University Library

MINERVA ACCESS

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Cao, L;Sutcliffe, W;Van Tonder, R;Bernlochner, FU;Adachi, I;Aihara, H;Al Said, S;Asner, DM;Atmacan, H;Aushev, T;Ayad, R;Babu, V;Bauer, M;Behera, P;Belous, K;Bennett, J;Bessner, M;Bhardwaj, V;Bilka, T;Biswal, J;Bonvicini, G;Bozek, A;Bračko, M;Browder, TE;Campajola, M;Červenkov, D;Chang, MC;Chang, P;Chekelian, V;Chen, A;Cheon, BG;Chilikin, K;Cho, HE;Cho, K;Cho, SJ;Choi, SK;Choi, Y;Choudhury, S;Cinabro, D;Cunliffe, S;Das, S;Dash, N;De Nardo, G;Di Capua, F;Dingfelder, J;DoleŽal, Z;Dong, TV;Dubey, S;Eidelman, S;Epifanov, D;Ferber, T;Ferlewicz, D;Frey, A;Fulsom, BG;Garg, R;Gaur, V;Garmash, A;Giri, A;Goldenzweig, P;Guan, Y;Hadjivasiliou, C;Hara, T;Hartbrich, O;Hayasaka, K;Hayashii, H;Hedges, MT;Hernandez Villanueva, M;Hou, WS;Hsu, CL;Iijima, T;Inami, K;Ishikawa, A;Itoh, R;Iwasaki, M;Iwasaki, Y;Jacobs, WW;Jang, EJ;Jeon, HB;Jia, S;Jin, Y;Joo, CW;Joo, KK;Kang, KH;Karyan, G;Kawasaki, T;Kichimi, H;Kiesling, C;Kim, BH;Kim, CH;Kim, DY;Kim, HJ;Kim, KH;Kim, SH;Kim, YK;Kinoshita, K;Kodyš, P;Konno, T;Korobov, A;Korpar, S;Kotchetkov, D

Title:
Measurements of partial branching fractions of inclusive B →xuℓ+vℓ decays with hadronic tagging

Date:
2021-07-01

Citation:
Cao, L., Sutcliffe, W., Van Tonder, R., Bernlochner, F. U., Adachi, I., Aihara, H., Al Said, S., Asner, D. M., Atmacan, H., Aushev, T., Ayad, R., Babu, V., Bauer, M., Behera, P., Belous, K., Bennett, J., Bessner, M., Bhardwaj, V., Bilka, T. ,... Kotchetkov, D. (2021). Measurements of partial branching fractions of inclusive B →xuℓ+vℓ decays with hadronic tagging. Physical Review D, 104 (1), https://doi.org/10.1103/PhysRevD.104.012008.

Persistent Link:
https://hdl.handle.net/11343/288800

License:
CC-BY

Measurements of partial branching fractions of inclusive $B \to X_u \ell^+ \nu_\ell$ decays with hadronic tagging

L. Cao,[1,*] W. Sutcliffe,[1] R. Van Tonder,[1] F. U. Bernlochner,[1,†] I. Adachi,[18,14] H. Aihara,[88] S. Al Said,[81,36] D. M. Asner,[3] H. Atmacan,[7] T. Aushev,[20] R. Ayad,[81] V. Babu,[8] M. Bauer,[34] P. Behera,[25] K. Belous,[29] J. Bennett,[51] M. Bessner,[17] V. Bhardwaj,[23] T. Bilka,[5] J. Biswal,[33] G. Bonvicini,[92] A. Bozek,[61] M. Bračko,[48,33] T. E. Browder,[17] M. Campajola,[30,56] D. Červenkov,[5] M.-C. Chang,[10] P. Chang,[60] V. Chekelian,[49] A. Chen,[58] B. G. Cheon,[16] K. Chilikin,[43] H. E. Cho,[16] K. Cho,[38] S.-J. Cho,[94] S.-K. Choi,[15] Y. Choi,[79] S. Choudhury,[24] D. Cinabro,[92] S. Cunliffe,[8] S. Das,[47] N. Dash,[25] G. De Nardo,[30,56] F. Di Capua,[30,56] J. Dingfelder,[1] Z. Doležal,[5] T. V. Dong,[11] S. Dubey,[17] S. Eidelman,[4,65,43] D. Epifanov,[4,65] T. Ferber,[8] D. Ferlewicz,[50] A. Frey,[13] B. G. Fulsom,[67] R. Garg,[68] V. Gaur,[91] A. Garmash,[4,65] A. Giri,[24] P. Goldenzweig,[34] Y. Guan,[7] C. Hadjivasiliou,[67] T. Hara,[18,14] O. Hartbrich,[17] K. Hayasaka,[63] H. Hayashii,[57] M. T. Hedges,[17] M. Hernandez Villanueva,[51] W.-S. Hou,[60] C.-L. Hsu,[80] T. Iijima,[55,54] K. Inami,[54] A. Ishikawa,[18,14] R. Itoh,[18,14] M. Iwasaki,[66] Y. Iwasaki,[18] W. W. Jacobs,[26] E.-J. Jang,[15] H. B. Jeon,[41] S. Jia,[11] Y. Jin,[88] C. W. Joo,[35] K. K. Joo,[6] K. H. Kang,[41] G. Karyan,[8] T. Kawasaki,[37] H. Kichimi,[18] C. Kiesling,[49] B. H. Kim,[75] C. H. Kim,[16] D. Y. Kim,[78] H. J. Kim,[41] K.-H. Kim,[94] S. H. Kim,[75] Y.-K. Kim,[94] K. Kinoshita,[7] P. Kodyš,[5] T. Konno,[37] A. Korobov,[4,65] S. Korpar,[48,33] D. Kotchetkov,[17] E. Kovalenko,[4,65] P. Križan,[45,33] R. Kroeger,[51] P. Krokovny,[4,65] T. Kuhr,[46] M. Kumar,[47] R. Kumar,[71] K. Kumara,[92] A. Kuzmin,[4,65] Y.-J. Kwon,[94] K. Lalwani,[47] J. S. Lange,[12] I. S. Lee,[16] S. C. Lee,[41] P. Lewis,[1] C. H. Li,[44] J. Li,[41] L. K. Li,[7] Y. B. Li,[69] L. Li Gioi,[49] J. Libby,[25] K. Lieret,[46] Z. Liptak,[17,‡] D. Liventsev,[92,18] J. MacNaughton,[52] C. MacQueen,[50] M. Masuda,[87,72] T. Matsuda,[52] D. Matvienko,[4,65,43] M. Merola,[30,56] F. Metzner,[34] K. Miyabayashi,[57] R. Mizuk,[43,20] G. B. Mohanty,[82] T. J. Moon,[75] T. Mori,[54] M. Mrvar,[28] R. Mussa,[31] M. Nakao,[18,14] Z. Natkaniec,[61] A. Natochii,[17] L. Nayak,[24] M. Nayak,[84] M. Niiyama,[40] N. K. Nisar,[3] S. Nishida,[18,14] K. Nishimura,[17] S. Ogawa,[85] H. Ono,[62,63] Y. Onuki,[88] P. Oskin,[43] P. Pakhlov,[43,53] G. Pakhlova,[20,43] T. Pang,[70] S. Pardi,[30] C. W. Park,[79] H. Park,[41] S.-H. Park,[94] S. Patra,[23] S. Paul,[83,49] T. K. Pedlar,[95] R. Pestotnik,[33] L. E. Piilonen,[91] T. Podobnik,[45,33] V. Popov,[20] E. Prencipe,[21] M. T. Prim,[1] M. Ritter,[46] M. Röhrken,[8] A. Rostomyan,[8] N. Rout,[25] M. Rozanska,[61] G. Russo,[56] D. Sahoo,[82] Y. Sakai,[18,14] S. Sandilya,[24] A. Sangal,[7] L. Santelj,[45,33] T. Sanuki,[86] V. Savinov,[70] G. Schnell,[2,22] J. Schueler,[17] C. Schwanda,[28] A. J. Schwartz,[7] Y. Seino,[63] K. Senyo,[93] M. E. Sevior,[50] M. Shapkin,[29] C. Sharma,[47] C. P. Shen,[11] J.-G. Shiu,[60] F. Simon,[49] A. Sokolov,[29] E. Solovieva,[43] S. Stanič,[64] M. Starič,[33] Z. S. Stottler,[91] J. F. Strube,[67] T. Sumiyoshi,[90] M. Takizawa,[76,19,73] U. Tamponi,[31] K. Tanida,[32] F. Tenchini,[8] K. Trabelsi,[42] M. Uchida,[89] T. Uglov,[43,20] Y. Unno,[16] S. Uno,[18,14] P. Urquijo,[50] Y. Usov,[4,65] S. E. Vahsen,[17] G. Varner,[17] K. E. Varvell,[80] A. Vinokurova,[4,65] V. Vorobyev,[4,65,43] A. Vossen,[9] E. Waheed,[18] C. H. Wang,[59] E. Wang,[70] M.-Z. Wang,[60] P. Wang,[27] M. Watanabe,[63] S. Watanuki,[42] S. Wehle,[8] J. Wiechczynski,[61] E. Won,[39] X. Xu,[77] B. D. Yabsley,[80] W. Yan,[74] S. B. Yang,[39] H. Ye,[8] J. H. Yin,[39] C. Z. Yuan,[27] Y. Yusa,[63] Z. P. Zhang,[74] V. Zhilich,[4,65] V. Zhukova,[43] and V. Zhulanov[4,65]

