Measurement of the Ratio of the $B^0 \to D^{*-} \tau^+ \nu_\tau$ and $B^0 \to D^{*-} \mu^+ \nu_\mu$
Branching Fractions Using Three-Prong $\tau$-Lepton Decays

R. Aaij et al.*
(LHCb Collaboration)

(Received 8 November 2017; revised manuscript received 5 March 2018; published 25 April 2018)

The ratio of branching fractions $R(D^{*-}) \equiv B(B^0 \to D^{*-}\tau^+\nu_\tau)/B(B^0 \to D^{*-}\mu^+\nu_\mu)$ is measured using a data sample of proton-proton collisions collected with the LHCb detector at center-of-mass energies of 7 and 8 TeV, corresponding to an integrated luminosity of 3 fb$^{-1}$. For the first time, $R(D^{*-})$ is determined using the $\tau$-lepton decays with three charged pions in the final state. The $B^0 \to D^{*-}\tau^+\nu_\tau$ yield is normalized to that of the $B^0 \to D^{*-}\pi^+\pi^-\pi^+$ mode, providing a measurement of $B(B^0 \to D^{*-}\tau^+\nu_\tau)/B(B^0 \to D^{*-}\pi^+\pi^-\pi^+) = 1.97 \pm 0.13 \pm 0.18$, where the first uncertainty is statistical and the second systematic. The value of $B(B^0 \to D^{*-}\tau^+\nu_\tau) = (1.42 \pm 0.094 \pm 0.129 \pm 0.054)$% is obtained, where the third uncertainty is due to the limited knowledge of the branching fraction of the normalization mode. Using the well-measured branching fraction of the $B^0 \to D^{*-}\mu^+\nu_\mu$ decay, a value of $R(D^{*-}) = 0.291 \pm 0.019 \pm 0.026 \pm 0.013$ is established, where the third uncertainty is due to the limited knowledge of the branching fractions of the normalization and $B^0 \to D^{*-}\mu^+\nu_\mu$ modes. This measurement is in agreement with the standard model prediction and with previous results.

DOI: 10.1103/PhysRevLett.120.171802

In the standard model (SM) of particle physics, flavor-changing processes such as semileptonic decays of $b$ hadrons are mediated by a $W$ boson with universal coupling to leptons. Differences between the expected branching fraction of semileptonic decays into the three lepton families originate from the different masses of the charged leptons. Lepton universality can be violated in many extensions of the SM with nontrivial flavor structure. Since uncertainties due to hadronic effects cancel to a large extent, the SM prediction for the ratios between branching fractions of semitauonic decays of $B$ mesons relative to decays involving lighter lepton families, such as $R(D^{*-}) \equiv B(B^0 \to D^{*-}\tau^+\nu_\tau)/B(B^0 \to D^{*-}\mu^+\nu_\mu)$ and $R(D^{+}) \equiv B(B^- \to D^{+}(0)\tau^+\bar{\nu}_\tau)/B(B^- \to D^{+}(0)\mu^+\bar{\nu}_\mu)$, is known with an uncertainty at the percent level [1-4]. The inclusion of charge-conjugate modes is implied throughout. These decays therefore provide a sensitive probe of SM extensions with mass-dependent couplings, such as models with an extended Higgs sector [5], or leptoquarks [6,7].

Measurements of $R(D^0)$, $R(D^-)$, $R(D^{*-})$, and $R(D^{+})$ have been reported by the BABAR [8,9] and Belle [10,11] Collaborations in final states involving electrons or muons from the $\tau$ decay. The LHCb Collaboration published a determination of $R(D^{*-})$ [12], where the $\tau$ lepton was reconstructed using leptonic decays to a muon. The first simultaneous measurements of $R(D^{*-})$, $R(D^{+})$, and $\tau$ polarization, using $\tau$ decays with one charged hadron in the final state, has recently been published by the Belle Collaboration [13]. All these measurements yield values that are above the SM predictions with a combined significance of 3.9 standard deviations [14].

