PROCEEDINGS OF SCIENCE

New $B \to D^{(*)}\tau\nu$ result from Belle

Thomas Kuhr* **for the Belle Collaboration**
Ludwig-Maximilians-Universität München
E-mail: Thomas.Kuhr@lmu.de

We present an update measurement of the branching ratio of $B \to D^{(*)}\tau\nu$ relative to $B \to D^{(*)}\ell\nu$ with the full Belle dataset. A hadronic tagging method is employed to reconstruct the second B meson in the $Y(4S)$ event, which provides sufficient information to determine the invariant mass of all undetected particles. While this provides a good separation of signal and normalization modes, additional information is used to separate the signal from backgrounds with multiple undetected particles.

Flavor Physics & CP Violation 2015
May 25-29, 2015
Nagoya, Japan

*Speaker.

PoS(FPCP2015)015

1. Introduction

The ratios

$$R(D) = \frac{\mathscr{B}(\bar{B} \to D\tau^-\bar{\nu}_\tau)}{\mathscr{B}(\bar{B} \to D\ell^-\bar{\nu}_\ell)} \tag{1.1}$$

and

$$R(D^*) = \frac{\mathscr{B}(\bar{B} \to D^*\tau^-\bar{\nu}_\tau)}{\mathscr{B}(\bar{B} \to D^*\ell^-\bar{\nu}_\ell)}, \tag{1.2}$$

with $\ell = e$ or μ are sensitive to new physics contributions. In particular a charged Higgs boson as predicted in two-Higgs doublet models (2HDM) [1] could lead to measurable deviations of the branching fraction or kinematic distributions from the standard model (SM) expectation. The first observation of an exclusive semitauonic B decay was reported by Belle in 2007 [2] and subsequent measurements by BaBar and Belle [3] reported branching fractions above—yet consistent with—the SM predictions. A significant deviation of the $R(D)$ and $R(D^*)$ values from the SM prediction was reported by BaBar in 2012 [4]. This called for an independent confirmation by the LHCb [5] and Belle collaborations. This article presents a new measurement of $R(D)$ and $R(D^*)$ with the full Belle $\Upsilon(4S) \to B\bar{B}$ dataset of 711 fb^{-1} [6].

2. Reconstruction and Simulation

Candidates of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays are reconstructed in four samples with different D meson states: D^+ mesons are reconstructed in the decays to $K^-\pi^+\pi^+$, $K^0_S\pi^+$, $K^0_S\pi^+\pi^0$, and $K^0_S\pi^+\pi^+\pi^-$; D^0 mesons to $K^-\pi^+$, $K^-\pi^+\pi^+\pi^-$, $K^-\pi^+\pi^0$, $K^0_S\pi^0$, and $K^0_S\pi^+\pi^-$; D^{*+} mesons to $D^0\pi^+$ and $D^+\pi^0$; and D^{*0} mesons to $D^0\pi^0$ and $D^0\gamma$. The tau lepton is reconstructed in the leptonic modes $\tau^- \to e^-\bar{\nu}_e\nu_\tau$ and $\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$ so that the signal and normalization modes have the same detectable final state particles. This reduces systematic uncertainties in the ratio, but also requires a technique to distinguish the signal from the normalization mode.

We exploit the kinematics of $e^+e^- \to \Upsilon(4S) \to B\bar{B}$ by reconstructing the accompanying B meson, B_{tag}, in a hadronic decay mode [7]. In case of a correct reconstruction of both B mesons in the event the only remaining undetected particles are neutrinos. We require no further tracks or π^0 candidates in the event. The invariant mass squared of all undetected particles,

$$M^2_{\text{miss}} = (p_{\text{beam}} - p_{\text{tag}} - p_{D^{(*)}} - p_\ell)^2/c^2 \tag{2.1}$$

where p_{beam}, p_{tag}, $p_{D^{(*)}}$, and p_ℓ are the four-momenta of the colliding beam particles, the B_{tag} candidate, and the reconstructed signal-B daughters, respectively, peaks at zero for the normalization mode where only one neutrino is not reconstructed. The M^2_{miss} distribution of the signal is broader and peaks at higher values because of three undetected neutrinos and can therefore be well distinguished from $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays. Normalization mode decays are suppressed by requiring $q^2 = (p_B - p_{D^{(*)}})^2 > 4\ \text{GeV}^2/c^2$.

Samples of simulated signal and background events are used to optimize the selection and to determine shapes and (relative) yields for the fit described in Section 4. Corrections for the yields and kinematic distributions of correctly and wrongly identified B_{tag} mesons, fake leptons, D mesons, decays to higher excitation D meson states (D^{**}), and $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays are applied to get a satisfactory agreement between simulation and data.

3. Data Composition

We identify the following components in the selected data samples:

- The *lepton normalization* mode of correctly reconstructed $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays has a freely floating yield in the fit.
- The *lepton cross-feed* component originates from $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ decays where the π^0 or photon from the D^* decay is not reconstructed. Its $M^2_{\rm miss}$ distribution is shifted to higher values and broader than that of the lepton normalization, but still well distinguishable from the $\bar{B} \to D\tau^-\bar{\nu}_\tau$ decays. The yield of this component in the $D\ell^-$ samples is a free parameter of the fit.
- The yield of the correctly identified *tau signal*, $Y^{D^{+,0}\ell^-}_{\tau\,\rm signal}$, is given in the $D\ell^-$ samples by

$$Y^{D^{+,0}\ell^-}_{\tau\,\rm signal} = R(D) Y^{D^{+,0}\ell^-}_{\ell\,\rm norm} / (2 f^{D^{+,0}}), \tag{3.1}$$

  where $R(D)$ is the parameter of interest and floating in the fit, $Y^{D^{+,0}\ell^-}_{\ell\,\rm norm}$ is the yield of the lepton normalization, and $f^{D^{+,0}}$ is the efficiency ratio between lepton normalization and tau signal determined from simulation. A similar equation is used for the tau signal yield in the $D^*\ell^-$ samples with $R(D^*)$ as free parameter and taking the cross-feed components into account.
- The *tau cross-feed* where the π^0 or photon from the D^* of an otherwise correctly reconstructed $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decay is missed is constrained relative to the tau signal using the ratio of yields of lepton cross-feed to lepton normalization and a factor that accounts for the difference in the cross-feed ratios between light and tau leptons and is determined from simulation.
- Compared to the cross-feed mentioned above it rarely happens that a charged pion from a $D^{*+} \to D^0\pi^+$ decay is not reconstructed. The yield of this component is constrained relative to the lepton normalization in the $D^{*+}\ell^-$ sample.
- The yield of combinatorial background D and D^* candidates is determined from sidebands in the D mass and $D^* - D$ mass difference, respectively.
- The D^{**} *background* component consists of $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell(\nu_\tau\bar{\nu}_\tau)$ decays where one or more pions from the D^{**} decay are not reconstructed. As the knowledge of the branching fractions of these decays is limited their yield cannot be determined reliably from simulation and has to be a free parameter of the fit. Because the $M^2_{\rm miss}$ distribution of this background is similar to that of the signal a further observable is considered in the fit as described in the next section.
- The probability to misidentify a hadron as a lepton is low and well described by simulation. The yield of the fake lepton component is fixed in the fit.
- Decays of $\bar{B} \to D^{(*)}D_s^-$ with $D_s^- \to \ell^-\bar{\nu}_\ell(\nu_\tau\bar{\nu}_\tau)$ have the same signature as the signal or normalization mode. As the corresponding D_s^- branching fractions are well known the small yield of this component is fixed in the fit.

- A few remaining background events do not fit into the categories above. The yield of this component is taken from simulation.

4. Fit

As explained above, the lepton normalization, the lepton cross-feed, and the tau signal components can be statistically separated by a fit of the $M^2_{\rm miss}$ distribution. But other backgrounds, in particular the D^{**} background, have a $M^2_{\rm miss}$ shape similar to the signal. Therefore the sample is split at $M^2_{\rm miss} = 0.85$ GeV$^2/c^4$. The $M^2_{\rm miss}$ distribution is fitted in the the low $M^2_{\rm miss}$ region with smoothed histogram probability density functions. A neural network is trained to separate tau signal from background at high $M^2_{\rm miss}$, mainly the D^{**} background, for events with $M^2_{\rm miss} > 0.85$ GeV$^2/c^4$. The most powerful input variable is $E_{\rm ECL}$, the sum of the energies of clusters in the calorimeter not assigned to any of the reconstructed daughter particles of the two B mesons. While the $E_{\rm ECL}$ distribution peaks at zero for signal it has a broader shape for background where pions are missed in the reconstruction. Further neural network input variables are $M^2_{\rm miss}$, q^2, the lepton momentum in the center-of-mass frame (p^*_ℓ), the number of additional π^0 candidates with loose selection criteria, the cosine of the angle between the momentum and vertex displacement of the $D^{(*)}$ meson, and the decay-channel identifiers of the B and $D^{(*)}$ mesons. A monotonic transformation of the network output provides the $o'_{\rm NB}$ value whose distribution is parameterized with a bi-furcated Gaussian for each component. The relative yields in the low and high $M^2_{\rm miss}$ regions of each component are determined from simulation.

The 12 free parameters (lepton normalization yield in the four $D^{(*)}\ell^-$ samples, lepton cross-feed yield in the two $D\ell^-$ samples, D^{**} background yield in the four $D^{(*)}\ell^-$ samples, and the parameters $R(D)$ and $R(D^*)$) are determined in a simultaneous unbinned maximum likelihood fit assuming isospin symmetry. The fit result describes the data well as can be seen in Figures 1 to 5.

5. Cross Checks and Systematic Uncertainties

The analysis procedure is checked with a cross validation on simulated data and pseudo-experiments. It is verified that the choice of $M^2_{\rm miss}$ split value does not bias the result. The $M^2_{\rm miss}$ resolution model is checked on a $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ enriched sample. A sample enriched in $B \to D^{**}\ell^-\bar{\nu}_\ell$ decays is selected by requiring an additional π^0 in the signal B meson reconstruction. Fits of $M^2_{\rm miss}$, the missing mass squared calculated without the additional π^0, $E_{\rm ECL}$, and p^*_ℓ give consistent yields in each of the four $D^{(*)}\ell^-\pi^0$ samples and indicate that the simulation describes the data in the tested dimensions.

The dominant systematic uncertainty comes from the limited knowledge of the D^{**} background. It includes uncertainties on the decay model and branching fractions. Other significant contributions to the systematic uncertainty of $R(D)$ and $R(D^*)$ come from efficiency ratios used in the fit and determined from simulation. The systematic effect of the fit template shapes is evaluated by using alternative models and the largest effect is observed for $o'_{\rm NB}$ on $R(D)$. Small systematic error contributions come from the yields fixed in the fit and the lepton identification efficiency. The total systematic uncertainty is 7.1% for $R(D)$ and 5.2% for $R(D^*)$ with a correlation of -0.32 determined mainly with pseudo-experiments.

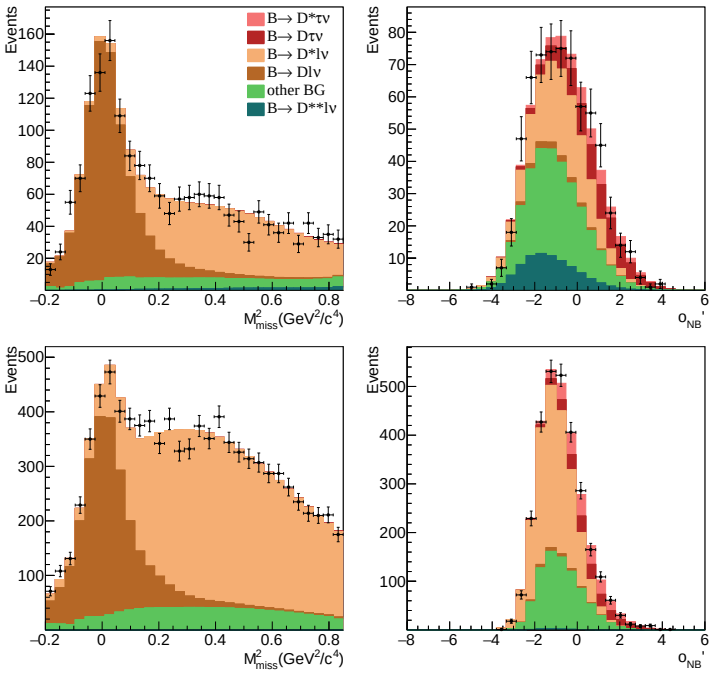

Figure 1: Fit projections and data points with statistical uncertainties in the $D^+\ell^-$ (top) and $D^0\ell^-$ (bottom) data samples. Left: $M^2_{\rm miss}$ distribution for $M^2_{\rm miss} < 0.85$ GeV$^2/c^4$; right: $o'_{\rm NB}$ distribution for $M^2_{\rm miss} > 0.85$ GeV$^2/c^4$.

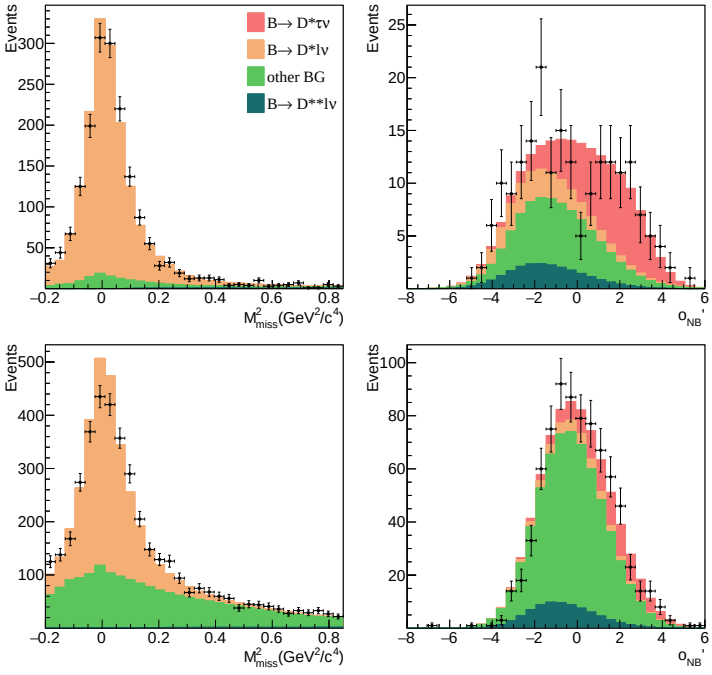

Figure 2: Fit projections and data points with statistical uncertainties in the $D^{*+}\ell^-$ (top) and $D^{*0}\ell^-$ (bottom) data samples. Left: $M^2_{\rm miss}$ distribution for $M^2_{\rm miss} < 0.85$ GeV$^2/c^4$; right: $o'_{\rm NB}$ distribution for $M^2_{\rm miss} > 0.85$ GeV$^2/c^4$.

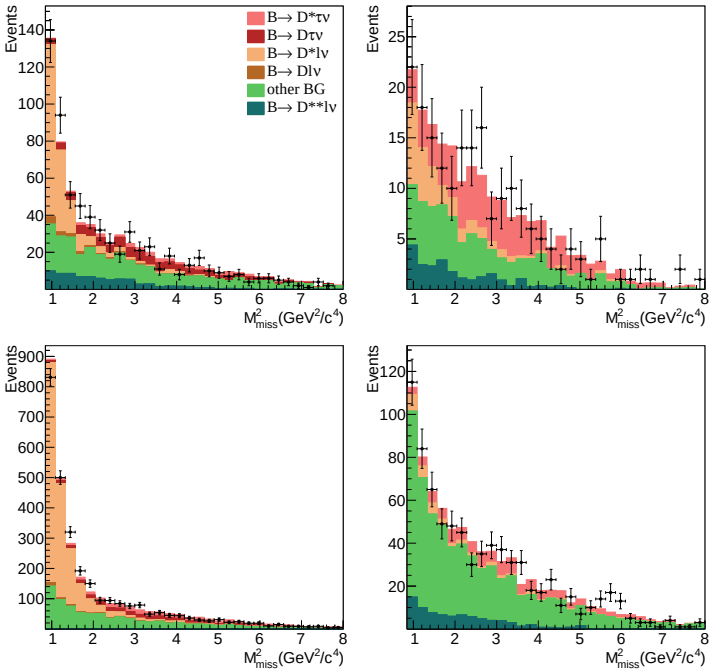

Figure 3: Projections of the fit results and data points with statistical uncertainties for the high M^2_{miss} region. Top left: $D^+\ell^-$; top right: $D^{*+}\ell^-$; bottom left: $D^0\ell^-$; bottom right: $D^{*0}\ell^-$.

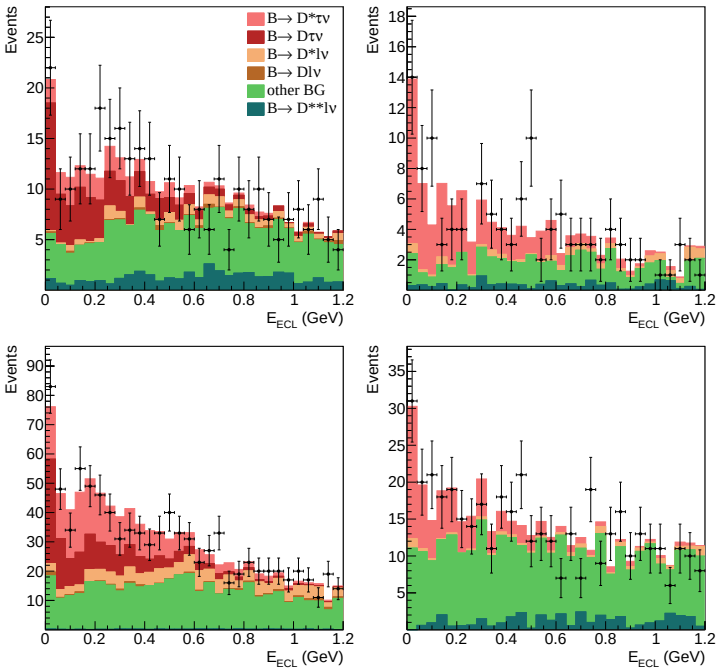

Figure 4: Projections of the fit results and data points with statistical uncertainties in a signal-enhanced region of $M^2_{\text{miss}} > 2.0$ GeV$^2/c^4$in the E_{ECL} dimension. Top left: $D^+\ell^-$; top right: $D^{*+}\ell^-$; bottom left: $D^0\ell^-$; bottom right: $D^{*0}\ell^-$.

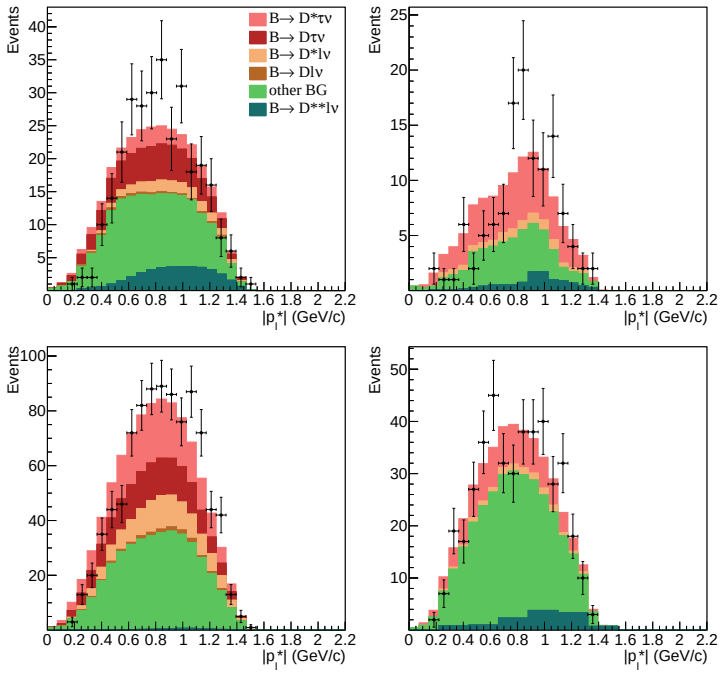

Figure 5: Projections of the fit results and data points with statistical uncertainties in a signal-enhanced region of $M^2_{\rm miss} > 2.0$ GeV$^2/c^4$in the p^*_ℓ dimension. Top left: $D^+\ell^-$; top right: $D^{*+}\ell^-$; bottom left: $D^0\ell^-$; bottom right: $D^{*0}\ell^-$.

6. Result and Discussion

The obtained relative branching fractions are

$$R(D) = 0.375 \pm 0.064(\text{stat.}) \pm 0.026(\text{syst.})$$
$$R(D^*) = 0.293 \pm 0.038(\text{stat.}) \pm 0.015(\text{syst.}) \,.$$

Figure 6 shows the result with uncertainties in the $R(D)$-$R(D^*)$ plane. The SM value and the BaBar result, both taken from Ref. [4], are within the 2σ contour. Although our measurement is consistent with the BaBar result and deviates from the SM prediction in the same direction it does not show a significant excess over the SM expectation.

To assess the compatibility of our data with a 2HDM of type II [1] we repeat the analysis assuming a $\tan\beta/m_{H^+}$ value of 0.5 c^2/GeV. As can be seen in Fig. 6 (right) the measured values of $R(D)$ and $R(D^*)$ are consistent with the predictions under this assumption. Also the background-subtracted and efficiency-corrected q^2 distributions shown in Fig. 7 agree with the expectation from the SM and the 2HDM of type II with $\tan\beta/m_{H^+} = 0.5$ c^2/GeV.

References

[1] M. Tanaka, Z. Phys. C **67** (1995) 321.

[2] A. Matyja *et al.* (Belle Collaboration), Phys. Rev. Lett. **99** (2007) 191807.

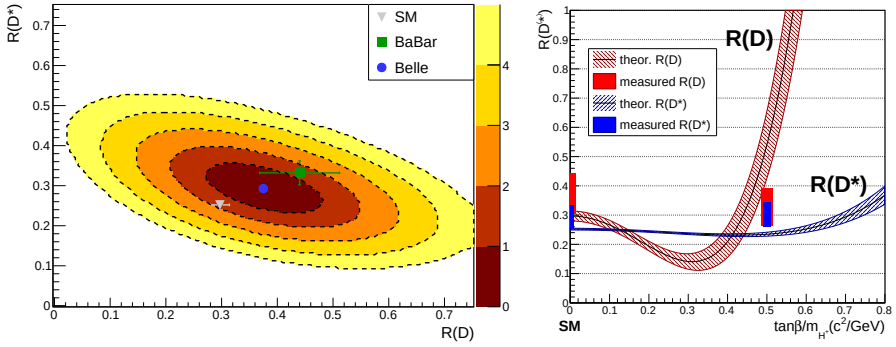

Figure 6: Left: Exclusion level of $R(D)$-$R(D^*)$ value assumptions in standard deviations, systematic uncertainties included. Right: Theoretical predictions with 1σ error ranges for $R(D)$ (red) and $R(D^*)$ (blue) for different values of $\tan\beta/m_{H^+}$ in the 2HDM of type II. This analysis' fit results for $\tan\beta/m_{H^+} = 0.5\ c^2/\text{GeV}$ and SM are shown with their 1σ ranges as red and blue bars with arbitrary width for better visibility.

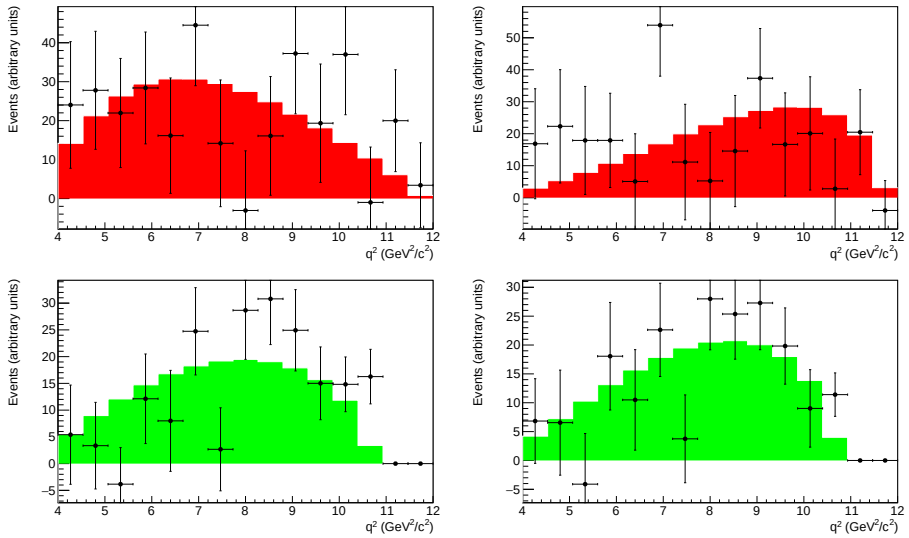

Figure 7: Background-subtracted q^2 distributions of the τ signal in the region of $M^2_{\text{miss}} < 0.85\ \text{GeV}^2/c^4$. The distributions are efficiency corrected and normalized to the fitted yield. The error bars show the statistical uncertainties. The histogram is the respective expected distribution from signal MC. Left: Standard Model result, right: Type-II 2HDM result with $\tan\beta/m_{H^+} = 0.5\ c^2/\text{GeV}$, top: $\bar{B} \to D\tau^-\bar{\nu}_\tau$, bottom: $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$.

[3] B. Aubert *et al.* (BaBar Collaboration), Phys. Rev. Lett. **100** (2008) 021801; I. Adachi *et al.* (Belle Collaboration), arXiv:0910.4301 [hep-ex]; A. Bozek *et al.* (Belle Collaboration), Phys. Rev. D **82** (2010) 072005.

[4] J. P. Lees *et al.* (BaBar Collaboration), Phys. Rev. Lett. **109** (2012) 101802; J. P. Lees *et al.* (BaBar Collaboration), Phys. Rev. D **88** (2013) 072012.

[5] R. Aaij *et al.* (LHCb Collaboration), arXiv:1506.08614 [hep-ex].

[6] M. Huschle *et al.* (Belle Collaboration), arXiv:1507.03233 [hep-ex].

[7] M. Feindt *et al.*, Nucl. Instr. and Meth. A **654** (2011) 1.