Study of $B^+ \to p\bar{n}\pi^0$

K.-N. Chu, Y.-R. Lin, M.-Z. Wang, I. Adachi, K. Adamczyk, H. Aihara, S. Al Said, D. M. Asner, H. Atmacan, V. Aulchenko, T. Aushev, R. Ayad, V. Babu, Sw. Banerjee, P. Behera, K. Belous, J. Bennett, M. Bessner, V. Bhardwaj, B. Bhuyan, T. Bilka, D. Biswas, D. Bodrov, G. Bonvicini, J. Borah, A. Bozek, M. Bračko, P. Branchini, T. E. Browder, A. Budano, M. Campajola, D. Červenkov, M.-C. Chang, P. Chang, A. Chen, B. G. Cheon, K. Chilikin, H. E. Cho, K. Cho, S.-J. Cho, S.-K. Choi, Y. Choi, S. Choudhury, D. Cinabro, G. De Nardo, G. De Pietro, R. Dhamija, F. Di Capua, J. Dingfelder, Z. Doležal, T. V. Dong, D. Epifanov, T. Ferber, D. Ferlewicz, B. G. Fulsom, V. Gaur, A. Garmash, A. Giri, P. Goldenzweig, B. Golob, E. Graziani, T. Gu, K. Gudkova, C. Hadjivasiliou, S. Halder, X. Han, K. Hayasaka, H. Hayashii, W.-S. Hou, C.-L. Hsu, K. Huang, T. Iijima, K. Inami, N. Ipsita, A. Ishikawa, R. Itoh, M. Iwasaki, W. W. Jacobs, E.-J. Jang, Y. Jin, K. K. Joo, D. Kalita, A. B. Kaliyar, T. Kawasaki, C. Kiesling, C. H. Kim, D. Y. Kim, K.-H. Kim, Y.-K. Kim, P. Kodyš, T. Konno, A. Korobov, S. Korpar, E. Kovalenko, P. Križan, P. Krokovny, T. Kuhr, R. Kumar, K. Kumara, A. Kuzmin, Y.-J. Kwon, Y.-T. Lai, T. Lam, J. S. Lange, S. C. Lee, C. H. Li, J. Li, L. K. Li, S. X. Li, Y. Li, Y. B. Li, L. Li Gioi, J. Libby, K. Lieret, D. Liventsev, T. Luo, M. Masuda, T. Matsuda, D. Matvienko, S. K. Maurya, F. Meier, M. Merola, F. Metzner, K. Miyabayashi, R. Mizuk, G. B. Mohanty, I. Nakamura, M. Nakao, Z. Natkaniec, A. Natochii, L. Nayak, N. K. Nisar, S. Nishida, Y. Onuki, P. Oskin, P. Pakhlov, G. Pakhlova, S. Pardi, H. Park, J. Park, A. Passeri, S. Paul, T. K. Pedlar, R. Pestotnik, L. E. Piilonen, T. Podobnik, E. Prencipe, M. T. Prim, A. Rostomyan, N. Rout, G. Russo, S. Sandilya, L. Santelj, V. Savinov, G. Schnell, C. Schwanda, Y. Seino, K. Senyo, M. E. Sevior, W. Shan, M. Shapkin, C. Sharma, C. P. Shen, J.-G. Shiu, B. Shwartz, F. Simon, A. Sokolov, E. Solovieva, M. Starič, Z. S. Stottler, J. F. Strube, M. Sumihama, K. Sumisawa, T. Sumiyoshi, M. Takizawa, U. Tamponi, K. Tanida, F. Tenchini, K. Trabelsi, M. Uchida, T. Uglov, Y. Unno, S. Uno, P. Urquijo, Y. Usov, S. E. Vahsen, R. van Tonder, G. Varner, K. E. Varvell, A. Vinokurova, A. Vossen, S. Watanuki, E. Won, X. Xu, B. D. Yabsley, W. Yan, S. B. Yang, J. Yelton, J. H. Yin, C. Z. Yuan, L. Yuan, Y. Yusa, Z. P. Zhang, V. Zhilich, and V. Zhukova

(Belle Collaboration)

We search for the tree-diagram dominated process $B^+ \to p\bar{n}\pi^0$, using a data sample of $772 \times 10^6 B\bar{B}$ pairs collected at the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB asymmetric-energy e^+e^- collider. This is the first search with the Belle detector for a decay mode including an antineutron. No significant signal is observed and a 90% credible upper limit on the branching fraction is set at 6.3×10^{-6}.