This Letter reports the first determination of $R(D^{*-})$ using the three-prong $\tau^+ \to \pi^+\pi^-\pi^0\bar{\nu}_\tau$ and $\tau^+ \to \pi^+\pi^-\pi^0\bar{\nu}_\mu$ decays. A more detailed description of this measurement is given in Ref. [15]. The $D^{*-}$ meson is reconstructed through the $D^{*-} \to D^0 (\to K^+\pi^-)\pi^-\pi^+$ decay chain. The visible final state consists of six charged tracks; neutral pions are ignored in this analysis. A data sample of proton-proton collisions, corresponding to an integrated luminosity of 3 fb$^{-1}$, collected with the LHCb detector at center-of-mass energies $\sqrt{s} = 7$ and 8 TeV is used.

In order to reduce experimental systematic uncertainties, the $B^0 \to D^{*-}\pi^+\pi^-\pi^+$ decay is chosen as a normalization channel. This leads to a measurement of the ratio

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

(1)
where $3\pi \equiv \pi^+\pi^-\pi^+$ and $N_{\text{sig}} (N_{\text{norm}})$ and $\epsilon_{\text{sig}} (\epsilon_{\text{norm}})$ are the yield and selection efficiency for the signal (normalization) channel, respectively. From this, $R(D^{-})$ is obtained as $R(D^{-}) = K(D^{-}) \times B(B^0 \to D^{-}3\pi)/B(B^0 \to D^-\pi^+\nu_\mu)$, where the branching fraction of the $B^0 \to D^-3\pi$ decay is taken as the weighted average of the measurements of Refs. [16–18] and that of the $B^0 \to D^-\pi^+\nu_\mu$ decay is taken from Ref. [14].

One of the key aspects of this analysis is the necessary suppression of the large background originating from $b$-hadron decays that include a $D^-$ meson, a $3\pi$ system, and any other unreconstructed additional particles, $X$. This is achieved by requiring that the position of the $3\pi$ vertex lies further away from the proton-proton interaction vertex than that of the $B^0$ vertex, as shown in Fig. 1. However, double-charm background processes, due to $B$-meson decays into a $D^-\pi$ and another charm hadron that subsequently decays into a final state containing three charged pions, are topologically similar to the signal. The largest contribution originates from $B \to D^-\pi^+X(X)$ decays, where $B$ denotes a $B^0$, $B^+$, or $B^0_s$ meson and the notation ($X$) is used when unreconstructed particles may be present in the decay chain. Double-charm backgrounds are suppressed by means of a multivariate algorithm [19] which exploits the differences in the decay dynamics and kinematics with respect to the signal process, together with different properties used by partial reconstruction algorithms.

The LHCb detector is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, described in detail in Refs. [20,21]. In the simulation, proton-proton collisions are generated using PYTHIA [22] with a specific LHCb configuration [23]. Decays of hadronic particles are described by EvtGen [24], in which final-state radiation is generated using PHOTOS [25]. The TAUOLA package [26] is used to simulate the decays of the $\tau$ lepton into $3\pi\bar{\nu}_{\tau}$ and $3\pi\nu_{\tau}$ final states, according to the resonance chiral Lagrangian model [27] with a tuning based on the results from the BABAR Collaboration [28]. The interaction of the generated particles with the detector and its response are implemented using the GEANT4 toolkit [29] as described in Ref. [30]. The signal decays are simulated using form factors that are derived from the heavy-quark effective theory [31]. The experimental values of the corresponding parameters are taken from Ref. [14], except for an unmeasured helicity-suppressed component, which is taken from Ref. [32].

The online event selection is performed by a trigger system [33], which consists of a hardware stage based on information from the calorimeter and muon systems followed by a software stage that performs a full event reconstruction. At the hardware stage, events are selected if either particles forming the signal candidate satisfy a requirement on transverse energy or particles other than those forming the signal candidate pass any trigger algorithm.