(Belle Collaboration)

[1]*University of Bonn, 53115 Bonn*
[2]*Department of Physics, University of the Basque Country UPV/EHU, 48080 Bilbao*
[3]*Brookhaven National Laboratory, Upton, New York 11973*
[4]*Budker Institute of Nuclear Physics SB RAS, Novosibirsk 630090*
[5]*Faculty of Mathematics and Physics, Charles University, 121 16 Prague*
[6]*Chonnam National University, Gwangju 61186*
[7]*University of Cincinnati, Cincinnati, Ohio 45221*
[8]*Deutsches Elektronen–Synchrotron, 22607 Hamburg*
[9]*Duke University, Durham, North Carolina 27708*
[10]*Department of Physics, Fu Jen Catholic University, Taipei 24205*
[11]*Key Laboratory of Nuclear Physics and Ion-beam Application (MOE) and Institute of Modern Physics, Fudan University, Shanghai 200443*
[12]*Justus-Liebig-Universität Gießen, 35392 Gießen*
[13]*II. Physikalisches Institut, Georg-August-Universität Göttingen, 37073 Göttingen*
[14]*SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193*
[15]*Gyeongsang National University, Jinju 52828*
[16]*Department of Physics and Institute of Natural Sciences, Hanyang University, Seoul 04763*
[17]*University of Hawaii, Honolulu, Hawaii 96822*
[18]*High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801*

[19]*J-PARC Branch, KEK Theory Center, High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801*
[20]*Higher School of Economics (HSE), Moscow 101000*
[21]*Forschungszentrum Jülich, 52425 Jülich*
[22]*IKERBASQUE, Basque Foundation for Science, 48013 Bilbao*
[23]*Indian Institute of Science Education and Research Mohali, SAS Nagar, 140306*
[24]*Indian Institute of Technology Hyderabad, Telangana 502285*
[25]*Indian Institute of Technology Madras, Chennai 600036*
[26]*Indiana University, Bloomington, Indiana 47408*
[27]*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049*
[28]*Institute of High Energy Physics, Vienna 1050*
[29]*Institute for High Energy Physics, Protvino 142281*
[30]*INFN—Sezione di Napoli, 80126 Napoli*
[31]*INFN—Sezione di Torino, 10125 Torino*
[32]*Advanced Science Research Center, Japan Atomic Energy Agency, Naka 319-1195*
[33]*J. Stefan Institute, 1000 Ljubljana*
[34]*Institut für Experimentelle Teilchenphysik, Karlsruher Institut für Technologie, 76131 Karlsruhe*
[35]*Kavli Institute for the Physics and Mathematics of the Universe (WPI), University of Tokyo, Kashiwa 277-8583*
[36]*Department of Physics, Faculty of Science, King Abdulaziz University, Jeddah 21589*
[37]*Kitasato University, Sagamihara 252-0373*
[38]*Korea Institute of Science and Technology Information, Daejeon 34141*
[39]*Korea University, Seoul 02841*
[40]*Kyoto Sangyo University, Kyoto 603-8555*
[41]*Kyungpook National University, Daegu 41566*
[42]*Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay*
[43]*P. N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow 119991*
[44]*Liaoning Normal University, Dalian 116029*
[45]*Faculty of Mathematics and Physics, University of Ljubljana, 1000 Ljubljana*
[46]*Ludwig Maximilians University, 80539 Munich*
[47]*Malaviya National Institute of Technology Jaipur, Jaipur 302017*
[48]*University of Maribor, 2000 Maribor*
[49]*Max-Planck-Institut für Physik, 80805 München*
[50]*School of Physics, University of Melbourne, Victoria 3010*
[51]*University of Mississippi, University, Mississippi 38677*
[52]*University of Miyazaki, Miyazaki 889-2192*
[53]*Moscow Physical Engineering Institute, Moscow 115409*
[54]*Graduate School of Science, Nagoya University, Nagoya 464-8602*
[55]*Kobayashi-Maskawa Institute, Nagoya University, Nagoya 464-8602*
[56]*Università di Napoli Federico II, 80126 Napoli*
[57]*Nara Women's University, Nara 630-8506*
[58]*National Central University, Chung-li 32054*
[59]*National United University, Miao Li 36003*
[60]*Department of Physics, National Taiwan University, Taipei 10617*
[61]*H. Niewodniczanski Institute of Nuclear Physics, Krakow 31-342*
[62]*Nippon Dental University, Niigata 951-8580*
[63]*Niigata University, Niigata 950-2181*
[64]*University of Nova Gorica, 5000 Nova Gorica*
[65]*Novosibirsk State University, Novosibirsk 630090*
[66]*Osaka City University, Osaka 558-8585*
[67]*Pacific Northwest National Laboratory, Richland, Washington 99352*
[68]*Panjab University, Chandigarh 160014*
[69]*Peking University, Beijing 100871*
[70]*University of Pittsburgh, Pittsburgh, Pennsylvania 15260*
[71]*Punjab Agricultural University, Ludhiana 141004*
[72]*Research Center for Nuclear Physics, Osaka University, Osaka 567-0047*
[73]*Meson Science Laboratory, Cluster for Pioneering Research, RIKEN, Saitama 351-0198*
[74]*Department of Modern Physics and State Key Laboratory of Particle Detection and Electronics, University of Science and Technology of China, Hefei 230026*
[75]*Seoul National University, Seoul 08826*

[76]Showa Pharmaceutical University, Tokyo 194-8543
[77]Soochow University, Suzhou 215006
[78]Soongsil University, Seoul 06978
[79]Sungkyunkwan University, Suwon 16419
[80]School of Physics, University of Sydney, New South Wales 2006
[81]Department of Physics, Faculty of Science, University of Tabuk, Tabuk 71451
[82]Tata Institute of Fundamental Research, Mumbai 400005
[83]Department of Physics, Technische Universität München, 85748 Garching
[84]School of Physics and Astronomy, Tel Aviv University, Tel Aviv 69978
[85]Toho University, Funabashi 274-8510
[86]Department of Physics, Tohoku University, Sendai 980-8578
[87]Earthquake Research Institute, University of Tokyo, Tokyo 113-0032
[88]Department of Physics, University of Tokyo, Tokyo 113-0033
[89]Tokyo Institute of Technology, Tokyo 152-8550
[90]Tokyo Metropolitan University, Tokyo 192-0397
[91]Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
[92]Wayne State University, Detroit, Michigan 48202
[93]Yamagata University, Yamagata 990-8560
[94]Yonsei University, Seoul 03722
[95]Luther College, Decorah, Iowa 52101

(Received 3 February 2021; accepted 25 May 2021; published 12 July 2021)

We present measurements of partial branching fractions of inclusive semileptonic $B \to X_u \ell^+ \nu_\ell$ decays using the full Belle dataset of 711 fb^{-1} of integrated luminosity at the $\Upsilon(4S)$ resonance and for $\ell = e, \mu$. Inclusive semileptonic $B \to X_u \ell^+ \nu_\ell$ decays are Cabibbo-Kobayashi-Maskawa (CKM) suppressed and measurements are complicated by the large background from CKM favored $B \to X_c \ell^+ \nu_\ell$ transitions, which have a similar signature. Using machine learning techniques, we reduce this and other backgrounds effectively, while retaining access to a large fraction of the $B \to X_u \ell^+ \nu_\ell$ phase space and high signal efficiency. We measure partial branching fractions in three phase-space regions covering about 31% to 86% of the accessible $B \to X_u \ell^+ \nu_\ell$ phase space. The most inclusive measurement corresponds to the phase space with lepton energies of $E_\ell^B > 1$ GeV, and we obtain $\Delta\mathcal{B}(B \to X_u \ell^+ \nu_\ell) = (1.59 \pm 0.07 \pm 0.16) \times 10^{-3}$ from a two-dimensional fit of the hadronic mass spectrum and the four-momentum-transfer squared distribution, with the uncertainties denoting the statistical and systematic error. We find $|V_{ub}| = (4.10 \pm 0.09 \pm 0.22 \pm 0.15) \times 10^{-3}$ from an average of four calculations for the partial decay rate with the third uncertainty denoting the average theory error. This value is higher but compatible with the determination from exclusive semileptonic decays within 1.3 standard deviations. In addition, we report charmless inclusive partial branching fractions separately for B^+ and B^0 mesons as well as for electron and muon final states. No isospin breaking or lepton flavor universality violating effects are observed.

