2017

# Search for B− → μ −ν¯ μ Decays at the Belle Experiment

A. Sibidanov et al.
*Belle Collaboration*

Ratnappuli L. Kulasiri
*Kennesaw State University*, rkulasir@kennesaw.edu

Belle Preprint 2017-17
KEK Preprint 2017-21



# Search for $B^- \to \mu^- \bar{\nu}_\mu$ Decays at the Belle Experiment

A. Sibidanov,[70,*] K. E. Varvell,[70] I. Adachi,[15,11] H. Aihara,[78] S. Al Said,[71,33] D. M. Asner,[61] T. Aushev,[49] R. Ayad,[71] V. Babu,[72] I. Badhrees,[71,32] S. Bahinipati,[18] A. M. Bakich,[70] V. Bansal,[61] E. Barberio,[46] P. Behera,[20] B. Bhuyan,[19] J. Biswal,[28] A. Bozek,[56] M. Bračko,[44,28] T. E. Browder,[14] D. Červenkov,[4] P. Chang,[55] V. Chekelian,[45] A. Chen,[53] B. G. Cheon,[13] K. Chilikin,[39,48] K. Cho,[34] S.-K. Choi,[12] Y. Choi,[69] D. Cinabro,[83] T. Czank,[76] N. Dash,[18] S. Di Carlo,[83] Z. Doležal,[4] Z. Drásal,[4] D. Dutta,[72] S. Eidelman,[3,59] D. Epifanov,[3,59] J. E. Fast,[61] T. Ferber,[7] B. G. Fulsom,[61] V. Gaur,[82] N. Gabyshev,[3,59] A. Garmash,[3,59] P. Goldenzweig,[30] D. Greenwald,[74] Y. Guan,[21,15] E. Guido,[26] J. Haba,[15,11] K. Hayasaka,[58] H. Hayashii,[52] M. T. Hedges,[14] S. Hirose,[50] W.-S. Hou,[55] C.-L. Hsu,[46] T. Iijima,[51,50] K. Inami,[50] G. Inguglia,[7] A. Ishikawa,[76] R. Itoh,[15,11] M. Iwasaki,[60] Y. Iwasaki,[15] W. W. Jacobs,[21] I. Jaegle,[8] H. B. Jeon,[37] Y. Jin,[78] K. K. Joo,[5] T. Julius,[46] J. Kahn,[41] A. B. Kaliyar,[20] K. H. Kang,[37] G. Karyan,[7] T. Kawasaki,[58] C. Kiesling,[45] D. Y. Kim,[67] J. B. Kim,[35] S. H. Kim,[13] Y. J. Kim,[34] K. Kinoshita,[6] P. Kodyš,[4] S. Korpar,[44,28] D. Kotchetkov,[14] P. Križan,[40,28] P. Krokovny,[3,59] T. Kuhr,[41] R. Kulasiri,[31] R. Kumar,[63] A. Kuzmin,[3,59] Y.-J. Kwon,[85] J. S. Lange,[9] I. S. Lee,[13] C. H. Li,[46] L. Li,[65] L. Li Gioi,[45] J. Libby,[20] D. Liventsev,[82,15] M. Lubej,[28] T. Luo,[62] M. Masuda,[77] T. Matsuda,[47] M. Merola,[25] K. Miyabayashi,[52] H. Miyata,[58] R. Mizuk,[39,48,49] G. B. Mohanty,[72] H. K. Moon,[35] T. Mori,[50] R. Mussa,[26] E. Nakano,[60] M. Nakao,[15,11] T. Nanut,[28] K. J. Nath,[19] Z. Natkaniec,[56] M. Nayak,[83,15] M. Niiyama,[36] N. K. Nisar,[62] S. Nishida,[15,11] S. Ogawa,[75] S. Okuno,[29] H. Ono,[57,58] P. Pakhlov,[39,48] G. Pakhlova,[39,49] B. Pal,[6] C.-S. Park,[85] C. W. Park,[69] H. Park,[37] S. Paul,[74] T. K. Pedlar,[42] R. Pestotnik,[28] L. E. Piilonen,[82] M. Ritter,[41] A. Rostomyan,[7] M. Rozanska,[56] Y. Sakai,[15,11] M. Salehi,[43,41] S. Sandilya,[6] Y. Sato,[50] V. Savinov,[62] O. Schneider,[38] G. Schnell,[1,17] C. Schwanda,[23] Y. Seino,[58] K. Senyo,[84] M. E. Sevior,[46] V. Shebalin,[3,59] C. P. Shen,[2] T.-A. Shibata,[79] J.-G. Shiu,[55] F. Simon,[45,73] A. Sokolov,[24] E. Solovieva,[39,49] M. Starič,[28] J. F. Strube,[61] J. Stypula,[56] M. Sumihama,[10] K. Sumisawa,[15,11] T. Sumiyoshi,[80] M. Takizawa,[66,16,64] U. Tamponi,[26,81] K. Tanida,[27] F. Tenchini,[46] K. Trabelsi,[15,11] M. Uchida,[79] S. Uehara,[15,11] T. Uglov,[39,49] Y. Unno,[13] S. Uno,[15,11] P. Urquijo,[46] C. Van Hulse,[1] G. Varner,[14] V. Vorobyev,[3,59] C. H. Wang,[54] M.-Z. Wang,[55] P. Wang,[22] M. Watanabe,[58] S. Watanuki,[76] E. Widmann,[68] E. Won,[35] Y. Yamashita,[57] H. Ye,[7] J. Yelton,[8] C. Z. Yuan,[22] Y. Yusa,[58] Z. P. Zhang,[65] V. Zhilich,[3,59] V. Zhukova,[39,48] V. Zhulanov,[3,59] and A. Zupanc[40,28]

(The Belle Collaboration)

[1] *University of the Basque Country UPV/EHU, 48080 Bilbao*
[2] *Beihang University, Beijing 100191*
[3] *Budker Institute of Nuclear Physics SB RAS, Novosibirsk 630090*
[4] *Faculty of Mathematics and Physics, Charles University, 121 16 Prague*
[5] *Chonnam National University, Kwangju 660-701*
[6] *University of Cincinnati, Cincinnati, Ohio 45221*
[7] *Deutsches Elektronen–Synchrotron, 22607 Hamburg*
[8] *University of Florida, Gainesville, Florida 32611*
[9] *Justus-Liebig-Universität Gießen, 35392 Gießen*
[10] *Gifu University, Gifu 501-1193*
[11] *SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193*
[12] *Gyeongsang National University, Chinju 660-701*
[13] *Hanyang University, Seoul 133-791*
[14] *University of Hawaii, Honolulu, Hawaii 96822*
[15] *High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801*
[16] *J-PARC Branch, KEK Theory Center, High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801*
[17] *IKERBASQUE, Basque Foundation for Science, 48013 Bilbao*
[18] *Indian Institute of Technology Bhubaneswar, Satya Nagar 751007*
[19] *Indian Institute of Technology Guwahati, Assam 781039*
[20] *Indian Institute of Technology Madras, Chennai 600036*
[21] *Indiana University, Bloomington, Indiana 47408*
[22] *Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049*

[23]*Institute of High Energy Physics, Vienna 1050*
[24]*Institute for High Energy Physics, Protvino 142281*
[25]*INFN - Sezione di Napoli, 80126 Napoli*
[26]*INFN - Sezione di Torino, 10125 Torino*
[27]*Advanced Science Research Center, Japan Atomic Energy Agency, Naka 319-1195*
[28]*J. Stefan Institute, 1000 Ljubljana*
[29]*Kanagawa University, Yokohama 221-8686*
[30]*Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, 76131 Karlsruhe*
[31]*Kennesaw State University, Kennesaw, Georgia 30144*
[32]*King Abdulaziz City for Science and Technology, Riyadh 11442*
[33]*Department of Physics, Faculty of Science, King Abdulaziz University, Jeddah 21589*
[34]*Korea Institute of Science and Technology Information, Daejeon 305-806*
[35]*Korea University, Seoul 136-713*
[36]*Kyoto University, Kyoto 606-8502*
[37]*Kyungpook National University, Daegu 702-701*
[38]*École Polytechnique Fédérale de Lausanne (EPFL), Lausanne 1015*
[39]*P.N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow 119991*
[40]*Faculty of Mathematics and Physics, University of Ljubljana, 1000 Ljubljana*
[41]*Ludwig Maximilians University, 80539 Munich*
[42]*Luther College, Decorah, Iowa 52101*
[43]*University of Malaya, 50603 Kuala Lumpur*
[44]*University of Maribor, 2000 Maribor*
[45]*Max-Planck-Institut für Physik, 80805 München*
[46]*School of Physics, University of Melbourne, Victoria 3010*
[47]*University of Miyazaki, Miyazaki 889-2192*
[48]*Moscow Physical Engineering Institute, Moscow 115409*
[49]*Moscow Institute of Physics and Technology, Moscow Region 141700*
[50]*Graduate School of Science, Nagoya University, Nagoya 464-8602*
[51]*Kobayashi-Maskawa Institute, Nagoya University, Nagoya 464-8602*
[52]*Nara Women's University, Nara 630-8506*
[53]*National Central University, Chung-li 32054*
[54]*National United University, Miao Li 36003*
[55]*Department of Physics, National Taiwan University, Taipei 10617*
[56]*H. Niewodniczanski Institute of Nuclear Physics, Krakow 31-342*
[57]*Nippon Dental University, Niigata 951-8580*
[58]*Niigata University, Niigata 950-2181*
[59]*Novosibirsk State University, Novosibirsk 630090*
[60]*Osaka City University, Osaka 558-8585*
[61]*Pacific Northwest National Laboratory, Richland, Washington 99352*
[62]*University of Pittsburgh, Pittsburgh, Pennsylvania 15260*
[63]*Punjab Agricultural University, Ludhiana 141004*
[64]*Theoretical Research Division, Nishina Center, RIKEN, Saitama 351-0198*
[65]*University of Science and Technology of China, Hefei 230026*
[66]*Showa Pharmaceutical University, Tokyo 194-8543*
[67]*Soongsil University, Seoul 156-743*
[68]*Stefan Meyer Institute for Subatomic Physics, Vienna 1090*
[69]*Sungkyunkwan University, Suwon 440-746*
[70]*School of Physics, University of Sydney, New South Wales 2006*
[71]*Department of Physics, Faculty of Science, University of Tabuk, Tabuk 71451*
[72]*Tata Institute of Fundamental Research, Mumbai 400005*
[73]*Excellence Cluster Universe, Technische Universität München, 85748 Garching*
[74]*Department of Physics, Technische Universität München, 85748 Garching*
[75]*Toho University, Funabashi 274-8510*
[76]*Department of Physics, Tohoku University, Sendai 980-8578*
[77]*Earthquake Research Institute, University of Tokyo, Tokyo 113-0032*
[78]*Department of Physics, University of Tokyo, Tokyo 113-0033*
[79]*Tokyo Institute of Technology, Tokyo 152-8550*
[80]*Tokyo Metropolitan University, Tokyo 192-0397*
[81]*University of Torino, 10124 Torino*
[82]*Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061*
[83]*Wayne State University, Detroit, Michigan 48202*
[84]*Yamagata University, Yamagata 990-8560*
[85]*Yonsei University, Seoul 120-749*

We report the result of a search for the decay $B^- \to \mu^- \bar{\nu}_\mu$. The signal events are selected based on the presence of a high momentum muon and the topology of the rest of the event showing properties of a generic $B$-meson decay, as well as the missing energy and momentum being consistent with the hypothesis of a neutrino from the signal decay. We find a 2.4 standard deviation excess above background including systematic uncertainties, which corresponds to a branching fraction of $\mathcal{B}(B^- \to \mu^- \bar{\nu}_\mu) = (6.46 \pm 2.22 \pm 1.60) \times 10^{-7}$ or a frequentist 90% confidence level interval on the $B^- \to \mu^- \bar{\nu}_\mu$ branching fraction of $[2.9, 10.7] \times 10^{-7}$. This result is obtained from a 711 fb$^{-1}$ data sample that contains $772 \times 10^6$ $B\bar{B}$ pairs, collected near the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB asymmetric-energy $e^+e^-$ collider.

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

In the Standard Model (SM), the branching fraction for the purely leptonic decay of a $B^-$ meson [1], assuming a massless neutrino, is:

$$\mathcal{B}(B^- \to \ell^- \bar{\nu}_\ell) = \frac{G_F^2 m_B m_\ell^2}{8\pi} \left(1 - \frac{m_\ell^2}{m_B^2}\right)^2 f_B^2 |V_{ub}|^2 \tau_B, \quad (1)$$

where $G_F$ is the Fermi constant, $m_B$ and $m_\ell$ are the masses of the $B$ meson and charged lepton, respectively, $f_B$ is the $B$-meson decay constant obtained from theory, $\tau_B$ is the lifetime of the $B$ meson and $V_{ub}$ is the CKM matrix element governing the coupling between $u$ and $b$ quarks. The FLAG [2] average of lattice QCD calculations gives $f_B = 0.186 \pm 0.004$ GeV, and the world-average value of $\tau_B$ is $1.638 \pm 0.004$ ps [3]. For the value of $|V_{ub}|$, we repeat the fit procedure described in Ref. [4], equipped with the most recent lattice QCD calculation by the FNAL/MILC collaborations [5] that provides a tight constraint on the hadronic form-factor $f_+(q^2)$ governing exclusive $\bar{B}^0 \to \pi^+ \ell^- \bar{\nu}_\ell$ decays. The form-factor parameters for $\bar{B}^0 \to \pi^+ \ell^- \bar{\nu}_\ell$ decay are also obtained with this procedure. The value of $|V_{ub}|$ thus obtained is $|V_{ub}| \times 10^3 = 3.736 \pm 0.142$ with fit quality $\chi^2 = 47.9$ for 45 degrees of freedom. Using these values as input parameters for Eq. 1, the expected branching fractions for $B^- \to \ell^- \bar{\nu}_\ell$ decays are displayed in Table I. Also shown in the Table are the expected event yields for $B^- \to \ell^- \bar{\nu}_\ell$ decays in the full Belle data set, where we use $\mathcal{B}(\Upsilon(4S) \to B^+ B^-) = 0.514 \pm 0.006$ [3].

TABLE I: The expected branching fractions and event yields in the full Belle data sample of $772 \times 10^6$ $B\bar{B}$ events for the decay $B^- \to \ell^- \bar{\nu}_\ell$.

| $\ell$ | $\mathcal{B}_{\rm SM}$ | $N_{\rm SM}^{\rm Belle}$ |
|---|---|---|
| $\tau$ | $(8.45 \pm 0.70) \times 10^{-5}$ | $(670 \pm 57) \times 10^2$ |
| $\mu$ | $(3.80 \pm 0.31) \times 10^{-7}$ | $301 \pm 25$ |
| $e$ | $(8.89 \pm 0.73) \times 10^{-12}$ | $0.0071 \pm 0.0006$ |

Due to the relatively small theoretical uncertainties within the SM framework, $B^- \to \ell^- \bar{\nu}_\ell$ decays are good candidates for testing SM predictions and searching for phenomena that might modify them. For instance, the effects of charged Higgs bosons in two-Higgs-doublet models of type-II [6], the R-parity-violating Minimal Supersymmetric Standard Model (MSSM) [7], or leptoquarks [8] may significantly change the $B^- \to \ell^- \bar{\nu}_\ell$ decay rates.

Moreover, by taking the ratios of purely leptonic $B^-$ decays, most of the input parameters in Eq. 1 cancel and very precise values are predicted. Predictions of the ratios $\mathcal{B}(B^- \to \tau^- \bar{\nu}_\tau)/\mathcal{B}(B^- \to e^- \bar{\nu}_e)$ and $\mathcal{B}(B^- \to \tau^- \bar{\nu}_\tau)/\mathcal{B}(B^- \to \mu^- \bar{\nu}_\mu)$ obtained within a general MSSM at large $\tan\beta$ [9] with heavy squarks [10] deviate from the SM expectations and the deviation can be as large as an order of magnitude in the grand unified theory framework [11].

There have been several searches for the decay $B^- \to \mu^- \bar{\nu}_\mu$ to date [12–16] and no evidence of the decay has been found, with the most stringent limit of $\mathcal{B}(B^- \to \mu^- \bar{\nu}_\mu) < 1.0 \times 10^{-6}$ at 90% confidence level set by the *BABAR* collaboration using an untagged method [14].

In this article, we present a search for the decay $B^- \to \mu^- \bar{\nu}_\mu$ that also uses the untagged method. This study is based on a 711 fb$^{-1}$ data sample that contains $(772 \pm 11) \times 10^6$ $B\bar{B}$ pairs, collected with the Belle detector at the KEKB asymmetric-energy $e^+e^-$ (3.5 on 8 GeV) collider [17] operating at the $\Upsilon(4S)$ resonance.

The Belle detector is a large-solid-angle magnetic spectrometer that consists of a silicon vertex detector (SVD), a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), a barrel-like arrangement of time-of-flight scintillation counters (TOF) and an electromagnetic calorimeter comprised of CsI(Tl) crystals (ECL) located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux-return yoke located outside of the coil is instrumented to detect $K_L^0$ mesons and to identify muons (KLM). The detector is described in detail elsewhere [18]. Two inner detector configurations were used. A 2.0 cm beampipe and a 3-layer silicon vertex detector were used for the first sample of $152 \times 10^6$ $B\bar{B}$ pairs, while a 1.5 cm beampipe, a 4-layer silicon detector and a small-cell inner drift chamber were used to record the remaining $620 \times 10^6$ $B\bar{B}$ pairs [19].

The data were collected at a center-of-mass energy of 10.58 GeV, corresponding to the $\Upsilon(4S)$ resonance. The size of the data sample is equivalent to an integrated luminosity of 711 fb$^{-1}$. We also utilise a sample of 79 fb$^{-1}$

collected below the $B\bar{B}$ threshold to characterize the contribution of the $e^+e^- \to q\bar{q}$ process, so-called continuum, where $q$ is either a $u$, $d$, $s$, or $c$ quark; this is one of the major backgrounds.

We use Monte Carlo (MC) samples based on the detailed detector geometry description implemented with the `GEANT3` package [20] to establish the analysis technique and study major backgrounds. Events with $B$-meson decays are generated using `EvtGen` [21]. The generated samples include $2 \times 10^6$ signal events, a sample of generic $B\bar{B}$ decays corresponding to ten times the integrated luminosity of the data, continuum corresponding to six times the data, $\bar{B} \to X_u \ell^- \bar{\nu}_\ell$ decays corresponding to twenty times the data, other $B$ decays with probability $\lesssim 4 \times 10^{-4}$ corresponding to fifty times the data, and $e^+e^- \to \tau^+\tau^-$ corresponding to five times the data, as well as other QED and two-photon processes with various multiples of the data. The simulation accounts for the evolution in background conditions and beam collision parameters. Final-state radiation from charged particles is modelled using the PHOTOS package [22].

MC samples for one of the largest backgrounds from $B$ decays, charmless semileptonic decays, are generated according to the number of $B\bar{B}$ pairs in data, scaled 20 times, assuming inclusive semileptonic branching fractions of $\mathcal{B}(\bar{B}^0 \to X_u^+ \ell^- \bar{\nu}_\ell) = 1.709 \times 10^{-3}$ and $\mathcal{B}(B^- \to X_u^0 \ell^- \bar{\nu}_\ell) = 1.835 \times 10^{-3}$. Samples with $\bar{B} \to \pi \ell^- \bar{\nu}_\ell$, $\bar{B} \to \rho \ell^- \bar{\nu}_\ell$, and $B^- \to \omega \ell^- \bar{\nu}_\ell$ decays are modelled using Light Cone Sum Rule form-factor predictions [23, 24]. Other decays to exclusive meson states are modelled using the updated quark model by Isgur-Scora-Grinstein-Wise [25]. The inclusive component of charmless semileptonic decays is modelled to leading order in $\alpha_s$ based on a prediction in the Heavy-Quark Expansion framework [26]. The fragmentation process of the resulting parton to the final hadron state is modelled using the PYTHIA6.2 package [27].

In addition, $8 \times 10^6$ $\bar{B} \to \pi \ell^- \bar{\nu}_\ell$ MC events are generated uniformly as a function of $q^2$. These events are reweighted to the most recent lattice QCD form-factor calculation, in order to decrease MC statistical fluctuations at high $q^2$ and to study the behavior of the fit procedure described below when form-factors are varied within uncertainties.

Finally, $10^6$ events of the three-body decay $B^- \to \mu^- \bar{\nu}_\mu \gamma$ are generated with photon energy above 25 MeV in the $B$ decay frame with the form-factor parameters $R = 3$ and $m_b = 5\,\mathrm{GeV}$ based on the work in Ref. [28].

The muon in $B^- \to \mu^- \bar{\nu}_\mu$ decay is monochromatic in the absence of radiation, with an energy of half the $B$-meson rest mass energy in the $B$-meson rest frame. In the $\Upsilon(4S)$ center-of-mass frame, where the $B$ meson is in motion, the boost smears the momentum of the muon, $p^*_\mu$, to the range (2.476, 2.812) GeV/$c$. We select well-reconstructed muon candidates in the wider region of (2.2, 4.0) GeV/$c$ to include enough data to validate the analysis procedure and estimate backgrounds. A blind analysis is performed with the $\Upsilon(4S)$ data in the $p^*_\mu$ interval (2.45, 2.85) GeV/$c$ excluded until the analysis procedure has been finalized. Signal muons are identified by a standard procedure based on their penetration range and degree of transverse scattering in the KLM detector with an efficiency of $\sim 90\%$ [29]. An additional selection is applied with information from the CDC, ECL, ACC, and TOF subdetectors, combined using an artificial neural network, to reject the charged-kaon muonic decay in flight. Background suppression of 33% is achieved by this procedure, with a signal-muon selection efficiency of 97%.

Charged particles, including the signal muon candidate, are required to originate from the region near the interaction point (IP) of the electron and positron beams. This region is defined by $|z_{\rm PCA}| < 2$ cm and $r_{\rm PCA} < 0.5$ cm, where $z_{\rm PCA}$ is the distance of the point of closest approach (PCA) from the IP along the $z$ axis (opposite the positron beam) and $r_{\rm PCA}$ is the distance from this axis in the transverse plane. The charged daughters of reconstructed long-lived neutral particles (converted $\gamma$, $K^0_S$, and $\Lambda$) are included in this list even if they fail the IP selection. All other charged particles are ignored. We discard the event if the total momentum of these particles exceeds 1.3 GeV/$c$ to suppress the background from mis-reconstructed long-lived neutral particles.

Each surviving track that is not classified as a long-lived neutral-particle daughter is assigned a unique identity. Electrons are identified using the ratio of the energy detected in the ECL to the track momentum, the ECL shower shape, position matching between the track and ECL cluster, the energy loss in the CDC, and the response of the ACC [30]. Muons are identified as described earlier for the signal muon candidates. Pions, kaons and protons are identified using the responses of the CDC, ACC, and TOF. In the expected momentum region for particles from $B$-meson decays, charged leptons are identified with an efficiency of about 75% while the probability to misidentify a pion as an electron (muon) is 1.9% (5%). Charged pions (kaons, protons) are selected with an efficiency of 86% (75%, 98%) and a pion (kaon, proton) misidentification probability of 6% (13%, 72%).

Photon candidates are selected using a polar-angle-dependent energy threshold chosen such that a photon with energy above (below) the threshold is more likely to originate from $B$-meson decay (calorimeter noise). In the barrel calorimeter, the energy threshold is about 40 MeV; in the forward and backward endcaps, it rises to 110 MeV and 150 MeV, respectively. Additionally, we require the total energy deposition in the calorimeter not associated with charged particles nor recognized as photons to be under 0.6 GeV.

The neutrino in $B^- \to \mu^- \bar{\nu}_\mu$ decay is not detected. The photons and surviving charged particles other than the signal muon should come from the companion $B$ me-

son in the $e^+e^- \to \Upsilon(4S) \to B^+B^-$ process. We select companion $B$ meson candidates that have invariant mass close to the nominal $B$-meson mass and total energy close to the nominal $B$-meson energy from the $\Upsilon(4S) \to B\bar{B}$ decay. These quantities are represented by the beam-constrained mass and energy

$$M_{\rm bc} = \sqrt{E^2_{\rm beam}/c^4 - |\sum_i \vec{\mathbf{p}}^*_i/c|^2}, \quad (2)$$

$$E_B = \sum_i \sqrt{(m_i c^2)^2 + |\vec{\mathbf{p}}^*_i c|^2}, \quad (3)$$

where $E_{\rm beam}$ is the beam energy in the $\Upsilon(4S)$ center-of-mass frame, and $\vec{\mathbf{p}}^*_i$ and $m_i$ are the center-of-mass frame momentum and mass, respectively, of the $i^{\rm th}$ particle that makes up the accompanying $B$-meson candidate. We retain events that satisfy $M_{\rm bc} > 5.1\,{\rm GeV}/c^2$ and $-3\,{\rm GeV} < E_B - E_{\rm beam} < 2\,{\rm GeV}$.

To exploit the jet-like structure of non-$B\bar{B}$ background, where particles tend to be produced collinearly, we define the direction $\hat{n}$ of the thrust axis by maximizing the quantity

$$\frac{\sum_i (\hat{n} \cdot \vec{\mathbf{p}}^*_i)^2}{\sum_i |\vec{\mathbf{p}}^*_i|^2}, \quad (4)$$

while satisfying the condition $\hat{n} \cdot (\sum_i \vec{\mathbf{p}}^*_i) > 0$. We require $\hat{n} \cdot \hat{p}^*_\mu > -0.8$, where $\hat{p}^*_\mu$ is the signal-muon direction, to remove muons collinear with the other particles in the event.

The missing energy of a neutrino from semileptonic decays of $B$ or $D$ mesons can be similar to that of the signal, and an excess of reconstructed charged leptons is a signature of these decays. We therefore require no more than one additional lepton in the event besides the signal muon.

The information from the KLM detector subsystem is also used to improve signal purity. We require no more than one $K^0_L$ cluster in the KLM and no $K^0_L$ clusters associated with ECL clusters. This selection rejects about 24% of background events and keeps about 90% of signal. The $K^0_L$ detection efficiency is calibrated using a $D^0 \to \phi K^0_S$ control sample.

The total signal selection efficiency for $B^- \to \mu^- \bar{\nu}_\mu$ decays is estimated at this stage to be around 38%, with an expected signal yield of $115 \pm 9$.

After all of the selections described above are applied, the remaining background is still more than three orders of magnitude larger than the expected signal yield. A multivariate data analysis is employed to further separate signal from background. We combine various kinematic parameters of an event into a single variable $o_{\rm nn}$ using an artificial neural network. We choose 14 input parameters that are uncorrelated with the absolute value of the muon momentum, and that collectively yield the best signal to background ratio. These parameters are five event-shape moments, the polar angle of the missing momentum vector, the angle between the thrust axis and the signal-muon direction, the energy difference $E_B - E_{\rm beam}$, the angle between the signal-muon direction and the thrust axis calculated using only photons, the angle between the momentum of the companion $B$ meson and the signal-muon direction, the $z$-axis distance between the signal muon's $z_{\rm PCA}$ and the reconstructed vertex of the companion $B$ meson, the square of the thrust as defined in Eq. (4), the sum of charges of charged particles in an event, and the polar angle of the muon momentum vector.

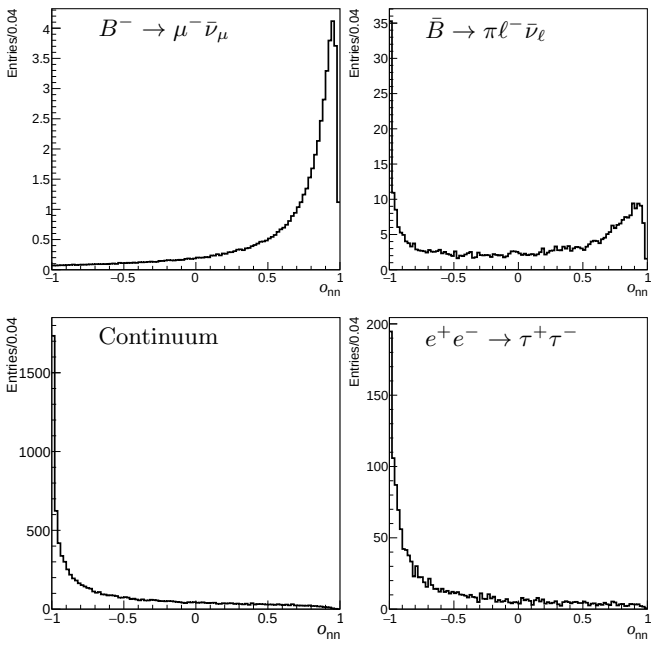


FIG. 1: The distributions of the neural network output variable for the signal and major background processes predicted by MC in the signal-enhanced region $2.644\,{\rm GeV}/c < p^*_\mu < 2.812\,{\rm GeV}/c$.

The employed configuration of the network consists of the input layer and two hidden layers having 56 and 28 neurons and the tanh activation function; in total, it has 2465 parameters to optimize. The MC sample is divided into equal training and testing parts with almost 2 million events in each. The distributions of the neural network output variables in the signal-enhanced momentum region are shown in Fig. 1. The only background components peaking in the signal region are $\bar{B} \to \pi \ell^- \bar{\nu}_\ell$ and, much less prominently, $\bar{B} \to \rho \ell^- \bar{\nu}_\ell$. All other major backgrounds decrease significantly approaching the $o_{\rm nn} \sim 1$ region and do not have a peaking behavior in the $o_{\rm nn}$ variable that can mimic the signal.

The signal yield is extracted by a binned maximum-likelihood fit in the $p^*_\mu$-$o_{\rm nn}$ plane using the method de-

scribed in Ref. [31], taking into account the uncertainty arising from the finite number of events in the template MC histograms. The fit region covers muon momenta from 2.2 to 4 GeV/$c$ with 50 MeV/$c$ bins and the full range of the $o_{\rm nn}$ variable from $-1$ to 1 with 0.04 bins. The region at high muon momentum $p^*_\mu$ and high $o_{\rm nn}$ is sparsely populated; to avoid bins with zero or a few events, which are undesirable for the fit method employed, we increased the bin size in this region. The fine binning in the signal region is preserved. After the rebinning, the $p^*_\mu$-$o_{\rm nn}$ histogram is reduced from 1800 to 1226 bins. The fit method tends to scale low-populated templates to improve the fit to data; because of this, background components with the predicted fraction of under 1% of the total number of events are fixed in the fit to the MC prediction. The fitted-yield components are the signal, $\bar{B} \to \pi\ell^-\bar{\nu}_\ell$, $\bar{B} \to \rho\ell^-\bar{\nu}_\ell$, the rest of the charmless semileptonic decays, $B\bar{B}$, $c\bar{c}$, $uds$, $\tau^+\tau^-$, and $e^+e^-\mu^+\mu^-$. The fixed-yield components are $\mu^+\mu^-$, $e^+e^-e^+e^-$, $e^+e^-u\bar{u}$, $e^+e^-s\bar{s}$, and $e^+e^-c\bar{c}$.

To obtain the signal branching fraction, we fit the ratio $R = N_{B\to\mu\bar{\nu}_\mu}/N_{B\to\pi\mu\bar{\nu}_\mu}$. This ratio also helps to reliably estimate the fit uncertainty. The result of the fit is $R = (1.66 \pm 0.57) \times 10^{-2}$, which is equivalent to a signal yield of $N_{B\to\mu\bar{\nu}_\mu} = 195 \pm 67$ and the branching fraction ratio of $\mathcal{B}(B^- \to \mu^-\bar{\nu}_\mu)/\mathcal{B}(\bar{B} \to \pi\ell^-\bar{\nu}_\ell) = (4.45 \pm 1.53_{\rm stat}) \times 10^{-3}$. This result can be compared to the MC prediction of this ratio $R_{\rm MC} = 114.6/11746 = 0.976 \times 10^{-2}$, obtained assuming $\mathcal{B}(B \to \mu\bar{\nu}_\mu) = 3.80 \times 10^{-7}$ and $\mathcal{B}(\bar{B} \to \pi\ell^-\bar{\nu}_\ell) = 1.45 \times 10^{-4}$ (the PDG average [3]). The fitted value of $R$ results in the branching fraction $\mathcal{B}(B \to \mu\bar{\nu}_\mu) = (6.46 \pm 2.22) \times 10^{-7}$, where the quoted uncertainty is statistical only. The statistical significance of the signal is $3.4\sigma$, determined from the likelihood ratio of the fits with a free signal component and with the signal component fixed to zero. The fit result of the reference process $\bar{B} \to \pi\ell^-\bar{\nu}_\ell$ agrees with the MC prediction to better than 10%. The projections of the fitted distribution in the signal-enhanced regions are shown in Fig. 2. The fit qualities of the displayed projections are $\chi^2/{\rm ndf} = 27.6/16$ (top panel) and $\chi^2/{\rm ndf} = 29.1/25$ (bottom panel), taking into account only data uncertainties.

The double ratio $R/R_{\rm MC}$ benefits from substantial cancellation of the systematic uncertainties from muon identification, lepton and neutral-kaon vetos and the companion $B$-meson decay mis-modelling, as well as partially cancelling trigger uncertainties and possible differences in the distribution of the $o_{\rm nn}$ variable.

In the signal region, the main background contribution comes from charmless semileptonic decays; in particular, the main components $\bar{B} \to \pi\ell^-\bar{\nu}_\ell$ and $\bar{B} \to \rho\ell^-\bar{\nu}_\ell$, which peak at high $o_{\rm nn}$ values, are carefully studied. With soft and undetected hadronic recoil, these decays are kinematically indistinguishable from the signal in an untagged analysis. For the $\bar{B} \to \pi\ell^-\bar{\nu}_\ell$ component,

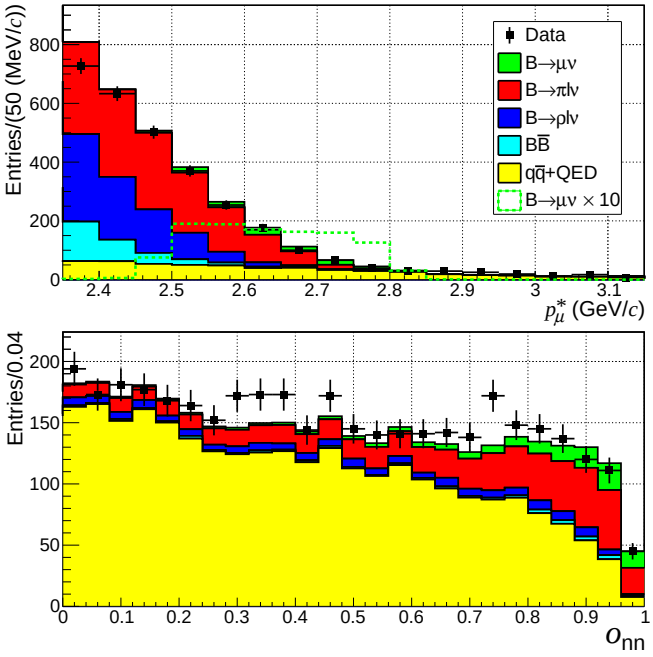


FIG. 2: Projections of the fitted distribution to data onto the histogram axes in the signal-enhanced regions $0.84 < o_{\rm nn}$ (top plot) and $2.6\,{\rm GeV}/c < p^*_\mu < 2.85\,{\rm GeV}/c$ (bottom plot).

we vary the form-factor shape within uncertainties obtained with the new lattice QCD result [5] and the procedure described in Ref. [4], which was used to estimate the value of $|V_{ub}|$. Since the form-factor is tightly constrained, the contribution to the systematic uncertainty from the $\bar{B} \to \pi\ell^-\bar{\nu}_\ell$ background is estimated to be only 0.9%. For the $\bar{B} \to \rho\ell^-\bar{\nu}_\ell$ component, the form-factors at high $q^2$ or high muon momentum have much larger uncertainties and several available calculations are employed [24, 25, 32], resulting in a systematic uncertainty of 12%.

The rare hadronic decay $B^- \to K^0_L\pi^-$, where $K^0_L$ is not detected and the high momentum $\pi$ is misidentified as a muon, is also indistinguishable from the signal decay and has a similar $o_{\rm nn}$ shape. This contribution is fixed in the fit and the signal yield difference, with and without the $B^- \to K^0_L\pi^-$ component, of 5.5% is taken as a systematic uncertainty since `GEANT3` poorly models $K^0_L$ interactions with materials.

The not-yet-discovered process $B^- \to \mu^-\bar{\nu}_\mu\gamma$ with a soft photon can mimic the signal decay. To estimate the uncertainty from this hypothetical background, we perform the fit with this contribution fixed to half of the best upper limit $\mathcal{B}(B^- \to \mu^-\bar{\nu}_\mu\gamma) < 3.4 \times 10^{-6}$ at 90% C.L. by Belle [33] and take the difference of 6% as the systematic uncertainty.

Previous studies [13, 14] did not characterize these backgrounds in a detailed manner, which could have led to a substantial underestimation of the systematic uncertainties.

In the region $p^*_\mu > 2.85$ GeV/$c$, where only continuum events are present, we observe an almost linearly growing data/fit difference with maximum deviation $\sim 20\%$ at $o_{\rm nn} \sim 1$. To estimate the uncertainty due to the level of data/MC agreement in the $o_{\rm nn}$ variable, we rescale linearly with $o_{\rm nn}$ the continuum histograms used in the fit and refit, obtaining a 15% lower value of $R$. For peaking components such as the signal $B^- \to \mu^-\bar{\nu}_\mu$ and the normalization decay $\bar{B} \to \pi\ell^-\bar{\nu}_\ell$, we use the fit/data ratio in the region $p^*_\mu < 2.5$ GeV/$c$ and apply it to the peaking components in the signal-region histograms ($B^- \to \mu^-\bar{\nu}_\mu$, $\bar{B} \to \pi\ell^-\bar{\nu}_\ell$ and $\bar{B} \to \rho\ell^-\bar{\nu}_\ell$). Refitting produces an 11% higher value of $R$. Simultaneously applying both effects leads to only a 2% shift in the refitted central value; thus, we include the individual deviations as systematic uncertainties in the continuum and signal peak descriptions.

In some cases, the signal muon and detected fraction of the particles from the companion $B$-meson decay do not provide enough particles for an event to be identified as a $B$-meson decay and hence to be recorded. The efficiency for recording these events is 84% as calculated using MC, and we take the event-recording uncertainty to be half of the inefficiency (8%) since it will be partially cancelled by taking the ratio with the normalization process $\bar{B} \to \pi\ell^-\bar{\nu}_\ell$.

The branching fraction of the normalization process $\bar{B} \to \pi\ell^-\bar{\nu}_\ell$ is known with 3.4% precision [3] and this is included as a systematic uncertainty.

The summary of the systematic uncertainties is shown in Table II. The total systematic uncertainty of 25% is obtained by summing the individual contributions discussed above in quadrature.

TABLE II: The summary of the systematic uncertainties for the branching fraction result.

| Source | Uncertainty (%) |
|---|---|
| $\bar{B} \to \pi\ell^-\bar{\nu}_\ell$ form-factor | 0.9 |
| $\bar{B} \to \rho\ell^-\bar{\nu}_\ell$ form-factor | 12 |
| $B^- \to K^0_L\pi^-$ | 5.5 |
| $B^- \to \mu^-\bar{\nu}_\mu\gamma$ | 6 |
| Continuum shape | 15 |
| Signal peak shape | 11 |
| Trigger | 8 |
| $\mathcal{B}(\bar{B} \to \pi\ell^-\bar{\nu}_\ell)$ | 3.4 |
| Total | 24.6 |

Incorporating systematic uncertainties, the final branching fraction for the signal decay is $\mathcal{B}(B^- \to \mu^-\bar{\nu}_\mu) = (6.46 \pm 2.22_{\rm stat} \pm 1.60_{\rm syst}) \times 10^{-7} = (6.46 \pm 2.74_{\rm tot}) \times 10^{-7}$. The accounted systematic uncertainties reduce the fit statistical signal significance from 3.4 to 2.4 standard deviations. A confidence interval using a frequentist approach based on Ref. [34] is evaluated with systematic uncertainties included and found to be $\mathcal{B}(B^- \to \mu^-\bar{\nu}_\mu) \in [2.9, 10.7] \times 10^{-7}$ at the 90% C.L., in agreement with the SM prediction $\mathcal{B}_{\rm SM}(B^- \to \mu^-\bar{\nu}_\mu) = (3.80 \pm 0.31) \times 10^{-7}$.

In conclusion, as a result of an untagged search with the full Belle $\Upsilon(4S)$ data set, we find a 2.4 standard deviation excess above background for the decay $B^- \to \mu^-\bar{\nu}_\mu$, with a measured branching fraction of $\mathcal{B}(B^- \to \mu^-\bar{\nu}_\mu) = (6.46 \pm 2.22_{\rm stat} \pm 1.60_{\rm syst}) \times 10^{-7}$ and a ratio of $\mathcal{B}(B^- \to \mu^-\bar{\nu}_\mu)/\mathcal{B}(\bar{B} \to \pi\ell^-\bar{\nu}_\ell) = (4.45 \pm 1.53_{\rm stat} \pm 1.09_{\rm syst}) \times 10^{-3}$. The 90% confidence interval for the obtained branching fraction in the frequentist approach is $\mathcal{B}(B^- \to \mu^-\bar{\nu}_\mu) \in [2.9, 10.7] \times 10^{-7}$. The forthcoming data from the Belle II experiment [35] should further improve the measurement.

We thank the KEKB group for the excellent operation of the accelerator; the KEK cryogenics group for the efficient operation of the solenoid; and the KEK computer group, the National Institute of Informatics, and the PNNL/EMSL computing group for valuable computing and SINET5 network support. We acknowledge support from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan, the Japan Society for the Promotion of Science (JSPS), and the Tau-Lepton Physics Research Center of Nagoya University; the Australian Research Council; Austrian Science Fund under Grant No. P 26794-N20; the National Natural Science Foundation of China under Contracts No. 10575109, No. 10775142, No. 10875115, No. 11175187, No. 11475187, No. 11521505 and No. 11575017; the Chinese Academy of Science Center for Excellence in Particle Physics; the Ministry of Education, Youth and Sports of the Czech Republic under Contract No. LTT17020; the Carl Zeiss Foundation, the Deutsche Forschungsgemeinschaft, the Excellence Cluster Universe, and the VolkswagenStiftung; the Department of Science and Technology of India; the Istituto Nazionale di Fisica Nucleare of Italy; National Research Foundation (NRF) of Korea Grants No. 2014R1A2A2A01005286, No. 2015R1A2A2A01003280, No. 2015H1A2A1033649, No. 2016R1D1A1B01010135, No. 2016K1A3A7A09005603, No. 2016R1D1A1B02012900; Radiation Science Research Institute, Foreign Large-size Research Facility Application Supporting project and the Global Science Experimental Data Hub Center of the Korea Institute of Science and Technology Information; the Polish Ministry of Science and Higher Education and the National Science Center; the Ministry of Education and Science of the Russian Federation and the Russian Foundation for Basic Research; the Slovenian Research Agency; Ikerbasque, Basque Foundation for Science and MINECO (Juan de la Cierva), Spain; the Swiss National Science Foundation; the Ministry of Education and the Ministry of Science and Technology of Taiwan; and the

U.S. Department of Energy and the National Science Foundation.

* now at University of Victoria, Victoria, British Columbia, Canada V8W 3P6


[1] Charge-conjugate decays are implied throughout this paper.

[2] S. Aoki *et al.*, Eur. Phys. J. C **77**, 112 (2017).

[3] C. Patrignani *et al.* [Particle Data Group], Chin. Phys. C **40**, 100001 (2016).

[4] A. Sibidanov *et al.* [Belle Collaboration], Phys. Rev. D **88**, 032005 (2013).

[5] J. A. Bailey *et al.* (Fermilab Lattice and MILC Collaborations), Phys. Rev. D **92**, 014024 (2015).

[6] W.-S. Hou, Phys. Rev. D **48**, 2342 (1993).

[7] S. Baek and Y. G. Kim, Phys. Rev. D **60**, 077701 (1999).

[8] H. Georgi and S. L. Glashow, Phys. Rev. Lett. **32**, 438 (1974).

[9] The parameter $\tan\beta$ is the ratio of the vacuum expectation values of the two Higgs fields, see e.g., A. Djouadi and J. Quevillon, JHEP **1310**, 028 (2013).

[10] G. Isidori and P. Paradisi, Phys. Lett. B **639**, 499 (2006).

[11] A. Filipuzzi and G. Isidori, Eur. Phys. J. C **64**, 55 (2009).

[12] Y. Yook *et al.* [Belle Collaboration], Phys. Rev. D **91**, 052016 (2015).

[13] N. Satoyama *et al.* [Belle Collaboration], Phys. Lett. B **647**, 67 (2007).

[14] B. Aubert *et al.* [*BABAR* Collaboration], Phys. Rev. D **79**, 091101 (2009).

[15] B. Aubert *et al.* [*BABAR* Collaboration], Phys. Rev. D **77**, 091104 (2008).

[16] B. Aubert *et al.* [*BABAR* Collaboration], Phys. Rev. D **81**, 051101 (2010).

[17] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res. Sect., A **499**, 1 (2003), and other papers included in this Volume; T.Abe *et al.*, Prog. Theor. Exp. Phys. (2013) 03A001 and following articles up to 03A011.

[18] A. Abashian *et al.* [Belle Collaboration], Nucl. Instrum. Methods Phys. Res. Sect. A **479**, 117 (2002); also see detector section in J. Brodzicka *et al.*, Prog. Theor. Exp. Phys. (2012) 04D001.

[19] Z. Natkaniec *et al.* [Belle SVD2 Group], Nucl. Instr. and Meth. A **560**, 1 (2006).

[20] R. Brun *et al.*, GEANT 3.21, CERN Report DD/EE/84-1 (1984).

[21] D. J. Lange, Nucl. Instrum. Meth. A **462**, 152 (2001).

[22] E. Barberio and Z. Wąs, Comput. Phys. Commun. **79**, 291 (1994).

[23] P. Ball and R. Zwicky, JHEP **0110**, 019 (2001).

[24] P. Ball and R. Zwicky, Phys. Rev. D **71**, 014029 (2005).

[25] D. Scora and N. Isgur, Phys. Rev. D **52**, 2783 (1995). N. Isgur, D. Scora, B. Grinstein and M. B. Wise, Phys. Rev. D **39**, 799 (1989).

[26] F. De Fazio and M. Neubert, JHEP **9906**, 017 (1999).

[27] T. Sjöstrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001).

[28] G. P. Korchemsky, D. Pirjol and T. M. Yan, Phys. Rev. D **61**, 114510 (2000).

[29] A. Abashian *et al.*, Nucl. Instrum. Meth. A **491**, 69 (2002).

[30] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima and T. Tsukamoto, Nucl. Instrum. Meth. A **485**, 490 (2002).

[31] R. J. Barlow and C. Beeston, Comput. Phys. Commun. **77**, 219 (1993).

[32] L. Del Debbio *et al.* [UKQCD Collaboration], Phys. Lett. B **416**, 392 (1998).

[33] A. Heller *et al.* [Belle Collaboration], Phys. Rev. D **91**, 112009 (2015).

[34] G. J. Feldman and R. D. Cousins, Phys. Rev. D **57**, 3873 (1998).

[35] T. Abe *et al.* [Belle-II Collaboration], arXiv:1011.0352 [physics.ins-det].