Search for the rare decay of $B^+ \rightarrow \ell^+ \nu \gamma$ with improved hadronic tagging

M. Gelb,1 F. U. Bernlochner,1 P. Goldenzwieg,1 F. Metzner,1 I. Adachi,19,15 H. Aihara,86 S. Al Said,80,36 D. M. Asner,4 H. Atmacan,76 T. Aslams,55 R. Ayad,80 V. Babu,81 I. Badhrees,80,35 V. Bansal,67 P. Behera,26 C. Beloño,4 B. Bhuyan,24 J. Biswal,31 A. Bobrov,45,33 M. Bracko,49,33 N. Braun,1 L. Cao,1 D. Cervenkov,6 V. Chekelian,50 A. Chen,5 B. G.cheon,17 K. Chilikin,44 K. Cho,38 S.-K. Choi,16 Y. Choi,78 S. Choudhury,25 D. Cinabro,90 S. Cuniliffe,9 N. Dashi,23 S. Di Carlo,42 Z. Doležal,16 T. Dong,19,15 S. Eidelman,55,44 D. Epifanov,5,65 J. E. Fast,67 T. Ferber,9 A. Frey,14 B. G. Fulsom,67 R. Garg,58 V. Gaur,89 N. Gabyshev,5,65 A. Garmash,5,65 J. Gemmler,4 A. Giri,24 D. Greenwald,20 S. Haba,19,15 T. Hara,19,15 K. Hayasaka,63 H. Hayashii,58 W.-S. Hou,61 K. Inami,56 G. Inguglia,9 A. Ishikawa,9 I. Adachi,19,15 M. Iwasaki,66 Y. Iwasaki,19,15 W. Jacobs,30 S. Jia,3 Y. Jin,86 D. Joffe,34 K. K. Joo,7 J. Kahn,46 A. B. Kaliyar,26 T. Kawasaki,37 H. Kichimi,19 C. Kiesling,50 D. Y. Kim,75 H. J. Kim,41 S. H. Kim,17 P. Kodyš,6 S. Korpar,19,33 D. Kotchetkov,18 P. Križan,45,33 R. Kroeger,52 P. Krokovny,5,65 T. Kuhr,46 R. Kulasiri,34 A. Kuzmin,5,65 Y.-J. Kwon,92 J. L. Lange,12 I. S. Lee,17 J. K. Lee,73 J. Y. Lee,73 S. C. Lee,41 Y. B. Li,69 L. Li Gioi,50 J. Libby,26 D. Livantsev,89,19 P.-C. Lu,61 M. Lubej,33 T. Luo,11 J. MacNaughton,1 M. Masuda,30,57 K. Miyabayashi,38 H. Miyata,63 R. Mizuk,4 B. G. Mohanty,81 T. Mori,56 M. Mrvar,39 R. Musa,19,15 K. J. Nath,24 Z. Natkuniec,62 M. Nayak,90,19 M. Niyyama,40 N. K. Nisar,80,19,15 S. Ogawa,83 P. Pakhlov,44,55 G. Pakhlov,44,55 B. Pal,5 S. Pardi,70 S. Paul,82 T. K. Pedlar,47 R. Pestotnik,33 E. Piilonen,89 V. Popov,14,55 E. Prencipe,44 M. Prim,46 M. Ritter,46 A. Rostomyan,4 G. Russo,30 D. Sahoo,41 Y. Sakai,19,15 M. Salehi,48,46 S. Sandilya,3 L. Santelj,19 T. Sanuki,84 V. Savinov,70 O. Schneider,43 G. Schnell,2,22 J. Schueler,18 C. Schwanda,9 Y. Seino,61 K. Senyo,91 O. Seon,19,56 M. E. Sevior,31 C. P. Shen,3 T.-A. Shibata,87 J.-G. Shiu,61 B. Shwartz,5,65 F. Simon,50 E. Solovieva,44,55 S. Stanic,64 M. Starić,33 J. F. Strube,67 M. Sumihama,13 T. Sunimiyoshi,88 W. Sutcliffe,1 M. Takizawa,74,20,71 K. Tanida,32 Y. Tao,10 F. Tenchini,9 M. Uchida,87 T. Uglow,44,55 Y. Unno,71 S. Uno,19,15 P. Urquijo,51 R. Van Tonder,1 G. Varner,18 K. E. Varvell,79 B. Wang,8 C. H. Wang,60 M.-Z. Wang,61 P. Wang,28 X. L. Wang,11 M. Watambe,63 S. Watamuki,46 E. Widmann,77 E. Won,39 H. Yamamoto,84 S. B. Yang,39 H. Ye,7 J. H. Yin,28 C. Z. Yuan,79 Y. Yusa,65 S. Zakharov,44,55 Z. P. Zhang,72 V. Zhilich,16 V. Zhukova,44 and V. Zhulanov45,65

(The Belle Collaboration)

1Institut für Experimentelle Teilchenphysik, Karlsruher Institut für Technologie, 76131 Karlsruhe
2University of the Basque Country UPV/EHU, 48080 Bilbao
3Beihang University, Beijing 100191
4Brookhaven National Laboratory, Upton, New York 11973
5Budker Institute of Nuclear Physics SB RAS, Novosibirsk 630090
6Faculty of Mathematics and Physics, Charles University, Prague 121 16 Prague
7Chonnam National University, Kwangju 660-701
8University of Cincinnati, Cincinnati, Ohio 45221
9Deutsches Elektronen-Synchrotron, 22607 Hamburg
10University of Florida, Gainesville, Florida 32611
11Key Laboratory of Nuclear Physics and Ion-beam Application (MOE) and Institute of Modern Physics, Pudan University, Shanghai 200443
12Justus-Liebig-Universität Gießen, 35392 Gießen
13Gifu University, Gifu 501-1193
14J. Physikalisches Institut, Georg-August-Universität Göttingen, 37073 Göttingen
15SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193
16Gyeongsang National University, Chinju 660-701
17Hanyang University, Seoul 133-791
18University of Hawaii, Honolulu, Hawaii 96822
19High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801
20J-PARC Branch, KEK Theory Center, High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801
21Forschungszentrum Jülich, 52425 Jülich
22IKERBASQUE, Basque Foundation for Science, 48013 Bilbao
23Indian Institute of Technology Kharagpur, Satya Nagar 751007
24Indian Institute of Technology Guwahati, Assam 781039
25Indian Institute of Technology Hyderabad, Telangana 502285
26Indian Institute of Technology Madras, Chennai 600036
We present the result of the search for the rare B meson decay of $B^+ \to \ell^+ \nu \gamma$ with $\ell = e, \mu$. For the search the full data set recorded by the Belle experiment of 711 fb$^{-1}$ integrated luminosity near the $T(4S)$ resonance is used. Signal candidates are reconstructed for photon energies $E_{\gamma}$ larger than 1 GeV using a novel multivariate tagging algorithm. The novel algorithm fully reconstructs the second $B$ meson produced in the collision using hadronic modes and was specifically trained to recognize the signal signature in combination with hadronic tag-side $B$ meson decays. This approach greatly enhances the performance. Background processes that can mimic this signature, mainly charmless semileptonic decays and continuum processes, are suppressed using multivariate methods. The number of signal candidates is determined by analyzing the missing mass squared distribution as inferred from the signal side particles and the kinematic properties of the tag-side $B$ meson. No significant excess over the background-only hypothesis is observed and upper limits on the partial branching fraction $\Delta B$ with $E_{\gamma} > 1$ GeV individually for electron and muon final states as well as for the average branching fraction of both lepton final states are reported. We find a Bayesian upper limit of $\Delta B(B^+ \to \ell^+ \nu \gamma) < 3.0 \times 10^{-6}$ at 90% CL and also report an upper limit on the first inverse moment of the light-cone distribution amplitude of the $B$ meson of $\lambda_B > 0.24 \text{GeV}$ at 90% CL.