The software trigger requires a two-, three-, or four-track secondary vertex with significant displacement from any primary proton-proton interaction vertex (PV) consistent with the decay of a $b$ hadron or a two-track vertex with a significant displacement from any PV consistent with a $D^0 \to K^+\pi^-$ decay. In both cases, at least one charged particle must have a transverse momentum $p_T > 1.7$ GeV/$c$ and must be inconsistent with originating from any PV. A multivariate algorithm [19] is used for the identification of secondary vertices consistent with the decay of a $b$ hadron, while secondary vertices consistent with the decay of a $D^0$ meson are identified using topological criteria.

In the offline selection, $D_0^0$, $D_+^0$, and $\tau$ candidates are selected based on kinematic, geometric, and particle identification criteria. Three charged pions are used to reconstruct $\tau$-decay candidates, including both the $\tau^+ \to 3\pi\nu_{\tau}$ and $\tau^+ \to 3\pi\bar{\nu}_{\tau}$ modes. The vertex position and the momentum of the $B^0$ candidate are determined through a fit to all reconstructed particles in the decay chain [34]. The difference of the positions of the $3\pi$ and the $B^0$ vertices along the beam direction, divided by its uncertainty, has to be greater than 4. This requirement suppresses the background due to $B \to D^{-}3\pi X$ decays by 3 orders of magnitude and has an efficiency of 35% for the signal. The normalization sample is selected by requiring the difference in the positions of the $D_0^0$ and $3\pi$ vertices along the beam direction, divided by its uncertainty, to be greater than 4.

Backgrounds due to partially reconstructed $B$-meson decays, where at least one additional particle originates from either the $3\pi$ vertex or the $B$ vertex, or from both, are suppressed by requiring a single $B^0$ candidate per event. In addition, a charged-particle isolation algorithm is applied as described in the following. Tracks other than those used for the signal candidate are considered if they have minimal

![FIG. 1. Topology of the signal decay. A requirement on the distance between the $3\pi$ and the $B^0$ vertices along the beam direction to be greater than 4 times its uncertainty is applied. For $B \to D^-3\pi(X)$ decays, the $3\pi$ vertex coincides with the $B$ vertex.](image-url)
requirements on the transverse momentum and are inconsistent with originating from any PV. If any of these tracks has an impact parameter significance with respect to either the \( B^0 \) or \( \tau \) vertex smaller than 5 standard deviations, the \( B^0 \) candidate is rejected. This criterion rejects 95% of candidates due to \( B \to D^{\ast\ast}\bar{D}^0(X) \) decays while retaining 80% of the signal decays. In addition, a neutral-particle isolation algorithm computes the multiplicities of reconstructed tracks and neutral particles, and the energy in the calorimeter system, contained in a cone centered around the direction of the \( \tau \) candidates. These variables are used as inputs of the multivariate classifier described below.

Variables such as the squared invariant mass of the \((\tau, \nu_\tau)\) pair, \( q^2 \), and the \( \tau \) decay time, \( t_\tau \), provide good discrimination between signal and background processes, but they depend on the momenta of the neutrinos in the final state of the \( B^0 \) decay. However, due to the presence of a single neutrino in the \( \tau \) decay, the momentum of the \( \tau \) lepton can be determined, up to a twofold ambiguity, from the momentum vector of the 3\( \pi \) system and the flight direction of the \( \tau \) candidate. The value of the \( \tau \) momentum is approximated by taking the average of the two solutions, as discussed in Ref. [15]. A similar strategy is used to compute the \( B^0 \) momentum. The \( B^0 \) rest frame variables are determined with sufficient accuracy to retain their discriminating power. A partial reconstruction is performed also under the background hypothesis where \( B^0 \to D^{\ast\ast}D^+_1(\to 3\pi N) \), with \( N \) denoting a neutral system. The variables describing decay kinematics, as reconstructed by this algorithm, differ between signal and background processes; a selected set is used as the input to the multivariate classifier described below.

