2430)] \mu^- \bar{\nu}_\mu$ [collectively referred to as $B \rightarrow D^{*+} \rightarrow D^{*+} \mu^- \bar{\nu}_\mu$], form factors are taken from Ref. [29]. The slope of the Isgur-Wise function [30,31] is included as a free parameter in the fit, with a constraint derived from fitting the $D^{*+} \mu^- \bar{\nu}_\mu$ control sample. This fit also serves to validate this choice of model for this background. Contributions from $B^0 \rightarrow [D'_1(2536), D'_2^0(2573)] \mu^- \bar{\nu}_\mu$ decays use a similar parametrization, keeping only the lowest-order terms. Semileptonic decays to heavier charged hadrons decaying as $D^{*+} \rightarrow D^{*+} \pi^\pm \pi^\mp$ and semitauonic decays $B \rightarrow [D_1(2420), D_2^0(2460), D'_1(2430)] \tau^- \bar{\nu}_\tau$ are modeled using the ISGW2 [32] parametrization. To improve the modeling for the former, a fit is performed to the $D^{*+} \mu^- \pi^+ \pi^- c$ control sample to generate an empirical correction to the $q^2$ distribution, as the resonances that contribute to this final state and their respective form factors are not known. The contribution of semimuonic decays to excited charm states amounts to approximately $12\%$ of the normalization mode in the fit to the signal sample.

An important background source is $B$ decays into final states containing two charged hadrons, $B \rightarrow D^{*+} H_s X$, followed by semileptonic decay of the charged hadron $H_s \rightarrow \mu^- \nu X$. This process occurs at a total rate of $6\% - 8\%$ relative to the normalization mode. The template for this process is generated using a simulated event sample of $B^+$ and $B^0$ decays, with an appropriate admixture of final states. Corrections to the simulated template are obtained by fitting the $D^{*+} \mu^- K^\mp$ control sample. A similar simulated sample is also used to generate kinematic distributions for final states containing a tertiary muon from $B \rightarrow D^{*+} D^- X$, with $D^- \rightarrow \tau^- \bar{\nu}_\tau$ and $\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$.

The kinematic distributions of hadrons misidentified as muons are derived based on the sample of $D^{*+} h^\pm$ candidates. Control samples of $D^{*+} (\Lambda)$ decays are used to determine the probabilities for a $\pi$ or $K (p)$ to be misidentified as a muon, and to generate a $3 \times 3$ matrix of probabilities for each species to satisfy the criteria for identification as a $\pi, K$, or $p$. These are used to determine the composition of the $D^{*+} h^\pm$ sample in order to model the background from hadrons misidentified as muons. Two methods are developed to handle the unfolding of the individual contributions of $\pi$, $K$, and $p$, which result in different values for $R(D'^+)$. The average of the two methods is taken as the nominal central value, and half the difference is assigned as a systematic uncertainty.