I. INTRODUCTION

The precision study of leptonic $B$ meson decays offers a unique path to test the Standard Model (SM) of particle physics: new heavy mediators or sterile neutrinos could contribute to the decay amplitudes and lead, for instance, to lepton flavor universality breaking effects. In the SM, however, leptonic transitions are helicity suppressed, making an observation in final states involving electrons or muons challenging. The helicity suppression can be lifted, if one considers final states which involve an additional photon in the final state, emitted for instance from the light $u$ quark. The decay rate for such $B^+ \to \ell^+ \nu \gamma$ processes [1] is suppressed by a factor of $\alpha_{em}$ and the resulting decay amplitude [2] may be expressed as

$$\mathcal{A}(B^+ \to \ell^+ \nu \gamma) = \frac{G_F V_{ub}}{4\sqrt{2}} \bar{u} \gamma_{\mu} \ell \gamma^\mu P_L \bar{\nu} \gamma_{\nu} P_L \bar{\ell} |B\rangle,$$

(1)

with the projection operator $P_L = (1 - \gamma_5)/2$, $G_F$ denoting Fermi’s constant, $\ell$ denoting either an electron or a muon field, $u$ and $\bar{u}$ as quark fields and $V_{ub}$ denoting the relevant Cabibbo-Kobayashi-Maskawa (CKM) [3, 4] matrix element of the transition. The hadronic transition can be fully described by two form factors parameterizing the axial-vector and vector hadronic currents, denoted by $F_A$ and $F_V$, respectively. The differential decay rate as a function of these two form factors and the photon energy $E_{\gamma}$ is given by

$$\frac{d\Gamma(B^+ \to \ell^+ \nu \gamma)}{dE_{\gamma}} = \alpha_{em} G_F^2 |V_{ub}|^2 m_B E_{\gamma}^3 \left(1 - \frac{2E_{\gamma}}{m_B}\right) \times \left( |F_V|^2 + |F_A + \frac{e_f f_B}{E_{\gamma}}|^2 \right),$$

(2)

with $e_f$ denoting the charge of the lepton and $m_B$ the $B$ meson mass. At high photon energies of $E_{\gamma} > 1$ GeV, both form factors can be expanded [5] as

$$F_V(E_{\gamma}) = \frac{e_f f_B m_B}{2E_{\gamma} \lambda_B (\mu)} R(E_{\gamma}, \mu) + \xi(E_{\gamma}) + \Delta \xi(E_{\gamma}),$$

(3)

$$F_A(E_{\gamma}) = \frac{e_f f_B m_B}{2E_{\gamma} \lambda_B (\mu)} R(E_{\gamma}, \mu) + \xi(E_{\gamma}) - \Delta \xi(E_{\gamma}),$$

(4)

with $f_B$ denoting the $B$ meson decay constant and $e_u$ the charge of the $u$ quark line. The factor $R(E_{\gamma}, \mu)$ accounts for photon emissions from the light spectator quark in the $B$ meson and is unity at tree level. The $\xi$ and $\Delta \xi$ terms are power suppressed by $1/m_B$ and $1/(2E_{\gamma})$ and contain a symmetry conserving and breaking part for both form factors $F_A$ and $F_V$. In Refs. [6, 7] the leading contributions to $\xi$ and $\Delta \xi$ were evaluated and predictions with an accuracy of $O(20\%)$ were presented. The parameter $\lambda_B$ is related to the first inverse moment of the leading-twist $B$ meson light-cone distribution amplitude $\phi_+(w)$ in the high energy limit, $\lambda_B^2 = \int_0^\infty dw \phi_+(w)$ with $w$ denoting the light-cone momentum. This parameter is of great relevance and serves as an important input to understand QCD factorization used to predict non-leptonic $B$ meson decays [5][10]. The partial branching of $B^+ \to \ell^+ \nu \gamma$ is expected to be of the order of $O(10^{-6})$ for photon energies of $E_{\gamma} > 1$ GeV and $\lambda_B$ in the range of several hundred MeV [5].

The search for the rare $B^+ \to \ell^+ \nu \gamma$ decay thus serves two purposes: first a limit or observation of the partial branching fraction can be used in combination with the decay rate Eq. (2) and input from theory for $\xi$ and $\Delta \xi$, as well as for $f_B$ and $V_{ub}$, to experimentally determine the value of $\lambda_B$. Second, with the Belle II experiment the $B^+ \to \ell^+ \nu \gamma$ decay rate could offer a future additional
path to determine $|V_{ub}|$ by using lattice calculations for $\lambda_B$. In this manuscript we present a novel method to determine $\lambda_B$, which uses the measured ratio of $B^+ \to \ell^+ \nu \gamma$ with respect to $B^+ \to \pi^0 \ell^+ \nu$. This cancels the explicit dependence of $V_{ub}$ on $\lambda_B$ and in the ratio several experimental uncertainties cancel.

The search in this manuscript supersedes the earlier result of Ref. [11], using an improved hadronic tagging and more accurate modeling of the charmless semileptonic backgrounds. The hadronic tagging is based on the Full Event Interpretation (FEI) [12], which uses multivariate methods trained to recognize the signal signature (consisting of a single photon with a lepton candidate) in conjunction with a hadronic $B$ meson decay. This signal-specific approach enhances the selection efficiency in comparison to the previous analysis by a factor of three, which results in an improvement of 1.9$\sigma$ for the expected significance for a given partial branching fraction of $\Delta B(B^+ \to \ell^+ \nu \gamma) = 5.0 \times 10^{-6}$ and with a similar ratio of signal over background events. Backgrounds that can mimic signal, most importantly $B^+ \to \pi^0 \ell^+ \nu$ and other charmless semileptonic decays, are suppressed with a dedicated $\pi^0$ and $\eta$ veto, which combines the signal photon with other calorimeter clusters to form a multivariate classifier. The signal is extracted by analyzing the missing mass squared distribution, $M^2_{miss}$, which for correctly reconstructed signal decays peaks around 0 GeV$^2$. Semileptonic background and event candidates from continuum processes are shifted towards positive values. The $B^+ \to \ell^+ \nu \gamma$ branching fraction is extracted simultaneously with a $B^+ \to \pi^0 \ell^+ \nu$ control sample, which constrains the most dominant background contribution. The extracted yields are converted into (partial) branching fractions to determine the ratio of $B^+ \to \ell^+ \nu \gamma$ to $B^+ \to \pi^0 \ell^+ \nu$ decays.

This manuscript is organized as follows: Section II details the analysis strategy and selection. Section III summarizes the statistical analysis of the $M^2_{miss}$ distribution and the limit setting procedure. In Section IV the calibration procedure of the multivariate tagging algorithm is summarized. Sections V and VI discuss the systematic uncertainties and the main results. The manuscript concludes with a summary in Section VII.

II. DATA SET AND ANALYSIS STRATEGY

The full Belle data set of $(772 \pm 10) \times 10^6$ $B$ meson pairs is analyzed, produced at the KEKB accelerator complex [13] at a center-of-mass energy of 10.58 GeV 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 Čerenkov 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 located outside of the coil is instrumented to detect $K_L^0$ mesons and to identify muons (KLM). A more detailed description of the detector can be found in Ref. [14].

All analyses steps are carried out using the Belle II analysis software framework [15] and the recorded Belle collision data and Monte Carlo (MC) simulated samples were converted using the tool described in Ref. [16]. MC samples of $B$ meson decays and non-resonant $e^+ e^- \to q\bar{q}$ with $q = u, d, s, c$ continuum processes are generated using the EvtGen generator [17], with sample sizes corresponding to approximately ten and six times the Belle collision data, respectively. The interactions of particles traversing the detector are simulated using Geant3 [18]. QED final state radiation (FSR) is simulated using the PHOTOS [19] package. The $B^+ \to \ell^+ \nu \gamma$ signal is simulated using the calculation of Ref. [20].

Charmless semileptonic decays are one of the main contributions to the background in the analysis, in particular $B^+ \to \pi^0 \ell^+ \nu$ with $\pi^0 \to \gamma\gamma$ can mimic the signal final state. The $B^+ \to \pi^0 \ell^+ \nu$ background is modeled using the BCL form factor parametrization [21] with central values and uncertainties from the global fit of Ref. [22]. The remaining charmless semileptonic background is modeled using a mix of resonant and non-resonant modes: the resonant contributions for $B^+ \to \omega \ell^+ \nu_s$, $B^+ \to \rho^0 \ell^+ \nu_s$, $B^0 \to \rho^- \ell^+ \nu_s$ are simulated according to a pole model documented in Ref. [17]. The contributions from $B^+ \to \eta \ell^+ \nu_s$, $B^+ \to \eta' \ell^+ \nu_s$, $B^+ \to f_2(1270) \ell^+ \nu_s$, and $B^+ \to b_1(1235) \ell^+ \nu_s$ are modeled using the ISGW2 model [23]. Non-resonant contributions are modeled using the DFN calculation [24] with a choice of its parameters $\lambda^{SF}_1$ and $\lambda^{SF}_2$ to approximatively reproduce the first and second moments of the inclusive $m_{\ell\nu\gamma}$ distribution. The fragmentation of the $X_u$ system is performed by Jetset [25]. Background processes from leptonic, semileptonic or hadronic $B$ meson decays are included in the simulation and all relevant decay branching fractions are corrected to correspond to the values of Ref. [26]. The efficiencies in the MC are corrected using data-driven methods and are described later.

The presence of a neutrino in $B^+ \to \ell^+ \nu \gamma$ decays prohibits the full reconstruction of the signal $B$ meson, and thus the entire $\Upsilon(4S)$ decay chain is considered to infer the missing momentum of the neutrino. First, a signal electron or muon and a photon candidate with at least 1 GeV in the laboratory frame are identified. The FEI algorithm then hierarchically reconstructs the rest of the event (ROE): The tag-side $B$ meson ($B_{tag}$) is reconstructed using 29 explicit hadronic decay channels, leading to $O(10000)$ final states. An optimized implementation of gradient-boosted decision trees (BDT) [27] is used for multivariate classification at the individual stages of the $B_{tag}$ reconstruction, which progresses by forming intermediate particles for $(J/\psi, \pi^0, K^0_S, D, D_s, D_s')$ from stable particle candidates for $(e^+, \mu^+, K^+, \pi^+, \gamma)$ to re-
construct $B$ candidates in six distinct stages. To each reconstructed candidate a signal probability $P_{\text{FEI}}$ is assigned, which is calculated by the respective classifier on the properties of the candidate (such as the invariant mass or vertex fit information) to discriminate signal from background candidates. A detailed description of the entire algorithm can be found in Ref. [12] and references therein. By reconstructing the signal side first, the FEI trains specifically to recognize $B^+ \to \ell^+ \nu \gamma$ decays in conjunction with hadronic $B$ meson decays. Using the kinematic information of the $B_{\text{tag}}$, the four-momenta of the signal side $B$ meson ($B_{\text{sig}}$) and thus the missing neutrino can be reconstructed in the center-of-mass frame as

$$p_{B_{\text{sig}}} = \begin{pmatrix} \sqrt{s}/2 \\ -\vec{p}_{B_{\text{tag}}} \end{pmatrix} \quad \text{and} \quad p_{\nu} = (p_{B_{\text{sig}}} - p_{\ell} - p_{\gamma}),$$

where $\sqrt{s}$ denotes the center-of-mass energy of the colliding $e^+ e^-$ pair, and $p_{\ell}$ and $\vec{p}_{\gamma}$ are the four- and three-momentum of a given particle $x$.

Only events with not more than 12 tracks are selected as the signal side consists of only one charged track and signal events typically have a low track multiplicity. Photons are identified as energy depositions in the calorimeter without an associated track. Only photons with an energy deposition of $E_\gamma > 100$ MeV, 150 MeV, and 50 MeV in the forward end-cap, backward end-cap and barrel part of the calorimeter, respectively, are considered. Charged tracks are required to have a distance of closest approach to the nominal interaction point transverse to and along the beam axis of $|d\tau| < 2$ cm and $|dz| < 4$ cm, respectively. Charged tracks are identified as electron or muon candidates by combining the information of multiple subdetectors into a likelihood ratio, the lepton identification ($\mathcal{L}_{\text{LID}}$). For electrons the identifying features are the ratio of the energy deposition in the ECL with respect to the reconstructed track momentum, the energy loss in the CDC, the shower shape in the ECL, the quality of the geometrical matching of the track to the shower position in the ECL, and the photon yield in the ACC [28]. Muon candidates are identified from charged track trajectories extrapolated to the outer detector. The identifying features are the difference between expected and measured penetration depth as well as the transverse deviation of KLM hits from the extrapolated trajectory [29]. Charged tracks are identified as pions or kaons using a likelihood classifier using information from the CDC, ACC, and TOF subdetectors.

Electrons can radiate sizable fractions of their kinetic energy through bremsstrahlung and FSR processes. Candidates for such photons are identified in the ECL using a cone of 5 degrees around the initial trajectory of the electron candidate, such that only photons radiated near the interaction region can be found. The bremsstrahlung and FSR photons are required to have an energy of $E_\gamma < 1$ GeV and if several photon candidates are identified, only the photon with the highest energy is used. The four-momentum of the signal side electron candidate is then corrected by adding the photon energy accordingly. The prompt signal-photon candidates from the $B^+ \to \ell^+ \nu \gamma$ decay are required to have $E_\gamma > 1$ GeV in the rest frame of the $B_{\text{sig}}$ meson. Further, signal-photon candidates must provide $R_{0/25} > 0.9$, which is defined as the ratio of the energies deposited in the 3 $\times$ 3 with respect to the energy deposited in the 5 $\times$ 5 CsI(Tl) crystals around the maximal energy deposition. $\mathcal{L}_{\text{LID}} > 0.8$ is required for electron and muon candidates. Since the efficiencies of the $\mathcal{L}_{\text{LID}}$ requirement can differ between MC and data, an efficiency correction is applied, measured on four-lepton and inclusive $B \to XJ/\psi (\to \ell^+ \ell^-)$ decays in bins of lepton momentum in the laboratory frame and polar angle. Furthermore, the invariant mass $M_{\ell\ell}$ of the reconstructed lepton-photon pair has to be within (1.0, 6.0) GeV.

Before the FEI tagging algorithm is applied, the corresponding tag-side is cleaned to remove events which do not allow for a reasonable tag-side reconstruction. Photon candidates must provide $R_{0/25} > 0.9$. In addition, a cut on the difference between the beam energy and the energy of the ROE calculated in the center-of-mass frame, $\Delta E_{\text{ROE}} < 2.0$ GeV, is applied. The beam-constrained mass $M_{\text{bc,ROE}} = \sqrt{(\sqrt{s}/2)^2 - \vec{p}_{\text{ROE}}^2}$ of the ROE calculated in the center-of-mass frame has to be larger than 4.8 GeV.

After the reconstruction of both $B_{\text{sig}}$ and $B_{\text{tag}}$, the $T(4S)$ candidate can be reconstructed. A best-candidate selection based on $P_{\text{FEI}}$ (of the $B_{\text{tag}}$ candidate) is performed, if more than one candidate per event is reconstructed. Further combinatoric background is removed with cuts on the reconstructed invariant mass of the $T(4S)$ candidate of $M_{T(4S)} \in (7.5, 10.5)$ GeV (considering the missing neutrino), and the difference between the beam energy and the energy of the $B_{\text{tag}}$ candidate, $\Delta E \in (-0.15, 0.1)$ GeV. The beam-constrained mass $M_{\text{bc}} = \sqrt{(\sqrt{s}/2)^2 - \vec{p}_{\text{bc}}^2}$ of the $B_{\text{tag}}$ candidate has to be within (5.27, 5.29) GeV. Only events with no unassigned tracks (either for $B_{\text{sig}}$ or $B_{\text{tag}}$) and $E_{\text{ECL}} \leq 0.9$ GeV, where $E_{\text{ECL}}$ is the sum of the remaining unassigned energy depositions in the ECL, are retained. Continuum background is removed by applying an additional cut $P_{\text{FEI}} > 0.01$, whose value was chosen by studying the $M_{\text{bc}}$ sideband, defined as $M_{\text{bc}} \in (5.24, 5.27)$ GeV.

The background from $B^+ \to \pi^0 \ell^+ \nu_\ell$ and $B^+ \to \eta \ell^+ \nu_\ell$ is peaking in $M_{\text{miss}}^2$ and is suppressed in a two-step procedure. First, the signal-side photon is combined with any other photon in the ROE to reconstruct a $\pi^0$ candidate. The event is vetoed if the candidate satisfies $M_{\pi^0} \in (110, 160)$ MeV, where $M_{\pi^0}$ is the invariant mass of the two photons combined. Further, a multivariate method is trained to suppress the remaining $B^+ \to \pi^0 \ell^+ \nu_\ell$ and $B^+ \to \eta \ell^+ \nu_\ell$ background using the following variables: the number of ECL cluster hits used for the signal-side photon reconstruction, $R_{0/25}$, the lateral distribution of the energy of the ECL cluster hits, the angle between...
the signal-side photon and the missing momentum \( \vec{p}_X \) calculated in the rest frame of the \( B_{\text{sig}} \), \( E_{\text{ECL}} \), and the energy asymmetry, revealing the asymmetry in energy distribution of the lepton and photon candidate of the \( B_{\text{sig}} \), calculated as

\[
A = \frac{E_\ell E_\gamma}{E_\ell + E_\gamma},
\]

To improve control over the normalization of the peaking background, control samples for \( B^+ \to \pi^0 \ell^+ \nu_\ell \), with \( \ell = e, \mu \), are reconstructed. The signal side selection is slightly adapted for the \( B^+ \to \pi^0 \ell^+ \nu_\ell \) selection: instead of a single photon with \( E_\gamma > 1 \text{ GeV} \), two photon candidates are combined to form a \( \pi^0 \) candidate and only events with an invariant mass of \( M_{\gamma\gamma} \in (115, 152) \text{ MeV} \) (corresponding to approximately \( \pm 3 \sigma \) in \( \pi^0 \) mass resolution), are retained. Both control samples and the \( B^+ \to \ell^+ \nu_\ell \gamma \) signal decays are analyzed simultaneously to extract the desired signal yields and to constrain the peaking \( B^+ \to \pi^0 \ell^+ \nu_\ell \) contaminations in the signal candidates.

For both the \( B^+ \to \ell^+ \nu_\ell \gamma \) and the \( B^+ \to \pi^0 \ell^+ \nu_\ell \) selections, non-resonant continuum processes are suppressed using a multivariate approach with the aforementioned implementation of BDT. The event topology for continuum processes differs from that of \( B \) meson decays. This can be exploited to suppress continuum events by using event shape variables, such as the magnitude of the thrust of final state particles forming the \( B_{\text{sig}} \) and ROE candidates, the angle between the \( B_{\text{sig}} \) and the \( z \)-axis and between the \( B_{\text{sig}} \) and the ROE, the reduced Fox-Wolfram moment \( R_2 \), the modified Fox-Wolfram moments \( m_3 \) and CLEO Cones \( \Omega \).

The cuts on the multivariate classifier for continuum and the peaking background suppression are simultaneously optimized with Punzi’s figure of merit \( c_2 \). After all selection steps, we obtain a signal reconstruction efficiency for \( B^+ \to \ell^+ \nu_\ell \gamma \) decays of 0.64\% (0.67\%) for the electron (muon) final state. On the normalization sample we obtain an efficiency of 0.38\% for both final states for \( B^+ \to \pi^0 \ell^+ \nu_\ell \) decays.

To discriminate the signal from background decays, the missing mass squared \( M_{\text{miss}}^2 \) of the event is calculated as

\[
M_{\text{miss}}^2 = m_\pi^2 = p_\pi^2 = (p_B_{\text{sig}} - p_\ell - p_X)^2,
\]

with \( p_X \) denoting \( p_\gamma \) for \( B^+ \to \ell^+ \nu_\ell \gamma \) signal events, and \( p_\pi \) for \( B^+ \to \pi^0 \ell^+ \nu_\ell \) normalization events, respectively. The signal and background yields are then obtained using the statistical analysis described in Section III. The analysis procedure is validated using two signal-depleted sidebands: an off-resonance sample, recorded 40 MeV below the \( T(4S) \) resonance, and the \( M_{\text{sig}} \) sideband were analyzed. Both showed good agreement between data and the MC expectation.

### III. STATISTICAL ANALYSIS AND LIMIT SETTING PROCEDURE

Signal and background yields are extracted using a binned maximum likelihood fit of the \( M_{\text{miss}}^2 \) distribution. For an individual channel, the likelihood function is constructed as

\[
\mathcal{L}_c = \prod_{i} \mathcal{P}(n_i; \nu_i),
\]

with \( \mathcal{P}(n_i; \nu_i) = \frac{\nu_i^{n_i}}{n_i!} e^{-\nu_i} \) denoting the Poisson distribution with \( n_i \) and \( \nu_i \) the number of observed and expected events in a given bin \( i \) of \( M_{\text{miss}}^2 \), respectively. Three different likelihood fits are carried out in this manuscript:

i. Semileptonic \( B^+ \to D^0 \ell^+ \nu_\ell \) decays are analyzed to determine a calibration factor for the FEI tagging efficiency. The selection and obtained calibration factors are further discussed in Section [IV]

ii. The branching fraction of \( B^+ \to \pi^0 \ell^+ \nu_\ell \) events is determined as a cross check of the FEI calibration procedure, cf. Section [IV]

iii. The \( B^+ \to \ell^+ \nu_\ell \gamma \) signal events are analyzed using a simultaneous fit to the \( B^+ \to \ell^+ \nu_\ell \gamma \) and \( B^+ \to \pi^0 \ell^+ \nu_\ell M_{\text{miss}}^2 \) distributions. A global likelihood function is constructed as

\[
\mathcal{L} = \prod_c \mathcal{L}_c \times \prod_k \tilde{G}(\theta_k)
\]

with \( c \) denoting the reconstructed event type corresponding to the four categories defined by the \( B^+ \to e^+ \nu_\ell \gamma, B^+ \to \mu^+ \nu_\ell \gamma, B^+ \to \pi^0 e^+ \nu_\ell \), and \( B^+ \to \pi^0 \mu^+ \nu_\ell \) channels. Further, \( \tilde{G}(\theta_k) \) denotes the standard normal distribution for nuisance parameters \( \theta_k \), which incorporate systematic uncertainties into the likelihood function. The various systematic uncertainties are further discussed in Section [IV].

The expected number of events in a given bin \( i \) of the \( M_{\text{miss}}^2 \) distribution and in a given category is constructed as

\[
\nu_i = \sum_j n_j f_{ij},
\]

with \( n_j \) the total number of events of type \( j \) and \( f_{ij} \) denoting the expected fraction of events of type \( j \) in the \( i \)-th bin. The fractions \( f_{ij} \) are obtained from the MC simulation and the event types for the \( B^+ \to D^0 \ell^+ \nu_\ell \) and \( B^+ \to \pi^0 \ell^+ \nu_\ell \) fits are further detailed in Sections [IV].

For the search for the rare \( B^+ \to \ell^+ \nu_\ell \gamma \) decay the yield of four event types are used as free parameters in the fit:

i. \( B^+ \to \ell^+ \nu_\ell \gamma \) signal events.
ii. $B^+ \to \pi^0 \ell^+ \nu_\ell$ normalization events.

iii. $B \to X_u \ell^+ \nu_\ell$ background events.

iv. Other $B$ meson or continuum background events.

The $B^+ \to \pi^0 \ell^+ \nu_\ell$ normalization mode is linked between the $B^+ \to \ell^+ \nu_\ell \gamma$ and $B^+ \to \pi^0 \ell^+ \nu_\ell$ categories and the global likelihood function $L$ is maximized to determine the estimates for the number of signal events. Confidence intervals (CI) are constructed using the profile likelihood method

$$
\lambda(\nu_j) = \log \frac{L(\nu_j, \hat{\nu}, \hat{\theta})}{L(\hat{\nu}, \hat{\nu}, \hat{\theta})},
$$

(10)

where $\hat{\nu}$, $\hat{\nu}$ and $\hat{\theta}$ are the values of the normalization of interest, the remaining normalizations, and nuisance parameters that unconditionally maximize the likelihood function while $\hat{\nu}$ and $\hat{\theta}$ are the values of the other normalizations and nuisance parameters which maximize the likelihood under the condition that the observable of interest is kept fixed at a given value $\nu_j$. In the asymptotic limit, approximate confidence intervals (CI) can be constructed using

$$
1 - CI = \int_{-2 \lambda(\nu_j)}^{\infty} f_\chi^2(x; 1 \text{ dof}) \, dx,
$$

(11)

with $f_\chi^2(x; 1 \text{ dof})$ denoting the $\chi^2$ distribution with one degree of freedom. In case of two parameters of interest a two-dimensional confidence level (CL) can be constructed via Eq. [11] by modifying $\lambda(\nu_j)$ to a likelihood ratio depending on two parameters, $\nu_j$ and $\nu_k$, $\lambda(\nu_j, \nu_k)$, and with $f_\chi^2$ correspondingly then having two degrees of freedom.

In case we observe no significant signal, we set a Bayesian limit by converting the likelihood Eq. [8] $L = L(n|\nu_j)$, into a probability density function $F$ of the parameter of interest $\nu_j$ using a flat prior $\pi(\nu_j)$ such that

$$
F(\nu_j | n) = \int_{0}^{\infty} \frac{L(n|\nu_j)\pi(\nu_j)}{\int_{0}^{\infty} L(n|\nu_j)\pi(\nu_j) \, d\nu_j}
$$

(12)

with $\pi(\nu_j)$ = constant for $\nu_j > 0$ and zero otherwise. In Eq. [12] $n$ denotes the vector of observed event yields in the given bins in all channels.

The fit procedure was validated using ensembles of pseudoexperiments generated with different input branching fractions for $B^+ \to \pi^0 \ell^+ \nu_\ell$ and $B^+ \to \pi^0 \ell^+ \nu_\ell$ decays. No biases or undercoverage of CI are observed.

IV. HADRONIC TAGGING EFFICIENCY CALIBRATION AND $B^+ \to \pi^0 \ell^+ \nu_\ell$ BRANCHING FRACTION

The multivariate classifiers that enter the hadronic tagging algorithm of the FEI are trained on simulated events. Due to imperfections in these simulations, e.g. due to inadequate modeling of hadronic decays or experimentally poorly constrained branching fractions, the tagging algorithm exhibits a different performance on simulated and recorded events. Due to this, the tagging algorithm has to be calibrated on data using well known processes. In this analysis, such a calibration is derived using three semileptonic $B$ channels with different multiplicities,

i. $B^+ \to D^0 (\to K^+ \pi^-) \ell^+ \nu_\ell$

ii. $B^+ \to D^0 (\to K^+ \pi^- \pi^0) \ell^+ \nu_\ell$

iii. $B^+ \to D^0 (\to K^+ \pi^- \pi^- \pi^-) \ell^+ \nu_\ell$

where $\ell = e, \mu$. The tag-side selection is identical with the nominal analysis. The signal side selects electron and muon candidates using the criteria detailed in Section[4]. Charged pion tracks are required to originate near the IP with $|dz| < 2 \text{ cm}$ and $|dz| < 4 \text{ cm}$. Charged kaons and pions are separated using a likelihood ratio, $\mathcal{P}_{K\pi}$, which combines the relevant information from the ACC, TOF and CDC subdetectors. We require $\mathcal{P}_{K\pi} < 0.4$ for pion and $\mathcal{P}_{K\pi} > 0.6$ for kaon candidates, respectively. Neutral pions are reconstructed from the combination of two photons with an invariant mass of $M_{\gamma\gamma} \in (117.8, 152) \text{ MeV}$. $D^0$ candidates are reconstructed combining the kaon candidates with the charged and neutral pion candidates. The reconstructed $D^0$ candidates are required to have $M_\ell \in (1.858, 1.872) \text{ GeV}$, $M_\ell \in (1.849, 1.879) \text{ GeV}$ and $M_\ell \in (1.854, 1.872) \text{ GeV}$ for the three channels i. - iii., respectively. Additional loose cuts are applied on
the beam-constrained mass of the $B_{\text{sig}}$ candidate (reconstructed from the $D^0$ and the lepton), $M_{bc} > 4.5$ GeV, and the cosine of the angle between the true $B$ meson (calculated from beam energy and momentum) and the reconstructed $D\ell$ system, $|\cos(\theta_{BD})| < 3.0$. An unconstrained vertex fit is applied on the $D$ and $B$ candidates and candidates with a $p$-value of the fit of $p_{\chi^2} > 0.01$ are retained.

The tag efficiency calibration factor $\epsilon$ is calculated by extracting the number of signal decays on data and comparing to the expected number of events from the MC simulation. The signal yield is determined using a binning maximum likelihood fit (cf. Section III) of the $M^2_{\text{miss}}$ distribution, reconstructed as

$$M^2_{\text{miss}} = m^2_{\text{miss}} = (p_{B_{\text{sig}}} - p_{\ell} - p_D)^2.$$

The obtained calibration factors of the three channels are shown in Fig. 1 and the global calibration factor is found to be $\epsilon = 0.825 \pm 0.014$ (stat.) $\pm 0.049$ (syst.).

To validate the found calibration factor, we measure the branching fraction of the $B^+ \rightarrow \pi^0 \ell^+ \nu_\ell$ decay and compare it to the current world average. We obtain $B(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell) = (7.8 \pm 0.6$ (stat.) $\times 10^{-5}$, which is in agreement with the average $B_{\text{PEG}}(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell) = (7.80 \pm 0.27) \times 10^{-5}$ of Ref. [20].

V. SYSTEMATIC UNCERTAINTIES

There are several systematic uncertainties that affect the measured yields and partial branching fractions: Table I summarizes the most important sources of uncertainty for the $B^+ \rightarrow \ell^+ \nu_\ell \gamma$ and $B^+ \rightarrow \pi^0 \ell^+ \nu_\ell$ branching fraction measurements.

The effect of all systematic uncertainties is directly incorporated into the likelihood Eq. 7 via the replacement of

$$\nu_j f_{ij} \rightarrow \nu_j f_{ij} \times \prod_k (1 + \theta_k \epsilon_{ijk}),$$

and

$$\nu_j f_{ij} \rightarrow \nu_j f_{ij} + \sum_k \theta_k \epsilon_{ijk},$$

for multiplicative and additive uncertainties, respectively. Nuisance parameters $\theta_k$ are constrained using standard normal distributions $G(\theta_k)$ in Eq. 8 for relative and absolute uncertainties $\epsilon_{ijk}$ of a source $k$ for a component $j$ and a given bin $i$. Systematic and statistical uncertainties are separated from each other using scans of the likelihood contour in which the systematic nuisance parameters are kept fixed at their best fit value.

The largest multiplicative systematic uncertainty on both branching fractions stems from the uncertainty on the tagging calibration (see the previous section). It is evaluated by shifting the central value of the combined correction factor according to its statistical and systematic uncertainty. This results in a relative uncertainty of $6.2\%$. The second largest uncertainty for $B^+ \rightarrow \pi^0 \ell^+ \nu_\ell$ is given by the statistical uncertainty on the signal reconstruction efficiency. Its uncertainty is evaluated using binomial uncertainties, following the prescription of Ref. [33]. Another large multiplicative uncertainty stems from the $L_{\text{LID}}$ efficiency, which is corrected in the simulation using data-driven methods. The statistical and systematic uncertainty on these correction factors are propagated and result in an uncertainty of $1.81\%$ and $1.97\%$ for $\Delta B(B^+ \rightarrow \ell^+ \nu_\ell \gamma)$ and $B(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell)$, respectively. The remaining two multiplicative uncertainties are from the number of $B\bar{B}$ pairs, used to convert the measured yield into (partial) branching fractions, and the uncertainty on reconstruction efficiency differences between the simulation and recorded collisions of charged tracks. The tracking efficiency differences are studied using $D^+ \rightarrow D^0 \pi$ decays with $D^0 \rightarrow \pi\pi K_S^0$ and $K_S^0 \rightarrow \pi^+ \pi^-$.

The uncertainty on $N_p_{\mu \pi}$ results in a relative error of $1.37\%$ and for the tracking efficiency an uncertainty of $0.35\%$ for the single signal side track is found.

| Source | $\Delta B(B^+ \rightarrow \ell^+ \nu_\ell \gamma)$ in $10^{-5}$ | $\Delta B(B^+ \rightarrow \ell^+ \nu_\ell \gamma)$ in $10^{-6}$ |
|--------|-----------------|-----------------|
| Calibration | $\pm 0.49$ | $\pm 0.09$ |
| Reconstruction efficiency | $\pm 0.20$ | $\pm 0.01$ |
| $L_{\text{LID}}$ efficiency | $\pm 0.16$ | $\pm 0.02$ |
| $N_p_{\mu \pi}$ | $\pm 0.11$ | $\pm 0.02$ |
| Tracking efficiency | $\pm 0.03$ | $\pm 0.0$ |
| Peaking background BDT | $\pm 0.02$ | $\pm 0.24$ |
| PDF templates | $\pm 0.08$ | $\pm 0.18$ |
| BCL model | $\pm 0.25$ | $\pm 0.01$ |
| Reconstructed tag channel | $\pm 0.01$ | $\pm 0.14$ |
| $B \rightarrow X, \ell^+ \nu_\ell$ | $\pm 0.02$ | $\pm 0.07$ |
| Signal model | $\pm 0.00$ | $\pm 0.03$ |
| Combined | $\pm 0.62$ | $\pm 0.36$ |

The largest additive systematic uncertainty for the $B^+ \rightarrow \ell^+ \nu_\ell \gamma$ partial branching fraction measurement stems from the systematic uncertainty assigned to the multivariate method that suppresses peaking background contributions. This uncertainty is evaluated by reweighting the MC samples to the distribution of the input variables used for the classification on data. The distribution which gives the largest deviation from the nominal result is used to estimate the uncertainty. The second largest additive uncertainty for the $B^+ \rightarrow \ell^+ \nu_\ell \gamma$ partial branching fraction measurement is due to limited MC statistics. The uncertainty is evaluated for each MC sample individually by producing a large ensemble of templates, where
(a) $B^+ \rightarrow e^+\nu_e\gamma$ final state

(b) $B^+ \rightarrow \mu^+\nu_{\mu}\gamma$ final state

(c) $B^+ \rightarrow \pi^0 e^+\nu_e$ final state

(d) $B^+ \rightarrow \pi^0 \mu^+\nu_{\mu}$ final state

**FIG. 2:** The post-fit $M^2_{\text{miss}}$ distributions for the simultaneous fit to the four categories are shown (cf. Section III). The individual fit components are shown as colored histograms, and the filled gray histogram shows their sum.

The numbers of entries are varied using a Poisson distribution. The templates of the ensemble are used to repeat the fit to estimate the total uncertainty. The largest additive systematic uncertainty for the $B^+ \rightarrow \pi^0 e^+\nu_e$ branching fraction is given by the uncertainty on the BCL form factors and is evaluated by variations using the covariance matrix from the global fit of Ref. [22].

The remaining additive uncertainties on both channels are evaluated as follows: The fraction of the individual channels in which the $B_{\text{tag}}$ is reconstructed differs between MC and data. To estimate the impact of this mismatch, the MC samples are corrected to the fraction in data of the reconstructed tag channels and the difference is taken as an estimation for the systematic uncertainty. In the fit, the individual branching fractions of charmless semileptonic background decay modes are kept fixed and modeled as a single floating background template. To estimate uncertainties due to slight shape differences in the $M^2_{\text{miss}}$ from these templates, we vary the decay branching fractions of $B^+ \rightarrow \omega \ell^+\nu_\ell$, $B^+ \rightarrow \rho^0 \ell^+\nu_\ell$, $B^0 \rightarrow \rho^- \ell^+\nu_\ell$, $B^+ \rightarrow \eta \ell^+\nu_\ell$, $B^+ \rightarrow \eta' \ell^+\nu_\ell$, and $B^0 \rightarrow \pi^- \ell^+\nu_\ell$ individually within their uncertainties [26]. The uncertainty on the $B^+ \rightarrow \ell^+\nu_\ell\gamma$ signal model is estimated by correcting the simulated events from the prediction of Ref. [20] to the state-of-the-art prediction of Ref. [5] and repeating the fit.

**VI. RESULTS**

Figure 2 shows the $M^2_{\text{miss}}$ distribution of the selected data events in the four categories of $B^+ \rightarrow e^+\nu_e\gamma$, $B^+ \rightarrow \mu^+\nu_{\mu}\gamma$, $B^+ \rightarrow \pi^0 e^+\nu_e$, and $B^+ \rightarrow \pi^0 \mu^+\nu_{\mu}$. The selected events are used to maximize the likelihood function Eq. 7 numerically, determining the four ($B^+ \rightarrow \ell^+\nu_\ell\gamma$) and
three \(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell\) event types detailed in Section \[III\].

The fitted \(B^+ \rightarrow \ell^+ \nu_\ell \gamma\) signal, \(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell\) normalization and other background contributions are shown as colored histograms and the summed signal plus background template is shown as a filled gray histogram. The observed partial branching fraction of \(B^+ \rightarrow \ell^+ \nu_\ell \gamma\) with \(E_\gamma > 1\) GeV is

\[
\Delta B(B^+ \rightarrow \ell^+ \nu_\ell \gamma) = (1.4 \pm 1.0 \pm 0.4) \times 10^{-6},
\]

(16)

where the first error is statistical and the second error contains all systematic uncertainties discussed in Section \[V\]. The significance over the background-only hypothesis for the \(B^+ \rightarrow \ell^+ \nu_\ell \gamma\) signal, as calculated using the likelihood ratio, is 1.4 standard deviations. The \(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell\) branching fraction is found to be

\[
B(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell) = (7.9 \pm 0.6 \pm 0.6) \times 10^{-5},
\]

(17)

and has better statistical precision than the measurement of Ref. \[34\]. A summary of all fit results, including fits of the individual electron and muon samples, is presented in Table \[III\] Figure 3a shows the two-dimensional likelihood ratio contours of \(-2\lambda\) for the combined measurement of \(B^+ \rightarrow \ell^+ \nu_\ell \gamma\) and \(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell\). The ellipses correspond to the given confidence level, including systematic uncertainties. Figure 3b shows the one-dimensional likelihood ratio scan and its conversion into a Bayesian PDF, which was obtained using a flat prior for the \(B^+ \rightarrow \ell^+ \nu_\ell \gamma\) measurement, see Section \[III\] for details.

![Fig. 3](image.png)

**FIG. 3:** Plot (a) shows the two-dimensional likelihood ratio contour \(-2\lambda\) for the combined measurement of \(B^+ \rightarrow \ell^+ \nu_\ell \gamma\) and \(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell\). The ellipses correspond to the given confidence level, including systematic uncertainties. Plot (b) shows the one-dimensional likelihood ratio contour and its conversion into a Bayesian PDF \(F(\nu_\ell \mid n)\) using a flat prior for the \(B^+ \rightarrow \ell^+ \nu_\ell \gamma\) measurement.

Due to the low significance of the measured \(B^+ \rightarrow \ell^+ \nu_\ell \gamma\) signal, we convert the likelihood into a Bayesian probability density function (PDF), with the procedure detailed in Section \[III\]. Figure 3b shows the one-dimensional likelihood ratio scan and the obtained Bayesian PDF, which was obtained using a flat prior in the partial branching fraction. The resulting limit for \(B^+ \rightarrow \ell^+ \nu_\ell \gamma\) at 90% CL is

\[
\Delta B(B^+ \rightarrow \ell^+ \nu_\ell \gamma) < 3.0 \times 10^{-6} \text{ at 90}\% \text{ CL}.\ (18)
\]

This provides a significantly more stringent limit than previous searches, and a summary of previous limits and individual limits for the electron and muon signal channel can be found in Table \[III\].

Using the \(B^+ \rightarrow \ell^+ \nu_\ell \gamma\) and \(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell\) branching fractions, the first inverse moment \(\lambda_B\) of the leading-twist \(B\) meson light-cone distribution amplitude \(\phi_+\) can be determined. Instead of directly using the measured \(B^+ \rightarrow \ell^+ \nu_\ell \gamma\) partial branching fraction, we use the theoretically well understood \(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell\) decay rate to derive a measurement of \(\lambda_B\) which is independent of \(V_{ub}\).

The value of \(\lambda_B\) is related to this ratio as

\[
R_\pi = \frac{\Delta B(B^+ \rightarrow \ell^+ \nu_\ell \gamma)}{B(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell)} = \frac{\Delta \Gamma(\lambda_B)}{\Gamma(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell)},
\]

(19)

with \(\Delta \Gamma(\lambda_B)\) denoting the partial decay rate as a function of \(\lambda_B\) with \(E_\gamma > 1\) GeV, and \(\Gamma(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell)\) denoting the total decay rate of \(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell\). Using the central values and the full experimental covariance we measure

\[
R_{\pi,\text{meas}} = (1.7 \pm 1.4) \times 10^{-2}.
\]

(20)

For the prediction of the \(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell\) decay rate, we use the global fit \[22\] of

\[1\] The statistical overlap with the previous measurement is unknown. Since the current result is not measured in bins of \(q^2\), the previous result should still be used for the determination of \(|V_{ub}|\) and world averages of the branching fraction.
TABLE II: Measured central values and the corresponding significance for the different channels.

| ℓ       | $|B(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell)(10^{-5})|$ | $\Delta B(B^+ \rightarrow \ell^+ \nu_\ell \gamma)(10^{-6})$ |
|---------|---------------------------------|---------------------------------|
| $\tau$  | $8.3^{+0.9}_{-0.8} \pm 0.9$     | $1.7^{+1.6}_{-1.4} \pm 0.7$    |
| $\mu$   | $7.2^{+0.8}_{-0.6} \pm 0.6$     | $1.6^{+1.4}_{-1.0} \pm 0.4$    |
| $e, \mu$| $7.5^{+0.9}_{-0.6} \pm 0.6$     | $1.4^{+1.0}_{-1.0} \pm 0.4$    |

TABLE III: Comparison to previous results of the partial branching fraction measurement. All limits correspond to the 90% CL.

| ℓ       | $\Delta B(B^+ \rightarrow \ell^+ \nu_\ell \gamma)$ limit (10^{-6}) | BaBar [23] | Belle [11] | This work |
|---------|---------------------------------------------------------------|------------|------------|-----------|
| $\tau$  | < 6.1                                                         | < 4.3      | < 4.3      |           |
| $\mu$   | < 3.4                                                         | < 3.4      | < 3.4      |           |
| $e, \mu$| < 14                                                          | < 3.5      | < 3.0      |           |

$\Gamma(B^+ \rightarrow \pi^0 \ell^+ \nu_\ell) = |V_{ub}|^2 \times (2.4 \pm 0.2) \times 10^{-12}$ GeV.

For the partial $B^+ \rightarrow \ell^+ \nu_\ell \gamma$ decay rate the predictions and uncertainties of Ref. [7] extrapolated to $E_\gamma > 1$ GeV are used. In Ref. [7], three different models are used to evaluate the dependence of the partial decay rate on the functional form of the light-cone distribution amplitude. Figure 4 shows the predicted and measured $R_\pi$ ratio as a function of $\lambda_B$. We solve Eq. 19 numerically and in Table IV the determined value of $\lambda_B$ for each of the three models are given, including the corresponding theoretical uncertainties of Ref. [7]. We use the shift in the central value between all three models to also quote a value of $\lambda_B$, whose uncertainty should incorporate the overall model dependence. For this we find

$$\lambda_B = 0.36^{+0.25+0.03+0.03}_{-0.08-0.03-0.03} \text{GeV} = 0.36^{+0.25}_{-0.09} \text{GeV},$$ (21)

where the first uncertainty is experimental, the second from the theoretical uncertainty on the $B^+ \rightarrow \ell^+ \nu_\ell \gamma$ prediction of Ref. [7] and the $B^+ \rightarrow \pi^0 \ell^+ \nu_\ell$ uncertainty from Ref. [22], and the third uncertainty is due to the light-cone distribution amplitude model dependence. We further obtain a one-sided limit of

$$\lambda_B > 0.24 \text{ GeV}$$ (22)

at 90% CL.

Note, that these estimates might suffer from additional uncertainties from the extrapolation to $E_\gamma > 1$ GeV. Further details can be found in Ref. [7].

VII. SUMMARY

In this manuscript, an improved search for the radiative leptonic decay $B^+ \rightarrow \ell^+ \nu_\ell \gamma$ on the full Belle data set recorded at the $Y(4S)$ resonance is presented. The results improve the previous analysis by our collaboration and increase the signal efficiency by a factor of three. In addition, the description of the important $B^+ \rightarrow \pi^0 \ell^+ \nu_\ell$ background was improved, by analyzing simultaneously $B^+ \rightarrow \pi^0 \ell^+ \nu_\ell$ signal events and using the global fit result of Ref. [22] to describe its form factors. The large improvement in sensitivity stems from employing a newly developed tagging algorithm developed for the Belle II experiment, the Full Event Interpretation [12]. Although this drastically improves the sensitivity, no significant signal of $B^+ \rightarrow \ell^+ \nu_\ell \gamma$ decays is observed. As it is not possible to determine the statistical overlap with the previous Belle result, this work supersedes Ref. [11].

The determined partial branching fraction for $B^+ \rightarrow \ell^+ \nu_\ell \gamma$ decays with photon energies $E_\gamma > 1$ GeV in the $B_{\text{sig}}$ rest frame is found to be

$$\Delta B(B^+ \rightarrow \ell^+ \nu_\ell \gamma) = (1.4 \pm 1.0 \pm 0.4) \times 10^{-6},$$ (23)

with a significance of 1.4 standard deviations over the background-only hypothesis. Using the likelihood contour and a flat prior, we determine a Bayesian upper
limit of

$$\Delta B(B^+ \to \ell^+ \nu \gamma) < 3.0 \times 10^{-6} ,$$  \hspace{1cm} (24)

at 90% confidence level.

In addition, we report an improved determination of the first inverse momentum $\lambda_B$ of the the light-cone distribution amplitude of the $B$ meson. It is done using a $V_{ub}$ independent way, by normalizing the measured partial branching fraction to the branching fraction of $B^+ \to \pi^0 \ell^+ \nu_\ell$. This reduces the experimental uncertainties, and the theoretical prediction of the total decay width of $B^+ \to \pi^0 \ell^+ \nu_\ell$ is well understood.

Using the result of Ref. \cite{7}, its associated uncertainties and an additional uncertainty to assess the model dependence, we obtain

$$\lambda_B = 0.36^{+0.25}_{-0.09} \text{ GeV} \hspace{1cm} (25)$$
or $\lambda_B > 0.24 \text{ GeV}$ at 90% CL.

The search of $B^+ \to \ell^+ \nu \gamma$ is limited by the available data set and its sensitivity will be greatly enhanced by the upcoming Belle II experiment. The anticipated data set of 50 ab$^{-1}$ will greatly reduce the experimental uncertainties on $\lambda_B$. Together with recent developments in lattice QCD calculations reducing the theoretical uncertainties on $\lambda_B$, the feasibility of $B^+ \to \ell^+ \nu \gamma$ increases as an alternative channel to measure $|V_{ub}|$ to provide a consistency check of the SM.

\section*{VIII. ACKNOWLEDGMENTS}

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, and the Pacific Northwest National Laboratory (PNNL) Environmental Molecular Sciences Laboratory (EMSL) computing group for strong computing support; and the National Institute of Informatics, and Science Information NETwork 5 (SINET5) for valuable network edge support from MEXT, JSPS and Nagoya’s TL-PRC (Japan); ARC (Australia); FWF (Austria); NSFC and CCEPP (China); MSMT (Czechia); CZF, DFG, EXC153, and VS (Germany); DST (India); INFN (Italy); MOE, MSIP, NRF, R$^2$ST, FLRFS project and GSDC of KISTI and KREONET/GLORIAD; the Polish Ministry of Science and Higher Education and the National Science Center; the Grant of the Russian Federation Government, Agreement No. 14.W03.31.0026; the Slovenian Research Agency; Ikerbasque, Basque Foundation for Science, Basque Government (No. IT96-16) 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.

We thank the KEKB group for excellent operation of the accelerator; the KEK cryogenics group for efficient solenoid operations; and the KEK computer group, the NII, and PNNL/EMSL for valuable computing and SINET5 network support. We acknowledge support from MEXT, JSPS and Nagoya’s TL-PRC (Japan); ARC (Australia); FWF (Austria); NSFC and CCEPP (China); MSMT (Czechia); CZF, DFG, EXC153, and VS (Germany); DST (India); INFN (Italy); MOE, MSIP, NRF, R$^2$ST, FLRFS project and GSDC of KISTI and KREONET/GLORIAD (Korea); MNiSW and NCN (Poland); MES (Russia); ARRS (Slovenia); Ikerbasque and MINECO (Spain); SNSF (Switzerland); MOE and MOST (Taiwan); and DOE and NSF (USA).

\begin{thebibliography}{99}
\bibitem{1} Throughout this paper, the inclusion of the charge-conjugate decay mode is implied unless otherwise stated.
\bibitem{2} Natural units are used throughout this paper with $c = 1$.
\bibitem{3} M. Kobayashi and T. Maskawa, Prog. Theor. Phys. 49, 652 (1973), doi:10.1143/PTP.49.652
\bibitem{4} N. Cabibbo, Phys. Rev. Lett. 10, 531 (1963), doi:10.1103/PhysRevLett.10.531
\bibitem{5} M. Beneke and J. Rohrwild, Eur. Phys. J. C
\end{thebibliography}
13

[1] Y. M. Wang and Y. L. Shen, JHEP **1805**, 184 (2018) doi:10.1007/JHEP05(2018)184 [arXiv:1803.06667 [hep-ph]].

[2] M. Beneke, V. M. Braun, Y. Ji and Y. B. Wei, JHEP **1807**, 154 (2018) doi:10.1007/JHEP07(2018)154 [arXiv:1804.04962 [hep-ph]].

[3] M. Beneke, G. Buchalla, M. Neubert and C. T. Sachrajda, Nucl. Phys. B **591**, 313 (2000) doi:10.1016/S0550-3213(00)00559-9 [hep-ph/0006124].

[4] M. Beneke and T. Feldmann, Nucl. Phys. B **592**, 3 (2001) doi:10.1016/S0550-3213(00)00585-X [hep-ph/0008255].

[5] A. G. Grozin and M. Neubert, Phys. Rev. D **55**, 272 (1997) doi:10.1103/PhysRevD.55.272 [hep-ph/9607366].

[6] A. Heller et al. (Belle Collaboration), Phys. Rev. D **91**, 112009 (2015) doi:10.1103/PhysRevD.91.112009 [arXiv:1504.05831 [hep-ex]].

[7] T. Keck et al., arXiv:1807.08680 [hep-ex].

[8] 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 (2013) and references therein.

[9] D. J. Lange, Nucl. Instrum. Meth. A **462**, 152 (2001). doi:10.1016/S0168-9002(01)00089-4

[10] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978). The modified Fox-Wolfram moments (SFW) used by Belle are described in K. Abe et al. (Belle Collaboration), Phys. Rev. Lett. **87**, 101801 (2001) and K. Abe et al. (Belle Collaboration), Phys. Lett. B **511**, 151 (2001).

[11] G. Punzi, eConf C **030908**, MODT002 (2003) [physics/0308063].

[12] M. Paterno, doi:10.2172/15017262

[13] A. Sibidanov et al. (Belle Collaboration), Phys. Rev. D **88**, 032005 (2013) doi:10.1103/PhysRevD.88.032005 [arXiv:1306.2781 [hep-ex]].

[14] B. Aubert et al. (BaBar Collaboration), Phys. Rev. D **80**, 111105 (2009) doi:10.1103/PhysRevD.80.111105 [arXiv:0907.1681 [hep-ex]].