The dominant double-charm background process \( B \to D^{\ast\ast}D^+_1(X) \) is reduced by taking into account the resonant structure of the 3\( \pi \) system. The \( \tau^+ \) lepton decays to 3\( \pi \) final states predominantly through the \( a_1(1260)^+ \to \rho^0\pi^+ \) decay. By contrast, the \( D^+_1 \) meson decays to 3\( \pi \) final states predominantly through the \( \eta \) and \( \eta' \) resonances. These and other features are exploited by means of a boosted decision tree (BDT) [35,36], as described in Ref. [15]. The BDT response in the simulation is validated using three control samples: a \( B \to D^{\ast\ast}D^+_1(X) \) data sample, which is obtained by using partial reconstruction under the background hypothesis; a \( B \to D^{\ast\ast}\bar{D}^0(X) \) data sample, with the subsequent \( \bar{D}^0 \to K^-3\pi \) decay, which is obtained by removing the charged-particle isolation criterion and requiring a particle satisfying kaon identification criteria with an origin at the 3\( \pi \) vertex; and a \( B \to D^{\ast\ast}D^+_1(X) \) data sample, with \( D^{\ast\ast} \to K^-\pi^+\pi^+ \), which is obtained replacing the negative pion with a candidate identified as a kaon. For all these samples, good agreement between the data and simulation is observed in the distributions of the variables used in the BDT. These control samples are also used to correct the simulation to reproduce the expected distributions of the fit variables in data.

The yield of the normalization mode is determined by fitting the invariant mass distribution of the \( D^{\ast\ast}3\pi \) system around the known \( B^0 \) mass [37] for candidates in the normalization sample. The fitting function of the normalization channel is the sum of a Gaussian function and a Crystal Ball function [38]. An exponential function is used for the combinatorial background. All parameters are floating in the fit. A total of \( N_{\text{norm}} = 17660 \pm 158 \) candidates are found, where a small contribution of \( 151 \pm 22 B^0 \to D^{\ast\ast}D^+_1(\to 3\pi) \) decays has been accounted for in the yield and uncertainty. The latter component is estimated by fitting the 3\( \pi \) mass distribution for candidates with a reconstructed \( B^0 \) mass in a window around the known value.

The signal yield is obtained from a three-dimensional binned fit to the data, in a region of the BDT output enriched in signal decays. The fit dimensions are \( q^2, t_\tau \), and the BDT output. Several components enter in the fit: in particular, a signal component which also accounts for higher-mass charm-meson states; background components due to \( B \to D^{\ast\ast}D^+_1(X), \ B \to D^+D^+(X), \) and \( B \to D^+D^0(X) \) decays; a residual contribution from \( B \to D^{\ast\ast}3\pi X \) decays; and a combinatorial background.

The signal template is the sum of terms, due to \( \tau^+ \to 3\pi\nu_\tau \) and \( \tau^+ \to 3\pi\nu_\tau \) decays, where the relative ratio between these components is fixed according to their branching fractions and simulation-derived selection efficiencies. A contribution due to \( B \to D^{\ast\ast}\tau^+\nu_\tau \) decays, where \( D^{\ast\ast} \) denotes \( P \)-wave charm mesons or any higher mass states, with the \( D^{\ast\ast} \) being produced in the \( D^{\ast\ast} \) decay chain, is also related to the signal yield through a proportionality factor derived from Ref. [39]. A data sample where the narrow \( D^0(2420) \) and \( D^0(2460) \) resonances are reconstructed in their \( D^+\pi \) decays is used to validate the simulation.

The background originating from decays of \( B \) mesons into \( D^{\ast\ast}D^+_1(X) \) final states is divided into contributions from \( B^0 \to D^{\ast\ast}\bar{D}^0, B^0 \to D^+D^+X, B^0 \to D^+D^+X_{01}(2317), B^0 \to D^+D^+X(2400), B \to D^{\ast\ast}D^+_1X, \) and \( B^0 \to D^+D^+_1X \) decays. The relative yield of each of these processes is constrained in the final fit using the results of an auxiliary fit, shown in Fig. 2, to the \( D^{\ast\ast}3\pi \) invariant mass. The fit is performed on a control sample of data obtained by reconstructing the \( D^+_1 \) decay.