I. INTRODUCTION

Since the first observation of the charmless baryonic B decay, $B^+ \to p\bar{p}K^+$ [1], many other similar three-body B decays have been found [2]. However, these decays predominantly proceed through the $b \to s$ penguin process, except for $B^+ \to p\bar{p}\pi^+$ [1] which is dominated by tree-diagram processes. One feature of these decays is that the baryon-antibaryon system has invariant mass near threshold

[3]. Recently, the LHCb collaboration has reported evidence of direct CP violation in $B^+ \to p\bar{p}K^+$ [4], indicating that both the $b \to u$ contribution and the interference between penguin and tree processes are sizable. For the charmed baryonic B decays, the CLEO collaboration has observed $B^0 \to p\bar{n}D^{*-}$ [5] with a much larger branching fraction than that of $B^0 \to p\bar{p}\bar{D}^{*0}$ [6]. The latter is believed to proceed via internal W emission, with color suppression in the formation of the final state. These findings inspire our search for $B^+ \to p\bar{n}\pi^0$ since it contains the external $W^+ \to p\bar{n}$ process that is not color suppressed. It is also interesting to compare its decay branching fraction to that of $B^+ \to p\bar{p}\pi^+$, $(1.62 \pm 0.20) \times 10^{-6}$ [1], and $B^0 \to p\bar{p}\pi^0$, $(5.0 \pm 1.9) \times 10^{-7}$ [7]. The comparison will shed more light on the $b \to u$ decay process. Noting an order-of-magnitude difference in the

branching fractions of $B^0 \to p\bar{n}D^{*-}$ and $B^0 \to p\bar{p}\bar{D}^{*0}$, we expect the branching fraction of $B^+ \to p\bar{n}\pi^0$ to also be 1 order-of-magnitude larger than that of $B^0 \to p\bar{p}\pi^0$ i.e. of order 10^{-6}. It will also be interesting to examine the invariant mass of the $p\bar{n}$ system, as the external $W^+ \to p\bar{n}$ emission may induce a flat distribution, as opposed to a sharp one near the threshold, which is seen in other charmless baryonic B decays [2].

II. DATA SAMPLE AND SIMULATION

We report a study of $B^+ \to p\bar{n}\pi^0$ using the full $\Upsilon(4S)$ dataset collected by the Belle detector [8,9] at the asymmetric-energy e^+ (3.5 GeV) e^- (8 GeV) KEKB collider [10,11]. This is the first search of a decay mode with an antineutron in the final state at Belle. The data sample used in this study corresponds to an integrated luminosity of 711 fb^{-1}, which contains $(772 \pm 11) \times 10^6$ $B\bar{B}$ pairs produced at the $\Upsilon(4S)$ resonance. The Belle detector surrounds the interaction point of KEKB. It 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 (ACC), a barrel-like arrangement of time-of-flight (TOF) scintillation counters, and an electromagnetic calorimeter (ECL) comprised of CsI(Tl) crystals located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux return located outside of the coil is instrumented to detect K_L^0 mesons and identify muons. Here, we reconstruct only $B^+ \to p\bar{n}\pi^0$ and not the charge conjugate mode $B^- \to \bar{p}n\pi^0$ due to the low detection efficiency of neutrons.

For the study of $B^+ \to p\bar{n}\pi^0$, samples simulated with the Monte Carlo (MC) technique are used to choose the signal selection criteria and to estimate the signal reconstruction efficiency. These samples are generated with EvtGen [12], and the detector response is simulated by GEANT3 [13]. We generate the signal MC sample by a phase space model reweighted with the $p\bar{n}$ mass distribution to follow the $p\bar{p}$ mass distribution of $B^+ \to p\bar{p}\pi^+$ [1]. The background samples include continuum events ($e^+e^- \to u\bar{u}, d\bar{d}, s\bar{s}$, and $c\bar{c}$), generic B decays ($b \to c$) and rare B decays ($b \to u, d, s$). These simulated background samples are 6 times larger than the integrated luminosity of the accumulated Belle data.