Combinatorial backgrounds are classified based on whether or not a genuine $D^{*+} \rightarrow D^0 \pi^+$ decay is present. Wrong-sign $D^0 \pi^- \mu^- \bar{\nu}_\mu$ combinations are used to determine the component with misreconstructed or false $D^{*+}$ candidates. The size of this contribution is constrained by fitting the $\Delta m$ distribution of $D^{*+} \mu^- \bar{\nu}_\mu$ candidates in the full $\Delta m$ region. The contribution from correctly reconstructed $D^{*+}$ candidates combined with $\mu^-$ from unrelated $b$ hadron decays is determined from wrong-sign $D^{*+} \mu^- \pi^-$ combinations. The size of this contribution is constrained by use of the mass region $5280 < m(D^+ \mu^-) < 10000$ MeV$/c^2$, which determines the expected ratio of $D^{*+} \mu^-$ to $D^{*+} \mu^+ \pi^-$ yields. In both cases, the contributions of misidentified muons are subtracted when generating the kinematic distributions for the fit.
The binned $m_{\text{miss}}^{2}$, $E_{\mu}^{*}$, and $q^{2}$ distributions in data are fit using a maximum likelihood method with three dimensional templates representing the signal, the normalization, and the background sources. To avoid bias, the procedure is developed and finalized without knowledge of the resulting value of $R(B^{+} \to D^{+} \pi^{+} \mu \nu)$. The templates extend over the kinematic region $-2 < m_{\text{miss}}^{2} < 10$ GeV$^2$/c$^4$ in 40 bins, $100 < E_{\mu}^{*} < 2500$ MeV in 30 bins, and $-0.4 < q^{2} < 12.6$ GeV$^2$/c$^4$ in 4 bins. The fit extracts the relative contributions of signal and normalization modes and their form factors; the relative yields of each of the $\bar{B} \to D^{*}(\to D^{+} \pi^{+}) \mu \nu$ and their form factors; the relative yields of $\bar{B}^{0} \to D^{*+}(\to D^{+} K_{S}^{0}) \mu \bar{\nu}_{\mu}$.
and $\bar{B} \to D^{\ast\ast}(\to D^{\ast+}\pi\pi)\mu^-\bar{\nu}$ decays; the relative yield of $\bar{B} \to D^{\ast+}H_c(\to \mu\nu X')X$ decays; the yield of reconstituted $D^{\ast+}$ and combinatorial backgrounds; and the background yield from hadrons misidentified as muons separately above and below $|p_T| = 10$ GeV. Uncertainties in the shapes of the templates due to the finite number of simulated events, which are therefore uncorrelated bin-to-bin, are incorporated directly into the likelihood using the Beeston-Barlow “light” procedure [33]. The fit includes shape uncertainties with bin-to-bin correlations (e.g. form factor uncertainties) via interpolation between nominal and alternative histograms. Control samples for partially reconstituted backgrounds (i.e. $D^{\ast+}\nu\pi^-$, $D^{\ast+}\nu\pi^+\pi^-$, and $D^{\ast+}\mu^-K^\pm$) are fit independently from the fit to the signal sample. Since the selections used for these control samples include inverting the isolation requirement used to select the signal sample, this method allows for the determination of the corrections to the $\bar{B} \to D^{\ast+}H_c(\to \mu\nu X')X$ and $\bar{B} \to D^{\ast+}\pi\nu\mu^-\bar{\nu}_\mu$ backgrounds with negligible influence from the signal and normalization events. The results are validated with an independently developed alternative fit. In this second approach, control samples are fit simultaneously with the signal sample with correction parameters allowed to vary, allowing correlations among parameters to be incorporated exactly. This fit also forgoes the use of interpolation in favor of reweighting the simulated samples and recomputing the kinematic distributions for each value of the corresponding parameters. The two fits are extensively cross-checked and give consistent results.

The results of the fit to the signal sample are shown in Fig. 1. Values of the $\bar{B} \to D^{\ast+}\mu^-\bar{\nu}_\mu$ form factor parameters determined by the fit agree with the current world average values. The fit finds $363,000 \pm 1600 \bar{B} \to D^{\ast+}\mu^-\bar{\nu}_\mu$ decays in the signal sample, and an uncorrected ratio of yields $N(\bar{B} \to D^{\ast+}\tau^-\bar{\nu}_\tau)/N(\bar{B} \to D^{\ast+}\mu^-\bar{\nu}_\mu) = (4.54 \pm 0.46) \times 10^{-2}$. Accounting for the $\tau^- \to \mu^-\nu\mu_\tau$ branching fraction [25] and the ratio of efficiencies results in $R(D^\ast) = 0.336 \pm 0.034$, where the uncertainty includes the statistical uncertainty, the uncertainty due to form factors, and the statistical uncertainty in the kinematic distributions used in the fit. As the signal yield is large, this uncertainty is dominated by the determination of various background yields in the fit and their correlations with the signal, which are as large as $-0.68$ in the case of $\bar{B} \to D^{\ast+}H_c(\to \mu\nu X')X$.

Systematic uncertainties on $R(D^\ast)$ are summarized in Table I. The uncertainty in extracting $R(D^\ast)$ from the fit (model uncertainty) is dominated by the statistical uncertainty of the simulated samples; this contribution is estimated via the reduction in the fit uncertainty when the sample statistical uncertainty is not considered in the likelihood. The systematic uncertainty from the kinematic shapes of the background from hadrons misidentified as muons is taken to be half the difference in $R(D^\ast)$ using the two unfolding methods. Form factor parameters are included in the likelihood as nuisance parameters, and represent a source of systematic uncertainty. The total uncertainty on $R(D^\ast)$ estimated from the fit therefore incorporates these sources. To separate the statistical uncertainty and the contribution of the form factor uncertainty, the fit is repeated with form factor parameters fixed to their best-fit values, and the reduction in uncertainty is used to determine the contribution from the form factor uncertainties. The systematic uncertainty from empirical corrections to the kinematic distributions of $\bar{B} \to D^{\ast+}(\to D^{\ast+}\pi\pi)\mu^-\bar{\nu}_\mu$ and $\bar{B} \to D^{\ast+}H_c(\to \mu\nu X')X$ backgrounds is also computed based on fixing the relevant parameters to their best fit values, as described above. The contribution of $\bar{B} \to D^{\ast+}(\to D^{\ast+}\pi\pi)\tau^-\bar{\nu}_\tau$, $\bar{B} \to D^{\ast+}(\to D^{\ast+}\pi\pi)\tau^-\bar{\nu}_\tau$, and $\bar{B} \to D^{\ast+}(D_s \to \nu\nu)X$ events is fixed to 12% of the corresponding semimuonic modes, with half of this yield assigned as a systematic uncertainty on $R(D^\ast)$. Similarly the contribution of $\bar{B} \to D^{\ast+}D_s^-(\to \tau^-\bar{\nu}_\tau)$ decays is fixed using known branching fractions [25], and 30% changes in the nominal value are taken as a systematic uncertainty. Corrections to the modeling of variables related to the pointing of the $D^0$ candidates to the PV are needed to derive the kinematic distributions for the fit. These corrections are derived from a comparison of simulated $\bar{B} \to D^{\ast+}\mu^-\bar{\nu}_\mu$ events with a pure $\bar{B} \to D^{\ast+}\mu^-\bar{\nu}_\mu$ data sample, and a systematic uncertainty is assigned by computing an alternative set of corrections using a different selection for this data subsample.

The expected yield of $D^{\ast+}\mu^-$ candidates compared to $D^{\ast+}\mu^+$ candidates (used to model the combinatorial background) varies as a function of $m(D^{\ast+}\mu^\pm)$. The size

### Table I. Systematic uncertainties in the extraction of $R(D^\ast)$.

<table>
<thead>
<tr>
<th>Model uncertainties</th>
<th>Absolute size ($\times 10^{-2}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulated sample size</td>
<td>2.0</td>
</tr>
<tr>
<td>Misidentified $\mu$ template shape</td>
<td>1.6</td>
</tr>
<tr>
<td>$B^0 \to D^{\ast+}(\tau^-/\mu^-)\bar{\nu}$ form factors</td>
<td>0.6</td>
</tr>
<tr>
<td>$\bar{B} \to D^{\ast+}H_c(\to \mu\nu X')X$ shape corrections</td>
<td>0.5</td>
</tr>
<tr>
<td>$B(\bar{B} \to D^{\ast+}\tau^-\bar{\nu}<em>\tau)/B(\bar{B} \to D^{\ast+}\mu^-\bar{\nu}</em>\mu)$</td>
<td>0.5</td>
</tr>
<tr>
<td>$\bar{B} \to D^{\ast+}(\to D^{\ast+}\pi\pi)\mu^-\bar{\nu}_\mu$ form factors</td>
<td>0.4</td>
</tr>
<tr>
<td>Corrections to simulation</td>
<td>0.4</td>
</tr>
<tr>
<td>Combinatorial background shape</td>
<td>0.3</td>
</tr>
<tr>
<td>$\bar{B} \to D^{\ast+}(\to D^{\ast+}\pi\pi)\mu^-\bar{\nu}_\mu$ form factors</td>
<td>0.3</td>
</tr>
<tr>
<td>$\bar{B} \to D^{\ast+}(D_s \to \nu\nu)X$ fraction</td>
<td>0.1</td>
</tr>
<tr>
<td>Total model uncertainty</td>
<td>2.8</td>
</tr>
<tr>
<td>Normalization uncertainties</td>
<td></td>
</tr>
<tr>
<td>Absolute size ($\times 10^{-2}$)</td>
<td></td>
</tr>
<tr>
<td>Simulated sample size</td>
<td>0.6</td>
</tr>
<tr>
<td>Hardware trigger efficiency</td>
<td>0.6</td>
</tr>
<tr>
<td>Particle identification efficiencies</td>
<td>0.3</td>
</tr>
<tr>
<td>Form factors</td>
<td>0.2</td>
</tr>
<tr>
<td>$B(\tau^- \to \mu^-\nu\mu_\tau)$</td>
<td>$&lt; 0.1$</td>
</tr>
<tr>
<td>Total normalization uncertainty</td>
<td>0.9</td>
</tr>
<tr>
<td>Total systematic uncertainty</td>
<td>3.0</td>
</tr>
</tbody>
</table>
of this effect is estimated in the $5280 < m(D^{+}\mu^\tau) < 10000\text{ MeV}/c^2$ region and the uncertainty is propagated as a systematic uncertainty on $R(D^*)$.

