Measurement of the branching fraction of $B \rightarrow D(\ast)\pi \nu$ at Belle using hadronic tagging in fully reconstructed events

A. Vossen et al.
Belle Collaboration

Follow this and additional works at: https://digitalcommons.kennesaw.edu/facpubs

Part of the Physics Commons

Recommended Citation
et al., A. Vossen, "Measurement of the branching fraction of $B \rightarrow D(\ast)\pi \nu$ at Belle using hadronic tagging in fully reconstructed events" (2018). Faculty Publications. 4181.
https://digitalcommons.kennesaw.edu/facpubs/4181

This Article is brought to you for free and open access by DigitalCommons@Kennesaw State University. It has been accepted for inclusion in Faculty Publications by an authorized administrator of DigitalCommons@Kennesaw State University. For more information, please contact digitalcommons@kennesaw.edu.
Measurement of the branching fraction of $B \to D(\ast)\pi \ell \nu$ at Belle using hadronic tagging in fully reconstructed events


(The Belle Collaboration)

1University of the Basque Country UPV/EHU, 48080 Bilbao
2Beihang University, Beijing 100191
3Brookhaven National Laboratory, Upton, New York 11973
4Budker Institute of Nuclear Physics SB RAS, Novosibirsk 630090
5Faculty of Mathematics and Physics, Charles University, 121 16 Prague
6University of Cincinnati, Cincinnati, Ohio 45221
7Deutsches Elektronen–Synchrotron, 22607 Hamburg
8Duke University, Durham, North Carolina 27708
9University of Florida, Gainesville, Florida 32611
10Justus-Liebig-Universität Gießen, 35392 Gießen
11Gifu University, Gifu 501-1193
12II. Physikalisches Institut, Georg-August-Universität Göttingen, 37073 Göttingen
13SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193
14Gyeongsang National University, Chinju 660-701
15Hanyang University, Seoul 133-791
16University of Hawaii, Honolulu, Hawaii 96822
17High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801
18J-PARC Branch, KEK Theory Center, High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801
19IKERBASQUE, Basque Foundation for Science, 48013 Bilbao
20Indian Institute of Technology Bhubaneswar, Satya Nagar 751007
21Indian Institute of Technology Guwahati, Assam 781039
22Indian Institute of Technology Hyderabad, Telangana 502285
23Indian Institute of Technology Madras, Chennai 600036
24Indiana University, Bloomington, Indiana 47408
25Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049
26Institute of High Energy Physics, Vienna 1050
27INFN - Sezione di Napoli, 80126 Napoli
28INFN - Sezione di Torino, 10125 Torino
29Advanced Science Research Center, Japan Atomic Energy Agency, Naka 319-1195
30J. Stefan Institute, Ljubljana
We report a measurement of the branching fraction of the decay \( B \to D^{(*)} \pi \ell \nu \). The analysis uses \( 7.72 \times 10^6 \) \( B \bar{B} \) pairs produced in \( e^+e^- \to \Upsilon(4S) \) data recorded by the Belle experiment at the KEKB asymmetric-energy \( e^+e^- \) collider. The tagging \( B \) meson in the decay is fully reconstructed in a hadronic decay mode. On the signal side, we reconstruct the decay \( B \to D^{(*)} \pi \ell \nu \) (\( \ell = e, \mu \)). The measured branching fractions are 

\[
\begin{align*}
B(B^+ \to D^- \pi^+ \ell^+ \nu) &= [4.55 \pm 0.27 \text{ (stat.)} \pm 0.39 \text{ (syst.)}] \times 10^{-3}, \\
B(B^0 \to D^0 \pi^- \ell^+ \nu) &= [4.05 \pm 0.36 \text{ (stat.)} \pm 0.41 \text{ (syst.)}] \times 10^{-3}, \quad B(B^+ \to D^{*-} \pi^0 \ell^- \nu) = [6.03
\end{align*}
\]


\[ \pm 0.43 \text{ (stat.)} \pm 0.38 \text{ (syst.)} \times 10^{-3}, \text{ and } \mathcal{B}(B^0 \rightarrow \bar{D}^0 \pi^- \ell^+ \nu) = [6.46 \pm 0.53 \text{ (stat.)} \pm 0.52 \text{ (syst.)}] \times 10^{-3}. \] These are in good agreement with the current world average values.

PACS numbers: 12.15.Hh, 13.20.He, 14.40.Nd, 12.38.Gc

I. INTRODUCTION

Semileptonic decays of B mesons are an important tool for precision measurements of CKM matrix elements and precision tests of the electroweak sector of the standard model. An important recent development was the observation of a more than 3\(\sigma\) deviation between the standard model expectation for \(R(D^{(*)})\) \([1, 2]\) and the combined experimental results from Babar \([3, 4]\), Belle \([5-7]\) and LHCb \([8, 9]\). Here, \(R(D^{(*)})\) is defined as the ratio of the branching fraction \((B)\) of \(B \rightarrow D^{(*)}\pi\ell\nu\) and \(B \rightarrow D^{(*)}\ell\nu\), \((\ell = e, \mu)\). We report on a new measurement of \(B \rightarrow D^{(*)}\pi\ell\nu\), which is important as a background for \(B \rightarrow D^{(*)}\pi\ell\nu\) decays, and in its own right, as a vehicle to understand high-multiplicity semi-leptonic B decays. The process \(B \rightarrow D^{(*)}\pi\ell\nu\) proceeds predominantly via \(B \rightarrow (D^{**} \rightarrow D^{(*)}\pi)\ell\nu\), where \(D^{**}\) is an orbitally excited \((L = 1)\) charmed meson. The \(D^{**}\) mass-spectrum is grouped into two pairs having light-quark total angular momentum \(j_q = \frac{1}{2}\) and \(j_q = \frac{3}{2}\) \([10]\). All states can decay via \(D^{**} \rightarrow D^{*}\pi\), while the \(2^+\) state can also decay via \(D^{**} \rightarrow D\pi\). Since the \(D^{**}\) masses are not far from threshold, and the \(j_q = \frac{3}{2}\) have a significant D-wave component, these states are narrow and were observed with a typical width of about 20 MeV \([11, 12]\). On the other hand, the states with \(j_q = \frac{1}{2}\) decay mainly via S-wave and are therefore expected to be broad resonances with a width of several hundred MeV \([10, 14]\). Compared to previous measurements of \(B(B \rightarrow D^{(*)}\pi\ell\nu)\) at Belle \([11]\), the analysis presented in this report benefits from the use of the full Belle dataset, containing 772 \(\times 10^6\) \(B\bar{B}\) pairs, recorded at the \(\Upsilon(4S)\) resonance and an improved hadronic-tagging method.

II. EXPERIMENTAL APPARATUS

The Belle experiment \([15]\) at the KEKB storage ring \([16]\) recorded about 1 \(\text{ab}^{-1}\) of \(e^+e^-\) annihilation data. The data were taken mainly at the \(\Upsilon(4S)\) resonance at \(\sqrt{s} = 10.58\) GeV, but also at \(\Upsilon(1S)\) to \(\Upsilon(5S)\) resonances and at \(\sqrt{s} = 10.52\) GeV. The Belle instrument used in this analysis includes the central drift chamber (CDC) and the silicon vertex detector, which provides precision tracking for tracks in the polar-angle range \(17.0^0 < \theta_{\text{lab}} < 150.0^0\), and the electromagnetic calorimeter (ECL) covering the same range. The polar angle \(\theta_{\text{lab}}\) is measured with respect to the \(z\)-axis, which is anti-parallel to the \(e^+\) beam. Charged particle identification is performed using specific ionization measurements in the CDC, time-of-flight information from the interaction point (IP) to a barrel of scintillators, light yield in an array of aerogel Cherenkov counters in the barrel and the forward endcap, as well as a muon- and \(K^0_L\)-identification system in the return yoke of the superconducting solenoid, which provides a 1.5 T magnetic field.

A. \(B_{\text{tag}}\) selection

A neural-network-based multivariate classifier, as implemented in the \texttt{NeuroBayes} package \([21, 22]\), is used to fully reconstruct \(B\) mesons that decay hadronically. The algorithm considers 17 final states for charged \(B\) candidates and 15 final states for neutral \(B\) candidates. Incorporating the subsequent hadronic decays and \(J/\psi\) leptonic decays, the algorithm investigates 1104 different decay topologies. The output variable \(a_{\text{tag}}\) of the algorithm takes a value between 0 and 1, with larger values corresponding to a higher likelihood that a \(B\) meson was correctly reconstructed.

We select events with \(\log(a_{\text{tag}}) > -3.5\). For each \(B_{\text{tag}}\), we impose a requirement on the difference between the measured center-of-mass (CM) energy and its nominal value of \(|\Delta E| = |E_{B_{\text{tag}}} - E_{\text{CM}}| < 0.18\) GeV, and on the beam constrained mass of \(M_{bc} = \)
\(\sqrt{(E_{CM}/c^2)^2 - (\vec{P}_{B_{tag}}/c)^2} > 5.27 \text{ GeV}/c^2\). Here, \(E_{B_{tag}}\) and \(\vec{P}_{B_{tag}}\) are the energy and momentum of the tagged \(B\) candidate.

Differences in the tagging efficiency between data and MC have been observed \[23\]. These depend on the tag-side reconstruction and the value of \(\sigma_{tag}\). We use a calibration derived in Ref. \[23\], which uses a control sample \(B \rightarrow X_{ll}v\) decays on the signal side. Based on this calibration, we assign an event-by-event weight based on the reconstructed \(B_{tag}\) decay mode and value of \(\sigma_{tag}\) to equalize the efficiency of the tagging algorithm between data and MC.

### B. \(B_{sig}\) reconstruction

Having selected the \(B_{tag}\) in this way, the signal side \(B_{sig}\) is then reconstructed with the charged tracks and photons in the event that are not part of the \(B_{tag}\) decay chain. Charged tracks are identified using the Belle particle identification (PID) \[24\]. We accept electrons in the laboratory frame polar-angle range \(17^\circ < \theta_e < 150^\circ\), and muons in the range \(25^\circ < \theta_\mu < 145^\circ\), where the relevant subsystems of the Belle PID have acceptance for these particles. To recover energy lost by bremsstrahling of electrons, we add the 4-vector of the closest \(\gamma\) found within \(5^\circ\) of an identified electron. Charged tracks that cannot be unambiguously identified are treated as pions. We reconstruct \(\pi^0\) candidates from pairs of photons, each of which satisfies a minimum energy requirement of 50 MeV, 75 MeV or 100 MeV in the barrel (32° < \(\theta_\pi\) < 130°), the forward endcap (17° < \(\theta_\pi\) < 32°) or the backward endcap (130° < \(\theta_\pi\) < 150°), respectively. We require the reconstructed mass to lie in the range \(0.12 \text{ GeV}/c^2 < M_{\pi^0}\) < 0.15 \text{ GeV}/c^2\), which corresponds to about five times the measured resolution around the nominal mass. To reduce overlap in the \(\pi^0\) candidate list, we sort them according to the most energetic daughter photon (and then, if needed, the second most energetic daughter) and remove any pion that shares photons with one that appears earlier in this list. We reconstruct \(K_{S}^{0}\) mesons from \(\pi^+\pi^-\) pairs. We require the two-pion invariant mass to lie in the range 0.482-0.514 \text{ GeV}/c^2 (about four times the experimental resolution around the nominal mass \[20\]). Different selections are applied, depending on the momentum of the \(K_{S}^{0}\) candidate in the laboratory frame \[25\]: For low \((p < 0.5 \text{ GeV}/c)\), medium \((0.5 \leq p \leq 1.5 \text{ GeV}/c)\), and high momentum \((p > 1.5 \text{ GeV}/c)\) candidates, we require the impact parameters of the pion daughters in the transverse plane (perpendicular to the beam) to be greater than 0.05 cm, 0.03 cm, and 0.02 cm, respectively. The angle in the transverse plane between the vector from the interaction point to the \(K_{S}^{0}\) vertex and the \(K_{S}^{0}\) flight direction is required to be less than 0.3 rad, 0.1 rad, and 0.03 rad for low, medium, and high momentum candidates, respectively; the separation distance along the \(z\) axis of the two pion trajectories at their closest approach must be below 0.8 cm, 1.8 cm, and 2.4 cm, respectively. Finally, for medium (high) momentum \(K_{S}^{0}\) candidates, we require the flight length in the transverse plane to be greater than 0.08 cm (0.22 cm).

Using the reconstructed pions and kaons, we reconstruct \(D\) mesons in the channels \(D^0 \rightarrow K^-\pi^+\pi^-\), \(D^0 \rightarrow K^0_{S}\pi^+\pi^-\), \(D^0 \rightarrow K^-K^+\), \(D^0 \rightarrow K^0_{S}\pi^0\), \(D^+ \rightarrow K^0_{S}\pi^+\), \(D^+ \rightarrow K^0_{S}\pi^0\), \(D^+ \rightarrow K^0_{S}\pi^+\pi^-\), \(D^+ \rightarrow K^-\pi^+\pi^-\), \(D^+ \rightarrow K^-K^+\). Here and throughout this report, the charge-conjugated modes are implied. We require a maximum difference of 3\(\sigma\) between the reconstructed mass and the nominal \(D\) mass. This corresponds to 15 MeV for all modes except the \(D^0 \rightarrow K^-\pi^+\pi^-\) channel where the corresponding value is 25 MeV. Using the \(D\) candidates, we reconstruct \(D^+\) mesons in the channels \(D^{*0} \rightarrow D^0\pi^0\), \(D^{*+} \rightarrow D^+\pi^0\), and \(D^{*+} \rightarrow D^0\pi^+\). The maximal difference allowed between the reconstructed mass and the nominal value is 3 MeV, which again corresponds to 3\(\sigma\). For both the \(D\) and \(D^+\) reconstruction, we perform a mass-vertex constrained fit and discard candidates for which this fit fails. We require that no additional charged track be in the event other than the decay products of the \(B_{tag}\), \(D^{(*)}\), the lepton, and the signal’s bachelor pion. Furthermore, we require the lepton and bachelor pion to be positively identified. We require that the pion, lepton and \(D^{(*)}\) meson form a overall charge neutral system with \(B_{tag}\). We also require \(M_{D^{(*)}\pi}\) to be less than 3 \text{ GeV}/c^2 and larger than 2.05 \text{ GeV}/c^2. There is the possibility of signal overlap, i.e. the non-tag final state may be combined in different signal states. This overlap fraction is about 5%. In such cases, we select at most one \(B_{sig}\) candidate per event using two criteria. First, we prefer \(D^*\) over \(D\) in the final state since, otherwise, we would have an extra \(\pi^0\) in the event, leading to additional missing energy. Second, we select the \(D^{(*)}\) whose reconstructed mass is closer to its nominal value. The requirements described above for \(M_{bc}, \Delta E, \sigma_{tag}\), and \(M_{D^{(*)}\pi}\) are determined by maximizing the figure of merit \(S/\sqrt{S+B}\) using MC simulation; here, \(S\) and \(B\) are the signal and background yields, respectively.

### C. Extraction of the branching fraction

The branching fractions are determined by fitting the \(M_{\gamma}^2 = (p_{e^+} + p_{e^-}) - p_{B_{tag}} - p_{D^{(*)}} - p_{\pi} - p_{\nu} - p_{\bar{p}}\)^2 / c^2 spectrum. Here, \((p_{e^+} + p_{e^-})\) is the sum of the four-momenta of the colliding beam particles and the other terms are the four-momenta of the indicated final-state particles. We fit the spectrum with probability density function (PDF) templates derived from simulation to extract the yields; then we determine \(B\), using the ratios of the fitted yields to those in the original MC and the branching fractions used in MC.

The agreement of the simulations with data is checked by comparing the sidebands \([-1 \text{ (GeV/c)^2} < M_{\gamma}^2 < -0.5 \text{ (GeV/c)^2}]\) and \([2 \text{ (GeV/c)^2} < M_{\gamma}^2 < 3.5 \text{ (GeV/c)^2}]\) and the signal region for events that were discarded for
failing to form a charge-neutral system. The reduced \( \chi^2 \), obtained by comparing the difference between data and MC, for these tests is 1.02, showing that the agreement of data and MC is good.

For the channels \( B^+ \rightarrow D^{(*)} \pi^+ \ell^+ \nu \) and \( B^0 \rightarrow D^{(*)} \pi^- \ell^- \nu \), we consider the following components in the MC:

- \( B \rightarrow D \pi \ell \nu \)
- \( B \rightarrow D^* \pi \ell \nu \)
- \( B \rightarrow D^{(*)} \pi \pi \ell \nu \)
- \( B \rightarrow D^{(*)} \ell \nu \)
- \( \text{other } B \bar{B} \)
- \( \text{continuum contributions.} \)

Since \( B \rightarrow D^* \pi \ell \nu \) contributes also as feed-down to \( B \rightarrow D \pi \ell \nu \) with a known ratio, we fit simultaneously the \( B \rightarrow D \pi \ell \nu \) and \( B \rightarrow D^* \pi \ell \nu \) channels. Charged and neutral \( B \) channels are fitted separately.

The simulation sample corresponds to five times the integrated luminosity of the data. With the given statistics, not all templates can be determined precisely enough for a stable fit. We therefore float only the \( B \rightarrow D \pi \ell \nu \), the \( B \rightarrow D^* \pi \ell \nu \) and the continuum yields. The contribution from “other \( B \bar{B} \)” is not small; however, the shape is very similar to the continuum contribution and, given the agreement of the data and simulation in the sidebands, it is reasonable to fix this contribution to the MC prediction. We use a binned extended maximum likelihood fit to extract the yields. The range of the fit is \(-0.3 \text{ (GeV}/c^2)^2 < M_B^2 < 2.0 \text{ (GeV}/c^2)^2 \) with 140 bins for the \( B^+ \rightarrow D^+ \pi^+ \ell^+ \nu \) and \( B^0 \rightarrow D^0 \pi^- \ell^- \nu \) channels. For the \( B^+ \rightarrow D^{*+} \pi^+ \ell^+ \nu \) and \( B^0 \rightarrow D^{*0} \pi^- \ell^- \nu \) channels, we use a range of \(-0.3 \text{ (GeV}/c^2)^2 < M_B^2 < 0.6 \text{ (GeV}/c^2)^2 \) with 54 bins. In the given \( M^2 \) ranges, we select 1566, 438, 3750, and 87 candidates for the \( B^+ \rightarrow D^+ \pi^+ \ell^+ \nu \), \( B^+ \rightarrow D^{*+} \pi^+ \ell^+ \nu \), \( B^0 \rightarrow D^{*0} \pi^- \ell^- \nu \), and \( B^0 \rightarrow D^{*0} \pi^- \ell^- \nu \) channels, respectively. Figure 1 shows the result of the fit to the combined \( B^+ \rightarrow D^+ \pi^+ \ell^+ \nu \) and \( B^+ \rightarrow D^{*+} \pi^+ \ell^+ \nu \) channels and Fig. 2 for the combined \( B^0 \rightarrow D^0 \pi^- \ell^- \nu \) and \( B^0 \rightarrow D^{*0} \pi^- \ell^- \nu \) channels. The \( \chi^2/\text{Ndf} \) value for the \( B^+ \) and \( B^0 \) mode fits is 1.1 and 1.2, respectively. Ndf refers to the number of degrees of freedom in the fit. Since the counts for some entries in the fitted histograms are small, we use the equivalent quantity for Poisson statistics (see, e.g., Eq. (40.16) in Ref. [20]). Tables I and II summarize the fit results.

We check that the fits are unbiased and give the expected uncertainty by fitting ensembles of simulated events generated by sampling from the fitting templates. We plot the resulting residuals, fit them to a normal distribution, and check the mean and standard deviation. Finally, we correct for the fact that our efficiency in

<table>
<thead>
<tr>
<th>Source</th>
<th>yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>( B^+ \rightarrow D^+ \pi^+ \ell^+ \nu )</td>
<td>515 ± 31</td>
</tr>
<tr>
<td>( B^+ \rightarrow D^{*+} \pi^+ \ell^+ \nu )</td>
<td>571 ± 40</td>
</tr>
<tr>
<td>Continuum</td>
<td>444 ± 136</td>
</tr>
<tr>
<td>Other ( B \bar{B} ) (fixed)</td>
<td>360</td>
</tr>
<tr>
<td>Other semi-leptonic ( B ) decays (fixed)</td>
<td>114</td>
</tr>
</tbody>
</table>

TABLE I. Results for the combined fit \( B^+ \rightarrow D^+ \pi^+ \ell^+ \nu \) and \( B^+ \rightarrow D^{*+} \pi^+ \ell^+ \nu \).

\( M_B \) is not constant. Since the shape of \( M_B \), determined by the poorly-known widths and relative branching fractions of the \( D^{*+} \) mesons, might be different in data and simulation, the non-constant efficiency may introduce an overall efficiency difference between data and simulation. We use a quadratic function to fit the efficiency for each channel after determining that higher-order polynomials do not improve the fit quality significantly. Then we determine the shape of \( M_B \) in data by subtracting the background components determined from simulation using the \( B \) determined from our fit to \( M_B^2 \). Comparing the integrated efficiency in data and simulation for the signal, we determine overall-efficiency calibration factors of 1.008 ± 0.007 for \( B^+ \rightarrow D^+ \pi^+ \ell^+ \nu \), 0.983 ± 0.006 for \( B^0 \rightarrow D^0 \pi^- \ell^- \nu \), 0.997 ± 0.002 for \( B^+ \rightarrow D^{*+} \pi^+ \ell^+ \nu \), and 0.98 ± 0.01 for \( B^0 \rightarrow D^{*0} \pi^- \ell^- \nu \).

D. Determination of systematics

There are three main sources of systematic uncertainties for our measurement: uncertainties in the simulation of our detector and underlying physics process, the statistical uncertainties of our fitting templates and the uncertainty of the efficiency correction based on the \( M_B \) shapes in data and MC. For all three of these sources, our strategy to determine the systematic uncertainty is to use a MC approach that is based on running 1000 ensembles of simulated events, where the source of the systematic uncertainty is varied as described below for each source. We check that the refitted branching fraction in question follows a normal distribution and use the standard deviation of this distribution as our systematic uncertainty.

For the uncertainties of the simulation of the detector, we consider the uncertainty in the determination of the
FIG. 1. (Color online) Binned extended maximum likelihood of the MC templates to the data for the combined fit to $B^+ \to D^- \pi^+ \ell^+ \nu$ (left) and $B^+ \to D^+ \pi^+ \ell^+ \nu$ (right). The data is shown with error bars. The legend in the left panel indicates each component in the fit. The dots at the bottom of each panel show the pulls between the data and the fit. For better visibility, we doubled the bin width for this plot.

FIG. 2. (Color online) Binned extended maximum likelihood of the MC templates to the data for the combined fit to $B^0 \to D^0 \pi^- \ell^+ \nu$ (left) and $B^0 \to D^{*0} \pi^- \ell^+ \nu$ (right). The data is shown with error bars. The legend in the left panel indicates each component in the fit. The dots at the bottom of each panel show the pulls between the data and the fit. For better visibility, we doubled the bin width for this plot.

correction factors of the simulation of the PID discussed earlier as well as the uncertainty on the tracking efficiency. Similarly, for the underlying physical processes, we consider the uncertainty of the $D$ and $B$ meson branching fractions and the $D^*$ and $D^{**}$ form factors. Furthermore, we consider the uncertainty of the calibration of the tagging algorithm, the uncertainty on the total number of $B \bar{B}$ pairs, and the uncertainty on the branching fractions of $\Upsilon(4S)$ to $B^+B^-$ and $B^0\bar{B}^0$. These sources of uncertainty of the simulation of the detector and un-
derlying physical processes are described in more detail in Ref. [25]. Since it is reasonable to assume that the sources of uncertainty follow a normal distribution, we draw for each ensemble of simulated events, source, and kinematic bin a new weight from a normal distribution with the corresponding width. This is then used to do an event-by-event weighting of the ensemble of simulated events. The advantage of this method is that correlations among the different sources for uncertainties as well as the dependence on the event kinematics are taken into account. By repeating this exercise while varying only one source at a time, we estimate the relative contributions of each source to the systematics. This decomposition is shown in Tables III and IV. We omit the uncertainties due to the $K_S^0$ efficiencies and the $D^*$ form factors because these are consistent with zero relative to the tabulated uncertainties.

From Tables III and IV, the combined systematic uncertainty on the branching fraction by varying all sources simultaneously are 8.3% for $B^+ \to D^- \pi^+ \ell^+ \nu$, 9.7% for $B^0 \to D^0 \pi^- \ell^+ \nu$, 5.8% for $B^+ \to D^{*-} \pi^+ \ell^+ \nu$, and 7.2% for $B^0 \to D^{*0} \pi^- \ell^+ \nu$.

We estimate the systematic uncertainty propagated from the statistical uncertainty of the fitting templates to be 1.9%, 2.6%, 3.2%, and 3.5% for the $B^+ \to D^- \pi^+ \ell^+ \nu$, $B^+ \to D^{*-} \pi^+ \ell^+ \nu$, $B^0 \to D^0 \pi^- \ell^+ \nu$, and $B^0 \to D^{*0} \pi^- \ell^+ \nu$ channels, respectively. These values are estimated using 10000 ensembles of simulated events for which we vary the templates using Poisson statistics. Finally, the uncertainty on the detector-efficiency dependence of $M_{D^{(*)}\pi}$ is estimated by varying the $M_{D^{(*)}\pi}$ spectrum for each channel within Poisson statistics and adding the difference of the average efficiency between the ±68% boundaries of the fit to the efficiency versus $M_{D^{(*)}\pi}$. The resulting uncertainty propagated to the branching fraction of interest is below 1% for each channel. The final systematic uncertainties on the branching fraction from all sources discussed above correspond to 8.6% for $B^+ \to D^- \pi^+ \ell^+ \nu$, 6.4% for $B^+ \to D^{*-} \pi^+ \ell^+ \nu$, 10.3% for $B^0 \to D^0 \pi^- \ell^+ \nu$, and 8.0% for $B^0 \to D^{*0} \pi^- \ell^+ \nu$.

### IV. RESULTS AND CONCLUSION

Using the combined fits, including the correction and systematics from the $M_{D^{(*)}\pi}$ efficiency, simulation uncertainties and statistical uncertainty of the templates, we obtain the following values for the branching fractions:

- $\mathcal{B}(B^+ \to D^- \pi^+ \ell^+ \nu) = [4.55 \pm 0.27 \text{ (stat.)} \pm 0.39 \text{ (syst.)}] \times 10^{-3}$,
- $\mathcal{B}(B^0 \to D^0 \pi^- \ell^+ \nu) = [4.05 \pm 0.36 \text{ (stat.)} \pm 0.41 \text{ (syst.)}] \times 10^{-3}$,
- $\mathcal{B}(B^+ \to D^{*-} \pi^+ \ell^+ \nu) = [6.03 \pm 0.43 \text{ (stat.)} \pm 0.38 \text{ (syst.)}] \times 10^{-3}$,
- $\mathcal{B}(B^0 \to D^{*0} \pi^- \ell^+ \nu) = [6.46 \pm 0.53 \text{ (stat.)} \pm 0.52 \text{ (syst.)}] \times 10^{-3}$.

These are within one standard deviation of the current world-average values [20] with the exception of $B^0 \to D^{*0} \pi^- \ell^+ \nu$, which deviates by $1.7\sigma$. These supersede the previous Belle result [11]. The total uncertainties on our measurement are slightly better than the current world-average for the channels $B^0 \to D^0 \pi^- \ell^+ \nu$ and $B^0 \to D^{*0} \pi^- \ell^+ \nu$, whereas they are the same for the channels $B^+ \to D^- \pi^+ \ell^+ \nu$ and $B^+ \to D^{*-} \pi^+ \ell^+ \nu$. A potential extension to this work would be to confirm the recent observation of $B \to D^{(*)} \pi \pi \ell \nu$ by BaBar [26] as well as to analyze the $M_{D^{(*)}\pi}$ distribution to extract the branching fractions and widths of the different $D^*$ mesons. Here, there are still some discrepancies between the Belle [11] and BaBar [13] measurements.
ACKNOWLEDGMENTS

We thank the KEKB group for the excellent operation of the accelerator; the KEK cryogenics group for the efficient operation of the solenoid; and the KEK computer group, the National Institute of Informatics, and the Pacific Northwest National Laboratory (PNNL) Environmental Molecular Sciences Laboratory (EMSL) computing group for valuable computing and Science Information NETwork 5 (SINET5) network support. We acknowledge support from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan, the Japan Society for the Promotion of Science (JSPS), and the Tau-Lepton Physics Research Center of Nagoya University; the Australian Research Council; Austrian Science Fund under Grant No. P 26794-N20; the National Natural Science Foundation of China under Contracts No. 11435013, No. 11475166, No. 11575017, No. 11675187, No. 11705209; Key Research Program of Frontier Sciences, Chinese Academy of Sciences (CAS), Grant No. QYZDJ-SSW-SLH011; the CAS Program of Frontier Sciences, Chinese Academy of Sciences (CAS), Grant No. QYZDJ-SSW-SLH011; the Ministry of Education, Youth and Sports of the Czech Republic under Contract No. LTT17020; the Carl Zeiss Foundation, the Deutsche Forschungsgemeinschaft, the Excellence Cluster Universe, and the VolkswagenStiftung; the Department of Science and Technology of India; the Istituto Nazionale di Fisica Nucleare of Italy; National Research Foundation (NRF) of Korea Grants No. 2014R1A2A2A01005286, No.2015R1A2A2A01003280, No. 2015H1A2A1033649, No. 2016R1D1A1B01010135, No. 2016K1A3A7A9005 603, No. 2016R1D1A1B02012900; Radiation Science Research Institute, Foreign Large-size Research Facility Application Supporting project and the Global Science Experimental Data Hub Center of the Korea Institute of Science and Technology Information; the Polish Ministry of Science and Higher Education and the National Science Center; the Ministry of Education and Science of the Russian Federation and the Russian Foundation for Basic Research; the Slovenian Research Agency; Ikerbasque, Basque Foundation for Science, Basque Government (No. IT956-16) and Ministry of Economy and Competitiveness (MINECO) (Juan de la Cierva), Spain; the Swiss National Science Foundation; the Ministry of Education and the Ministry of Science and Technology of Taiwan; and the United States Department of Energy and the National Science Foundation.