III. EVENT RECONSTRUCTION AND SELECTION

We require protons to originate within a 2.0 cm region along the beam and within a 0.3 cm region on the transverse plane around the interaction region. To identify protons, we utilize the likelihood information determined for each particle type by the CDC, TOF and ACC. We identify a track as a proton when $\frac{L_p}{L_p + L_K} > 0.6$ and $\frac{L_p}{L_p + L_\pi} > 0.6$, where L_p, L_K and L_π are likelihoods for protons, charged kaons and charged pions, which are the same selection criteria as

Ref. [1]. The π^0 is reconstructed from two photons. Each photon is an ECL cluster unmatched with any charged tracks, with a minimum energy in the laboratory frame of 0.05 GeV. To reduce combinatorial background, the π^0 energy is required to be larger than 1.2 GeV and the reconstructed mass is required to be in the range $0.111 < M_{\gamma\gamma} < 0.151$ GeV/c^2, which corresponds to about a ± 3.0 standard deviation (σ) window. We then perform a mass-constrained fit to the nominal π^0 mass [2] in order to improve the resolution of the reconstructed π^0 four-momentum.

Antineutrons deposit more energy in ECL than γ if the annihilation process occurs, and we use this feature to identify antineutrons. We pick up clusters in ECL, not matched with charged tracks. In order to identify $\bar{n}$, a deep learning application programming interface, Keras [14], is used to construct a multivariate analyzer for antineutron selection ($\bar{n}$ MVA). This contains five hidden layers, each with 20 nodes using a rectified linear unit [15]. We optimize the $\bar{n}$ MVA using adaptive moment estimation [16] and cross-entropy loss function. There are five input parameters for the deep learning package to distinguish $\bar{n}$ candidates from γ candidates: the total deposited energy of an ECL cluster, E_{cluster}, the highest deposited energy among all crystals in the cluster, E_{highest}, the number of hits in the cluster, N_{hits}, the standard deviation of the lateral deposited energy among all crystals, clusterLAT, and the ratio between the energy sum of the 3×3 and 5×5 crystals centered on the crystal with the largest deposited energy, $\frac{E_9}{E_{25}}$. The training process of $\bar{n}$ MVA is as follows. Antineutron and γ candidates from a subset of the generic B decays MC samples are chosen according to the generator information, out of which 81% is used for training, 9% is used for validation, and 10% is used for testing. Training is stopped when the loss function calculated out of the validation set stops rising for three epochs. The receiver operating characteristic (ROC) curve of the MVA over the training sample is shown at Fig. 1. The output of the $\bar{n}$ MVA, $C_{\bar{n}}$, ranges from 0 to $+1$, where the value is close to $+1$ for $\bar{n}$-like candidates and 0 for γ-like candidates. We then require $E_{\text{cluster}} > 0.5$ GeV and $C_{\bar{n}} > 0.7$ for $B^+ \to p\bar{n}\pi^0$ with a signal selection efficiency of 86.7% and a background rejection rate of 84.5%. We use $\Delta E = E_{\text{recon}}^* - E_{\text{beam}}^*$ to identify B decays, where E_{recon}^* and E_{beam}^* are the reconstructed B energy and the beam energy measured in the $\Upsilon(4S)$ rest frame, respectively. We cannot directly measure the $\bar{n}$ energy and momentum. Assuming the $\bar{n}$ originates from the e^+e^- interaction point, its momentum direction can be obtained by the energy-weighted position of the ECL cluster. We then constrain the B and $\bar{n}$ to their nominal masses [2] and use the measured momentum and energy of the p and π^0 to determine the $\bar{n}$ momentum value and evaluate E_{recon}^*. The constraints lead to an effective threshold, $\Delta E > -0.01$ GeV, as the energy of the B meson

ROC for nbarMVA
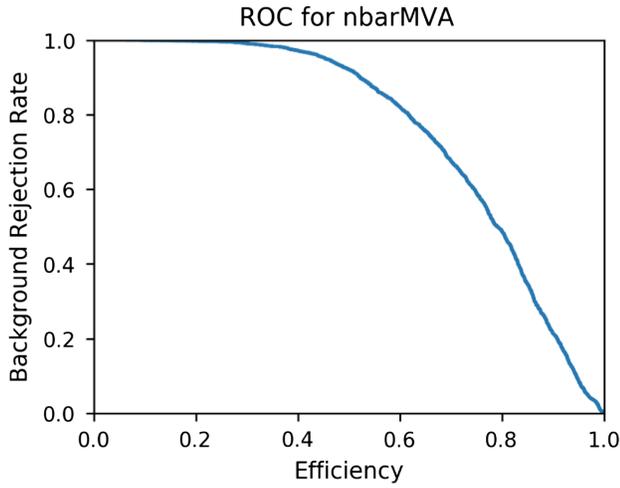
FIG. 1. The ROC curve of $\bar{n}$ MVA where the signal (background) refers to $\bar{n}$ (γ) identified by the generator information. The horizontal axis is the signal efficiency (true positive rate) and the vertical axis is the background rejection rate ($1 -$ false positive rate).

cannot be lower than the nominal mass, which is about 0.01 GeV lower than the beam energy. We keep B candidates with $-0.01 < \Delta E < 0.05$ GeV. We allow only one B candidate in each event. We choose the candidate with the smallest value of χ^2 from a fit of the B vertex position, based on the proton track, a virtual π^0 track constructed from the interaction point and the π^0 momentum vector, and a π^0 mass constraint. This fit does not include the $\bar{n}$ candidate: if there is more than one $\bar{n}$ candidate, we choose one at random. From the MC study, the fraction of $B^+ \to p\bar{n}\pi^0$ MC events with multiple B candidates is 9.5%, mostly due to the double counting of $\bar{n}$ candidates. Excluding double counting $\bar{n}$ candidates, the multiple candidate selection removes 0.3% of the $B^+ \to p\bar{n}\pi^0$ signal.

From the MC simulation, continuum events are the dominant background source in the candidate region ($-0.01 < \Delta E < 0.05$ GeV). Variables describing event topology are used to distinguish spherical $B\bar{B}$ events from jetlike continuum events. We use a neural network package, NeuroBayes [17], to separate the B signal from the continuum background. There are 28 input parameters in the neural network, of which 23 parameters are modified Fox-Wolfram moments [18]. The remaining five parameters are the separation between the B candidate vertex and the accompanying B vertex along the longitudinal direction; the angle between the B flight direction and the beam axis in the $\Upsilon(4S)$ rest frame; the angle between the thrust axes of the B candidate vertex and the accompanying B candidate in the $\Upsilon(4S)$ rest frame; the sphericity [19] of the event calculated in the $\Upsilon(4S)$ rest frame; and the B flavor tagging quality parameter [20]. We build a neural network and train it for continuum suppression using MC samples of $B^+ \to p\bar{n}\pi^0$ and continuum background. The output of the neural

network, C_{nb}, ranges from -1 to $+1$, where the value is close to $+1$ for $B\bar{B}$-like and -1 for continuum-like events. We require C_{nb} to be larger than 0.9. The continuum background is still dominant after the selection. The signal efficiency is 4.4% after all selections. The contribution of other B decays is negligible in the candidate region except for $B^+ \to p\bar{\Lambda}\pi^0$ with $\bar{\Lambda} \to \bar{n}\pi^0$.

IV. SIGNAL AND BACKGROUND EXTRACTION

To extract the $B^+ \to p\bar{n}\pi^0$ yield for events in the candidate region, we perform an extended unbinned likelihood fit to variables ΔE and C_{nbtr}, where C_{nbtr} is transformed from C_{nb} as

$$C_{nbtr} = \ln\frac{1.0 - C_{nb}}{C_{nb} - 0.9}. \tag{1}$$

These variables show no correlation in MC, and are treated as uncorrelated in the following. The likelihood is defined by

$$\mathcal{L} = \frac{e^{-\Sigma_{j=1}^3(n_j)}}{N!}\prod_{i=1}^{N}\sum_{j=1}^{3}(n_j P_j(C_{nbtr}^i, \Delta E^i)), \tag{2}$$

where N is the total number of events, n_j is the yield for each component, i denotes the event index, j stands for the component index (signal, $B^+ \to p\bar{\Lambda}\pi^0$, other background), and P_j represents the probability density function (PDF). To model the signal distributions, we use a sum of bifurcated Gaussian and Gaussian functions for both ΔE and C_{nbtr}. The signal distributions in ΔE and C_{nbtr} are calibrated with $B^+ \to p\bar{p}\pi^+$ and $B^+ \to p\bar{p}K^+$ events where the momentum of $\bar{p}$ is also calculated from the mass constraints. We can then obtain a ΔE signal distribution similar to that of in $B^+ \to p\bar{n}\pi^0$ and calibrate it by comparing the shape difference between MC and data. In addition, we can also measure their branching fractions, where $\mathcal{B}(B^+ \to p\bar{p}\pi^+) = (1.61 \pm 0.17 \pm 0.05) \times 10^{-6}$ and $\mathcal{B}(B^+ \to p\bar{p}K^+) = (5.5 \pm 0.4) \times 10^{-6}$, corresponding to the previous study in Belle [1], thus validating the whole analysis procedure. For the background, we use a second-order threshold function for ΔE, whose form is $f(x) \propto (x - x_0)^p \exp(c_1 x + c_2 x^2)$ for $x > x_0$ and 0 for $x \le x_0$, where c_1, c_2, p and x_0 are shape parameters whose initial values are determined by fitting to the continuum MC sample but are floated when fitting to real data, except for x_0. We use the sum of a bifurcated Gaussian and a Gaussian to describe C_{nbtr}. The ΔE distribution of the $B^+ \to p\bar{\Lambda}\pi^0$ background peaks at zero, but is wider than the distribution for signal. We model the ΔE distribution with a histogram function based on MC, and the sum of a bifurcated Gaussian and a Gaussian with a common mean for C_{nbtr}. We fix the shapes of ΔE and C_{nbtr} for the signal and $B^+ \to p\bar{\Lambda}\pi^0$ and allow the yields of signal, background and all
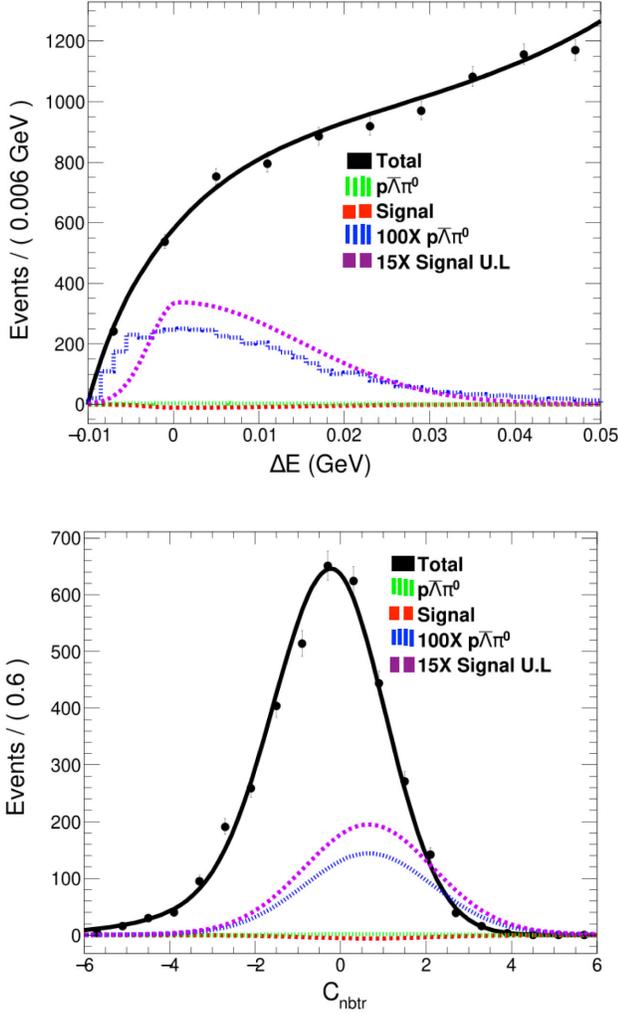
FIG. 2. Fit results of $B^+ \to p\bar{n}\pi^0$ projected onto ΔE (top, with $-6 < C_{nbtr} < 6$) and C_{nbtr} (bottom, with $-0.01 < \Delta E < 0.01$ GeV). The black solid line represents the total fit result. The green dotted line represents fixed $B^+ \to p\bar{\Lambda}\pi^0$ distribution. The red dashed line represents the signal distribution. The blue dotted line represents fixed $B^+ \to p\bar{\Lambda}\pi^0$ distribution for a branching fraction of 3×10^{-4}, 100 times larger to the fixed branching fraction. The violet dashed line represents the hypothetical signal distribution for a branching fraction of 9.5×10^{-5}, 15 times larger to the upper limit we calculate.

other PDF shape parameters of background except for the threshold to be floated. The yield of $B^+ \to p\bar{\Lambda}\pi^0$ is fixed at 15.0 according to the estimation based on the Belle previous study [21].

The fit result is shown in Fig. 2. We obtain a signal yield of -28.7 ± 49.0 and a background of $9950.7^{+100.7}_{-99.0}$. Since the signal yield is not significant, we set an upper limit on the branching fraction using a Bayesian technique, with a flat prior on the branching fraction. Systematic uncertainties, described below, are taken into account by smearing the likelihood function with a Gaussian whose width is the total systematic uncertainty. We use the efficiency obtained
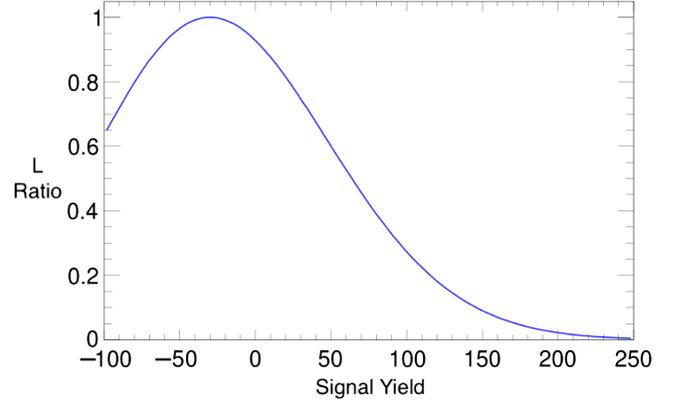
FIG. 3. Smeared likelihood ratio distribution with different fixed signal yields.

from the MC simulation and fit result, and obtain an upper limit at 90% credibility on the $B^+ \to p\bar{n}\pi^0$ branching fraction of 6.3×10^{-6} by searching out the point covering 90% of total area under the smeared likelihood distribution with different fixed signal yields from 0. Figure 3 shows the likelihood ratio distribution with different fixed signal yields. We also add components in Fig. 2 to show the distributions of signal with a hypothetical branching fraction corresponding to 15 times the upper limit we calculate and $B^+ \to p\bar{\Lambda}\pi^0$ with the yield 100 times.

V. SYSTEMATIC UNCERTAINTIES

Sources of the systematic uncertainties for the branching fraction of $B^+ \to p\bar{n}\pi^0$ are summarized in Table I. The uncertainty on the number of $B\bar{B}$ pairs is 1.4%. We calculate the signal efficiency using the reweighted phase-space model of $B^+ \to p\bar{n}\pi^0$, and estimate the efficiency uncertainty to be 2.6% by measuring the difference of signal efficiency with the uniform phase-space distribution. By using the partially reconstructed $D^{*+} \to D^0\pi^+$ with $D^0 \to \pi^+\pi^- K_S^0$ events, the uncertainty due to the charged-track reconstruction efficiency is estimated to be 0.4% per track. We use a

TABLE I. Table of systematic uncertainties (%) for the branching fraction of $B^+ \to p\bar{n}\pi^0$, all considered independent.

Uncertainties	$B^+ \to p\bar{n}\pi^0$
$N_{B\bar{B}}$	1.4
Decay model	2.6
Tracking	0.4
p identification	0.3
π^0 reconstruction	2.3
$\bar{n}$ selection	6.0
Continuum suppression	1.2
ΔE, C_{nbtr} shape	9.1
Sum	11.7
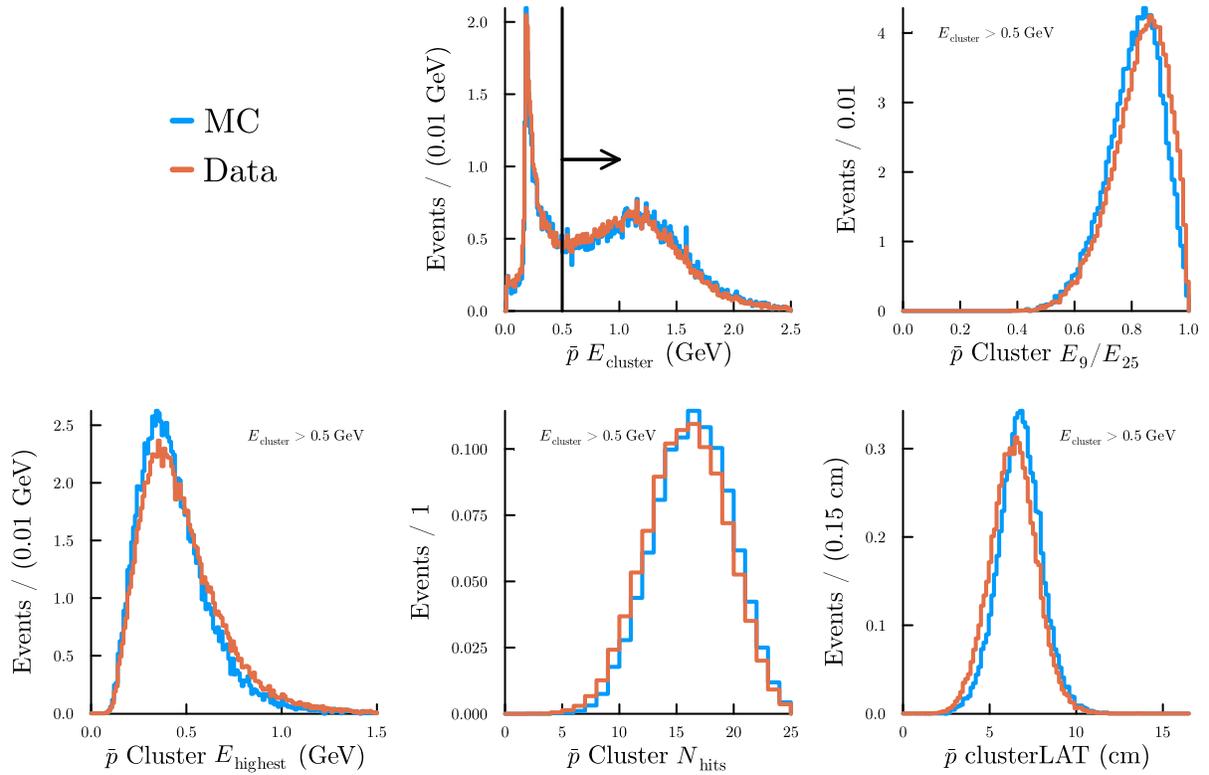
FIG. 4. Discrepancy of the training variables of the $\bar{n}$ MVA between MC and data, inspected with the sideband-subtracted $\bar{\Lambda} \to \bar{p}\pi^+$.

$\Lambda \to p\pi^-$ sample to calibrate the MC proton identification efficiency and assign an uncertainty of 0.3% for $B^+ \to p\bar{n}\pi^0$. For π^0 reconstruction, we estimate its uncertainty to be 2.3% by using a $\tau^- \to \pi^-\pi^0\nu$ data sample [22]. The uncertainty due to the fixed normalization of the $B^+ \to p\bar{\Lambda}\pi^0$ component is found to be negligible [21].

We calibrate the efficiency of the $\bar{n}$ selection with a controlled sample, $\bar{\Lambda} \to \bar{p}\pi^+$. We consider two efficiencies separately. The first is the detection efficiency of an anti-nucleus cluster in ECL, $\epsilon_{\text{detection}}$. The second is the efficiency of the MVA cut of such cluster, $\epsilon_{\text{selection}}$. By extrapolating from the $\bar{p}$ track into ECL and selecting clusters that pass the $\bar{n}$ MVA selection, the two efficiencies can be studied simultaneously with the $\bar{\Lambda} \to \bar{p}\pi^+$ sample, i.e.,

$$\epsilon_{\text{detection}} \cdot \epsilon_{\text{selection}} = \frac{N(\bar{\Lambda} \text{ whose clusters pass } \bar{n} \text{ selection})}{N(\bar{\Lambda})}.$$

The efficiency corrections are tabulated in 10×12 bins of the momentum (12 regions) and $\cos\theta$ (ten regions) of $\bar{p}$. The MC efficiency is corrected using the table, and the statistical uncertainty of the tabulated corrections is 0.4%. The effect of smearing due to $\bar{n}$ momentum and $\cos\theta$ resolution is 0.1%. The dominant uncertainty for $\bar{n}$ selection is related to the efficiency difference between $\bar{n}$ and $\bar{p}$ in MC. In response to the ignorance of low-momentum efficiency difference which

is model dependent, we assign twice of the discrepancy as the systematic uncertainty which amounts to 6.0%.

The discrepancy of the training variables of the $\bar{n}$ MVA between data and MC, inspected with the sideband-subtracted $\bar{\Lambda} \to \bar{p}\pi^+$ controlled sample, is shown in Fig. 4.

The decay mode $B^0 \to \bar{D}^0\pi^0$ with $\bar{D}^0 \to K^+\pi^-$ is used to estimate the systematic error due to continuum suppression. We estimate the signal yields and signal efficiencies before and after the selection criteria of $C_{\text{nb}} > 0.9$. The difference of the efficiencies between the data and MC, 1.2%, is taken as the systematic uncertainty due to continuum suppression. The decay modes $B^+ \to p\bar{p}\pi^+$ and $B^+ \to p\bar{p}K^+$ are used to estimate the systematic error due to the shapes of ΔE and C_{nbtr}, where the magnitude of the $\bar{p}$ momentum is derived from its direction, as the $\bar{n}$ momentum is reconstructed in $B^+ \to p\bar{n}\pi^0$. The uncertainty due to fixing these shapes is examined by repeating the fit with each parameter varied by one standard deviation from its nominal value. The resulting difference from the nominal fit is taken as the systematic uncertainty. After linearly summing up uncertainties of all parameters for ΔE and C_{nbtr}, the uncertainty for ΔE and C_{nbtr} shapes is evaluated to be 9.0%. In addition, since the control mode does not include π^0 in the final state, we estimated the effect of π^0 to ΔE by using another control sample, $B^0 \to \bar{D}^0\pi^0$, where we apply the same method of shape uncertainty evaluation on its ΔE distribution. We get 1.5% as the result. After merging it to the uncertainty, we get 9.1% as the total shape uncertainty.

The assumption of no correlation between ΔE and C_{nbtr} is examined by replacing the PDF of B signal events with the corresponding 2-D histogram function; and the associated uncertainty is found to be negligible.

VI. SUMMARY

In summary, we report a decay upper limit of 6.3×10^{-6} for $B^+ \to p\bar{n}\pi^0$ at 90% credibility. Since $B^+ \to p\bar{n}\pi^0$ is not found, this study provides no evidence for the contribution of the external $W^+ \to p\bar{n}$ process. In order to understand the whole picture of rare baryonic B decays, other modes including an $\bar{n}$ such as $B^0 \to p\bar{n}\pi^-$, $B^0 \to p\bar{n}K^-$ and $B^+ \to p\bar{n}\bar{D}^0$ should be studied.

[1] J.-T. Wei *et al.* (Belle Collaboration), Phys. Lett. B **659**, 80 (2008).

[2] R. Workman *et al.* (Particle Date Group), Prog. Theor. Exp. Phys. **2022**, 083C01 (2022).

[3] A. J. Bevan, B. Golob, Th. Mannel, S. Prell, and B. D. Yabsley, Eur. Phys. J. C **74**, 3026 (2014).

[4] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. Lett. **113**, 141801 (2014).

[5] S. Anderson *et al.* (CLEO Collaboration), Phys. Rev. Lett. **86**, 2732 (2001).

[6] P. del Amo Sanchez *et al.* (Belle Collaboration), Phys. Lett. B **659**, 80 (2008).

[7] B. Pal *et al.* (Belle Collaboration), Phys. Rev. D **99**, 091104 (2019).

[8] A. Abashian *et al.* (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 117 (2002).

[9] J. Brodzicka *et al.*, Prog. Theor. Exp. Phys. **2012**, 04D001 (2012).

[10] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res., Sect. A **499**, 1 (2003).

[11] T. Abe *et al.*, Prog. Theor. Exp. Phys. **2013**, 03A001 (2013).

[12] D. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).

[13] R. Brun *et al.*, CERN Report No. DD/EE/84-1, 1984.

[14] F. Chollet *et al.*, Keras, https://keras.io (2015).

[15] A. F. Agarap, arXiv:1803.08375.

[16] D. P. Kingma and J. Ba, arXiv:1412.6980.

[17] M. Feindt and U. Kerzel, Nucl. Instrum. Methods Phys. Res., Sect. A **559**, 190 (2006).

[18] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978). The modified moments used in this paper are described in S. H. Lee *et al.* (Belle Collaboration), Phys. Rev. Lett. **91**, 261801 (2003).

[19] J. Bjorken and S. Brodsky, Phys. Rev. D **1**, 1416 (1970).

[20] H. Kakuno *et al.* (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **533**, 516 (2004).

[21] M.-Z. Wang *et al.* (Belle Collaboration), Phys. Rev. D **76**, 052004 (2007).

[22] S. Ryu *et al.* (Belle Collaboration), Phys. Rev. D **89**, 072009 (2014).