Exclusive $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ Decays with Hadronic Full Event Interpretation Tagging in 34.6 fb$^{-1}$ of Belle II Data

F. Abudinén,47 I. Adachi,24,21 R. Adak,18 K. Adamczyk,72 P. Ahlborg,109 J. K. Ahn,54 H. Aihara,127 N. Akopov,133 A. Aloisio,97,40 F. Ameli,44 L. Andricek,63 N. Anh Ky,37,14 D. M. Asner,3 H. Atmacan,111 V. Aulchenko,4,74 T. Aushev,26 V. Aushev,88 T. Aziz,89 V. Babu,12 S. Bacher,72 S. Baehrt,51 S. Bahinipati,28 A. M. Bakich,126 P. Bambade,106 Sw. Banerjee,116 S. Bansal,79 M. Barrett,24 G. Batignani,100,43 J. Baudot,107 M. Bauer,51 A. Beaulieu,129 J. Becker,51 P. K. Behera,31 M. Bender,59 J. V. Bennett,120 E. Bernieri,45 F. U. Bernlochner,109 M. Bertemes,34 M. Bessner,113 S. Bettarini,100,43 V. Bhardwaj,27 B. Bhuyan,29 F. Bianchi,103,46 T. Bilka,7 S. Bilokin,50 D. Biswas,116 A. Bobrov,4,74 A. Bondar,4,74 G. Bonvicini,131 A. Bozek,72 M. Bračko,118,87 P. Branchini,45 N. Braun,51 R. A. Briere,5 T. E. Browder,113 D. N. Brown,116 A. Budano,45 L. Burmistrov,106 S. Bussino,102,45 M. Campajola,97,40 L. Cao,109 G. Caria,119 G. Casarosa,100,43 C. Cecchi,99,42 D. Červenkov,7 M.-C. Chang,17 P. Chang,70 R. Cheaib,110 V. Chekelian,62 Y. Q. Chen,123 Y.-T. Chen,70 B. G. Cheon,23 K. Chilikin,57 K. Chirapatpimol,8 H.-E. Cho,23 K. Cho,53 S.-J. Cho,134 S.-K. Choi,22 S. Choudhury,30 D. Cinabro,131 L. Corona,100,43 L. M. Cremaldi,120 D. Cuesta,107 S. Cunliffe,12 T. Czank,128 N. Dash,31 F. Dattola,12 E. De La Cruz-Burelo,6 G. De Nardo,97,40 M. De Nuccio,113 G. De Pietro,45 R. de Sangro,39 T. Destefanis,109 M. Destefanis,103,46 S. Dev,91 A. De Yta-Hernandez,6 A. Di Canto,3 F. Di Capua,97,40 S. Di Carlo,106 J. Dingfelder,109 Z. Doležal,7 I. Domínguez Jiménez,96 T. V. Dong,18 K. Dort,50 D. Dossett,119 S. Dubey,113 S. Duell,109 G. Dujany,107 S. Eidelman,4,57,74 M. Eliaševičius,109 D. Epifanov,4,74 J. E. Fast,78 T. Ferber,12 D. Ferlewicz,119 G. Finocchiaro,39 S. Fiore,44 P. Fischer,114 A. Fodor,64 F. Forti,100,43 A. Frey,19 M. Friedl,34 B. G. Fulsom,78 M. Gabriel,62 N. Gabyshev,4,74 E. Ganiev,104,47 M. Garcia-Hernandez,5 R. Garg,79 A. Garmash,4,74 V. Gaur,130 A. Gaz,66,67 U. Gebauer,19 M. Gelb,51 A. Gellrich,12 J. Gemmer,51 T. Geßler,50 D. Getzkow,50 R. Giordano,97,40 A. Giri,30 A. Glazov,42 B. Gobbo,47 R. Godang,124 P. Goldenzweig,51 B. Golob,115,87 P. Gomis,38 P. Grace,108 W. Gradl,49 E. Graziani,45 D. Greenwald,90 Y. Guan,111 C. Hadjivasiliou,78 S. Halder,89 K. Hara,24,21 T. Hara,24,21 O. Hartbrich,113 T. Hauth,51 K. Hayasaka,73 H. Hayashii,69 C. Hearty,110,36 M. Heck,51 M. T. Hedges,113 I. Heredia de la Cruz,6,11 M. Hernández Villanueva,120 A. Hershenhorn,110 T. Higuchi,128 E. C. Hill,110 H. Hirata,66 M. Hoek,49 M. Hohmann,119 S. Hollitting,108 T. Hotta,77 C.-L. Hsu,126 Y. Hu,35 K. Huang,70 T. Iijima,66,67 K. Inami,66 G. Inguglia,34 J. Irakkathil Jabbar,51 A. Ishikawa,24,21 R. Itoh,24,21 M. Iwasaki,76 Y. Iwasaki,24
McGill University, Montréal, Québec, H3A 2T8, Canada
Moscow Physical Engineering Institute, Moscow 115409, Russian Federation
Graduate School of Science, Nagoya University, Nagoya 464-8602, Japan
Kobayashi-Maskawa Institute, Nagoya University, Nagoya 464-8602, Japan
Institute for Advanced Research, Nagoya University, Nagoya 464-8602, Japan
Nara Women’s University, Nara 630-8506, Japan
Department of Physics, National Taiwan University, Taipei 10617, Taiwan
National United University, Miao Li 36003, Taiwan
H. Niewodniczanski Institute of Nuclear Physics, Krakow 31-342, Poland
Niigata University, Niigata 950-2181, Japan
Novosibirsk State University, Novosibirsk 630090, Russian Federation
Okinawa Institute of Science and Technology, Okinawa 904-0495, Japan
Osaka City University, Osaka 558-8585, Japan
Research Center for Nuclear Physics, Osaka University, Osaka 567-0047, Japan
Pacific Northwest National Laboratory, Richland, Washington 99352, U.S.A.
Panjab University, Chandigarh 160014, India
Peking University, Beijing 100871, China
Punjab Agricultural University, Ludhiana 141004, India
Meson Science Laboratory, Cluster for Pioneering Research, RIKEN, Saitama 351-0198, Japan
Seoul National University, Seoul 08826, South Korea
Showa Pharmaceutical University, Tokyo 194-8543, Japan
Soochow University, Suzhou 215006, China
Soongsil University, Seoul 06978, South Korea
J. Stefan Institute, 1000 Ljubljana, Slovenia
Taras Shevchenko National Univ. of Kiev, Kiev, Ukraine
Tata Institute of Fundamental Research, Mumbai 400005, India
Department of Physics, Technische Universität München, 85748 Garching, Germany
Tel Aviv University, School of Physics and Astronomy, Tel Aviv, 69978, Israel
Toho University, Funabashi 274-8510, Japan
Department of Physics, Tohoku University, Sendai 980-8578, Japan
Tokyo Institute of Technology, Tokyo 152-8550, Japan
Tokyo Metropolitan University, Tokyo 192-0397, Japan
Universidad Autonoma de Sinaloa, Sinaloa 80000, Mexico
Dipartimento di Scienze Fisiche, Università di Napoli Federico II, I-80126 Napoli, Italy
Dipartimento di Fisica e Astronomia, Università di Padova, I-35131 Padova, Italy
Dipartimento di Fisica, Università di Perugia, I-06123 Perugia, Italy
Dipartimento di Fisica, Università di Pisa, I-56127 Pisa, Italy
Università di Roma “La Sapienza,” I-00185 Roma, Italy
Dipartimento di Matematica e Fisica, Università di Roma Tre, I-00146 Roma, Italy
Dipartimento di Fisica, Università di Torino, I-10125 Torino, Italy
Dipartimento di Fisica, Università di Trieste, I-34127 Trieste, Italy
Université de Montréal, Physique des Particules, Montréal, Québec, H3C 3J7, Canada
Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France
Université de Strasbourg, CNRS, IPHC, UMR 7178, 67037 Strasbourg, France
Department of Physics, University of Adelaide, Adelaide, South Australia 5005, Australia
University of Bonn, 53115 Bonn, Germany
Abstract

We present the results of the re-discovery of the decay $B^0 \to \pi^- \ell^+ \nu_\ell$ in 34.6 fb$^{-1}$ of Belle II data using hadronic $B$-tagging via the Full Event Interpretation algorithm. We observe 21 signal events on a background of 155 in a fit to the distribution of the square of the missing mass, $M_{\text{miss}}^2$, with a significance of 5.69$\sigma$, and determine a total branching fraction of $(1.58 \pm 0.43_{\text{stat}} \pm 0.07_{\text{sys}}) \times 10^{-4}$. 
1. INTRODUCTION

Since the start of its first physics run in 2019, the Belle II detector has collected over 74 fb$^{-1}$ of data from electron-positron collisions. This early data has been invaluable for investigating the performance of the detector and the analysis software, and has also led to the re-discovery of many physics processes.

In this paper, we study rare semi-leptonic decays of the form $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$, where $\ell = e, \mu$ \cite{12}. These decays are considered golden modes for precise determinations of the absolute value of the CKM-matrix element $|V_{ub}|$. The integrated luminosity collected at present is too small to provide a competitive measurement. We present the re-discovery of the decay $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ in 34.6 fb$^{-1}$ of Belle II data via hadronic $B$-tagging provided by the Full Event Interpretation (FEI) algorithm \cite{1}.

2. THE BELLE II DETECTOR

The structure of the Belle II detector is described in detail in Ref. \cite{2}. The inner-most layers are known collectively as the vertex detector, or VXD, and are dedicated to the tracking of charged particles and the precise determination of particle decay vertices. The VXD is comprised of two layers of silicon pixel sensors surrounded by four layers of silicon strip detectors. The central drift chamber (CDC) surrounds the VXD, encompassing the barrel region of the detector, and is primarily responsible for the reconstruction of charged tracks and the determination of their momenta.

Particle identification is provided by two independent Cherenkov-imaging components, the Time-Of-Propagation (TOP) counter and the Aerogel Ring-Imaging Cherenkov detector (ARICH), located in the barrel and forward endcap regions of the detector, respectively. The electromagnetic calorimeter (ECL) encases all of the previous layers, and consists of a total of 8736 scintillator crystals used primarily for the determination of the energies of charged tracks and photons. A superconducting solenoid surrounds the inner components and provides the 1.5T magnetic field required by the various sub-detectors. Finally, the $K_{L}^0$ and muon detector (KLM) forms the outermost detector layer aimed at the detection of $K_{L}^0$ mesons and muons.

3. DATASETS

The amount of data studied for this analysis corresponds to an integrated luminosity of 34.6 fb$^{-1}$. To simulate signal and background, fully simulated Monte Carlo (MC) samples of charged and neutral decays of pairs of $B$-mesons, as well as continuum $e^+e^- \rightarrow q\bar{q}$ ($q = u, d, s, c$) were used, with all samples generated alongside beam background effects including beam scattering and radiative processes. Table I lists the number of events used for each of the MC components. These samples correspond to a total integrated luminosity of 200 fb$^{-1}$.

In addition to the generic MC, dedicated $B \rightarrow X_u \ell \nu$ signal samples were used, where $X_u$ is a hadronic system resulting from the quark flavor transition $b \rightarrow u$, for the in-depth study of the signal decays and related backgrounds, and included the implementation of the hybrid modelling technique for $b \rightarrow u$ transitions described in Ref. \cite{3}. Each $B^+ \rightarrow X_u \ell \nu$ and $B^0 \rightarrow X_u \ell \nu$ sample consisted of a total of 50 million resonant (R) events containing the relevant exclusive decays as well as 50 million non-resonant (I) events corresponding
TABLE I: Number of MC events used for the analysis, equivalent to 200 fb$^{-1}$.

| $N_{\text{events used}} \times 10^6$ |
|---|
| $B^+ B^-$ | 108.0 |
| $B^0 \bar{B}^0$ | 102.0 |
| $c\bar{c}$ | 265.8 |
| $u\bar{u}$ | 321.0 |
| $s\bar{s}$ | 76.6 |
| $d\bar{d}$ | 80.2 |

to the inclusive component, simulated using the BLNP heavy-quark-effective-theory-based model [4]. These samples were then combined together and the eFFORT tool [5] was used to calculate a hybrid weight per event in three-dimensional bins of the generated lepton energy in the $B_{\text{sig}}$-frame, $E_B^\ell$, the four-momentum transfer to the leptonic system, $q^2$, and the mass of the hadronic system containing an up-quark, $M_X$, such that $H_i = R_i + w_i I_i$. The number of total hybrid events per bin, $H_i$, is the sum of the number of resonant events $R_i$ and the number of inclusive events $I_i$ scaled down by the appropriate weight $w_i$.

For the analysis, the $B \to X_u \ell \nu$ events from the generic 200 fb$^{-1}$ MC sample were removed and replaced with the equivalent amount of this hybrid re-weighted MC.

4. FULL EVENT INTERPRETATION

The Full Event Interpretation (FEI) [4] is a machine learning algorithm developed for tagged analysis at Belle II. It supports both hadronic and semi-leptonic tagging, reconstructing $B$-mesons across more than 4000 individual decay chains. The algorithm utilises the FastBDT software package that trains a series of multi-variate classifiers for each tagging channel via a number of stochastic gradient-boosted decision trees [6]. The training is performed in a hierarchical manner with final-state particles being reconstructed first from detector information. The decay channels are then built up from these particles as illustrated in Figure 1 with the reconstruction of the $B$-mesons performed last. For each $B$-meson tag candidate reconstructed by the FEI, a value of the final multi-variate classifier output, the SignalProbability, is assigned. The SignalProbability is distributed between 0 and 1, representing candidates identified as being background-like and signal-like, respectively.

Official FEI skims of both data and MC are produced centrally by the Belle II collaboration, and are available for use in analyses. These include both hadronic and semileptonic skims, and involve the application of the FEI together with a number of loose selections that aim to reduce the sample sizes with little to no loss of signal events.

For the hadronic FEI, the minimum number of charged particle tracks per event satisfying certain quality criteria is set to three. The vast majority of $B$-meson decay chains corresponding to the hadronic FEI channels include at least three charged particles, and such a cut is useful at suppressing background from non-$BB$ events. The criteria chosen include track parameter cuts to ensure close proximity to the interaction point (IP), with the distance from the center of the detector along the $z$-axis (corresponding to the solenoidal-field axis) and in the transverse plane satisfying $|z_0| < 2.0$ cm and $|d_0| < 0.5$ cm, respectively. A
FIG. 1: The hierarchical structure of the Full Event Interpretation tagging algorithm.

Detector information including tracks and energy deposits in the electromagnetic calorimeter (ECL) and $K^0_L$ and muon detector (KLM) is first used to reconstruct and train a classifier for each final-state particle. Intermediate particles are then reconstructed in stages to form the parent $B$-mesons, each with an associated $\text{SignalProbability}$. 

A minimum threshold $p_t > 0.1$ GeV is placed on the track transverse momentum \[13\]. Similar restrictions are applied to the ECL clusters in the event, with at least three clusters required within the polar angle acceptance of the CDC, $0.297 < \theta < 2.618$, that satisfy a minimum energy threshold $E > 0.1$ GeV. The total visible energy per event is required to be at least 4 GeV. The total energy deposited in the ECL is restricted to $2 \text{GeV} < E_{\text{ECL}} < 7 \text{GeV}$, however, to suppress events with an excess of energy deposits due to beam background.

The FEI typically results in many $B_{\text{tag}}$ candidates per event. The number of these candidates is reduced with selections on the beam-constrained mass, $M_{bc}$, and energy difference, $|\Delta E|$, with these defined as follows:

$$M_{bc} = \sqrt{\frac{E_{\text{beam}}^2}{4} - \vec{p}_{B_{\text{tag}}}^2}, \quad \Delta E = E_{B_{\text{tag}}} - \frac{E_{\text{beam}}}{2},$$

where $E_{\text{beam}}$ is the centre-of-mass (CMS) energy of the $e^+e^-$ system, 10.58 GeV, and $\vec{p}_{B_{\text{tag}}}$ and $E_{B_{\text{tag}}}$ are the $B_{\text{tag}}$ momentum and energy in the CMS frame, respectively. The cuts applied during the hadronic FEI skim are $M_{bc} > 5.24$ GeV and $|\Delta E| < 0.2$ GeV.

Finally, a loose cut on the $B_{\text{tag}}$ classifier output, $\text{SignalProbability} > 0.001$, provides further background rejection with little signal loss.
5. EVENT SELECTION AND ANALYSIS STRATEGY

For this analysis, the distribution of the square of the missing mass, $M^2_{\text{miss}}$, was the variable chosen for the determination of the $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ yield. We define the four-momentum of the signal $B$-meson in the CMS frame as follows,

$$p_{B_{\text{sig}}} \equiv (E_{B_{\text{sig}}}, \mathbf{p}_{B_{\text{sig}}}) = \left( \frac{m_{\Upsilon(4S)}}{2}, -\mathbf{p}_{B_{\text{tag}}} \right),$$

where $m_{\Upsilon(4S)}$ is the mass of the $\Upsilon(4S)$ meson as recorded by the PDG [7]. We set the energy of $B_{\text{sig}}$ to be half of the $\Upsilon(4S)$ rest mass, and take the $B_{\text{sig}}$ momentum to be the negative $B_{\text{tag}}$ momentum. We then define the missing four-momentum as

$$p_{\text{miss}} \equiv (E_{\text{miss}}, \mathbf{p}_{\text{miss}}) = p_{B_{\text{sig}}} - p_Y,$$

where $Y$ is a pseudo-particle formed from the combination of the pion and lepton four-momenta. The square of the missing momentum can then simply be defined as $M^2_{\text{miss}} \equiv p^2_{\text{miss}}$.

The signal selection criteria were determined using simulation, with data blinded in the signal region of $M^2_{\text{miss}} \leq 1.0$ GeV$^2$ and the agreement between data and MC was investigated in the $M^2_{\text{miss}}$ sideband. The event selections applied follow closely those from a 2013 study [8] of exclusive $B \rightarrow X_u \ell \nu$ decays using the full 711 fb$^{-1}$ Belle dataset with hadronic tagging. All selections were applied in addition to the hadronic FEI skim cuts detailed in the previous section.

At the event-level, a loose selection on the second normalised Fox-Wolfram moment [9] was applied to suppress continuum background, with $R^2 < 0.4$. To reject incorrectly reconstructed $B_{\text{tag}}$ candidates, the tag-side beam-constrained mass cut was tightened to $M_{bc} > 5.27$ GeV. The $B_{\text{tag}}$ candidate having the highest value of the SignalProbability classifier output was retained in each event.

For the reconstructed electrons and muons, cuts on the track impact parameters were applied to select those originating close to the interaction point (IP). These consisted of cuts on the $z$-axis and transverse-plane distances from the IP of $|dz| < 5$ cm and $dr < 2$ cm, respectively. Only those leptons within the acceptance of the CDC were selected. Electrons and muons were identified through cuts on the particle identification variables provided by the Belle II Analysis Software Framework (BASF2) [10]. These variables describe the likelihood that a charged track corresponds to a specified final-state particle against all other possible particles, and are defined between 0 (unlikely) and 1 (likely). The likelihoods are built from a combination of the information returned from all of the individual sub-detectors. The relevant variables and their selections were $\text{electronicID} > 0.9$ and $\text{muonID} > 0.9$. A minimum threshold on the lab-frame momentum was placed on the reconstructed leptons, with $p_{\text{lab}} > 0.3$ GeV for electrons and $p_{\text{lab}} > 0.6$ GeV for muons. The four-momentum of the reconstructed electrons was also corrected in order to account for Bremsstrahlung radiation. For any energy deposit in the ECL not associated with a charged particle track, it was checked whether the distance between this cluster and the cluster associated with an electron track was less than 0.5 mm. If this condition was met and the cluster energy was between 0.2 and 1 times the energy of the track assigned to the electron, a weight was calculated for the relation between the electron and photon. In the case of multiple
photons satisfying these conditions, only the photon with the weight corresponding to the closest relation was taken, its four-momentum was added to the electron, and the photon was excluded from the rest of the event. Finally, a single lepton was kept per event with the highest value of the electronID or muonID.

For the reconstructed charged pions, similar impact parameter cuts were applied as those for the leptons, with \( dr < 2 \text{ cm} \) and \( |dz| < 4 \text{ cm} \). The charged pion tracks were similarly only selected within the CDC acceptance, with the additional condition that a minimum number of 20 hits was recorded by the CDC. The sign of the charge of the reconstructed pion was explicitly required to be opposite that of the lepton. A cut on the relevant particle identification variable was also applied, with \( \text{pionID} > 0.5 \).

Combining the centre-of-mass frame four-momenta of the reconstructed pion and lepton together into the pseudo-particle \( Y \), a selection on the cosine of the angle between the flight direction of the signal \( B \)-meson and the \( Y \), \( \cos(\theta_{BY}) \), was made. A value of \( |\cos(\theta_{BY})| < 1 \) is expected if only a neutrino is missing in the reconstruction. However, to account for resolution effects and to avoid introducing potential bias in the background \( M_{\text{miss}}^2 \) distributions, this requirement was loosened to \( |\cos(\theta_{BY})| < 3 \). To ensure that the reconstructed lepton and pion tracks originated from the same vertex, a cut on the difference between the \( z \)-coordinates of both tracks was applied, with \( |z_\ell - z_\pi| < 1 \text{ mm} \).

For this analysis, a single \( \Upsilon(4S) \) candidate with the lowest value of the \( M_{\text{miss}}^2 \) was retained per event. All remaining tracks and clusters in the ECL after the reconstruction of the \( \Upsilon(4S) \) were combined into a single object known as the rest-of-event (ROE). Events in which additional tracks satisfying the conditions \( dr < 2 \text{ cm} \), \( |dz| < 5 \text{ cm} \) and \( p_t > 0.2 \text{ GeV} \) remained after the reconstruction of the \( \Upsilon(4S) \) were excluded. The CMS energies in the ROE corresponding to deposits of neutral particles in the ECL were summed for those clusters satisfying energy cuts of \( E > 0.1 \text{ GeV} \), \( E > 0.09 \text{ GeV} \) and \( E > 0.16 \text{ GeV} \) for the forward end-cap, barrel and backward end-cap regions respectively. A minimum threshold on this energy was placed to account for the neutrino, with \( E_{\text{miss}} > 0.3 \text{ GeV} \). The extra energy was also required to be below a maximum value of \( E_{\text{residual}} < 1.0 \text{ GeV} \).

6. RESULTS

Before unblinding the dataset, a toy MC study was performed to evaluate the expected significance of the \( B^0 \to \pi^- \ell^+ \nu_\ell \) signal in 34.6 fb\(^{-1}\). To evaluate the number of events expected in data, the number of total MC events normalised to this integrated luminosity and passing all selections was scaled down by a hadronic FEI calibration factor of 0.8301. This factor was applied in order to account for the difference in the tag-side reconstruction efficiency of the FEI between data and MC. An independent study was performed in order to evaluate this factor through fitting the lepton momentum spectrum in \( B \to X \ell \nu \) decays and taking the ratio of signal events in data and MC.

Template Probability Density Functions (PDFs) were built from signal and background \( M_{\text{miss}}^2 \) distributions in MC, with signal candidates defined as those reconstructed from MC events containing \( B^0 \to \pi^- \ell^+ \nu_\ell \) decays. The various backgrounds included the cross-feeds from \( B^0 \to \rho^- \ell^+ \nu_\ell \) and other \( B \to X_u \ell \nu \) decays, as well as candidates reconstructed from other generic \( B \bar{B} \) and continuum events. Due to low statistics, these background components were combined together into a single background PDF. Each MC template was weighted by a set of corrections to account for the difference in lepton identification efficiencies and pion and kaon fake-rates between MC and data. These corrections were evaluated per event.
based on the magnitude of the lab-frame momentum $p$ and polar angle $\theta$ of the reconstructed lepton tracks.

A set of 500 toy MC samples was generated from the combined signal and background template distributions. Each toy consisted of a $M^2_{\text{miss}}$ distribution with a total number of events defined within a range of Poissonian uncertainty on the expected number of data events. Under the signal + background hypothesis, extended unbinned maximum likelihood template fits to the $M^2_{\text{miss}}$ distributions of each toy were performed using the RooFit framework \cite{11}, with all background components combined into a single background template. A Gaussian PDF was fitted to the signal yields obtained from the fits to each toy sample and is shown in Figure 2. The mean and standard deviation of the fitted Gaussian is listed in Table II together with the expected yield as determined from the MC. The mean signal yield agrees with the predicted yield within uncertainty.

Additional extended maximum likelihood fits were then performed to the same toy MC samples under the background-only hypothesis. The likelihood ratio $\lambda$ between both fits was computed for each toy:

$$\lambda = \frac{L_B}{L_{S+B}},$$

where $L_B$ and $L_{S+B}$ are the maximised likelihoods returned by the fits to the background-only and signal + background hypotheses, respectively.

A significance estimator $\Sigma$ was subsequently defined for each toy sample based on the likelihood ratio:

$$\Sigma = \sqrt{-2 \ln \lambda}.$$

Figure 5 shows the resultant distribution of $\Sigma$ for the decays to charged pions. The median of the distribution corresponds to a significance of $5.63\sigma$.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig2.png}
\caption{Distribution of the fitted signal yields derived from fits to the $M^2_{\text{miss}}$ distributions of 500 toy MC $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ samples.}
\end{figure}
The data was subsequently unblinded in the signal region of \( M_{\text{miss}}^2 \leq 1.0 \text{ GeV}^2 \), with the resultant unfitted distribution displayed in Figure 3. For charged particle tracks in data, a factor of 1.00056 was applied to the track momentum to correct for differences between data and MC. In addition to the lepton ID corrections, each MC component was scaled down by the hadronic FEI calibration factor of 0.8301 in order to account for the difference in the FEI reconstruction efficiency between data and MC. Fairly good agreement was observed across the \( M_{\text{miss}}^2 \) distribution, including in the signal region. A clear signal peak can be seen at \( M_{\text{miss}}^2 = 0 \) for both data and MC, with all other backgrounds peaking at higher values of \( M_{\text{miss}}^2 \).

The signal and background MC templates shown in Figure 3 were then used to perform an extended unbinned maximum likelihood fit to the \( M_{\text{miss}}^2 \) distribution in data under the signal + background hypothesis. Figure 4 displays this fitted distribution and the signal yield obtained is listed in Table II. The observed yield in data agrees well with the predicted yield from MC.

The observed significance was similarly evaluated by performing an additional fit to data under the background-only hypothesis. A significance of \( 5.69 \sigma \) was observed, as shown in Figure 5.

Finally, the branching fraction for the \( B^0 \rightarrow \pi^- \ell^+ \nu \ell \) decay was extracted using the following formula:

\[
B(B^0 \rightarrow \pi^- \ell^+ \nu \ell) = \frac{N_{\text{data}}^{\text{sig}}(1 + f_{+0})}{4 \times C_F \times N_{BB} \times \epsilon},
\]

where \( N_{\text{data}}^{\text{sig}} \) is the fitted signal yield obtained from data, \( f_{+0} \) is the ratio between the
branching fractions of the decays of the \( \Upsilon(4S) \) meson to pairs of charged and neutral \( B \)-mesons \[^7\], \( CF_{FEI} \) is the FEI calibration factor, \( N_{BB} \) is the number of \( B \)-meson pairs counted in the current dataset, and \( \epsilon \) is the reconstruction efficiency. The factor of 4 present in the denominator accounts for the two \( B \)-mesons in the \( \Upsilon(4S) \) decay and the reconstruction of both light lepton flavors.

The signal efficiency was calculated from the ratio of signal events present in MC before and after all analysis selections, and was determined to be \((0.216 \pm 0.001)\%\). The values of the above parameters together with the measured branching fraction are summarised in Table III. The branching fraction agrees well with the world average value, \((1.50 \pm 0.06) \times 10^{-4}\) \[^7\].

![FIG. 4: Post-fit \( M^2_{\text{miss}} \) distribution from 34.6 fb\(^{-1}\) of data.](image)

| Predicted yield from MC | Fitted mean \( \mu \) from MC | Fitted standard deviation \( \sigma \) from MC | Observed fitted yield in data |
|-------------------------|-------------------------------|---------------------------------|-------------------------------|
| 19.83                   | 20.01 \( \pm \) 0.26         | 5.79 \( \pm \) 0.19             | 20.79 \( \pm \) 5.68         |

**TABLE II:** Mean and standard deviation of the Gaussian PDF fit to the distributions of signal yields from 500 toy MC samples. The numbers of expected and observed signal events in data are also listed.

**7. SYSTEMATIC UNCERTAINTIES**

A number of sources of systematic uncertainty were identified for this analysis and evaluated for the branching fraction measurement. The relative uncertainties for each source,
FIG. 5: Distribution of the estimator of expected significance for $B^0 \rightarrow \pi^- \ell^+ \nu$ decays across 500 toy MC samples. The observed significance in data is also plotted.

|                |             |
|----------------|-------------|
| $N_{\text{data}}^{\text{sig}}$ | $20.79 \pm 5.68$ |
| $f_{+0}$       | $1.058 \pm 0.024$ |
| $\text{CF}_{\text{FEI}}$     | $0.8301 \pm 0.0286$ |
| $N_{B\bar{B}}$   | $(37.711 \pm 0.602) \times 10^6$ |
| $\epsilon$      | $(0.216 \pm 0.001)$% |
| $\mathcal{B}(B^0 \rightarrow \pi^- \ell^+ \nu)$ | $(1.58 \pm 0.43_{\text{stat}} \pm 0.07_{\text{sys}}) \times 10^{-4}$ |

TABLE III: Measured branching fraction of $B^0 \rightarrow \pi^- \ell^+ \nu$ decays using $34.6 \text{ fb}^{-1}$ of data. The values of the parameters used in the measurement are also given.

in percent, are summarised in Table IV and included:

- $f_{+0}$: We combine the errors on the world averages for the branching fractions $\mathcal{B}(\Upsilon(4S) \rightarrow B^+ B^-)$ and $\mathcal{B}(\Upsilon(4S) \rightarrow B^0 \bar{B}^0)$ and calculate the relative uncertainty on the fraction $f_{+0}$.

- **FEI calibration**: The given uncertainty on the calibration factor for the hadronic FEI was determined taking into account multiple systematic effects in the fitting to the lepton momentum spectrum of $B \rightarrow X \ell \nu$ decays. These included uncertainties on both the branching fractions and form factors of the various semileptonic components of $B \rightarrow X \ell \nu$, the lepton ID efficiency and fake rate uncertainties, tracking uncertainties, and template uncertainties as a result of MC statistics. The relative uncertainty on the calibration factor forms the dominant source of systematic uncertainty for this analysis.

- $N_{B\bar{B}}$: The uncertainty on the number of $B\bar{B}$ events in the present dataset includes systematic effects due to uncertainties on the luminosity, beam energy spread and shift, tracking efficiency and the selection efficiency of $B\bar{B}$ events.
• **Reconstruction efficiency**: We represent the uncertainty on the reconstruction efficiency with a binomial error dependent on the size of the MC samples used for the analysis.

• **Tracking**: We assign a flat systematic uncertainty of 0.80% for each charged track. We assume complete correlation between the uncertainties of the lepton and pion tracks and obtain a total relative uncertainty of 1.60%.

• **Lepton identification**: The lepton efficiencies and pion fake rates are evaluated in bins of the lepton momentum $p$ and polar angle $\theta$, each with statistical and systematic uncertainties. The effect of these uncertainties on the signal reconstruction efficiency was determined through generating 500 variations on the nominal correction weights via Gaussian smearing. The relative uncertainty was then taken from the spread on the values of the reconstruction efficiency over all variations.

| Source of systematic uncertainty | % of $\mathcal{B}$ |
|---------------------------------|---------------------|
| $f_{+0}$                        | 1.17                |
| FEI calibration                 | 3.45                |
| $N_{BB}$                        | 1.60                |
| Reconstruction efficiency $\epsilon$ | 0.46                |
| Tracking                        | 1.60                |
| Lepton ID                       | 1.05                |
| **Total**                       | **4.44**            |

TABLE IV: Sources of systematic uncertainties and their percentages of the total measured branching fraction.

We do not currently include systematic uncertainties due to modelling $B \rightarrow X_u \ell \nu$ decays, as these are expected to be sufficiently small in comparison with the other sources listed for the current dataset.

8. SUMMARY

In summary, we have presented a re-discovery of the semi-leptonic decay $B^0 \rightarrow \pi^- \ell^+ \nu_\ell$ via hadronic tagging in 34.6 fb$^{-1}$ of Belle II data, with an observed significance of 5.69$\sigma$. A branching fraction of $(1.58 \pm 0.43_{\text{stat}} \pm 0.07_{\text{sys}}) \times 10^{-4}$ was measured for this decay, in agreement with the world average [7].

Acknowledgements

We thank the SuperKEKB group for the excellent operation of the accelerator; the KEK cryogenics group for the efficient operation of the solenoid; and the KEK computer group for on-site computing support. This work was supported by the following funding sources: Science Committee of the Republic of Armenia Grant No. 18T-1C180; Australian Research
Council and research grant Nos. DP180102629, DP170102389, DP170102204, DP150103061, FT130100303, and FT130100018; Austrian Federal Ministry of Education, Science and Research, and Austrian Science Fund No. P 31361-N36; Natural Sciences and Engineering Research Council of Canada, Compute Canada and CANARIE; Chinese Academy of Sciences and research grant No. QYZDJ-SSW-SLH011, National Natural Science Foundation of China and research grant Nos. 11521505, 11575017, 11675166, 11761141009, 11705209, and 11975076, LiaoNing Revitalization Talents Program under contract No. XLYC1807135, Shanghai Municipal Science and Technology Committee under contract No. 19ZR1403000, Shanghai Pujiang Program under Grant No. 18PJ1401000, and the CAS Center for Excellence in Particle Physics (CCEPP); the Ministry of Education, Youth and Sports of the Czech Republic under Contract No. LTT17020 and Charles University grants SVV 260448 and GAUK 404316; European Research Council, 7th Framework PIEF-GA-2013-622527, Horizon 2020 Marie Sklodowska-Curie grant agreement No. 700525 ‘NIOBE,’ and Horizon 2020 Marie Sklodowska-Curie RISE project JENNIFER2 grant agreement No. 822070 (European grants); L’Institut National de Physique Nucléaire et de Physique des Particules (IN2P3) du CNRS (France); BMBF, DFG, HGF, MPG, AvH Foundation, and Deutsche Forschungsgemeinschaft (DFG) under Germany’s Excellence Strategy – EXC2121 “Quantum Universe” – 390833306 (Germany); Department of Atomic Energy and Department of Science and Technology (India); Israel Science Foundation grant No. 2476/17 and United States-Israel Binational Science Foundation grant No. 2016113; Istituto Nazionale di Fisica Nucleare and the research grants BELLE2; Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research grant Nos. 16H03968, 16H03993, 16H06492, 16K05323, 17H01133, 17H05405, 18K03621, 18H03710, 18H05226, 19H00682, 26220706, and 26400255, the National Institute of Informatics, and Science Information NETwork 5 (SINET5), and the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan; National Research Foundation (NRF) of Korea Grant Nos. 2016R1A1A1B01010135, 2016R1A1B02012900, 2018R1A2B3003643, 2018R1A6A1A06024970, 2018R1A1A1B0704294, 2019K1-A3A7A09033840, and 2019R1I1A3A01058933, Radiation Science Research Institute, Foreign Large-size Data Hub Center of the Korea Institute of Science and Technology Information and KREONET/GLORIAD; Universiti Malaya RU grant, Akademi Sains Malaysia and Ministry of Education Malaysia; Frontiers of Science Program contracts FOINS-296, CB-221329, CB-236394, CB-254409, and CB-180023, and SEP-CINVESTAV research grant 237 (Mexico); the Polish Ministry of Science and Higher Education and the National Science Center; the Ministry of Science and Higher Education of the Russian Federation, Agreement 14.W03.31.0026; University of Tabuk research grants S-1440-0321, S-0256-1438, and S-0280-1439 (Saudi Arabia); Slovenian Research Agency and research grant Nos. J1-9124 and P1-0135; Agencia Estatal de Investigacion, Spain grant Nos. FPA2014-55613-P and FPA2017-84445-P, and CIDEVENT/2018/020 of Generalitat Valenciana; Ministry of Science and Technology and research grant Nos. MOST106-2112-M-002-005-MY3 and MOST107-2119-M-002-035-MY3, and the Ministry of Education (Taiwan); Thailand Center of Excellence in Physics; TUBITAK ULAKBIM (Turkey); Ministry of Education and Science of Ukraine; the US National Science Foundation and research grant Nos. PHY-1807007 and PHY-1913789, and the US Department of Energy and research grant Nos. DE-AC06-76RL01830, DE-SC0007983, DE-SC0009824, DE-SC0009973, DE-SC0010073, DE-SC0010118, DE-SC0010504, DE-SC0011784, DE-SC0012704; and the National Foundation for Science and Technology Development (NAFOSTED) of Vietnam under contract No
References

[1] T. Keck et al., The Full Event Interpretation – An exclusive tagging algorithm for the Belle II experiment, Comput. Softw. Big Sci. 3 (2019) arXiv:1807.08680 [hep-ex]
[2] T. Abe, Belle II Collaboration, Belle II Technical Design Report, arXiv:1011.0352 [physics.ins-det]
[3] D. J. Ramirez, C. and G. Burdman, Semileptonic $b \rightarrow u$ Decay, Phys. Rev. D41 (1990)
[4] N. M. Lange, B. O. and G. Paz, Theory of Charmless Inclusive B Decays and the Extraction of $V_{ub}$, Phys. Rev. D72 (2005) 073006.
[5] M. Prim, b2-hive/eFFORT v0.1.0, July, 2020. https://doi.org/10.5281/zenodo.3965699
[6] T. Keck, FastBDT: A speed-optimized and cache-friendly implementation of stochastic gradient-boosted decision trees for multivariate classification, arXiv:1609.06119 [cs.LG].
[7] M. Tanabashi et al., Particle Data Group, Review of Particle Physics, Phys. Rev. D98 (2018) 030001.
[8] A. Sibidanov et al., Belle Collaboration, Study of Exclusive $B \rightarrow X_u \ell \nu$ Decays and Extraction of $|V_{ub}|$ using Full Reconstruction Tagging at the Belle Experiment, Phys. Rev. D88 (2013) 032005 arXiv:1306.2781 [hep-ex]
[9] G. C. Fox and S. Wolfram, Observables for the Analysis of Event Shapes in $e^+e^-$ Annihilation and Other Processes, Phys. Rev. Lett. 41 (1978) 1581.
[10] T. Kuhr et al., The Belle II Core Software, Comput. Softw. Big Sci. 3 (2019)
[11] W. Verkerke and D. P. Kirkby, The RooFit toolkit for data modeling, eConf C0303241 (2003) MOLT007.
[12] Charge conjugate processes are implied for all quoted decays of $B$-mesons throughout this paper.
[13] Throughout this paper, we use the convention $c = 1$ for the speed of light.