The \( D^+_1 \) decay model used in the simulation does not accurately describe the data because of the limited knowledge of the \( D^+_1 \) decay amplitude to \( 3\pi X \) final states. Therefore, the contribution of the background from \( D^+_1 \) decays is determined from the data in a control region, selected by the BDT output, where this background is abundant. In this region, the distributions of the minimum and maximum invariant masses of the oppositely charged pions, \( \min[m(\pi^+\pi^-)] \) and \( \max[m(\pi^+\pi^-)] \), the invariant mass of the same-charge pion pair, and that of the 3\( \pi \) system are fitted simultaneously in order to determine the contributions from different \( D^+_1 \) final states. These are grouped in four categories. The first (second) includes \( D^+_1 \) decays into \( \eta\pi \) or \( \eta\rho \) (\( \eta'\pi \) or \( \eta'\rho \)), where at least one pion
The subdivided into two contributions, depending on whether the final fit are then recomputed by taking from the simulation for these decays in the BDT output region considered in the elsewhere. The contribution of the former background is from the π(π′) decay. The third category contains D^+_s decays where at least one pion originates from another intermediate resonance such as an ω or φ meson, D^+_s → 3πX decays where none of the three pions originates from an intermediate resonance, and D^+_s → τ^+ (→ 3πντ), ντ, decays. The fourth category consists of backgrounds without D^+_s mesons. Figure 3 shows, as an example, the distribution of \( \min(m(\pi^+\pi^-)) \) and the resulting fit components. The results obtained by the fit in this region of BDT output are used to compute weights for each D^+_s decay mode, to be applied to the simulation. The templates used for these decays in the BDT output region considered in the final fit are then recomputed by taking from the simulation the relative proportion between the yields in the two regions of the BDT output for each decay mode.

Background originating from \( B \to D^{*-}D^0X \) decays is subdivided into two contributions, depending on whether the 3π system originates from the same D^0 vertex or one pion originates from the D^0 vertex and the other two from elsewhere. The contribution of the former background is constrained by the yield obtained from the \( B \to D^{*-}D^0(X) \) control sample. The template shape is also validated using this control sample. The yield of the other \( B \to D^{*-}D^0X \) background component is a free parameter in the fit, while its shape is taken from the simulation. The yield of the \( B \to D^{*-}3\pi X \) background is also a free parameter. The template shape is validated using the corresponding control sample. A residual background from \( B \to D^{*-}3\pi X \) modes is included in the fit. The yields of these components are constrained by those measured from a data sample enriched with \( B \to D^{*-}3\pi X \) decays in which the distance of the B vertex from the PV exceeds that of the 3π.

The combinatorial background is divided into two contributions, depending on whether the background contains a real \( D^{*-} \to D^0\pi^- \) decay chain or not. In the first case, the \( D^{*-} \) and the 3π systems are required to originate from different B decays. The templates for this background are taken from the simulation. A sample of candidates where the \( D^{*-} \) and the 3π systems have the same charge is used to normalize the data and simulation in the region where the \( D^{*-}3\pi \) mass is above the known B mass. The background not including a real \( D^{*-} \) decay chain is parametrized and constrained using candidates outside a window around the known \( D^0 \) mass.

The results of the fit are shown in Fig. 4. The global \( \chi^2 \) of the fit is 1.15 per degree of freedom, after taking into account the statistical fluctuation in the simulation templates. The signal yield is corrected for a small bias of 40 candidates, due to the finite size of the templates from the simulation, as detailed below, giving \( N_{\text{sig}} = 1296 \pm 86 \) candidates. The result

\[
K(D^{*-}) = 1.97 \pm 0.13(\text{stat}) \pm 0.18(\text{syst})
\]

determined from Eq. (1), where the efficiencies for events within LHCb acceptance are \( (0.39 \times 10^{-3}) \) and \( (1.36 \times 10^{-3}) \) for signal and normalization modes, respectively, are taken from the simulation, and an effective sum \( (13.81 \pm 0.07)\% \) of the branching fractions for the \( \tau^+ \to 3\pi\nu_{\tau} \) and \( \tau^+ \to 3\pi\rho^0\nu_{\tau} \) decays is used to account for the different selection efficiencies between the two modes and small feeddown from other \( \tau \) decays. A correction factor 1.056 ± 0.025 has also been applied to account for discrepancies between the data and simulation and for a small feeddown contribution from \( B^+_s \to D^{*-}\tau^+\nu_{\tau} \) decays, where \( D^{*-} \to D^0K^0 \).

The branching fraction

\[
B(B^0 \to D^{*-}\tau^+\nu_{\tau}) = [1.42 \pm 0.094(\text{stat}) \\
\pm 0.129(\text{syst}) \pm 0.054(\text{ext})] \times 10^{-2}
\]

is obtained by using \( B(B^0 \to D^{*-}3\pi) = (7.214 \pm 0.28) \times 10^{-3} \), the weighted average of the LHCb [16], BABAR [17], and Belle [18] measurements. Finally, the ratio of branching fractions...
R(D)/C3 − ð = 0.291/C6 0.019 (stat)/C6 0.026 (syst)/C6 0.013 (ext)
is obtained by using B(B0 → D−μνμ) = (4.88 ± 0.10) × 10−2 from Ref. [14]. In both results, the third uncertainty is
due to the limited knowledge of the external branching fractions.
Systematic uncertainties on R(D) are reported in Table I. The uncertainty due to the limited size of the simulated samples is computed by repeatedly sampling each template with a bootstrap procedure, performing the fit, and taking the standard deviation of the results obtained.

Empty bins in the templates used in the fit also introduce a positive bias of 3% in the determination of the signal yield. This corresponds to a correction of 40 candidates, with an uncertainty of 1.3%. The limited size of the simulated samples also contributes to the systematic uncertainty on the efficiencies for signal and normalization modes.

The systematic uncertainty associated with the signal decay model derives from the limited knowledge of the form factors and the τ polarization, from possible contributions from other τ decay modes, and from the relative branching fractions and selection efficiencies of τ+ → 3π0ντ and τ+ → 3πντ decays. Uncertainties due to the
knowledge of the $D^{*+}\tau^+\nu_\tau$ contribution to the signal yield are estimated using a control sample, where one additional charged pion originating from the $B$ vertex is identified. The observed yield of the narrow $D_s(2420)^0$ resonance is used to infer a 40\% uncertainty on the yield of $D_s^{*+}\tau^+\nu_\tau$ decays relative to that of the signal. A systematic uncertainty is also assigned to take into account the feeddown from $B^0$ decays into $D_s^{*-}\tau^+\nu_\tau$.

The uncertainty due to the knowledge of the $D_s^+$ decay model is estimated by repeatedly varying the correction factors of the templates within their uncertainties, as determined from the associated control sample, and performing the fit. The spread of the fit results is assigned as the corresponding systematic uncertainty. The template shapes of the $D_s^{*-}D_s^{*+}, D_s^{*-}D^0,$ and $D^+D^-+X$ backgrounds depend on the dynamics of the corresponding decays. Empirical variations of the kinematic distribution are performed, and the spread of the fit results is taken as a systematic uncertainty. A similar procedure is applied to the template for the combinatorial background. Other sources of systematic uncertainty arise from the inaccuracy on the yields of the various background contributions and from the limited knowledge of the normalization modeling and the resonant structure of the residual background due to $B \to D_s^{*-}3\pi X$ decays.

Systematic effects on the efficiencies for signal and normalization partially cancel in the ratio. The trigger efficiency depends on the distributions of the decay time of the $3\pi$ system and the invariant mass of the $D_s^{*-}3\pi$ system. These distributions differ between the signal and normalization modes, and the difference of the trigger efficiency for these two decays is taken into account.

In conclusion, the first measurement of $\mathcal{R}(D_s^{*-})$ with three-prong $\tau$ decays has been performed by using a technique that is complementary to all previous measurements of this quantity and offers the possibility to study other $B$-hadron decay modes in a similar way. The result, $\mathcal{R}(D_s^{*-}) = 0.291 \pm 0.019\text{(stat)} \pm 0.026\text{(syst)} \pm 0.013\text{(ext)}$, is one of the most precise single measurements performed so far. It is 1.1 standard deviations higher than the SM calculation (0.252 $\pm$ 0.003) of Ref. [1] and consistent with previous determinations. An average of this measurement with the LHCb result using $\tau^+ \to \mu^+\nu_\mu\bar{\nu}_\tau$ decays [12], accounting for small correlations due to form factors, $\tau$ polarization, and $D^{*+}\tau^+\nu_\tau$ feeddown, gives $\mathcal{R}(D_s^{*-}) = 0.31 \pm 0.016\text{(stat)} \pm 0.021\text{(syst)}$, consistent with the world average and 2.2 standard deviations above the SM prediction. The overall status of $\mathcal{R}(D)$ and $\mathcal{R}(D_s^{*})$ measurements is reported in Ref. [14]. After the inclusion of this result, the combined discrepancy of $\mathcal{R}(D)$ and $\mathcal{R}(D_s^{*})$ determinations with the SM prediction is 4.1 standard deviations.

We express our gratitude to our colleagues in the CERN accelerator institutes. We acknowledge support from CERN and from the national agencies: CAPES, CNPq, FAPERJ, and FINEP (Brazil); MOST and NSFC (China); CNRS/IN2P3 (France); BMBF, DFG, and MPG (Germany); INFN (Italy); NWO (Netherlands); MINISW and NCN (Poland); MEN/IFA (Romania); MinES and FASO (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); NSF (USA). We acknowledge the computing resources that are provided by CERN, IN2P3 (France), KIT and DESY (Germany), INFN (Italy), SURF (Netherlands), PIC (Spain), GridPP (United Kingdom), RRCKI and Yandex LLC (Russia), CSCS (Switzerland), IFIN-HH (Romania), CBPF (Brazil), PL-GRID (Poland), and OSC (USA). We are indebted to the communities behind the multiple open-source software packages on which we depend. Individual groups or members have received support from AvH Foundation (Germany), EPLANET, Marie Sklodowska-Curie Actions, and ERC (European Union), ANR, Labex P2IO, ENIGMASS and OCEVU, and Région Auvergne-Rhône-Alpes (France), RFBR and Yandex LLC (Russia), GVA, XuntaGal, and GENCAT (Spain), and Herchel Smith Fund, the Royal Society, the English-Speaking Union, and the Leverhulme Trust (United Kingdom).


[11] Y. Sato et al. (Belle Collaboration), Measurement of the branching ratio of $B^{0} \rightarrow D^{+}\tau^{-}\bar{\nu}_{\tau}$ relative to $B^{0} \rightarrow D^{+}\ell^{-}\bar{\nu}_{\ell}$ decays with a semileptonic tagging method, Phys. Rev. D 94, 072007 (2016).


[16] R. Aaij et al. (LHCb Collaboration), Study of $B^{0} \rightarrow D^{+}\pi^{-}\pi^{-}\pi^{+}$ and $B^{0} \rightarrow D^{+}\pi^{+}\pi^{-}\pi^{-}$ decays, Phys. Rev. D 87, 092001 (2013).


[18] G. Majumder et al. (Belle Collaboration), Observation of $B^{0} \rightarrow D^{*-}(5\pi)^{+}$, $B^{+} \rightarrow D^{*-}(4\pi)^{+}$, and $B^{+} \rightarrow D^{(*)0}(5\pi)^{+}$, Phys. Rev. D 70, 111103 (2004).


[27] J. M. Nugent, T. Przedziński, P. Roig, O. Shekhovtsova, and Z. Was, Resonance chiral Lagrangian currents and experimental data for $\tau^{-} \rightarrow \pi^{-}\pi^{0}\pi^{0} \nu_{\tau}$, Phys. Rev. D 88, 093012 (2013).


54 Oliver Lodge Laboratory, University of Liverpool, Liverpool, United Kingdom
55 Imperial College London, London, United Kingdom
56 School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom
57 Department of Physics, University of Oxford, Oxford, United Kingdom
58 Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
59 University of Cincinnati, Cincinnati, Ohio, USA
60 University of Maryland, College Park, Maryland, USA
61 Syracuse University, Syracuse, New York, USA
62 Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil
   (associated with Institution Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil)
63 University of Chinese Academy of Sciences, Beijing, China
   (associated with Institution Center for High Energy Physics, Tsinghua University, Beijing, China)
64 School of Physics and Technology, Wuhan University, Wuhan, China
   (associated with Institution Center for High Energy Physics, Tsinghua University, Beijing, China)
65 Institute of Particle Physics, Central China Normal University, Wuhan, Hubei, China
   (associated with Institution Center for High Energy Physics, Tsinghua University, Beijing, China)
66 Departamento de Física, Universidad Nacional de Colombia, Bogota, Colombia
   (associated with Institution LPNHE, Université Pierre et Marie Curie, Université Paris Diderot, CNRS/IN2P3, Paris, France)
67 Institut für Physik, Universität Rostock, Rostock, Germany
   (associated with Institution Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany)
68 National Research Centre Kurchatov Institute, Moscow, Russia
   (associated with Institution Institute of Theoretical and Experimental Physics (ITEP), Moscow, Russia)
69 National Research Tomsk Polytechnic University, Tomsk, Russia
   (associated with Institution Institute of Theoretical and Experimental Physics (ITEP), Moscow, Russia)
70 Instituto de Física Corpuscular, Centro Mixto Universidad de Valencia - CSIC, Valencia, Spain
   (associated with Institution ICCUB, Universitat de Barcelona, Barcelona, Spain)
71 Van Swinderen Institute, University of Groningen, Groningen, The Netherlands
   (associated with Institution Nikhef National Institute for Subatomic Physics, Amsterdam, The Netherlands)

a Also at Università di Ferrara, Ferrara, Italy.
b Also at LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain.
c Also at Laboratoire Leprince-Ringuet, Palaiseau, France.
d Also at Università di Milano Bicocca, Milano, Italy.
e Also at Università di Modena e Reggio Emilia, Modena, Italy.
f Also at Novosibirsk State University, Novosibirsk, Russia.
g Also at Università di Cagliari, Cagliari, Italy.
h Also at Università di Bologna, Bologna, Italy.
i Also at Università di Roma Tor Vergata, Roma, Italy.
j Also at Università di Genova, Genova, Italy.
k Also at Scuola Normale Superiore, Pisa, Italy.
l Also at Università di Bari, Bari, Italy.
m Also at Università degli Studi di Milano, Milano, Italy.
n Also at Universidade Federal do Triângulo Mineiro (UFTM), Uberaba-MG, Brazil.
o Also at AGH - University of Science and Technology, Faculty of Computer Science, Electronics and Telecommunications, Kraków, Poland.
p Also at Università di Padova, Padova, Italy.
q Also at Iligan Institute of Technology (IIT), Iligan, Philippines.
r Also at Hanoi University of Science, Hanoi, Viet Nam.
s Also at Università di Pisa, Pisa, Italy.
t Also at Università della Basilicata, Potenza, Italy.
u Also at Università di Roma La Sapienza, Roma, Italy.
v Also at Università di Urbino, Urbino, Italy.
w Also at P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia.