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

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

(The Belle Collaboration)

[1]*School of Physics, University of Sydney, New South Wales 2006*
[2]*High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801*
[3]*SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193*
[4]*Department of Physics, University of Tokyo, Tokyo 113-0033*
[5]*Department of Physics, Faculty of Science, University of Tabuk, Tabuk 71451*
[6]*Department of Physics, Faculty of Science, King Abdulaziz University, Jeddah 21589*
[7]*Pacific Northwest National Laboratory, Richland, Washington 99352*
[8]*Moscow Institute of Physics and Technology, Moscow Region 141700*
[9]*Tata Institute of Fundamental Research, Mumbai 400005*
[10]*King Abdulaziz City for Science and Technology, Riyadh 11442*
[11]*Indian Institute of Technology Bhubaneswar, Satya Nagar 751007*
[12]*School of Physics, University of Melbourne, Victoria 3010*
[13]*Indian Institute of Technology Madras, Chennai 600036*
[14]*Indian Institute of Technology Guwahati, Assam 781039*
[15]*J. Stefan Institute, 1000 Ljubljana*
[16]*H. Niewodniczanski Institute of Nuclear Physics, Krakow 31-342*
[17]*University of Maribor, 2000 Maribor*
[18]*University of Hawaii, Honolulu, Hawaii 96822*
[19]*Faculty of Mathematics and Physics, Charles University, 121 16 Prague*
[20]*Department of Physics, National Taiwan University, Taipei 10617*
[21]*Max-Planck-Institut für Physik, 80805 München*
[22]*National Central University, Chung-li 32054*
[23]*Hanyang University, Seoul 133-791*
[24]*P.N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow 119991*
[25]*Moscow Physical Engineering Institute, Moscow 115409*
[26]*Korea Institute of Science and Technology Information, Daejeon 305-806*

[27]Gyeongsang National University, Chinju 660-701
[28]Sungkyunkwan University, Suwon 440-746
[29]Wayne State University, Detroit, Michigan 48202
[30]Department of Physics, Tohoku University, Sendai 980-8578
[31]Budker Institute of Nuclear Physics SB RAS, Novosibirsk 630090
[32]Novosibirsk State University, Novosibirsk 630090
[33]Deutsches Elektronen–Synchrotron, 22607 Hamburg
[34]Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
[35]Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, 76131 Karlsruhe
[36]Department of Physics, Technische Universität München, 85748 Garching
[37]Indiana University, Bloomington, Indiana 47408
[38]INFN—Sezione di Torino, 10125 Torino
[39]Niigata University, Niigata 950-2181
[40]Nara Women's University, Nara 630-8506
[41]Graduate School of Science, Nagoya University, Nagoya 464-8602
[42]Kobayashi-Maskawa Institute, Nagoya University, Nagoya 464-8602
[43]Osaka City University, Osaka 558-8585
[44]University of Florida, Gainesville, Florida 32611
[45]Kyungpook National University, Daegu 702-701
[46]Chonnam National University, Kwangju 660-701
[47]Ludwig Maximilians University, 80539 Munich
[48]Soongsil University, Seoul 156-743
[49]Korea University, Seoul 136-713
[50]University of Cincinnati, Cincinnati, Ohio 45221
[51]Faculty of Mathematics and Physics, University of Ljubljana, 1000 Ljubljana
[52]Kennesaw State University, Kennesaw, Georgia 30144
[53]Punjab Agricultural University, Ludhiana 141004
[54]Yonsei University, Seoul 120-749
[55]Justus-Liebig-Universität Gießen, 35392 Gießen
[56]University of Science and Technology of China, Hefei 230026
[57]University of Pittsburgh, Pittsburgh, Pennsylvania 15260
[58]Earthquake Research Institute, University of Tokyo, Tokyo 113-0032
[59]University of Miyazaki, Miyazaki 889-2192
[60]INFN—Sezione di Napoli, 80126 Napoli
[61]Kyoto University, Kyoto 606-8502
[62]Toho University, Funabashi 274-8510
[63]Kanagawa University, Yokohama 221-8686
[64]Nippon Dental University, Niigata 951-8580
[65]Luther College, Decorah, Iowa 52101
[66]University of Malaya, 50603 Kuala Lumpur
[67]École Polytechnique Fédérale de Lausanne (EPFL), Lausanne 1015
[68]University of the Basque Country UPV/EHU, 48080 Bilbao
[69]IKERBASQUE, Basque Foundation for Science, 48013 Bilbao
[70]Institute of High Energy Physics, Vienna 1050
[71]Yamagata University, Yamagata 990-8560
[72]Beihang University, Beijing 100191
[73]Tokyo Institute of Technology, Tokyo 152-8550
[74]Excellence Cluster Universe, Technische Universität München, 85748 Garching
[75]Institute for High Energy Physics, Protvino 142281
[76]Gifu University, Gifu 501-1193
[77]Tokyo Metropolitan University, Tokyo 192-0397
[78]Showa Pharmaceutical University, Tokyo 194-8543
[79]J-PARC Branch, KEK Theory Center, High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801
[80]Theoretical Research Division, Nishina Center, RIKEN, Saitama 351-0198
[81]University of Torino, 10124 Torino
[82]Advanced Science Research Center, Japan Atomic Energy Agency, Naka 319-1195
[83]National United University, Miao Li 36003
[84]Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049
[85]Stefan Meyer Institute for Subatomic Physics, Vienna 1090

(Received 12 December 2017; revised manuscript received 1 May 2018; published 18 July 2018)

We report the results of a search for the rare, purely leptonic decay $B^- \to \mu^- \bar{\nu}_\mu$ performed with a 711 fb^{-1} data sample that contains 772×10^6 $B\bar{B}$ pairs, collected near the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB asymmetric-energy e^+e^- collider. 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 $[2.9, 10.7] \times 10^{-7}$.

DOI: 10.1103/PhysRevLett.121.031801

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_ℓ are the masses of the B meson and charged lepton, respectively, f_B is the B-meson decay constant obtained from theory, τ_B is the lifetime of the B meson, and V_{ub} is the Cabibbo-Kobayashi-Maskawa matrix element governing the coupling between the u and b quarks. The FLAG [2] average of lattice QCD calculations gives $f_B = 0.186 \pm 0.004$ GeV, the world-average value of τ_B is 1.638 ± 0.004 ps [3], and the value of $|V_{ub}|$ is $(3.736 \pm 0.142) \times 10^{-3}$, obtained by the fit procedure described in Ref. [4], equipped with the most recent lattice QCD calculation by the FNAL and MILC collaborations [5]. Using these values as input parameters for Eq. (1), the expected branching fraction for $B^- \to \mu^- \bar{\nu}_\mu$ is $(3.80 \pm 0.31) \times 10^{-7}$ and the event yield in the full Belle data set is 301 ± 25 (for both charges).

Because of 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 [14]. Searches for the $B^- \to \tau^- \bar{\nu}_\tau$ decay by the Belle [17,18] and *BABAR* [19,20] experiments have found evidence for the decay with an average PDG [3] branching fraction of $\mathcal{B}(B^- \to \tau^- \bar{\nu}_\tau) = (1.09 \pm 0.24) \times 10^{-4}$, consistent with the SM prediction.

We present a search for the decay $B^- \to \mu^- \bar{\nu}_\mu$ using the *untagged* method. In such a method, the candidate decay products of the other B meson are required to satisfy generic kinematic requirements, consistent with the B-decay hypothesis. This study is based on a 711 fb^{-1} data sample that contains $(772 \pm 11) \times 10^6$ $B\bar{B}$ pairs, collected at a center-of-mass energy of $\sqrt{s} = 10.58$ GeV, corresponding to the $\Upsilon(4S)$ resonance, with the Belle detector at the KEKB asymmetric-energy e^+e^- (3.5 on 8 GeV) collider [21].

The Belle detector is a large-solid-angle magnetic spectrometer that consists of a silicon vertex detector, a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACCs), a barrel-like arrangement of time-of-flight (TOF) scintillation counters, and an electromagnetic calorimeter composed 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 [22].

The event sample obtained at the $\Upsilon(4S)$ resonance contains not only a large sample of $\Upsilon(4S) \to B\bar{B}$ events but also a background arising from so-called continuum processes: annihilation into lighter fermions $e^+e^- \to q\bar{q}$ ($q = u$, d, s, c, and τ, μ) and two-photon production $e^+e^- \to e^+e^-\gamma^*\gamma^*$, $\gamma^*\gamma^* \to q\bar{q}$. To characterize the

contribution of these events in our search, which is substantial, we use a 79 fb^{-1} sample collected 40 MeV below the $B\bar{B}$ production threshold.

We use Monte Carlo (MC) samples based on the detailed detector geometry description implemented with the GEANT3 package [23] to establish the analysis technique and study major backgrounds. Events with B-meson decays are generated using EVTGEN [24]. The generated samples include 2×10^6 signal events, a sample of generic $B\bar{B}$ decays corresponding to 10 times the integrated luminosity of the data, $c\bar{c}$ production as well as $u\bar{u}$, $d\bar{d}$, and $s\bar{s}$ (or, for short, uds) corresponding to 6 times the data, $\bar{B} \to X_u \ell^- \bar{\nu}_\ell$ decays corresponding to 20 times the data, other B decays with probability $\lesssim 4 \times 10^{-4}$ corresponding to 50 times the data, and $e^+e^- \to \tau^+\tau^-$ corresponding to 5 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 over the course of the data-taking period. Final-state radiation from charged particles is modeled using the PHOTOS package [25].

In addition, 8×10^6 MC events of one of the largest backgrounds, $\bar{B} \to \pi \ell^- \bar{\nu}_\ell$, 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 $m_b = 5$ GeV$/c^2$ based on the work in Ref. [26].

The muon in $B^- \to \mu^- \bar{\nu}_\mu$ decay is monochromatic in the absence of radiation, with a momentum in the B-meson rest frame $p_\mu^B \approx m_B/2$. In the $\Upsilon(4S)$ center-of-mass frame, where the B meson is in motion, the boost smears the momentum of the muon, p_μ^*, 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. To reduce potential bias, the $\Upsilon(4S)$ data in the p_μ^* interval $(2.45, 2.85)$ GeV$/c$ was not considered until the analysis procedure was 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 ~90% [27]. 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 γ, K_S^0, and Λ) are included in this list even if they fail the IP selection. All other charged particles are ignored when constructing the B-meson kinematic variables. We discard the event if the total momentum of these ignored particles exceeds 1.3 GeV$/c$ to suppress the background from misreconstructed 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 the ECL cluster, the energy loss in the CDC, and the response of the ACC [28]. 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 of misidentifying 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 end caps, it rises to 110 and 150 MeV, respectively. Additionally, we require the total energy deposition in the calorimeter not associated with charged particles or 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 meson 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

$$M_{\rm bc} = \sqrt{E_{\rm beam}^2/c^4 - \left|\sum_i \vec{\mathbf{p}}_i^*/c\right|^2}, \tag{2}$$

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

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 ith particle that makes up the accompanying B-meson candidate. We retain events that satisfy $M_{\rm bc} > 5.1$ GeV$/c^2$ and -3 GeV $< E_B - E_{\rm beam} < 2$ GeV.

To exploit the jetlike structure of the continuum background, where particles tend to be produced collinearly, we define the direction $\hat{n}$ of the thrust axis by maximizing the quantity

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

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 and oppositely directed with respect to 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. The KLM cannot measure the K_L^0 energy—only the interaction position—and this can lead to an incorrect estimation of the missing energy to be attributed to the signal neutrino. We require no more than one K_L^0 cluster in the KLM and no K_L^0 clusters associated with ECL clusters. This selection rejects about 24% of the background events and keeps about 90% of the signal. The K_L^0 detection efficiency is calibrated using a $D^0 \to \phi K_S^0$, $\phi \to K_S^0 K_L^0$ control sample.

The total signal selection efficiency for $B^- \to \mu^- \bar{\nu}_\mu$ decays is estimated to be around 38% at this stage, with an expected signal yield of 115 ± 9. However, the remaining background is still more than 3 orders of magnitude larger than the expected signal yield. A multivariate data analysis is therefore employed to further separate signal from background. We combine various kinematic variables of an event into a single variable $o_{\rm nn}$ using an artificial neural network. We choose 14 input variables that are uncorrelated with the absolute value of the muon momentum and that collectively yield the best signal to background ratio. These variables 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.

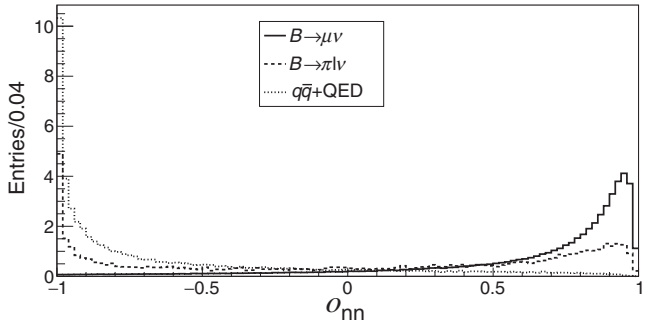

FIG. 1. The distributions of the neural network output variable for the signal and major background processes predicted by MC in the signal-enhanced region 2.644 GeV$/c < p_\mu^* < 2.812$ GeV$/c$ (the relative normalization of these distributions is arbitrary).

The distributions of the neural network output variable in the signal-enhanced region in p_μ^* 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_μ^*-$o_{\rm nn}$ plane using the method described in Ref. [29], 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 5 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_μ^* 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_μ^*-$o_{\rm nn}$ 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 $\mu^+\mu^-$, $e^+e^-e^+e^-$, $e^+e^-u\bar{u}$, $e^+e^-d\bar{d}$, $e^+e^-s\bar{s}$, and $e^+e^-c\bar{c}$. The $\bar{B} \to \pi \ell^- \bar{\nu}_\ell$ component is composed of both $B^- \to \pi^0 \ell^- \bar{\nu}_\ell$ and $\bar{B}^0 \to \pi^+ \ell^- \bar{\nu}_\ell$ decays, with the ratio fixed by isospin symmetry and assuming $\mathcal{B}(\Upsilon(4S) \to B^+B^-) = 0.514$ since, in our untagged analysis, they are indistinguishable.

To obtain the signal branching fraction, we fit for 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

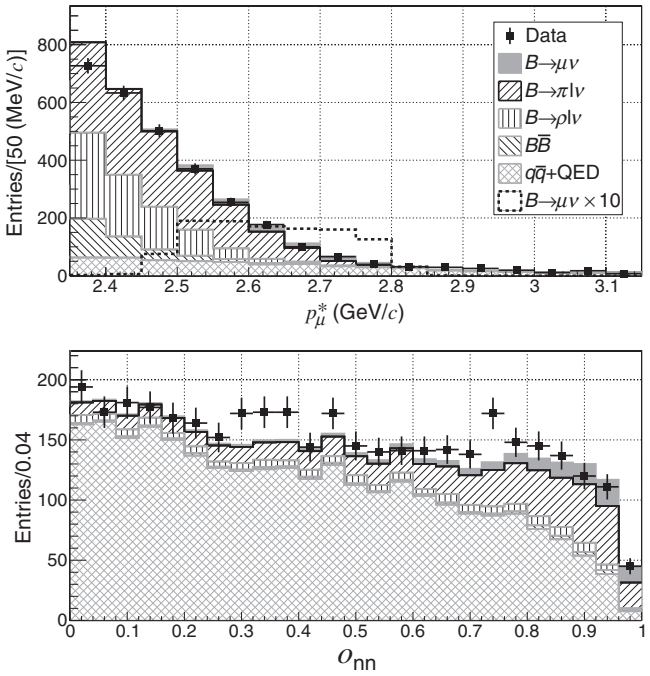

FIG. 2. Projections of the fitted distribution to data onto the histogram axes in the signal-enhanced regions (top panel) $0.84 < o_{\rm nn}$ and (bottom panel) 2.6 GeV/c $< p^*_\mu <$ 2.85 GeV/c.

the MC prediction of this ratio $R_{\rm MC} = 114.6/11746 = 0.976 \times 10^{-2}$, obtained assuming $\mathcal{B}(B\to\mu\bar{\nu}_\mu)=3.80\times10^{-7}$ and $\mathcal{B}(\bar{B}^0 \to \pi^+\ell^-\bar{\nu}_\ell) = 1.45 \times 10^{-4}$ (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 3.4σ, 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 $\chi^2/\text{ndf} = 27.6/16$ (top panel) and $\chi^2/\text{ndf} = 29.1/25$ (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 mismodeling, as well as partially canceling 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, 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 [30–32], resulting in a systematic uncertainty of 12%. The $\bar{B} \to \rho\ell^-\bar{\nu}_\ell$ decays are a significant part of the background in the low muon momentum region, and form-factor mismodeling may lead to a worse description of the data in this region.

The rare hadronic decay $B^- \to K^0_L\pi^-$, where K^0_L is not detected and the high momentum π 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.

In the region $p^*_\mu > 2.85$ GeV/c, where only continuum events are present, we observe an almost linearly growing data-fit difference as a function of $o_{\rm nn}$ with a maximum deviation of ~20% at $o_{\rm nn} \sim 1$. To estimate the potential bias due to this dependence, 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-to-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 canceled 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 total systematic uncertainty of 25% is obtained by summing the individual contributions discussed above in quadrature.

Incorporating systematic uncertainties, the final branching fraction for the signal decay is $\mathcal{B}(B^- \to \mu^- \bar{\nu}_\mu) = (6.46 \pm 2.22_{\rm stat} \pm 1.60_{\rm syst}) \times 10^{-7} = (6.46 \pm 2.74_{\rm tot}) \times 10^{-7}$. This result supersedes the previous Belle untagged search [13]. The 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., consistent with the SM prediction $\mathcal{B}_{\rm SM}(B^- \to \mu^- \bar{\nu}_\mu) = (3.80 \pm 0.31) \times 10^{-7}$.

In conclusion, we have performed an untagged search for the process $B^- \to \mu^- \bar{\nu}_\mu$ using the full Belle $\Upsilon(4S)$ data sample, finding a 2.4 standard deviation excess above background, 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 Pacific Northwest National Laboratory (PNNL) Environmental Molecular Sciences Laboratory (EMSL) computing group for valuable computing and Science Information NETwork 5 (SINET5) network support. We acknowledge support from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan, the Japan Society for the Promotion of Science (JSPS), and the Tau-Lepton Physics Research Center of Nagoya University; the Australian Research Council including Grants No. DP180102629, No. DP170102389, No. DP170102204, No. DP150103061, No. FT130100303; Austrian Science Fund under Grant No. P 26794-N20; the National Natural Science Foundation of China under Contracts No. 11435013, No. 11475187, No. 11521505, No. 11575017, No. 11675166, No. 11705209; Key Research Program of Frontier Sciences, Chinese Academy of Sciences (CAS), Grant No. QYZDJ-SSW-SLH011; the CAS Center for Excellence in Particle Physics (CCEPP); Fudan University Grants No. JIH5913023, No. IDH5913011/003, No. JIH5913024, No. IDH5913011/002; 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, and No. 2016R1D1A1B02012900; Radiation Science Research Institute, Foreign Large-size Research Facility Application Supporting project and the Global Science Experimental Data Hub Center of the Korea Institute of Science and Technology Information; the Polish Ministry of Science and Higher Education and the National Science Center; the Ministry of Education and Science of the Russian Federation and the Russian Foundation for Basic Research; the Slovenian Research Agency; Ikerbasque, Basque Foundation for Science, Basque Government (No. IT956-16) and Ministry of Economy and Competitiveness (MINECO) (Juan de la Cierva), Spain; the Swiss National Science Foundation; the Ministry of Education and the Ministry of Science and Technology of Taiwan; and the United States Department of Energy and the National Science Foundation.

*Present address: University of Victoria, Victoria, British Columbia, Canada V8W 3P6.

[1] Charge-conjugate decays are implied throughout this Letter.
[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] see, e.g., A. Djouadi and J. Quevillon, J. High Energy Phys. 10 (2013) 028.
[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] B. Kronenbitter *et al.* (Belle Collaboration), Phys. Rev. D **92**, 051102 (2015).
[18] K. Hara *et al.* (Belle Collaboration), Phys. Rev. Lett. **110**, 131801 (2013).
[19] J. P. Lees *et al.* (*BABAR* Collaboration), Phys. Rev. D **88**, 031102 (2013).
[20] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **81**, 051101 (2010).
[21] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res., Sect. A **499**, 1 (2003), and references therein;

T.Abe *et al.*, Prog. Theor. Exp. Phys. (**2013**) 03A001 and following articles up to 03A011.
[22] A. Abashian *et al.* (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 117 (2002); also see the detector section in J. Brodzicka *et al.*, Prog. Theor. Exp. Phys. (**2012**) 04D001.
[23] R. Brun *et al.*, GEANT3, CERN Report No. DD/EE/84-1, 1984.
[24] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).
[25] E. Barberio and Z. Wąs, Comput. Phys. Commun. **79**, 291 (1994).
[26] G. P. Korchemsky, D. Pirjol, and T. M. Yan, Phys. Rev. D **61**, 114510 (2000).
[27] A. Abashian *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **491**, 69 (2002).
[28] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima, and T. Tsukamoto, Nucl. Instrum. Methods Phys. Res., Sect. A **485**, 490 (2002).
[29] R. J. Barlow and C. Beeston, Comput. Phys. Commun. **77**, 219 (1993).
[30] P. Ball and R. Zwicky, Phys. Rev. D **71**, 014029 (2005).
[31] 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).
[32] L. Del Debbio, J. M Flynn, L. Lellouch, and J. Nieves (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.