Uncertainties in converting the fitted ratio of signal and normalization yields into $R(D^*)$ (normalization uncertainties) come from the finite statistical precision of the simulated samples used to determine the efficiency ratio, and several other sources. The efficiency of the hardware triggers obtained in simulation differs between magnet polarities and between $\text{Pythia}$ versions—the midpoint of the predictions is taken as the nominal value and the range of variation is taken as a systematic uncertainty on the efficiency ratio. Particle identification efficiencies are applied to simulation based on binned $J/\psi \rightarrow \mu^+\mu^-$ and $D^0 \rightarrow K^-\pi^+$ control samples, which introduces a systematic uncertainty that is estimated by binning the control samples differently and by comparing to simulated particle identification. The signal and normalization form factors vary with $\theta$, and between $\text{Pythia}$ versions (syst). The finite statistical precision of the simulated samples used to determine the efficiency ratio, and between $\text{Pythia}$ versions (syst). We are also thankful for the computing resources and the access to software R&D tools provided by Yandex LLC (Russia). Individual groups or members have received support from EPLANET, Marie Sklodowska-Curie Actions and ERC (European Union), Conseil général de Haute-Savoie, Labex ENIGMASS and OCEVU, Région Auvergne (France), RFBR (Russia), XuntaGal and GENCAT (Spain), Royal Society and Royal Commission for the Exhibition of 1851 (United Kingdom).

[5] A. Bozek et al. (Belle Collaboration), Observation of $B^+ \rightarrow D^{0}\tau^+\nu_\tau$ and evidence for $B^+ \rightarrow D^0\tau^+\nu_\tau$ at Belle, Phys. Rev. D 82, 072005 (2010).
H.H. Wills Physics Laboratory, University of Bristol, Bristol, United Kingdom
Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom
Department of Physics, University of Warwick, Coventry, United Kingdom
STFC Rutherford Appleton Laboratory, Didcot, United Kingdom
School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom
School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom
Oliver Lodge Laboratory, University of Liverpool, Liverpool, United Kingdom
Imperial College London, London, United Kingdom
School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom
Department of Physics, University of Oxford, Oxford, United Kingdom
Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
University of Cincinnati, Cincinnati, Ohio, USA
University of Maryand, College Park, Maryland, USA
Syracuse University, Syracuse, New York, USA
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil
(Institut für Physik, Universität Rostock, Rostock, Germany)
(Institut de Particule de BASE (IPB), Université de Genève, Geneva, Switzerland)
(Instituto de Física Corpuscular (IFIC), Universitat de Valencia-CSIC, Valencia, Spain)
(Instituto de Física Corpuscular de Barcelona, Barcelona, Spain)
Yandex School of Data Analysis, Moscow, Russia (associated with Institute of Theoretical and Experimental Physics (ITEP), Moscow, Russia)
Instituto de Física Corpuscular (IFIC), Universitat de Valencia-CSIC, Valencia, Spain
(associated with Universitat de Barcelona, Barcelona, Spain)
Van Swinderen Institute, University of Groningen, Groningen, The Netherlands
(associated with Nikhef National Institute for Subatomic Physics, Amsterdam, The Netherlands)

† Deceased.
Also at Università di Ferrara, Ferrara, Italy.
Also at Università di Modena e Reggio Emilia, Modena, Italy.
Also at Università di Milano Bicocca, Milano, Italy.
Also at LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain.
Also at Università di Bologna, Bologna, Italy.
Also at Università di Roma Tor Vergata, Roma, Italy.
Also at Università di Genova, Genova, Italy.
Also at Scuola Normale Superiore, Pisa, Italy.
Also at Università di Cagliari, Cagliari, Italy.
Also at Politecnico di Milano, Milano, Italy.
Also at Universidade Federal do Triângulo Mineiro (UFTM), Uberaba-MG, Brazil.
Also at AGH—University of Science and Technology, Faculty of Computer Science, Electronics and Telecommunications, Kraków, Poland.
Also at Università di Padova, Padova, Italy.
Also at Hanoi University of Science, Hanoi, Viet Nam.
Also at Università di Bari, Bari, Italy.
Also at Università degli Studi di Milano, Milano, Italy.
Also at Università di Roma La Sapienza, Roma, Italy.
Also at Università di Pisa, Pisa, Italy.
Also at Università di Urbino, Urbino, Italy.
Also at P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia.