



BELLE2-CONF-PH-2022-017
October 9, 2022

Determination of $|V_{ub}|$ from untagged $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays using 2019-2021 Belle II data

(The Belle II Collaboration)

F. Abudinén, I. Adachi, K. Adamczyk, L. Aggarwal, P. Ahlburg, H. Ahmed, J. K. Ahn, H. Aihara, N. Akopov, A. Aloisio, F. Ameli, L. Andricsek, N. Anh Ky, D. M. Asner, H. Atmacan, V. Aulchenko, T. Aushev, V. Aushev, T. Aziz, V. Babu, S. Bacher, H. Bae, S. Baehr, S. Bahinipati, A. M. Bakich, P. Bambade, Sw. Banerjee, S. Bansal, M. Barrett, G. Batignani, J. Baudot, M. Bauer, A. Baur, A. Beaubien, A. Beaulieu, J. Becker, P. K. Behera, J. V. Bennett, E. Bernieri, F. U. Bernlochner, V. Bertacchi, M. Bertemes, E. Bertholet, M. Bessner, S. Bettarini, V. Bhardwaj, B. Bhuyan, F. Bianchi, T. Bilka, S. Bilokin, D. Biswas, A. Bobrov, D. Bodrov, A. Bolz, A. Bondar, G. Bonvicini, A. Bozek, M. Bračko, P. Branchini, N. Braun, R. A. Briere, T. E. Browder, D. N. Brown, A. Budano, L. Burmistrov, S. Bussino, M. Campajola, L. Cao, G. Casarosa, C. Cecchi, D. Červenkov, M.-C. Chang, P. Chang, R. Cheaib, P. Cheema, V. Chekelian, C. Chen, Y. Q. Chen, Y. Q. Chen, Y.-T. Chen, B. G. Cheon, K. Chilikin, K. Chirapatpimol, H.-E. Cho, K. Cho, S.-J. Cho, S.-K. Choi, S. Choudhury, D. Cinabro, L. Corona, L. M. Cremaldi, S. Cunliffe, T. Czank, S. Das, N. Dash, F. Dattola, E. De La Cruz-Burelo, S. A. De La Motte, G. de Marino, G. De Nardo, M. De Nuccio, G. De Pietro, R. de Sangro, B. Deschamps, M. Destefanis, S. Dey, A. De Yta-Hernandez, R. Dhamija, A. Di Canto, F. Di Capua, S. Di Carlo, J. Dingfelder, Z. Doležal, I. Domínguez Jiménez, T. V. Dong, M. Dorigo, K. Dort, D. Dossett, S. Dreyer, S. Dubey, S. Duell, G. Dujany, P. Ecker, S. Eidelman, M. Eliachevitch, D. Epifanov, P. Feichtinger, T. Ferber, D. Ferlewicz, T. Fillinger, C. Finck, G. Finocchiaro, P. Fischer, K. Flood, A. Fodor, F. Forti, A. Frey, M. Friedl, B. G. Fulsom, M. Gabriel, A. Gabrielli, N. Gabyshev, E. Ganiev, M. Garcia-Hernandez, R. Garg, A. Garmash, V. Gaur, A. Gaz, U. Gebauer, A. Gellrich, J. Gemmler, T. Geßler, G. Ghevodnyan, G. Giakoustidis, R. Giordano, A. Giri, A. Glazov, B. Gobbo, R. Godang, P. Goldenzweig, B. Golob, P. Gomis, G. Gong, P. Grace, W. Gradl, S. Granderath, E. Graziani, D. Greenwald, T. Gu, Y. Guan, K. Gudkova, J. Guilliams, C. Hadjivasiliou, S. Halder, K. Hara, T. Hara, O. Hartbrich, K. Hayasaka, H. Hayashii, S. Hazra, C. Hearty, M. T. Hedges, I. Heredia de la Cruz, M. Hernández Villanueva, A. Hershenhorn, T. Higuchi, E. C. Hill, H. Hirata, M. Hoek, M. Hohmann, S. Hollitt, T. Hotta, C.-L. Hsu, K. Huang, T. Humair, T. Iijima, K. Inami, G. Inguglia, N. Ipsita, J. Irakkathil Jabbar, A. Ishikawa, S. Ito, R. Itoh, M. Iwasaki, Y. Iwasaki, S. Iwata, P. Jackson, W. W. Jacobs, D. E. Jaffe, E.-J. Jang, M. Jeandron, H. B. Jeon, Q. P. Ji, S. Jia, Y. Jin, C. Joo, K. K. Joo, H. Junkerkalefeld, I. Kadenko, J. Kahn, H. Kakuno, M. Kaleta, A. B. Kaliyar, J. Kandra, K. H. Kang, S. Kang, P. Kapusta, R. Karl, G. Karyan, Y. Kato, H. Kawai, T. Kawasaki, C. Ketter, H. Kichimi, C. Kiesling, C.-H. Kim, D. Y. Kim, H. J. Kim, K.-H. Kim, K. Kim,

S.-H. Kim, Y.-K. Kim, Y. Kim, T. D. Kimmel, H. Kindo, K. Kinoshita, C. Kleinwort, B. Knysh, P. Kodyš, T. Koga, S. Kohani, K. Kojima, I. Komarov, T. Konno, A. Korobov, S. Korpar, N. Kovalchuk, E. Kovalenko, R. Kowalewski, T. M. G. Kraetzschmar, F. Krinner, P. Križan, R. Kroeger, J. F. Krohn, P. Krokovny, H. Krüger, W. Kuehn, T. Kuhr, J. Kumar, M. Kumar, R. Kumar, K. Kumara, T. Kumita, T. Kunigo, M. Künzel, S. Kurz, A. Kuzmin, P. Kvasnička, Y.-J. Kwon, S. Lacaprara, Y.-T. Lai, C. La Licata, K. Lalwani, T. Lam, L. Lanceri, J. S. Lange, M. Laurenza, K. Lautenbach, P. J. Laycock, R. Leboucher, F. R. Le Diberder, I.-S. Lee, S. C. Lee, P. Leitl, D. Levit, P. M. Lewis, C. Li, L. K. Li, S. X. Li, Y. B. Li, J. Libby, K. Lieret, J. Lin, Z. Liptak, Q. Y. Liu, Z. A. Liu, D. Liventsev, S. Longo, A. Loos, A. Lozar, P. Lu, T. Lueck, F. Luetticke, T. Luo, C. Lyu, C. MacQueen, M. Maggiora, R. Maiti, S. Maity, R. Manfredi, E. Manoni, A. Manthei, S. Marcello, C. Marinas, L. Martel, A. Martini, L. Massaccesi, M. Masuda, T. Matsuda, K. Matsuoka, D. Matvienko, J. A. McKenna, J. McNeil, F. Meggendorfer, F. Meier, M. Merola, F. Metzner, M. Milesi, C. Miller, K. Miyabayashi, H. Miyake, H. Miyata, R. Mizuk, K. Azmi, G. B. Mohanty, N. Molina-Gonzalez, S. Moneta, H. Moon, T. Moon, J. A. Mora Grimaldo, T. Morii, H.-G. Moser, M. Mrvar, F. J. Müller, Th. Muller, G. Muroyama, C. Murphy, R. Mussa, I. Nakamura, K. R. Nakamura, E. Nakano, M. Nakao, H. Nakayama, H. Nakazawa, A. Narimani Charan, M. Naruki, Z. Natkaniec, A. Natochii, L. Nayak, M. Nayak, G. Nazaryan, D. Neverov, C. Niebuhr, M. Niiyama, J. Ninkovic, N. K. Nisar, S. Nishida, K. Nishimura, M. H. A. Nouxman, K. Ogawa, S. Ogawa, S. L. Olsen, Y. Onishchuk, H. Ono, Y. Onuki, P. Oskin, F. Otani, E. R. Oxford, H. Ozaki, P. Pakhlov, G. Pakhlova, A. Paladino, T. Pang, A. Panta, E. Paoloni, S. Pardi, K. Parham, H. Park, S.-H. Park, B. Paschen, A. Passeri, A. Pathak, S. Patra, S. Paul, T. K. Pedlar, I. Peruzzi, R. Peschke, R. Pestotnik, F. Pham, M. Piccolo, L. E. Piilonen, G. Pinna Angioni, P. L. M. Podesta-Lerma, T. Podobnik, S. Pokharel, L. Polat, V. Popov, C. Praz, S. Prell, E. Prencipe, M. T. Prim, M. V. Purohit, H. Purwar, N. Rad, P. Rados, S. Raiz, A. Ramirez Morales, R. Rasheed, N. Rauls, M. Reif, S. Reiter, M. Remnev, I. Ripp-Baudot, M. Ritter, M. Ritzert, G. Rizzo, L. B. Rizzuto, S. H. Robertson, D. Rodríguez Pérez, J. M. Roney, C. Rosenfeld, A. Rostomyan, N. Rout, M. Rozanska, G. Russo, D. Sahoo, Y. Sakai, D. A. Sanders, S. Sandilya, A. Sangal, L. Santelj, P. Sartori, Y. Sato, V. Savinov, B. Scavino, M. Schnepf, M. Schram, H. Schreeck, J. Schueler, C. Schwanda, A. J. Schwartz, B. Schwenker, M. Schwickardi, Y. Seino, A. Selce, K. Senyo, I. S. Seong, J. Serrano, M. E. Sevier, C. Sfienti, V. Shebalin, C. P. Shen, H. Shibuya, T. Shillington, T. Shimasaki, J.-G. Shiu, B. Shwartz, A. Sibidanov, F. Simon, J. B. Singh, S. Skambraks, J. Skorupa, K. Smith, R. J. Sobie, A. Soffer, A. Sokolov, Y. Soloviev, E. Solovieva, S. Spataro, B. Spruck, M. Starič, S. Stefkova, Z. S. Stottler, R. Stroili, J. Strube, J. Stypula, Y. Sue, R. Sugiura, M. Sumihama, K. Sumisawa, T. Sumiyoshi, W. Sutcliffe, S. Y. Suzuki, H. Svidras, M. Tabata, M. Takahashi, M. Takizawa, U. Tamponi, S. Tanaka, K. Tanida, H. Tanigawa, N. Taniguchi, Y. Tao, P. Taras, F. Tenchini, R. Tiwary, D. Tonelli, E. Torassa, N. Toutounji, K. Trabelsi, I. Tsaklidis, T. Tsuboyama, N. Tsuzuki, M. Uchida, I. Ueda, S. Uehara, Y. Uematsu, T. Ueno, T. Uglov, K. Unger, Y. Unno, K. Uno, S. Uno, P. Urquijo, Y. Ushiroda, Y. V. Usov, S. E. Vahsen, R. van Tonder, G. S. Varner, K. E. Varvell, A. Vinokurova, L. Vitale, V. Vobbilisetti, V. Vorobyev, A. Vossen, B. Wach, E. Waheed, H. M. Wakeling, K. Wan, W. Wan Abdullah, B. Wang, C. H. Wang,

E. Wang, M.-Z. Wang, X. L. Wang, A. Warburton, M. Watanabe, S. Watanuki, J. Webb, S. Wehle, M. Welsch, C. Wessel, J. Wiechczynski, P. Wieduwilt, H. Windel, E. Won, L. J. Wu, X. P. Xu, B. D. Yabsley, S. Yamada, W. Yan, S. B. Yang, H. Ye, J. Yelton, J. H. Yin, M. Yonenaga, Y. M. Yook, K. Yoshihara, T. Yoshinobu, C. Z. Yuan, Y. Yusa, L. Zani, Y. Zhai, J. Z. Zhang, Y. Zhang, Y. Zhang, Z. Zhang, V. Zhilich, J. Zhou, Q. D. Zhou, X. Y. Zhou, V. I. Zhukova, V. Zhulanov, and R. Žlebčík

(The Belle II Collaboration)

Abstract

We present an analysis of the charmless semileptonic decay $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$, where $\ell = e, \mu$, from 198.0 million pairs of $B\bar{B}$ mesons recorded by the Belle II detector at the SuperKEKB electron-positron collider. The decay is reconstructed without identifying the partner B meson. The partial branching fractions are measured independently for $B^0 \rightarrow \pi^- e^+ \nu_e$ and $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ as functions of q^2 (momentum transfer squared), using 3896 $B^0 \rightarrow \pi^- e^+ \nu_e$ and 5466 $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ decays. The total branching fraction is found to be $(1.426 \pm 0.056 \pm 0.125) \times 10^{-4}$ for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays, where the uncertainties are statistical and systematic, respectively. By fitting the measured partial branching fractions as functions of q^2 , together with constraints on the nonperturbative hadronic contribution from lattice QCD calculations, the magnitude of the Cabibbo-Kobayashi-Maskawa matrix element V_{ub} , $(3.55 \pm 0.12 \pm 0.13 \pm 0.17) \times 10^{-3}$, is extracted. Here, the first uncertainty is statistical, the second is systematic and the third is theoretical.

1. INTRODUCTION

The Cabibbo-Kobayashi-Maskawa (CKM) matrix elements are fundamental parameters of the Standard Model (SM) of particle physics [1]. The magnitude of the matrix element V_{ub} can be determined by measuring the differential decay rate of $B \rightarrow X_u \ell \nu$ events, which is proportional to $|V_{ub}|^2$. Here, X_u is a charmless hadronic final state and ℓ is a light charged lepton. One method to measure $|V_{ub}|$ is inclusive, where no specific X_u final state is reconstructed, but rather the sum of all possible final states is analysed. This is in contrast to the exclusive method in which a specific final state is studied. The two methods have complementary uncertainties introduced by theoretical QCD descriptions. In inclusive decays, these descriptions involve the calculation of the total semileptonic rate, while in the exclusive case they take the form of parameterizations of the low-energy strong interactions (form factors). However, the results for $|V_{ub}|$ obtained by the two methods differ significantly [2]. In order to resolve this disagreement, further measurements using novel approaches are beneficial.

The decay $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ (with charge conjugation implied throughout) is experimentally and theoretically the most reliable mode to measure $|V_{ub}|$ through an exclusive channel at the B Factories. Here, we present a study of this decay using data from the Belle II detector located at the SuperKEKB electron-positron collider at KEK, in Japan. The signal decay is reconstructed from decays of B^0 mesons in $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B^0 \bar{B}^0$ events. The reconstruction method employed here is called untagged since the signal lepton and pion candidates are selected without prior reconstruction (tagging) of the partner B meson. This leads to a high signal efficiency, but also a low purity due to higher combinatorial background from the partner B meson and increased backgrounds from e^+e^- collisions that produce light quark pair (continuum) events.

The events are separated into six disjoint intervals (bins) of squared momentum transfer from the B meson to the pion, q^2 . After selecting signal events and suppressing backgrounds, the signal yields are extracted from an extended likelihood fit to the binned two-dimensional distribution of the energy difference ΔE , and the beam-constrained mass M_{bc} , both defined below, in bins of q^2 . Theoretical form-factor predictions from lattice QCD are combined with the measured differential branching fractions to determine $|V_{ub}|$ [3].

2. DETECTOR, DATA SET AND SIMULATION

2.1. Belle II detector

The Belle II detector [4] is located at the SuperKEKB [5] asymmetric-energy electron-positron collider, running at or slightly below the $\Upsilon(4S)$ resonance energy. The detector is composed of subdetectors arranged around the collision point in a cylindrical geometry. The longitudinal direction is defined by the z -axis, which points approximately along the electron-beam direction, while the transverse plane is defined by the x - and y -axes, which point radially out of the detector ring. A superconducting magnet producing a uniform 1.5 T magnetic field oriented along the detector axis is installed outside all but the last subdetector. Particles produced at the interaction point first traverse the vertex detector

composed of a two-layer pixel detector and a four-layer silicon strip detector that measure positions of decays into charged particles. Currently, however, the second layer of the pixel detector covers only one sixth of the azimuthal angle.

Trajectories of charged particles (tracks) are further determined from a central drift chamber (CDC) that has a polar angle acceptance of 17–150°. We use two Cherenkov particle-identification subdetectors: the time-of-propagation counter is located around the barrel region of the detector and the aerogel ring-imaging Cherenkov detector is located in the forward endcap. Photon- and electron-energy measurements in the barrel region and the endcaps are performed using the electromagnetic calorimeter, which is composed of CsI(Tl) crystals. Surrounding the magnet is the resistive-plate-based K_L^0 and muon detector, which also doubles as a flux return for the magnet.

Using specific-ionization energy-loss information provided by the CDC and information from the two particle-identification subdetectors and the K_L^0 and muon detector, charged particles of different masses are distinguished and particle-identification (ID) variables are constructed. These are normalized ratios of likelihoods for one charged-particle hypothesis to the sum of all possible charged-particle likelihoods.

The detector performance is monitored using control samples; we correct for time-dependent changes such as shifts in the magnetic field map or the photon energy scale. In addition, the efficiencies and misidentification probabilities of pions, kaons, electrons and muons are monitored separately by charge in bins of particle momentum and polar angle using well-known physics control modes, such as $J/\psi \rightarrow \ell^+\ell^-$ and $D^* \rightarrow D^0[K\pi]\pi$.

2.2. Data set

The primary data set used in this analysis is collected at a center-of-mass (CM) energy of $\sqrt{s} = 10.58$ GeV, corresponding to the mass of the $\Upsilon(4S)$ resonance. A data set corresponding to an integrated luminosity of 189 fb⁻¹ is collected at this CM energy, equivalent to a sample of (198.0 ± 3.0) million $\Upsilon(4S) \rightarrow B\bar{B}$ events. In addition, we use a sample corresponding to 18 fb⁻¹ of off-resonance collision data, collected at a CM energy 60 MeV below the $\Upsilon(4S)$ resonance, to describe background from continuum processes. These include $q\bar{q}$ continuum background, i.e., $e^+e^- \rightarrow u\bar{u}$, $d\bar{d}$, $s\bar{s}$ and $c\bar{c}$. The off-resonance data also contain other continuum backgrounds, such as $e^+e^- \rightarrow \tau^+\tau^-$, and two-photon processes, where $e^+e^- \rightarrow e^+e^-\ell^+\ell^-$.

2.3. Monte Carlo simulation

Simulated Monte Carlo (MC) samples of continuum and $\Upsilon(4S) \rightarrow B\bar{B}$ events, corresponding to an integrated luminosity of 1 ab⁻¹ are used. We use these to identify efficient background-discriminating variables and to form fit templates for signal extraction. In addition, MC samples of signal $B^0 \rightarrow \pi^-\ell^+\nu_\ell$ decays are used to obtain the reconstruction efficiencies and study the key kinematic distributions. In order to describe the composition of the $B \rightarrow X_u\ell\nu$ decays, 50 million events each of resonant and nonresonant $B^0 \rightarrow X_u\ell\nu$, and $B^\pm \rightarrow X_u\ell\nu$ are simulated.

The $B\bar{B}$ events are generated using the event generator `EvtGen` [6], while the $q\bar{q}$ continuum events are generated using `PYTHIA` [7]. Tau-pair events are generated using `KKMC` [8], and their decays are handled using `TAUOLA` [9]. The simulation of two-photon processes is performed using `AAFH` [10]. Backgrounds induced by the beams are mixed into the MC samples using simulations of beam-induced backgrounds [11]. The propagation of the particles through the detector and the resulting interactions are simulated using `Geant4` [12] and final-state radiation is modeled using `PHOTOS` [13]. All recorded and simulated events are handled with the analysis software framework `basf2` [14].

The branching-fraction values of the $B \rightarrow X\ell\nu$ decays assumed in the simulation are obtained from current world averages [15], in combination with assumed isospin symmetry, following the procedure in Ref. [16] for $B \rightarrow X_c\ell\nu$ decays. The remaining difference between the sum of the exclusive $B \rightarrow X_c\ell\nu$ decay branching fractions and the measured total branching fraction accounts for approximately 4% of the $B \rightarrow X_c\ell\nu$ decays and is assumed to be saturated by $B \rightarrow D^{(*)}\eta\ell\nu$ decays. For $B \rightarrow X_u\ell\nu$ decays, the difference between the sum of the exclusive branching fractions and the measured total branching fraction is assumed to be saturated by nonresonant $B \rightarrow X_u\ell\nu$ decays with multiple pions in the final state following Ref. [17]. The nonresonant $B \rightarrow X_u\ell\nu$ decays account for approximately 80% of the $B \rightarrow X_u\ell\nu$ decays.

The nonresonant $B \rightarrow X_u\ell\nu$ composition is described by implementing a hybrid model [18], following closely the method in Ref. [19]. This approach combines the exclusive and nonresonant decay rates in bins of X_u particle mass m_X , the rest-frame lepton energy E_ℓ , and q^2 , in order to recover the inclusive rates.

For the form factors of $B \rightarrow D\ell\nu$ and $B \rightarrow D^*\ell\nu$ decays, we use the Boyd-Grinstein-Lebed parameterization [20] with central values from Ref. [21] and [22], respectively. For the $B \rightarrow \rho\ell\nu$ and $B \rightarrow \omega\ell\nu$ form factors we choose the Bourrely-Caprini-Lellouch (BCL) parameterization [23] with central values from Ref. [24]. Finally, for the form-factor description of the $B \rightarrow \eta^{(\prime)}\ell\nu$ decays a light-cone sum rule calculation is chosen [25].

3. EVENT RECONSTRUCTION AND SELECTION

3.1. Event reconstruction

We begin signal reconstruction by identifying track candidates that pass certain quality criteria. The extrapolated trajectories must pass within 1 cm (3 cm) of the interaction point transverse (parallel) to the beam. Furthermore, charged particles must have a transverse momentum greater than 0.05 GeV (we use natural units with $c = 1$ throughout) and polar angles within the CDC acceptance, in order to suppress misreconstructed track candidates. We discard events with fewer than five tracks satisfying the above criteria.

In the remaining events, we select lepton candidates from among the selected tracks by requiring their CM momentum p_ℓ^* to satisfy $1.0 < p_\ell^* < 3.2$ GeV. Electron and muon candidates are required to have an electron or muon likelihood greater than 0.9, respectively. The average electron (muon) efficiency is 91 (93)%. The hadron misidentification rates are 0.2% for the electron and 3.3% for the muon selection, respectively. Electron-candidate four-momenta are corrected for bremsstrahlung by adding to them the four-momenta of photons

identified within a cone around the electron direction.

Signal-pion candidates are selected from all remaining tracks that pass within 2 cm (4 cm) of the interaction point transverse (parallel) to the beam. They are required to have a charge opposite to the lepton candidate and to have a pion likelihood greater than 0.1 in order to suppress misidentification of kaons as pions. To improve the particle-identification performance, we require the pion candidates to leave at least 20 random measurement points (hits) in the CDC. The resulting average pion efficiency is 84% with a kaon misidentification rate of 7%.

We reduce backgrounds by removing candidates with kinematic properties inconsistent with the signal B decay. Under the assumption that only a single massless particle is not included in the event reconstruction, the angle between the B meson and the combination of the lepton and pion candidates, denoted Y , is fully determined,

$$\cos \theta_{BY} = \frac{2E_B E_Y - m_B^2 - m_Y^2}{2p_B p_Y}, \quad (1)$$

where E_Y , p_Y , and m_Y are the energy, magnitude of the three-momentum, and invariant mass of the Y , respectively. The energy E_B and the magnitude of the three-momentum of the B meson p_B are calculated from the beam properties, and m_B is the mass of the B meson [15]. For correctly reconstructed signal decays, $\cos \theta_{BY}$ must lie between -1 and 1 . However, to allow for a background-dominated region, we set a looser requirement of $|\cos \theta_{BY}| < 2.2$. The event shape discriminates between $B\bar{B}$ and continuum events. Thus we require the second normalized Fox-Wolfram moment [26] to be less than 0.4.

We obtain the kinematic properties of the neutrino by assuming that the sum of the remaining tracks and electromagnetic energy-depositions (clusters) in the event, called the rest of event (ROE), represents the partner B meson. From energy and momentum conservation, we construct a missing four-momentum in the CM frame,

$$(E_{\text{miss}}^*, \vec{p}_{\text{miss}}^*) = (E_{\Upsilon(4S)}, \vec{p}_{\Upsilon(4S)}) - \left(\sum_i E_i^*, \sum_i \vec{p}_i^* \right), \quad (2)$$

where E_i^* and \vec{p}_i^* correspond to the CM energy and momentum of the i th track or cluster in the event, respectively. This yields the neutrino momentum, $\vec{p}_\nu^* = \vec{p}_{\text{miss}}^*$, and energy, $E_\nu^* = |\vec{p}_\nu^*| = |\vec{p}_{\text{miss}}^*|$.

Since all reconstructed tracks and clusters contribute to the resolution of the neutrino momentum, obtaining a pure ROE is critical. To reduce the impact of clusters from beam-induced backgrounds, acceptance losses, or other effects, we impose further criteria. We only consider clusters that are within the CDC acceptance and have transverse momenta in forward, barrel, and backward directions greater than 0.03, 0.04, and 0.06 GeV, respectively. The transverse momentum of a cluster is defined using the energy and the location with respect to the interaction point of the energy deposition. The clusters are required to be detected within 200 ns of the collision time, which is approximately five times the mean timing resolution of the calorimeter. The clusters also have to consist of more than a single calorimeter crystal. In addition to removing background particles from the ROE, we must account for particles that may escape undetected. To reduce the impact of these events, we require that the polar angle of the missing momentum in the laboratory frame θ_{miss} is within the CDC acceptance.

3.2. Signal extraction variables

We calculate q^2 from $q^2 = (p_B - p_\pi)^2$, and thus need a way to estimate the B momentum vector. One existing method, called the *Diamond Frame* [27], takes the weighted average of four possible \vec{p}_B vectors uniformly distributed in azimuthal angle on the cone defined by $\cos\theta_{BY}$ using weights of $\sin^2\theta_B$, which expresses the prior probability of the B flight direction in $\Upsilon(4S)$ decays with respect to the beam axis. A second method, called the *ROE method* [28], assumes the signal B vector to be the vector on the $\cos\theta_{BY}$ cone that is closest to antiparallel to the ROE momentum vector. We introduce a new method that combines these two by multiplying the Diamond Frame weights by $\frac{1}{2}(1 - \vec{p}_B \cdot \vec{p}_{\text{ROE}}/(|\vec{p}_B||\vec{p}_{\text{ROE}}|))$ and averaging over ten vectors uniformly distributed on the cone. We adopt this combined method because, in simulation, it assigns reconstructed signal candidates to the correct q^2 bin more often than other methods do, leading to a reduction in the bin migrations of up to 2% and resolutions in q^2 ranging from 0.14–0.45 GeV². We divide B candidates into six q^2 bins with the following labels: $q1 : q^2 \in [0, 4]$, $q2 : [4, 8]$, $q3 : [8, 12]$, $q4 : [12, 16]$, $q5 : [16, 20]$, $q6 : [20, \infty]$ GeV².

Using ROE information, two further variables that test the consistency of a candidate with a signal B decay are the beam-constrained mass, defined as

$$M_{bc} = \sqrt{E_{\text{beam}}^{*2} - |\vec{p}_B^*|^2} = \sqrt{\left(\frac{\sqrt{s}}{2}\right)^2 - |\vec{p}_B^*|^2} \quad (3)$$

and the energy difference, defined as

$$\Delta E = E_B^* - E_{\text{beam}}^* = E_B^* - \frac{\sqrt{s}}{2}, \quad (4)$$

where E_{beam}^* , E_B^* and \vec{p}_B^* are the single-beam energy, reconstructed B energy, and reconstructed B momentum all determined in the $\Upsilon(4S)$ rest frame, respectively. The reconstructed B energy (momentum) is given by the sum of the reconstructed energies (momenta) of the signal lepton and pion and the inferred neutrino energy (momentum) described above. We define a fit region in ΔE and M_{bc} , corresponding to $|\Delta E| < 0.95$ GeV and $5.095 < M_{bc} < 5.295$ GeV.

4. BACKGROUND SUPPRESSION

4.1. Background categories

The sample can be separated into two main categories: $B\bar{B}$ events and non- $B\bar{B}$ events. For the $B\bar{B}$ events we define a subcategory that combines signal and combinatorial signal events. In combinatorial signal either the pion or lepton candidate is incorrectly identified. We further split the $B\bar{B}$ background into the two largest semileptonic backgrounds, $B \rightarrow X_c \ell \nu$ and $B \rightarrow X_u \ell \nu$. The remaining $B\bar{B}$ events are combined into an *other $B\bar{B}$* category, mainly composed of candidates with misidentified leptons or with leptons from secondary decays. The non- $B\bar{B}$ events are combined into a continuum background category, which contains $q\bar{q}$ and other continuum events.

4.2. Boosted decision trees

In order to further reduce the $B\bar{B}$ and $q\bar{q}$ backgrounds we train boosted decision trees (BDTs) using the FastBDT methodology [29], which differentiate between signal and each background category separately. Since the background composition is different in each q^2 bin, we train the classifier and optimize the selection separately for each q^2 bin. This results in a total of 12 BDTs. In each training we combine the electron and muon modes, since they have similar background compositions. We use equal amounts of signal and background simulation events, corresponding to 200 fb^{-1} of $q\bar{q}$ simulated data, and 600 fb^{-1} of $B\bar{B}$ simulated data, due to the low $B\bar{B}$ background retention.

Two input variables, $\cos\theta_{BY}$ and θ_{miss} , are common to the $q\bar{q}$ and $B\bar{B}$ suppression. Further input variables in the $q\bar{q}$ suppression BDT are based on the thrust axis, which is the axis that maximizes the sum of the projected momenta of all charged particles in the event. Distinct thrust axes can be defined for the signal B and the ROE. Their magnitudes, and the angle between the two axes $\cos\theta_T$ serve as input variables. In addition, three cones with opening angles of 10, 20, and 30° centered around the thrust axis are defined, and the momentum flow into each of the three cones is added as an input variable [30]. The input variables of the $B\bar{B}$ suppression BDTs are the number of tracks, the angle between the lepton and pion momenta, and the momentum of the ROE. In addition, a vertex fit to the pion and lepton candidates is performed using TreeFit [31] and the χ^2 probability is used as a discriminating feature. The remaining two variables are the cosines of the angles between the signal B momentum vector and the vector connecting its fitted vertex to the interaction point in the plane parallel and perpendicular to the beam axis, respectively. The shared input variable $\cos\theta_{BY}$ is one of the most discriminating variables in both the suppression of continuum and $B\bar{B}$ events. The $\cos\theta_T$ and the χ^2 vertex fit probability also provide high discriminating power in the suppression of continuum and $B\bar{B}$ backgrounds, respectively. We identify the optimal selection criterion on the output classifier within each q^2 bin by maximizing the ratio between the number of signal events and the square root of the sum of the number of signal and background events, as predicted by simulation.

After selecting on the output classifiers, the continuum background contains a significant amount of residual two-photon background, especially $ee\tau\tau$ events, in the electron mode. Therefore, for the electron mode only, we train another BDT common to all q^2 bins, to separate signal from $ee\tau\tau$ events. We use a sample corresponding to 2 ab^{-1} of $ee\tau\tau$ simulated data and an equivalent amount of signal events. The input variables include the total charge in the event, θ_{miss} , the polar angle of the lepton, the number of clusters, the mass of the ROE, and ξ_z . Here, the variable ξ_z is calculated as the sum of the momenta parallel to the beam axis divided by the sum of the energies of all charged particles in the laboratory frame. None of the selected input variables are strongly correlated with ΔE or M_{bc} .

4.3. Best candidate selection and efficiencies

At this stage, an average of 1.02 candidates remain per selected event. In events with multiple candidates, we randomly select one and discard the rest. The overall signal efficiencies range from 8% to 17% in the electron mode and from 13% to 23% in the muon mode, depending on the q^2 bin.

5. SIGNAL EXTRACTION

5.1. Continuum treatment

Because of the small off-resonance sample size, using the data directly as a fit template introduces large systematic uncertainties. Instead, we weight the simulated continuum candidates in order to use the resulting template during signal extraction. We therefore compare the q^2 shapes of simulated continuum data and off-resonance data. We observe a normalization difference in the electron mode, and to a lesser extent in the muon mode. After correcting the simulated continuum data for the total normalization, any residual difference in the q^2 spectrum is resolved by correcting the q^2 spectrum bin-by-bin. This approach relies on the assumption that the difference between off-resonance data and the simulated continuum sample is independent of ΔE and M_{bc} . Any observed deviation from this assumption is treated as a systematic uncertainty and explained in Section 6.

5.2. Fit setup

We extract the signal by performing an extended likelihood fit to the binned two-dimensional distribution of ΔE and M_{bc} in bins of q^2 . In total we have $5 (\Delta E) \times 4 (M_{bc}) \times 6 (q^2) = 120$ bins. The distributions of ΔE and M_{bc} integrated over the six q^2 bins for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays are shown in Figure 1. The likelihood to be maximized is

$$\mathcal{L}(s_j, b_k) = \prod_i \text{Poisson}(N_i | \sum_j s_{ij} + \sum_k b_{ik}), \quad (5)$$

where N_i is the observed number of events in bin i , s_{ij} is the number of events in bin i of signal fit template j , and b_{ik} is the number of events in bin i of background fit template k . The combinatorial signal is included in the signal yield. The backgrounds are split into four templates according to the description in Section 4.4.1. In addition, there is one independent signal template for each of the six q^2 bins, leading to a total of ten templates. The templates are constructed from 1 ab^{-1} of simulated data, where we use the weighted simulated continuum events to extract the continuum template. We use a Gaussian penalty factor to constrain the continuum yield to the scaled off-resonance yield.

5.3. Fit results

The fit projections of ΔE and M_{bc} in each q^2 bin are shown in Figure 2. The correlations between the component yields are all lower than 0.7. The highest observed correlations occur between the $B \rightarrow X_c \ell \nu$ and continuum yields. In the higher q^2 bins, the signal becomes increasingly correlated to the $B \rightarrow X_u \ell \nu$ yield.

The q^2 bin migrations in signal range from 3.7–9.8%. For correctly reconstructed signal they are largest in the lowest q^2 bin and decrease at higher q^2 . The inclusion of combinatorial signal, however, also results in larger bin migrations in the highest q^2 bin. We correct the signal yields obtained from the fit for the migrations by applying the inverse detector

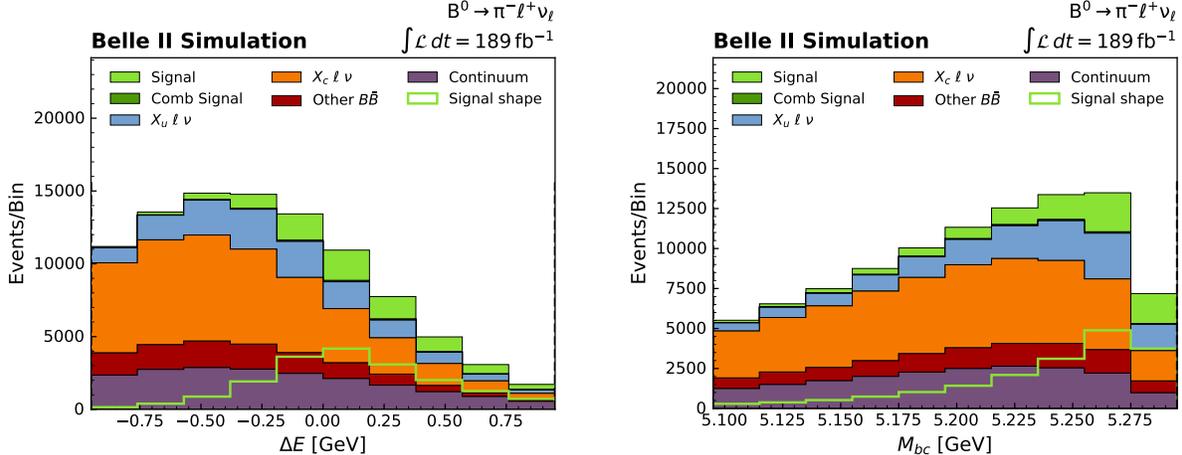


FIG. 1: Simulated distributions of (left) ΔE and (right) M_{bc} integrated over the six q^2 bins for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays.

response matrix [32]. The corrected yields are given in Table I. The $B^0 \rightarrow \pi^- e^+ \nu_e$ signal yields are lower than the $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ signal yields. This can partly be attributed to lower signal efficiencies due to the additional selection on the two-photon background suppression BDT.

TABLE I: Signal yields corrected for bin migrations in each q^2 bin with statistical and systematic uncertainties. The boundaries of the q^2 bins are given in the text above.

| q^2 bin | $B^0 \rightarrow \pi^- e^+ \nu_e$ | $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ |
|-----------|-----------------------------------|---------------------------------------|
| $q1$ | $426 \pm 49 \pm 73$ | $749 \pm 74 \pm 400$ |
| $q2$ | $927 \pm 56 \pm 133$ | $1076 \pm 61 \pm 164$ |
| $q3$ | $856 \pm 64 \pm 96$ | $1238 \pm 77 \pm 128$ |
| $q4$ | $577 \pm 64 \pm 64$ | $819 \pm 77 \pm 71$ |
| $q5$ | $497 \pm 66 \pm 60$ | $775 \pm 84 \pm 75$ |
| $q6$ | $613 \pm 73 \pm 227$ | $809 \pm 86 \pm 164$ |

6. SYSTEMATIC UNCERTAINTIES

The fractional uncertainties on the corrected signal yields in each q^2 bin from various sources of systematic uncertainty are shown in Table II. All systematic uncertainties are evaluated using the same approach. For each source of uncertainty, we vary the templates 1000 times by sampling from Gaussian distributions of the central values. For example, to evaluate the uncertainties due to the $B \rightarrow \rho \ell \nu$ form factors, we sample 1000 alternative $B \rightarrow X_u \ell \nu$ distributions by assuming the form-factor parameter uncertainties follow Gaussian distributions. We create 1000 simplified simulated data (toy) distributions by adding the resulting variations to any unaffected templates. We then fit the nominal templates to the

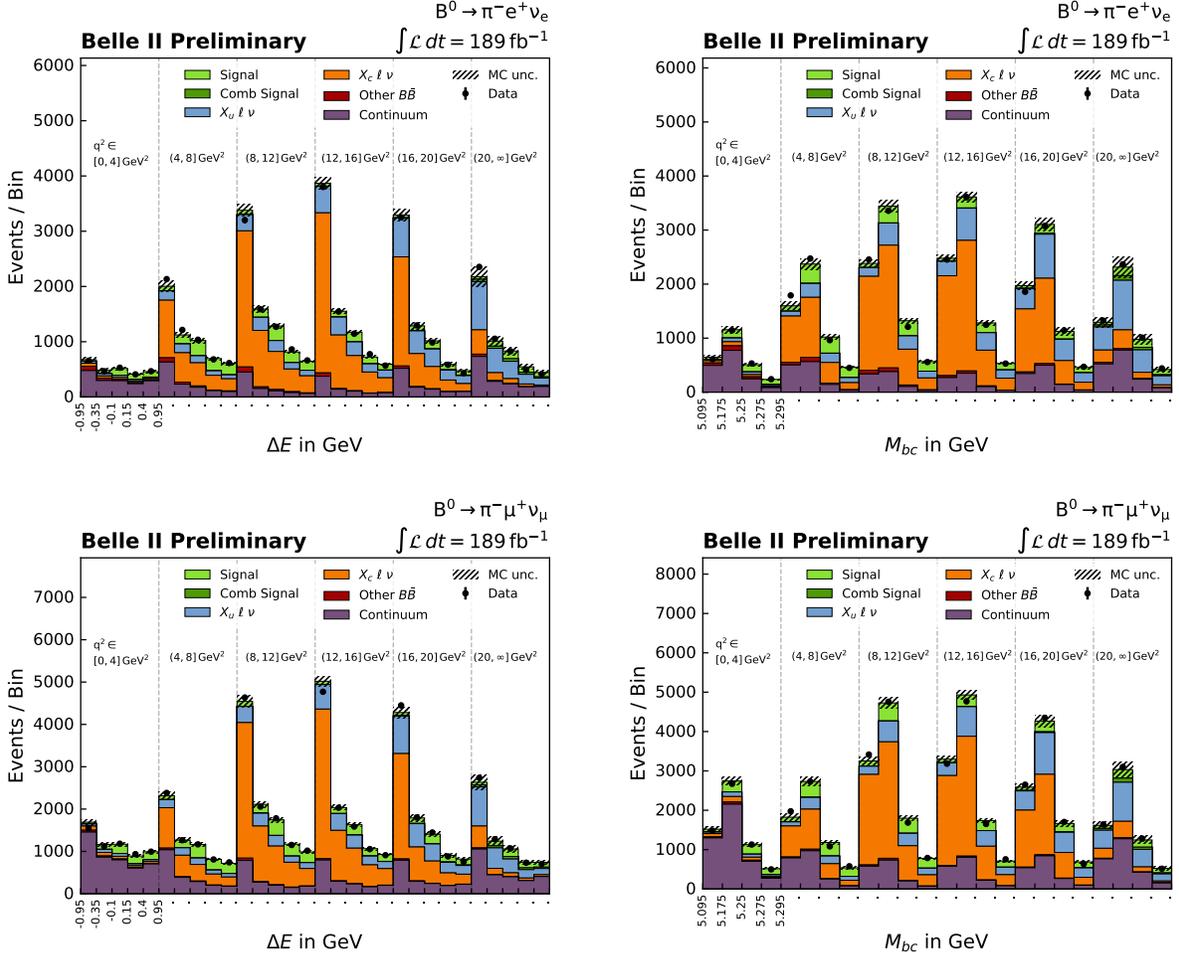


FIG. 2: Distributions of (left) ΔE and (right) M_{bc} in the six q^2 bins for (top) $B^0 \rightarrow \pi^- e^+ \nu_e$ and (bottom) $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ candidates reconstructed in Belle II data with fit projections overlaid.

toy distributions and obtain a covariance matrix for each source of uncertainty using Pearson correlation [33].

The largest contribution to the uncertainty in the $B \rightarrow X_u \ell \nu$ background template comes from the uncertainty in the $B \rightarrow \rho \ell \nu$ form-factor parameters. The effect of the $B \rightarrow \rho \ell \nu$ form-factor and branching fraction uncertainties is included in the $B \rightarrow \rho \ell \nu$ category in Table II. We also evaluate the uncertainties due to the $B \rightarrow \pi \ell \nu$, $B \rightarrow \omega \ell \nu$, $B \rightarrow \eta \ell \nu$, and $B \rightarrow \eta' \ell \nu$ form factors, and obtain uncertainties shown under the $B \rightarrow X_u \ell \nu$ category. This category also includes the effects of uncertainties of the exclusive and inclusive $B \rightarrow X_u \ell \nu$ branching fractions, except for the $B \rightarrow \rho \ell \nu$ branching fraction. The $B \rightarrow X_c \ell \nu$ category in Table II includes the effects of the uncertainties of the $B \rightarrow D \ell \nu$ and $B \rightarrow D^* \ell \nu$ form-factor parameters, and the exclusive and inclusive $B \rightarrow X_c \ell \nu$ branching fractions.

The detector uncertainties include uncertainties arising from the tracking efficiency and the corrections to the lepton- and pion-identification efficiencies. The effect of having a small sample of simulated data is also considered.

Two sources of uncertainty are included in the continuum category in Table II. One is the uncertainty on the signal yields due to the limited off-resonance sample size. The resulting uncertainties range from 3% to 11% in the intermediate q^2 bins and dominate the uncertainties in the highest and lowest q^2 bins, in which the continuum background is largest. The other source of uncertainty originates from the assumption that the continuum reweighting in q^2 bins is independent of ΔE and M_{bc} . To estimate this, we parameterize the shape difference between simulated continuum and off-resonance data as a linear function in ΔE . We fit this parameterization to the observed difference and obtain central values and a covariance for the two parameters. We then follow the same procedure used in the evaluation of the other systematic uncertainties. We sample 1000 varied continuum templates, generate the corresponding toy distributions, and by fitting the nominal templates to the toy distributions obtain refitted yields. If the difference between the nominal fit central value and the mean value of the refitted yields is significant, we take this difference as an additional uncertainty on the yield. This introduces large uncertainties to the highest and lowest q^2 bins in the electron and muon modes, respectively.

We consider additional systematic uncertainties from the number of $B\bar{B}$ pairs $N_{B\bar{B}}$, and the branching fraction of $\Upsilon(4S) \rightarrow B^0\bar{B}^0$, $f = 0.486 \pm 0.006$ [15]. These uncertainties do not affect the yields, but they contribute a 2% systematic uncertainty to the branching fraction results.

TABLE II: Summary of fractional uncertainties on the yields.

| Source | $B^0 \rightarrow \pi^- e^+ \nu_e$ | | | | | | $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ | | | | | |
|----------------------------|-----------------------------------|------|------|------|------|------|---------------------------------------|------|------|------|------|------|
| | q1 | q2 | q3 | q4 | q5 | q6 | q1 | q2 | q3 | q4 | q5 | q6 |
| Detector | 1.2 | 1.0 | 1.1 | 1.4 | 2.3 | 2.4 | 2.3 | 3.2 | 3.3 | 1.2 | 1.9 | 3.8 |
| MC sample size | 4.0 | 2.0 | 2.4 | 2.8 | 3.9 | 5.6 | 3.9 | 2.0 | 2.3 | 2.7 | 3.4 | 4.8 |
| Continuum | 13.1 | 5.5 | 4.4 | 7.8 | 10.5 | 33.9 | 53.3 | 8.8 | 3.2 | 4.5 | 8.0 | 11.4 |
| $B \rightarrow \rho l \nu$ | 9.5 | 12.5 | 9.7 | 6.9 | 3.4 | 12.9 | 8.7 | 11.6 | 8.6 | 6.3 | 3.3 | 14.3 |
| $B \rightarrow X_u l \nu$ | 3.3 | 1.9 | 2.1 | 2.1 | 1.8 | 3.7 | 3.4 | 2.3 | 2.0 | 2.3 | 2.1 | 6.0 |
| $B \rightarrow X_c l \nu$ | 2.3 | 3.0 | 1.1 | 0.8 | 0.5 | 2.4 | 2.4 | 1.5 | 1.5 | 0.8 | 0.5 | 2.2 |
| Total syst. | 17.2 | 14.3 | 11.2 | 11.1 | 12.0 | 37.0 | 53.4 | 15.2 | 10.3 | 8.7 | 9.7 | 20.3 |
| Stat. | 10.2 | 6.01 | 6.86 | 8.08 | 10.3 | 13.2 | 10.4 | 6.0 | 6.4 | 7.8 | 9.7 | 13.4 |
| Total | 20.2 | 15.5 | 13.2 | 13.7 | 15.9 | 39.2 | 54.5 | 16.4 | 12.2 | 11.6 | 13.7 | 24.3 |

7. RESULTS

7.1. Branching fractions

The partial branching fraction in q^2 bin i is calculated using the corrected yield, N_i , and the corresponding signal efficiency, ϵ_i , from

$$\Delta\mathcal{B}_i = \frac{N_i}{2f\epsilon_i N_{B\bar{B}}}. \quad (6)$$

The results for $B^0 \rightarrow \pi^- e^+ \nu_e$ and $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ decays, $\Delta\mathcal{B}_{i,e}$ and $\Delta\mathcal{B}_{i,\mu}$, respectively, are given in Table III. We also provide an average over both decay channels, fully correlating common systematic uncertainties in Table III¹. The total branching fraction determined from the sum of the averaged partial branching fractions is

$$\mathcal{B}(B^0 \rightarrow \pi^- \ell^+ \nu_\ell) = (1.426 \pm 0.056(\text{stat}) \pm 0.125(\text{syst})) \times 10^{-4},^1$$

where the first uncertainty is statistical and the second is systematic.

TABLE III: $\Delta\mathcal{B}$ ($\times 10^4$) in each q^2 bin calculated from the corrected signal yields for the $B^0 \rightarrow \pi^- e^+ \nu_e$ and $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ modes. The averaged partial branching fractions for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ are also shown. The first uncertainty is statistical and the second is systematic.

| q^2 bin | $B^0 \rightarrow \pi^- e^+ \nu_e$ | $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ | $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ |
|-----------|-----------------------------------|---------------------------------------|---|
| $q1$ | $0.261 \pm 0.030 \pm 0.045$ | $0.281 \pm 0.028 \pm 0.150$ | $0.272 \pm 0.031 \pm 0.044$ |
| $q2$ | $0.298 \pm 0.018 \pm 0.043$ | $0.285 \pm 0.016 \pm 0.044$ | $0.290 \pm 0.013 \pm 0.040$ |
| $q3$ | $0.268 \pm 0.020 \pm 0.030$ | $0.290 \pm 0.018 \pm 0.030$ | $0.279 \pm 0.014 \pm 0.028$ |
| $q4$ | $0.196 \pm 0.022 \pm 0.022$ | $0.204 \pm 0.019 \pm 0.018$ | $0.199 \pm 0.015 \pm 0.017$ |
| $q5$ | $0.180 \pm 0.024 \pm 0.022$ | $0.194 \pm 0.021 \pm 0.019$ | $0.188 \pm 0.016 \pm 0.015$ |
| $q6$ | $0.200 \pm 0.024 \pm 0.074$ | $0.220 \pm 0.023 \pm 0.045$ | $0.198 \pm 0.031 \pm 0.042$ |

7.2. $|V_{ub}|$ determination

We extract $|V_{ub}|$ using χ^2 fits to the measured q^2 spectra. We include lattice QCD constraints on the eight BCL parameters from Ref. [3] as nuisance parameters. These constrain the shape and normalization of the relevant form factors entering the differential decay rate and allow for a determination of $|V_{ub}|$.

The χ^2 is defined as

$$\chi^2 = \sum_{i,j=1}^6 (\Delta\mathcal{B}_i - \Delta\Gamma_i \tau) C_{ij}^{-1} (\Delta\mathcal{B}_j - \Delta\Gamma_j \tau) + \chi_{\text{LQCD}}^2, \quad (7)$$

where C_{ij}^{-1} is the inverse total covariance matrix of the measured partial branching fractions in bin i , \mathcal{B}_i . The quantities $\Delta\Gamma_i$ contain the predictions for the partial decay rates in bin i , τ is the B^0 lifetime, and χ_{LQCD}^2 incorporates the constraints from the lattice calculation.

The obtained results are

$$\begin{aligned} |V_{ub}|_{B^0 \rightarrow \pi^- e^+ \nu_e} &= (3.60 \pm 0.18(\text{stat}) \pm 0.14(\text{syst}) \pm 0.18(\text{theo})) \times 10^{-3} \\ |V_{ub}|_{B^0 \rightarrow \pi^- \mu^+ \nu_\mu} &= (3.71 \pm 0.16(\text{stat}) \pm 0.15(\text{syst}) \pm 0.17(\text{theo})) \times 10^{-3}, \\ |V_{ub}|_{B^0 \rightarrow \pi^- \ell^+ \nu_\ell} &= (3.55 \pm 0.12(\text{stat}) \pm 0.13(\text{syst}) \pm 0.17(\text{theo})) \times 10^{-3}.^1 \end{aligned}$$

¹ Note: these results differ slightly from those presented at the 2022 conferences.

The value of $|V_{ub}|$ determined by fitting the averaged $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ partial branching fractions is lower than the $|V_{ub}|$ results from the $B^0 \rightarrow \pi^- e^+ \nu_e$ and $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ samples. This feature of the fit arises because the extraction of $|V_{ub}|$ is most sensitive to the high q^2 region, where the average partial branching fraction listed in Table III is also lower than the $B^0 \rightarrow \pi^- e^+ \nu_e$ and $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ partial branching fractions due to correlations between q^2 bins. The partial branching fractions of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ measured as a function of q^2 are shown in Figure 3. The fitted differential rate is also shown, and the one, two, and three standard-deviation uncertainty bands are given.

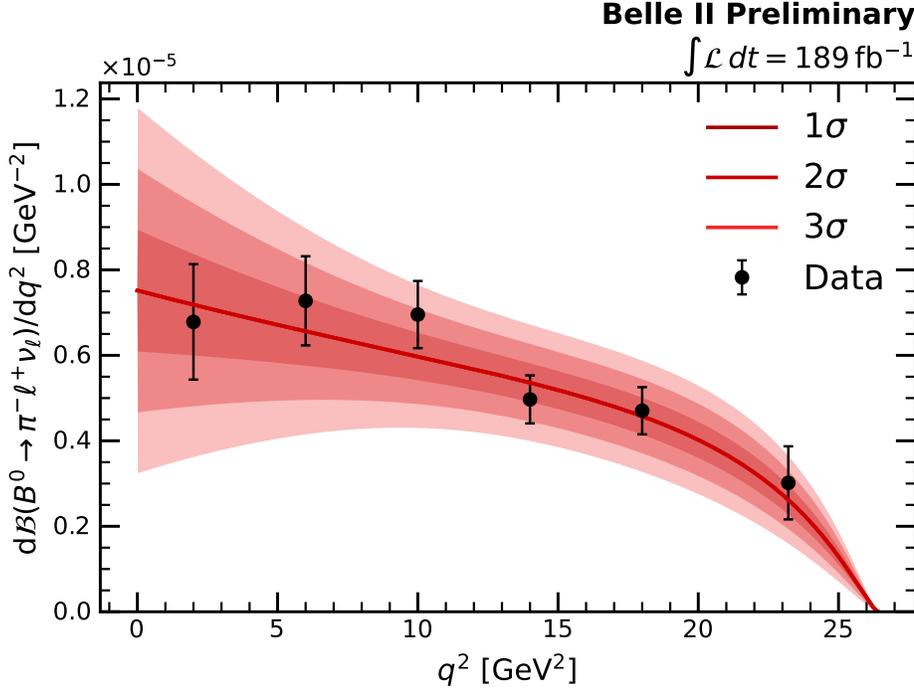


FIG. 3: Measured averaged partial branching fractions as a function of q^2 . The fitted differential rate is shown together with the one, two, and three standard-deviation uncertainty bands.

8. SUMMARY

We extract partial branching fractions for $B^0 \rightarrow \pi^- e^+ \nu_e$ and $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ decays reconstructed in an electron-positron collision data sample corresponding to 189 fb^{-1} collected by the Belle II experiment in 2019-2021. By averaging the results we obtain partial branching fractions for $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$. The total branching fraction result of $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ is found to be $(1.426 \pm 0.056(\text{stat}) \pm 0.125(\text{syst})) \times 10^{-4}$. This is consistent with the current world average of $(1.50 \pm 0.06) \times 10^{-4}$ [15]. Currently our results are limited by the size of the off-resonance data set. This uncertainty could be reduced by improvements in the simulation of continuum backgrounds. We also extract values of the CKM matrix-element magnitude $|V_{ub}|$ and from $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ decays obtain $(3.55 \pm 0.12(\text{stat}) \pm 0.13(\text{syst}) \pm 0.17(\text{theo})) \times 10^{-3}$. This agrees with the current value obtained by HFLAV [2] of $(3.70 \pm 0.10(\text{exp}) \pm 0.12(\text{theo})) \times 10^{-3}$.

9. ACKNOWLEDGEMENTS

We thank the SuperKEKB group for the excellent operation of the accelerator; the KEK cryogenics group for the efficient operation of the solenoid; the KEK computer group for on-site computing support; and the raw-data centers at BNL, DESY, GridKa, IN2P3, and INFN for off-site computing support.

-
- [1] M. Kobayashi and T. Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.
 - [2] Y. S. Amhis et al. Averages of b -hadron, c -hadron, and τ -lepton properties as of 2018. *Eur. Phys. J.*, C81:226, 2021.
 - [3] J. A. Bailey, A. Bazavov, et al. $|V_{ub}|$ from $B \rightarrow \pi \ell \nu$ decays and $(2 + 1)$ -flavor lattice QCD. *Physical Review D*, 92(1), Jul 2015.
 - [4] T. Abe et al. Belle II Technical Design Report. Nov 2010.
 - [5] K. Akai et al. SuperKEKB Collider. *Nucl. Instrum. Meth.*, A907:188–199, 2018.
 - [6] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152–155, 2001.
 - [7] T. Sjöstrand et al. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015.
 - [8] S. Jadach et al. The Precision Monte Carlo event generator KK for two fermion final states in e^+e^- collisions. *Comput. Phys. Commun.*, 130:260–325, 2000.
 - [9] S. Jadach et al. TAUOLA: A Library of Monte Carlo programs to simulate decays of polarized tau leptons. *Comput. Phys. Commun.*, 64:275–299, 1990.
 - [10] F. A. Berends and R. van Gulik. GaGaRes: A Monte Carlo generator for resonance production in two-photon physics. *Computer Physics Communications*, 144(1):82–103, 2002.
 - [11] A. Natchii et al. Beam background expectations for Belle II at SuperKEKB. 2022.
 - [12] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.
 - [13] E. Barberio et al. PHOTOS: A Universal Monte Carlo for QED radiative corrections in decays. *Comput. Phys. Commun.*, 66:115–128, 1991.
 - [14] T. Kuhr et al. The Belle II Core Software. *Comput. Softw. Big Sci.*, 3(1), 2019.
 - [15] P. A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8), 2020.
 - [16] F. U. Bernlochner and Z. Ligeti. Semileptonic $B_{(s)}$ decays to excited charmed mesons with e, μ, τ and searching for new physics with $R(D^{**})$. *Phys. Rev. D*, 95(1), 2017.
 - [17] B. O. Lange et al. Theory of charmless inclusive B decays and the extraction of $|V_{ub}|$. *Phys. Rev. D*, 72, 2005.
 - [18] C. Ramirez et al. Semileptonic $b \rightarrow u$ decays. *Phys. Rev. D*, 41:1496–1503, Mar 1990.
 - [19] M. T. Prim et al. Search for $B^+ \rightarrow \mu^+ \nu_\mu$ and $B^+ \rightarrow \mu^+ N$ with inclusive tagging. *Phys. Rev. D*, 101(3), 2020.
 - [20] C. G. Boyd et al. Constraints on Form Factors for Exclusive Semileptonic Heavy to Light Meson Decays. *Phys. Rev. Lett.*, 74:4603–4606, Jun 1995.
 - [21] R. Glattauer et al. Measurement of the decay $B \rightarrow D \ell \nu_\ell$ in fully reconstructed events and determination of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cb}|$. *Phys. Rev. D*, 93,

Feb 2016.

- [22] D. Ferlewicz et al. Revisiting fits to $B^0 \rightarrow D^{*-} \ell^+ \nu_\ell$ to measure $|V_{cb}|$ with novel methods and preliminary LQCD data at nonzero recoil. *Phys. Rev. D*, 103(7), 2021.
- [23] C. Bourrely et al. Model-independent description of $B \rightarrow \pi \ell \nu$ decays and a determination of $|V_{ub}|$. *Phys. Rev. D*, 79, Jan 2009.
- [24] F. U. Bernlochner et al. $B \rightarrow \rho \ell \nu$ and $\omega \ell \nu$ in and beyond the Standard Model: Improved predictions and $|V_{ub}|$. *Phys. Rev. D*, 104(3), 2021.
- [25] G. Duplancic and B. Melic. Form factors of $B, B_s \rightarrow \eta^{(\prime)}$ and $D, D_s \rightarrow \eta^{(\prime)}$ transitions from QCD light-cone sum rules. *JHEP*, 11:138, 2015.
- [26] G. C. Fox and S. Wolfram. Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes. *Phys. Rev. Lett.*, 41:1581–1585, Dec 1978.
- [27] B. Aubert et al. Measurements of the $B \rightarrow D^*$ form factors using the decay $\bar{B}^0 \rightarrow D^{*+} e^- \nu_e$. *Phys. Rev. D*, 74, Nov 2006.
- [28] E. Waheed et al. Measurement of the CKM Matrix Element $|V_{cb}|$ from $B^0 \rightarrow D^{*-} \ell^+ \nu_\ell$ at Belle. *Phys. Rev. D*, 100, Sep 2019.
- [29] T. Keck. FastBDT: A speed-optimized and cache-friendly implementation of stochastic gradient-boosted decision trees for multivariate classification, 2016.
- [30] D. M. Asner et al. Search for exclusive charmless hadronic B decays. *Physical Review D*, 53(3):1039–1050, Feb 1996.
- [31] J. F. Krohn et al. Global decay chain vertex fitting at Belle II. *Nucl. Instrum. Meth. A*, 976, 2020.
- [32] G. Cowan. The CERN statistics handbook. *Oxford University, Oxford, UK*, 1998.
- [33] J. Benesty, J. Chen, Y. Huang, and I. Cohen. Noise reduction in speech processing. *Springer*, pages 37–40, 2009.