DOI: 10.1103/PhysRevD.104.012008

I. INTRODUCTION

Precision measurements of the absolute value of the Cabibbo-Kobayashi-Maskawa (CKM) matrix element V_{ub} are important to challenge the Standard Model of particle physics (SM) [1,2]. In the SM, the CKM matrix is a 3×3 unitary matrix and responsible for the known charge-parity (CP)-violating effects in the quark sector [3]. There are indications of CP violation (CPV) in the neutrino sector [4], but it remains unclear if both sources of CP violation are sufficient to explain the matter dominance of today's Universe. This motivates the search for new sources of CP-violating phenomena. If such exist in the form of heavy exotic particles that couple to quarks in some form, their presence might alter the properties of measurements constraining the unitarity of the CKM matrix [5]. Precise measurements of $|V_{ub}|$ and the CKM angle $\gamma = \phi_3$ are imperative to isolate such effects, as their measurements

*cao@physik.uni-bonn.de
†florian.bernlochner@uni-bonn.de
‡Present address: Hiroshima University, Higashi-Hiroshima, Hiroshima 739-8530.

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP^3.

involve tree-level processes, which are expected to remain unaffected by new physics and thus provide an unbiased measure for the amount of CPV due to the Kobayashi-Maskawa (KM) mechanism [2] alone.

Charmless semileptonic decays of B mesons provide a clean avenue to measure $|V_{ub}|$, as their decay rate is theoretically better understood than purely hadronic transitions and their decay signature is more accessible than leptonic B meson decays. The existing measurements either focus on exclusive final states, with $B \to \pi \ell^+ \nu_\ell$ [6] and the ratio of $\Lambda_b \to p\mu^+\nu_\mu$ and $\Lambda_b \to \Lambda_c \mu^+ \nu_\mu$ [7] providing the most precise measurements to date, and measurements reconstructing the $B \to X_u \ell^+ \nu_\ell$ decay fully inclusively.[1] Central for both approaches are reliable predictions of the (partial) decay rates $\Delta\Gamma(B \to X_u \ell^+ \nu_\ell)$ (omitting the CKM factor) from theory to convert measured (partial or full) branching fractions, $\Delta\mathcal{B}(B \to X_u \ell^+ \nu_\ell)$, into measurements of $|V_{ub}|$ via

$$|V_{ub}| = \sqrt{\frac{\Delta\mathcal{B}(B \to X_u \ell^+ \nu_\ell)}{\tau_B \Delta\Gamma(B \to X_u \ell^+ \nu_\ell)}},$$

with τ_B denoting the B meson lifetime. For exclusive measurements, the nonperturbative parts of the decay rates can be reliably predicted by lattice QCD [8] or light-cone sum rules [9] and constrained by the measurements of the decay dynamics. The determination of $|V_{ub}|$ using inclusive decays is very challenging due to the large background from the CKM favored $B \to X_c \ell^+ \nu_\ell$ process. Both processes have a very similar decay signature in the form of a high-momentum lepton, a hadronic system, and missing energy from the neutrino that escapes detection. Figure 1 shows an illustration of both processes for a B^0 meson decay. A clear separation of the processes is only possible in kinematic regions where $B \to X_c \ell^+ \nu_\ell$ is kinematically forbidden. In these regions, however, nonperturbative shape functions enter the description of the decay dynamics, making predictions for the decay rates dependent on the precise modeling. These functions parametrize at leading order the Fermi motion of the b quark inside the B meson. Properties of the leading-order $\Lambda_{\rm QCD}/m_b$ shape function can be determined using the photon energy spectrum of $B \to X_s\gamma$ decays and moments of the lepton energy or hadronic invariant mass in semileptonic B decays [10–12], but the modeling of both the leading and subleading shape functions introduces large theory uncertainties on the decay rate. In the future, more model-independent approaches aim to directly measure the leading-order shape function [13,14].

As such methods are not yet realized, it is beneficial to extend the measurement region as much as possible into the

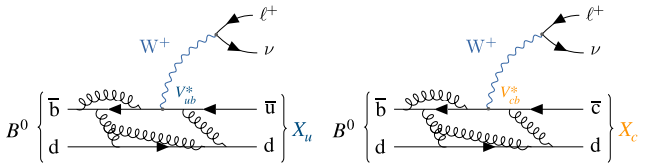


FIG. 1. The CKM suppressed and favored inclusive semileptonic processes $B \to X_u \ell^+ \nu_\ell$ (left) and $B \to X_c \ell^+ \nu_\ell$ (right) for a B^0 meson decay.

$B \to X_c \ell^+ \nu_\ell$ dominated phase space. This was done, e.g., by Refs. [15,16]. This reduces the theory uncertainties on the predicted partial rates [17–22], although making the measurement more prone to systematic uncertainties. This strategy is also adopted in the measurement described in this paper.

The corresponding world averages of $|V_{ub}|$ from both exclusive and inclusive determinations are [6]

$$|V_{ub}^{\rm excl}| = (3.67 \pm 0.09 \pm 0.12) \times 10^{-3}, \quad (1)$$

$$|V_{ub}^{\rm incl}| = (4.32 \pm 0.12^{+0.12}_{-0.13}) \times 10^{-3}. \quad (2)$$

Here the uncertainties are experimental and from theory. Both world averages exhibit a disagreement of about 3 standard deviations between them. This disagreement is limiting the reach of present-day precision tests of the KM mechanism and searches for loop-level new physics, see, e.g., Ref. [23] for a recent analysis. For a more complete review the interested reader is referred to Refs. [24,25].

One important experimental method to extend the probed $B \to X_u \ell^+ \nu_\ell$ phase space into regions dominated by $B \to X_c \ell^+ \nu_\ell$ transitions is the full reconstruction of the second B meson of the $e^+e^- \to \Upsilon(4S) \to B\bar{B}$ process. This process is referred to as "tagging" and allows for the reconstruction of the hadronic X system of the semileptonic process. In addition, the neutrino four-momentum can be reconstructed. Properties of both are instrumental to distinguish $B \to X_u \ell^+ \nu_\ell$ and $B \to X_c \ell^+ \nu_\ell$ processes. In this manuscript the reconstruction of the second B meson and the separation of $B \to X_u \ell^+ \nu_\ell$ from $B \to X_c \ell^+ \nu_\ell$ processes were carried out using machine learning approaches. Several neural networks were trained to identify correctly reconstructed tag-side B mesons. The distinguishing variables of the classification algorithm were carefully selected in order not to introduce a bias in the measured partial branching fractions. In addition, the modeling of backgrounds was validated in $B \to X_c \ell^+ \nu_\ell$ -enriched selections. We report the measurement of three partial branching fractions, covering 31%–86% of the accessible $B \to X_u \ell^+ \nu_\ell$ phase space. The measurement of fully differential distributions, which allow one to determine the leading and subleading shape functions, is left for future work.

The main improvement over the previous Belle result of Ref. [16] lies in the adoption of a more efficient tagging

[1]Charge conjugation is implied throughout this paper. In addition, $B \to X_u \ell^+ \nu_\ell$ is defined as the average branching fraction of charged and neutral B meson decays and $\ell = e$ or μ.

algorithm for the reconstruction of the second B meson and the improvements of the $B \to X_u \ell^+ \nu_\ell$ signal and $B \to X_c \ell^+ \nu_\ell$ background descriptions. In addition, the full Belle dataset of 711 fb^{-1} is analyzed and we avoid the direct use of kinematic properties of the candidate semileptonic decay in the background suppression. After the final selection we retain a factor of approximatively 1.8 times more signal events than the previous analysis with a ca. 20% improved purity.

The remainder of this manuscript is organized as follows: Section II provides an overview of the dataset and the simulated signal and background samples that were used in the analysis. Section III details the analysis strategy and reconstruction of the hadronic X system of the semileptonic decay. Section IV introduces the fit procedure used to separate $B \to X_u \ell^+ \nu_\ell$ signal from background contributions. Section V lists the systematic uncertainties affecting the measurements and Sec. VI summarizes sideband studies central to validate the modeling of the crucial $B \to X_c \ell^+ \nu_\ell$ background processes. Finally, Sec. VII shows the selected signal events and compares them with the expectation from simulation. In Sec. VIII the measured partial branching fractions and subsequent values of $|V_{ub}|$ are discussed. Section IX presents our conclusions.

II. DATASET AND SIMULATED SAMPLES

The analysis utilizes the full Belle dataset of $(772 \pm 10) \times 10^6$ B meson pairs, which were produced at the KEKB accelerator complex [26] with a center-of-mass energy of $\sqrt{s} = 10.58$ GeV corresponding to the $\Upsilon(4S)$ resonance. In addition, 79 fb^{-1} of collision events recorded 60 MeV below the $\Upsilon(4S)$ resonance peak are used to derive corrections and for cross-checks.

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 barrellike arrangement of time-of-flight scintillation counters (TOF), 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 located outside of the coil is instrumented to detect K_L^0 mesons (KLM) and to identify muons. A more detailed description of the detector and its layout and performance can be found in Ref. [27] and in references therein.

Charged tracks are identified as electron or muon candidates by combining the information of multiple subdetectors into a lepton identification likelihood ratio, $\mathcal{L}_{\mathrm{LID}}$. For electrons, the most important 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 can be identified from charged track trajectories extrapolated to the outer detector. The most important 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 ratio $\mathcal{L}_{\mathrm{K}/\pi\mathrm{ID}} = \mathcal{L}_{\mathrm{KID}}/(\mathcal{L}_{\mathrm{KID}} + \mathcal{L}_{\pi\mathrm{ID}})$. The most important identifying features of the kaon ($\mathcal{L}_{\mathrm{KID}}$) and pion ($\mathcal{L}_{\pi\mathrm{ID}}$) likelihoods for low-momentum particles with transverse momentum below 1 GeV in the laboratory frame are the recorded energy loss by ionization, $\mathrm{d}E/\mathrm{d}x$, in the CDC, and the time of flight information from the TOF. Higher-momentum kaon and pion classification relies on the Cherenkov light recorded in the ACC. In order to avoid the difficulties in understanding the efficiencies of reconstructing K_L^0 mesons, they are not explicitly reconstructed or used in this analysis.

Photons are identified as energy depositions in the ECL, vetoing clusters to which an associated track can be assigned. Only photons with an energy deposition of $E_\gamma > 100$, 150, and 50 MeV in the forward end cap, backward end cap, and barrel part of the calorimeter, respectively, are considered. We reconstruct π^0 candidates from photon candidates. The invariant mass is required to fall inside a window[2] of $m_{\gamma\gamma} \in [0.12, 0.15]$ GeV, which corresponds to about 2.5 times the π^0 mass resolution.

Monte Carlo (MC) samples of B meson decays and continuum processes ($e^+e^- \to q\bar{q}$ with $q = u, d, s, c$) are simulated using the EvtGen generator [30]. These samples are used to evaluate reconstruction efficiencies and acceptance, and to estimate background contaminations. The sample sizes used correspond to approximately ten and five times, respectively, the Belle collision data for B meson and continuum decays. The interactions of particles traversing the detector are simulated using GEANT3 [31]. Electromagnetic final-state radiation is simulated using the PHOTOS [32] package for all charged final-state particles. The efficiencies in the MC are corrected using data-driven methods to account for, e.g., differences in identification and reconstruction efficiencies.

The most important background processes are semileptonic $B \to X_c \ell^+ \nu_\ell$ decays and continuum processes, which both can produce high-momentum leptons in a momentum range similar to the $B \to X_u \ell^+ \nu_\ell$ process. The semileptonic background from $B \to X_c \ell^+ \nu_\ell$ decays is dominated by $B \to D \ell^+ \nu_\ell$ and $B \to D^* \ell^+ \nu_\ell$ decays. The $B \to D \ell^+ \nu_\ell$ decays are modeled using the Boyd-Grinstein-Lebed (BGL) parametrization [33] with form factor central values and uncertainties taken from the fit in Ref. [34]. For $B \to D^* \ell^+ \nu_\ell$ we use the BGL implementation proposed by Refs. [35,36] with form factor central values and uncertainties from the fit to the

[2]We use natural units: $\hbar = c = 1$.

measurement of Ref. [37]. Both backgrounds are normalized to the average branching fraction of Ref. [6] assuming isospin symmetry. Semileptonic $B \to D^{**}\ell^+\nu_\ell$ decays with $D^{**} = \{D_0^*, D_1^*, D_1, D_2^*\}$ denoting the four orbitally excited charmed mesons are modeled using the heavy-quark-symmetry-based form factors proposed in Ref. [38]. We simulate all D^{**} decays using masses and widths from Ref. [39]. For the branching fractions we adopt the values of Ref. [6] and correct them to account for missing isospin-conjugated and other established decay modes, following the prescription given in Ref. [38]. To correct for the fact that the measurements were carried out in the $D^{**0} \to D^{(*)+}\pi^-$ decay modes, we account for the missing isospin modes with a factor of

$$f_\pi = \frac{\mathcal{B}(\bar{D}^{**0} \to D^{(*)-}\pi^+)}{\mathcal{B}(\bar{D}^{**0} \to \bar{D}^{(*)}\pi)} = \frac{2}{3}. \tag{3}$$

The measurements of the $B \to D_2^*\ell\bar{\nu}_\ell$ in Ref. [6] are converted to only account for the $\bar{D}_2^{*0} \to D^{*-}\pi^+$ decay. To also account for $\bar{D}_2^{*0} \to D^-\pi^+$ contributions, we apply a factor of [39]

$$f_{D_2^*} = \frac{\mathcal{B}(\bar{D}_2^{*0} \to D^-\pi^+)}{\mathcal{B}(\bar{D}_2^{*0} \to D^{*-}\pi^+)} = 1.54 \pm 0.15. \tag{4}$$

The world average of $B \to D_1^*\ell\bar{\nu}_\ell$ given in Ref. [6] combines measurements, which show poor agreement, and the resulting probability of the combination is below 0.01%. Notably, the measurement of Ref. [40] is in conflict with the measured branching fractions of Refs. [41,42] and with the expectation of $\mathcal{B}(B \to D_1^*\ell\bar{\nu}_\ell)$ being of similar size than $\mathcal{B}(B \to D_0\ell\bar{\nu}_\ell)$ [43,44]. We perform our own average excluding Ref. [40] and use

$$\mathcal{B}(B^+ \to \bar{D}_1^{*0}(\to D^{*-}\pi^+)\ell^+\nu_\ell) = (0.28 \pm 0.06) \times 10^{-2}. \tag{5}$$

The world average of $B \to D_1\ell\bar{\nu}_\ell$ does not include contributions from prompt three-body decays of $D_1 \to D\pi\pi$. We account for these using a factor [45]

$$f_{D_1} = \frac{\mathcal{B}(\bar{D}_1^0 \to D^{*-}\pi^+)}{\mathcal{B}(\bar{D}_1^0 \to \bar{D}^0\pi^+\pi^-)} = 2.32 \pm 0.54. \tag{6}$$

We subtract the contribution of $D_1 \to D\pi\pi$ from the measured nonresonant plus resonant $B \to D\pi\pi\ell\bar{\nu}_\ell$ branching fraction of Ref. [46]. To account for missing isospin-conjugated modes of the three-hadron final states we adopt the prescription from Ref. [46], which calculates an average isospin correction factor of

$$f_{\pi\pi} = \frac{\mathcal{B}(\bar{D}^{**0} \to \bar{D}^{(*)0}\pi^+\pi^-)}{\mathcal{B}(\bar{D}^{**0} \to \bar{D}^{(*)}\pi\pi)} = \frac{1}{2} \pm \frac{1}{6}. \tag{7}$$

The uncertainty takes into account the full spread of final states [$f_0(500) \to \pi\pi$ or $\rho \to \pi\pi$ result in $f_{\pi\pi} = 2/3$ and $1/3$, respectively] and the nonresonant three-body decays ($f_{\pi\pi} = 3/7$). We further assume that

$$\begin{aligned} \mathcal{B}(\bar{D}_2^* \to \bar{D}\pi) + \mathcal{B}(\bar{D}_2^* \to \bar{D}^*\pi) &= 1, \\ \mathcal{B}(\bar{D}_1 \to \bar{D}^*\pi) + \mathcal{B}(\bar{D}_1 \to \bar{D}\pi\pi) &= 1, \\ \mathcal{B}(\bar{D}_1^* \to \bar{D}^*\pi) = 1, \quad \text{and} \quad \mathcal{B}(\bar{D}_0 \to \bar{D}\pi) &= 1. \end{aligned} \tag{8}$$

For the remaining $B \to D^{(*)}\pi\pi\ell^+\nu_\ell$ contributions we use the measured value of Ref. [46]. The remaining "gap" between the sum of all considered exclusive modes and the inclusive $B \to X_c\ell^+\nu_\ell$ branching fraction ($\approx 0.8 \times 10^{-2}$ or 7%–8% of the total $B \to X_c\ell^+\nu_\ell$ branching fraction) is filled in equal parts with $B \to D\eta\ell^+\nu_\ell$ and $B \to D^*\eta\ell^+\nu_\ell$ and we assume a 100% uncertainty on this contribution. We simulate $B \to D^{(*)}\pi\pi\ell^+\nu_\ell$ and $B \to D^{(*)}\eta\ell^+\nu_\ell$ final states assuming that they are produced by the decay of two broad resonant states D_{gap}^{**} with masses and widths identical to D_1^* and D_0. Although there is currently no experimental evidence for decays of charm $1P$ states into these final states or the existence of such an additional broad state (e.g., a $2S$) in semileptonic transitions, this description provides a better kinematic description of the initial three-body decay, $B \to D_{\text{gap}}^{**}\ell\bar{\nu}_\ell$, than, e.g., a model based on the equidistribution of all final-state particles in phase space. For the form factors we adapt Ref. [38].

Semileptonic $B \to X_u\ell^+\nu_\ell$ decays are modeled as a mixture of specific exclusive modes and nonresonant contributions. We normalize their corresponding branching fractions to the world averages from Ref. [39]: semileptonic $B \to \pi\ell^+\nu_\ell$ decays are simulated using the Bourrely-Caprini-Lellouch (BCL) parametrization [47] with form factor central values and uncertainties from the global fit carried out by Ref. [48]. The processes of $B \to \rho\ell^+\nu_\ell$ and $B \to \omega\ell^+\nu_\ell$ are modeled using the BCL form factor parametrization. We use the fit of Ref. [49], that combines the measurements of Refs. [50–52] with the light-cone sum rule predictions of Ref. [9] to determine a set of form factor central values and uncertainties. The processes of $B \to \eta\ell^+\nu_\ell$ and $B \to \eta'\ell^+\nu_\ell$ are modeled using the LCSR calculation of Ref. [53]. For the uncertainties we assume for these states that the pole parameters $\alpha^{+/0}$ and the form factor normalization $f_{B\eta}^+(0)$ at maximum recoil can be treated as uncorrelated. In addition to these narrow resonances, we simulate nonresonant $B \to X_u\ell^+\nu_\ell$ decays with at least two pions in the final state following the De Fazio and Neubert (DFN) model [54]. The triple differential rate of this model is a function of the four-momentum-transfer squared (q^2), the lepton energy (E_ℓ^B) in the B rest frame, and the hadronic invariant mass squared (M_X^2) of the X_u system at next-to-leading-order precision in the strong coupling constant α_s. This triple differential rate is convolved with a

nonperturbative shape function using an *ad hoc* exponential model. The free parameters of the model are the b quark mass in the Kagan-Neubert scheme [55], $m_b^{\mathrm{KN}} = (4.66 \pm 0.04)$ GeV and a nonperturbative parameter $a^{\mathrm{KN}} = 1.3 \pm 0.5$. The values of these parameters were determined in Ref. [56] from a fit to $B \to X_c \ell^+ \nu_\ell$ and $B \to X_s \gamma$ decay properties. At leading order, the nonperturbative parameter a^{KN} is related to the average momentum squared of the b quark inside the B meson and determines the second moment of the shape function. It is defined as $a^{\mathrm{KN}} = -3\bar{\Lambda}^2/\lambda_1 - 1$ with the binding energy $\bar{\Lambda} = m_B - m_b^{\mathrm{KN}}$ and the kinetic energy parameter λ_1. The hadronization of the parton-level $B \to X_u \ell^+ \nu_\ell$ DFN simulation is carried out using the JETSET algorithm [57], producing final states with two or more mesons. The inclusive and exclusive $B \to X_u \ell^+ \nu_\ell$ predictions are combined using a so-called hybrid approach, which is a method originally suggested by Ref. [58], and our implementation closely follows Ref. [59] and uses the library of Ref. [60]. To this end, we combine both predictions such that the partial branching fractions in the triple differential rate of the inclusive ($\Delta\mathcal{B}_{ijk}^{\mathrm{incl}}$) and combined exclusive ($\Delta\mathcal{B}_{ijk}^{\mathrm{excl}}$) predictions reproduce the inclusive values. This is achieved by assigning weights to the inclusive contributions w_{ijk} such that

$$\Delta\mathcal{B}_{ijk}^{\mathrm{incl}} = \Delta\mathcal{B}_{ijk}^{\mathrm{excl}} + w_{ijk} \times \Delta\mathcal{B}_{ijk}^{\mathrm{incl}}, \tag{9}$$

with i, j, k denoting the corresponding bin in the three dimensions of q^2, E_ℓ^B, and M_X,

$$\begin{aligned} q^2 &= [0, 2.5, 5, 7.5, 10, 12.5, 15, 20, 25]\ \mathrm{GeV}^2, \\ E_\ell^B &= [0, 0.5, 1, 1.25, 1.5, 1.75, 2, 2.25, 3]\ \mathrm{GeV}, \\ M_X &= [0, 1.4, 1.6, 1.8, 2, 2.5, 3, 3.5]\ \mathrm{GeV}. \end{aligned}$$

To study the model dependence of the DFN shape function, we also determine weights using the Bosch-Lange-Neubert-Paz (BLNP) model of Ref. [61] and treat the difference later as a systematic uncertainty. For the b quark mass in the shape-function scheme we use $m_b^{\mathrm{SF}} = 4.61$ GeV and $\mu_\pi^{2\,\mathrm{SF}} = 0.20$ GeV^2. Figures detailing the hybrid model construction can be found in Appendix A.

Table I summarizes the branching fractions for the signal and the important $B \to X_c \ell^+ \nu_\ell$ background processes that were used. Figure 2 shows the generator-level distributions and yields of $B \to X_c \ell^+ \nu_\ell$ and $B \to X_u \ell^+ \nu_\ell$ after the tag-side reconstruction (cf. Sec. III). The $B \to X_u \ell^+ \nu_\ell$ yields were scaled up by a factor of 50 to make them visible. A clear separation can be obtained at low values of M_X and high values of E_ℓ^B.

TABLE I. Branching fractions for $B \to X_u \ell^+ \nu_\ell$ and $B \to X_c \ell^+ \nu_\ell$ background processes that were used are listed. More details on the applied corrections can be found in the text. We neglect the small contribution from $B^+ \to D_s^{(*)} K^+ \ell^+ \nu_\ell$ which has a branching fraction of similar size as $B \to D\pi\pi\ell^+\nu_\ell$.

$\mathcal{B}$	Value B^+	Value B^0
$B \to X_u \ell^+ \nu_\ell$		
$B \to \pi \ell^+ \nu_\ell$	$(7.8 \pm 0.3) \times 10^{-5}$	$(1.5 \pm 0.06) \times 10^{-4}$
$B \to \eta \ell^+ \nu_\ell$	$(3.9 \pm 0.5) \times 10^{-5}$	⋯
$B \to \eta' \ell^+ \nu_\ell$	$(2.3 \pm 0.8) \times 10^{-5}$	⋯
$B \to \omega \ell^+ \nu_\ell$	$(1.2 \pm 0.1) \times 10^{-4}$	⋯
$B \to \rho \ell^+ \nu_\ell$	$(1.6 \pm 0.1) \times 10^{-4}$	$(2.9 \pm 0.2) \times 10^{-4}$
$B \to X_u \ell^+ \nu_\ell$	$(2.2 \pm 0.3) \times 10^{-3}$	$(2.0 \pm 0.3) \times 10^{-3}$
$B \to X_c \ell^+ \nu_\ell$		
$B \to D \ell^+ \nu_\ell$	$(2.5 \pm 0.1) \times 10^{-2}$	$(2.3 \pm 0.1) \times 10^{-2}$
$B \to D^* \ell^+ \nu_\ell$	$(5.4 \pm 0.1) \times 10^{-2}$	$(5.1 \pm 0.1) \times 10^{-2}$
$B \to D_0^* \ell^+ \nu_\ell$ ($\hookrightarrow D\pi$)	$(4.2 \pm 0.8) \times 10^{-3}$	$(3.9 \pm 0.7) \times 10^{-3}$
$B \to D_1^* \ell^+ \nu_\ell$ ($\hookrightarrow D^*\pi$)	$(4.2 \pm 0.8) \times 10^{-3}$	$(3.9 \pm 0.8) \times 10^{-3}$
$B \to D_1 \ell^+ \nu_\ell$ ($\hookrightarrow D^*\pi$)	$(4.2 \pm 0.3) \times 10^{-3}$	$(3.9 \pm 0.3) \times 10^{-3}$
$B \to D_2^* \ell^+ \nu_\ell$ ($\hookrightarrow D^*\pi$)	$(1.2 \pm 0.1) \times 10^{-3}$	$(1.1 \pm 0.1) \times 10^{-3}$
$B \to D_2^* \ell^+ \nu_\ell$ ($\hookrightarrow D\pi$)	$(1.8 \pm 0.2) \times 10^{-3}$	$(1.7 \pm 0.2) \times 10^{-3}$
$B \to D_1 \ell^+ \nu_\ell$ ($\hookrightarrow D\pi\pi$)	$(2.4 \pm 1.0) \times 10^{-3}$	$(2.3 \pm 0.9) \times 10^{-3}$
$B \to D\pi\pi\ell^+\nu_\ell$	$(0.6 \pm 0.6) \times 10^{-3}$	$(0.6 \pm 0.6) \times 10^{-3}$
$B \to D^*\pi\pi\ell^+\nu_\ell$	$(2.2 \pm 1.0) \times 10^{-3}$	$(2.0 \pm 1.0) \times 10^{-3}$
$B \to D\eta\ell^+\nu_\ell$	$(4.0 \pm 4.0) \times 10^{-3}$	$(4.0 \pm 4.0) \times 10^{-3}$
$B \to D^*\eta\ell^+\nu_\ell$	$(4.0 \pm 4.0) \times 10^{-3}$	$(4.0 \pm 4.0) \times 10^{-3}$
$B \to X_c \ell^+ \nu_\ell$	$(10.8 \pm 0.4) \times 10^{-2}$	$(10.1 \pm 0.4) \times 10^{-2}$

III. ANALYSIS STRATEGY, HADRONIC TAGGING, AND X RECONSTRUCTION

A. Neural-network-based tag-side reconstruction

We reconstruct collision events using the hadronic full reconstruction algorithm of Ref. [62]. The algorithm reconstructs one of the B mesons produced in the collision event using hadronic decay channels. We label such B mesons in the following as B_{tag}. Instead of attempting to reconstruct as many B meson decay cascades as possible, the algorithm employs a hierarchical reconstruction ansatz in four stages: At the first stage, neural networks are trained to identify charged tracks and neutral energy depositions as detector stable particles ($e^+, \mu^+, K^+, \pi^+, \gamma$), neutral π^0 candidates, or K_S^0 candidates. At the second stage, these candidate particles are combined into heavier meson candidates ($J/\psi, D^0, D^+, D_s$) and for each target final state a neural network is trained to identify probable candidates. In addition to the classifier output from the first stage, vertex fit probabilities of the candidate combinations and

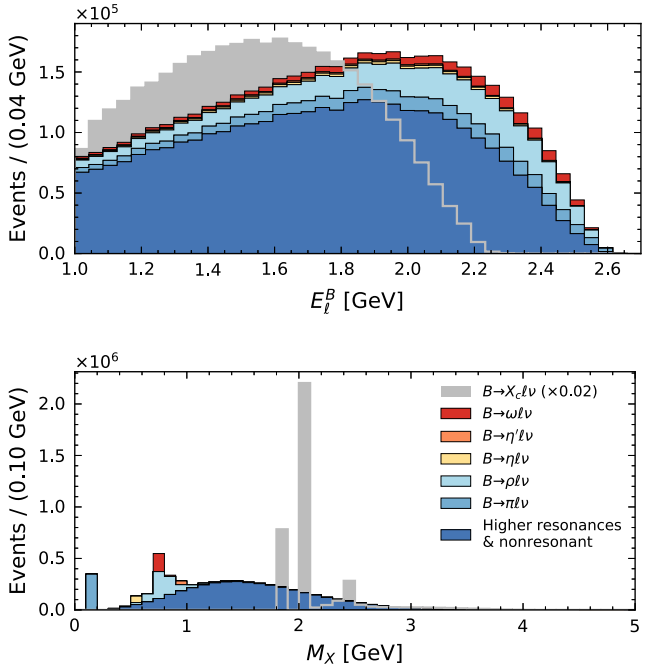


FIG. 2. The generator-level E_ℓ^B and M_X distributions of the CKM suppressed and favored inclusive semileptonic processes, $B \to X_u \ell^+ \nu_\ell$ (scaled up by a factor of 50) and $B \to X_c \ell^+ \nu_\ell$, respectively, are shown, using the models described in the text.

the full four-momentum of the combination are passed to the input layer. At the third stage, candidates for D^{*0}, D^{*+}, and D_s^* mesons are formed and separate neural networks are trained to identify viable combinations. The input layer aggregates the output classifiers from all previous reconstruction stages. The final stage combines the information from all previous stages to form $B_{\rm tag}$ candidates. The viability of such combinations is again assessed by a neural network that was trained to distinguish correctly reconstructed candidates from wrong combinations and whose output classifier score we denote by $\mathcal{O}_{\rm FR}$. Over 1104 decay cascades are reconstructed in this manner, achieving an efficiency of 0.28% and 0.18% for charged and neutral B meson pairs [63], respectively. Finally, the output of this classifier is used as an input and combined with a range of event shape variables to train a neural network to distinguish reconstructed B meson candidates from continuum processes. The output classifier score of this neural network is denoted as $\mathcal{O}_{\rm Cont}$. Both classifier scores are mapped to a range of [0, 1) signifying the reconstruction quality of poor to excellent candidates. We retain $B_{\rm tag}$ candidates that show at least moderate agreement based on these two outputs and require that $\mathcal{O}_{\rm FR} > 10^{-4}$ and $\mathcal{O}_{\rm Cont} > 10^{-4}$. Despite these relatively low values, knowledge of the charge and momentum of the decay constituents in combination with the known beam energy allows one to infer the flavor and four-momentum of the $B_{\rm tag}$ candidate. We require the $B_{\rm tag}$ candidates to have at least a beam-constrained mass of

$$M_{\rm bc} = \sqrt{E_{\rm beam}^2 - |\mathbf{p}_{\rm tag}|^2} > 5.27\ \text{GeV}, \tag{10}$$

with $\mathbf{p}_{\rm tag}$ denoting the momentum of the $B_{\rm tag}$ candidate in the center-of-mass frame of the colliding e^+e^- pair. Furthermore, $E_{\rm beam} = \sqrt{s}/2$ denotes half the center-of-mass energy of the colliding e^+e^- pair. The energy difference

$$\Delta E = E_{\rm tag} - E_{\rm beam} \tag{11}$$

is already used in the input layer of the neural network trained in the final stage of the reconstruction. Here $E_{\rm tag}$ denotes the energy of the $B_{\rm tag}$ candidate in the center-of-mass frame of the colliding e^+e^- pair. In each event a single $B_{\rm tag}$ candidate is then selected according to the highest $\mathcal{O}_{\rm FR}$ score of the hierarchical full reconstruction algorithm. All tracks and clusters not used in the reconstruction of the $B_{\rm tag}$ candidate are used to define the signal side.

B. Signal side reconstruction

The signal side of the event is reconstructed by identifying a well-reconstructed lepton with $E_\ell^B = |\mathbf{p}_\ell^B| > 1$ GeV in the signal B rest frame[3] using the likelihood mentioned in Sec. II. The signal B rest frame is calculated using the momentum of the $B_{\rm tag}$ candidate via

$$p_{\rm sig} = p_{e^+e^-} - \left(\sqrt{m_B^2 + |\mathbf{p}_{\rm tag}|^2}, \mathbf{p}_{\rm tag}\right), \tag{12}$$

with $p_{e^+e^-}$ denoting the four-momentum of the colliding electron-positron pair. Leptons from J/ψ and photon conversions in detector material are rejected by combining the lepton candidate with oppositely charged tracks (t) on the signal side and demanding that $m_{\ell t} > 0.14$ and $m_{et} \notin [3.05, 3.15]$ GeV or $m_{\mu t} \notin [3.06, 3.12]$ GeV. If multiple lepton candidates are present on the signal side, the event is discarded as multiple leptons are likely to originate from a double semileptonic $b \to c \to s$ cascade. For charged $B_{\rm tag}$ candidates, we demand that the charge assignment of the signal side lepton be opposite that of the $B_{\rm tag}$ charge. The hadronic X system is reconstructed from the remaining unassigned charged particles and neutral energy depositions. Its four-momentum is calculated as

$$p_X = \sum_i \left(\sqrt{m_\pi^2 + |\mathbf{p}_i|^2}, \mathbf{p}_i\right) + \sum_j (E_j, \mathbf{k}_j), \tag{13}$$

with $E_i = |\mathbf{k_i}|$ the energy of the neutral energy depositions and all charged particles with momentum $\mathbf{p_i}$ are assumed to be pions. With the X system reconstructed, we can also reconstruct the missing mass squared,

[3]We neglect the small correction of the lepton mass term to the energy of the lepton.

$$M^2_{\text{miss}} = (p_{\text{sig}} - p_X - p_\ell)^2, \tag{14}$$

which should peak at zero, $M^2_{\text{miss}} \approx m^2_\nu \approx 0\ \text{GeV}^2$, for correctly reconstructed semileptonic $B \to X_u \ell^+ \nu_\ell$ and $B \to X_c \ell^+ \nu_\ell$ decays. The hadronic mass of the X system is later used to discriminate $B \to X_u \ell^+ \nu_\ell$ signal decays from $B \to X_c \ell^+ \nu_\ell$ and other remaining backgrounds. It is reconstructed using

$$M_X = \sqrt{(p_X)^\mu (p_X)_\mu}. \tag{15}$$

In addition, we reconstruct the four-momentum-transfer squared q^2 as

$$q^2 = (p_{\text{sig}} - p_X)^2. \tag{16}$$

The resolution of both variables for $B \to X_u \ell^+ \nu_\ell$ is shown in Fig. 3 as residuals with respect to the generated values of q^2 and M_X. The resolution for M_X has a root-mean-square (rms) deviation of 0.47 GeV, but exhibits a large tail toward larger values. The distinct peak at 0 is from $B^0 \to \pi^- \ell^+ \nu_\ell$ and other low-multiplicity final states comprised of only charged pions. The four-momentum-transfer squared q^2 exhibits a large resolution, which is caused by a combination of the tag-side B and the X reconstruction. The rms deviation for q^2 is 1.59 GeV^2. The core resolution is dominated by the tagging resolution, whereas the large negative tail is dominated from the resolution of the reconstruction of the X system.

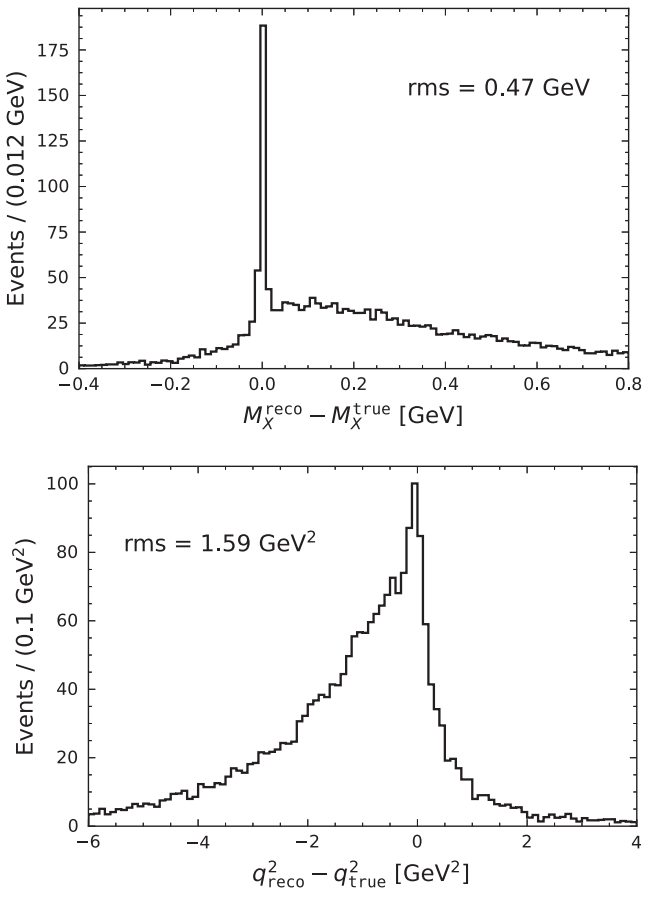


FIG. 3. The resolution of the reconstructed M_X and q^2 values for $B \to X_u \ell^+ \nu_\ell$ signal is shown as a residual with respect to the generated values.

C. Background suppression BDT

At this point in the reconstruction, the $B \to X_c \ell^+ \nu_\ell$ process completely dominates the selected events. To identify $B \to X_u \ell^+ \nu_\ell$, we combine several distinguishing features into a single discriminant. This is achieved by using a machine-learning-based classification with boosted decision trees (BDTs). Note that all momenta are in the center-of-mass frame of the colliding e^+e^- pair. These features are as follows:

1. M^2_{miss}: The average $B \to X_c \ell^+ \nu_\ell$ multiplicity is higher than $B \to X_u \ell^+ \nu_\ell$, broadening the missing mass squared distribution.
2. D^* veto: We search for low-momentum neutral and charged pions in the X system with $|\mathbf{p}_\pi| < 220$ MeV, compatible with a $D^* \to D\pi$ transition. The key idea of this is that, due to the small available phase space from the small mass difference between the D^* and D mesons, the flight direction of the slow pion is strongly correlated with the D^* momentum direction. The energy and momentum of a D^* candidate can thus be approximated as

$$E_{D^*} = \frac{m_{D^*}}{m_{D^*} - m_D} \times E_\pi,$$
$$\mathbf{p}_{D^*} = \mathbf{p}_\pi \times \frac{\sqrt{E^2_{D^*} - m^2_{D^*}}}{|\mathbf{p}_\pi|}, \tag{17}$$

with m_{D^*} and m_D denoting the D^* and D meson masses, respectively, and $E_\pi = \sqrt{m^2_\pi + |\mathbf{p}_\pi|^2}$ is the energy of the slow pion. Using the D^* candidate four-momentum $p_{D^*} = (E_{D^*}, \mathbf{p}_{D^*})$ we can calculate

$$M^2_{\text{miss},D^*} = (p_{\text{sig}} - p_{D^*} - p_\ell)^2,$$
$$\cos\theta_{B,D^*\ell} = \frac{2E_{\text{beam}}E_{D^*\ell} - m^2_B - m^2_{D^*\ell}}{2|\mathbf{p}_B||\mathbf{p}_{D^*\ell}|},$$
$$\cos\theta^* = \frac{\mathbf{p}_\ell \cdot \mathbf{p}_{D^*}}{|\mathbf{p}_\ell||\mathbf{p}_{D^*}|}, \tag{18}$$

with $p_{D^*\ell} = p_{D^*} + p_\ell = (E_{D^*\ell}, \mathbf{p}_{D^*\ell})$ and $|\mathbf{p}_B| = \sqrt{E^2_B - m^2_B}$. These three variables are used exclusively for events with charged and neutral slow pion candidates.
3. Kaons: We identify the number of K^+ candidates using the particle-identification likelihood, cf. Sec. II.

In addition, we reconstruct K_S^0 candidates from displaced tracks found in the X system.

4. $B_{\rm sig}$ vertex fit: The charmed mesons produced in $B \to X_c \ell^+ \nu_\ell$ transitions exhibit a longer lifetime than their charmless counterparts produced in $B \to X_u \ell^+ \nu_\ell$ decays. This can be exploited by carrying out a vertex fit using the lepton and all charged constituents, not identified as kaons, of the X system and we use its χ^2 value as a discriminator.
5. $Q_{\rm tot}$: The total event charge as calculated from the X system plus lepton on the signal and from the $B_{\rm tag}$ constituents. Due to the larger average multiplicity of $B \to X_c \ell^+ \nu_\ell$, the expected net zero event charge is more often violated in comparison to $B \to X_u \ell^+ \nu_\ell$ candidate events.

We use the BDT implementation of Ref. [64] and train a classifier $\mathcal{O}_{\rm BDT}$ with simulated $B \to X_u \ell^+ \nu_\ell$ and $B \to X_c \ell^+ \nu_\ell$ events, which we discard in the later analysis. Reference [64] uses optimized boosting and pruning procedures to maximize the classification performance. We choose a selection criteria on $\mathcal{O}_{\rm BDT}$ that rejects 98.7% of $B \to X_c \ell^+ \nu_\ell$ and retains 18.5% of $B \to X_u \ell^+ \nu_\ell$ signal. This working point was chosen by maximizing the significance of the most inclusive partial branching fraction, taking into account the full set of systematic uncertainties and the full analysis procedure. The stability of the result as a function of the BDT selection is further discussed in Sec. VIII.

Table II lists the efficiencies for signal and $B \to X_c \ell^+ \nu_\ell$ background for the $M_{\rm bc}$ and the BDT selections. Figure 4 shows the output classifier of the background suppression BDT for MC events and data. The classifier output shows good agreement between simulated and observed data, with the exception of the first two signal depleted bins. A comparison of the shape of all input variables for $B \to X_u \ell^+ \nu_\ell$ and $B \to X_c \ell^+ \nu_\ell$, and further MC events and data comparisons, can be found in Appendix B.

D. Tagging efficiency calibration

The reconstruction efficiency of the hadronic full reconstruction algorithm of Ref. [62] differs between simulated samples and the reconstructed data. This difference

TABLE II. The selection efficiencies for $B \to X_u \ell^+ \nu_\ell$ signal, $B \to X_c \ell^+ \nu_\ell$, and for data are listed after the reconstruction of the $B_{\rm tag}$ and lepton candidate. The nominal selection requirement on the BDT classifier $\mathcal{O}_{\rm BDT}$ is 0.85. The other two requirements were introduced to test the stability of the result, cf. Sec. VIII.

Selection	$B \to X_u \ell^+ \nu_\ell$ (%)	$B \to X_c \ell^+ \nu_\ell$ (%)	Data (%)
$M_{\rm bc} > 5.27\,{\rm GeV}$	84.8	83.8	80.2
$\mathcal{O}_{\rm BDT} > 0.85$	18.5	1.3	1.6
$\mathcal{O}_{\rm BDT} > 0.83$	21.9	1.7	2.1
$\mathcal{O}_{\rm BDT} > 0.87$	14.5	0.9	1.1

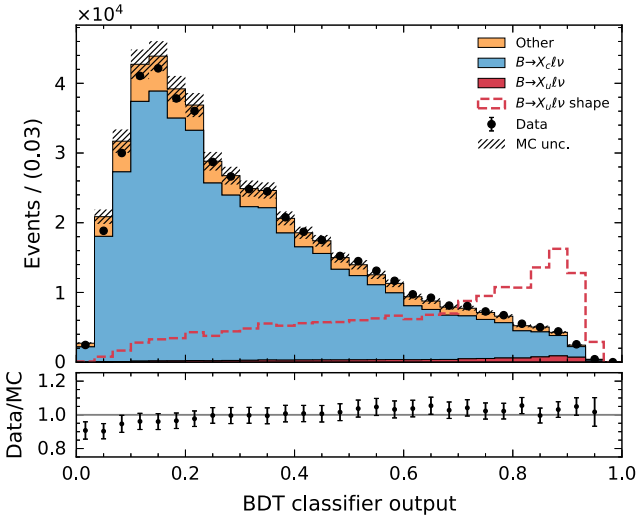


FIG. 4. The shape of the background suppression classifier $\mathcal{O}_{\rm BDT}$ is shown. MC sample is divided into $B \to X_u \ell^+ \nu_\ell$ signal, the dominant $B \to X_c \ell^+ \nu_\ell$ background, and all other contributions. To increase visibility, the $B \to X_u \ell^+ \nu_\ell$ component is shown with a scaling factor (red dashed line). The uncertainties on the MC events contain the full systematic errors and are further discussed in Sec. V.

mainly arises due to imperfections, e.g., in the simulation of detector responses, particle-identification (PID) efficiencies, or incorrect branching fractions in the reconstructed decay cascades. To address this, the reconstruction efficiency is calibrated using a data-driven approach and we follow closely the procedure outlined in Ref. [34]. We reconstruct full reconstruction events by requiring exactly one lepton on the signal side, and apply the same $B_{\rm tag}$ and lepton selection criteria outlined in the previous section. This $B \to X \ell^+ \nu_\ell$ enriched sample is divided into groups of subsamples according to the $B_{\rm tag}$ decay channel and the multivariate classifier output $\mathcal{O}_{\rm FR}$ used in the hierarchical reconstruction. Each of these groups of subsamples is studied individually to derive a calibration factor for the hadronic tagging efficiency: the calibration factor is obtained by comparing the number of inclusive semileptonic B meson decays, $N(B \to X \ell^+ \nu_\ell)$, in data with the expectation from the simulated samples, $N^{\rm MC}(B \to X \ell^+ \nu_\ell)$. The semileptonic yield is determined via a binned maximum likelihood fit using the lepton energy spectrum. To reduce the modeling dependence of the $B \to X \ell^+ \nu_\ell$ sample this is done in a coarse granularity of five bins. The calibration factor of each these groups of subsamples is given by

$$C_{\rm tag}(B_{\rm tag}{\rm mode}, \mathcal{O}_{\rm FR}) = \frac{N(B \to X \ell^+ \nu_\ell)}{N^{\rm MC}(B \to X \ell^+ \nu_\ell)}. \quad (19)$$

The free parameters in the fit are the yield of the semileptonic $B \to X \ell^+ \nu_\ell$ decays, the yield of backgrounds from fake leptons, and the yield of backgrounds from true leptons.

Approximately 1200 calibration factors are determined this way. The leading uncertainty on the C_{tag} factors is from the assumed $B \to X\ell^+\nu_\ell$ composition and the lepton PID performance, cf. Sec. V. We also apply corrections to the continuum efficiency. These are derived by using the off-resonance sample and comparing the number of reconstructed off-resonance events in data with the simulated on-resonance continuum events, correcting for differences in the selection.

IV. FITTING PROCEDURE

After the selection, we retain 9875 events. In order to determine the $B \to X_u\ell^+\nu_\ell$ signal yield and constrain all backgrounds, we perform a binned likelihood fit of these events in several discriminating variables. To reduce the dependence on the precise modeling of the $B \to X_u\ell^+\nu_\ell$ signal, we use coarse bins over regions that are very sensitive to the admixture of resonant and nonresonant decays, cf. Sec. II, and explore different variables for the signal extraction. The total likelihood function is constructed as the product of individual Poisson distributions $\mathcal{P}$,

$$\mathcal{L} = \prod_i^{\text{bins}} \mathcal{P}(n_i;\nu_i) \times \prod_k \mathcal{G}_k, \tag{20}$$

with n_i denoting the number of observed data events and ν_i the total number of expected events in a given bin i. Here, $\mathcal{G}_k$ are nuisance parameter (NP) constraints, whose role is to incorporate systematic uncertainties of a source k into the fit. Their construction is further discussed in Sec. V. The number of expected events in a given bin ν_i is estimated using simulated collision events and is given by

$$\nu_i = \sum_k^{\text{component}} f_{ik}\eta_k, \tag{21}$$

with η_k denoting the total number of events from a given fit component k, and f_{ik} denoting the fraction of such events being reconstructed in bin i as determined by the MC simulation. The three fit components we determine are as follows:

(a) Signal $B \to X_u\ell^+\nu_\ell$ events that fall inside a phase-space region for a partial branching fraction we wish to determine.

(b) Signal $B \to X_u\ell^+\nu_\ell$ events that fall outside said region if applicable. This component can have very similar shapes as other backgrounds. We thus constrain this component in all fits to its expectation using the world average of $\mathcal{B}(B \to X_u\ell^+\nu_\ell) = (2.13 \pm 0.30) \times 10^{-3}$ [39]. We also investigated different approaches: for instance linking this component with the component of (a). This leads to small shifts of $\mathcal{O}(0.3-1\%)$ of the reported partial branching fractions using this component.

(c) Background events; such are dominated by $B \to X_c\ell^+\nu_\ell$ and other decays that produce leptons in the final state (e.g., from $B \to h_1h_2$ and $h_2 \to h_3\ell^-\nu$ with h_1, h_2, and h_3 denoting hadronic final states). Other contributions are from misidentified lepton candidates and a small amount of continuum processes. A full description of all background processes is given in Sec. III.

We carry out five separate fits to measure three partial branching fractions, using different discriminating variables to determine the $B \to X_u\ell^+\nu_\ell$ yield. The fits and variables are as follows:

1. The hadronic mass, M_X: Signal is expected to predominantly populate the low hadronic mass region, whereas remaining $B \to X_c\ell^+\nu_\ell$ background will produce a sharp peak at around $M_X \approx 2$ GeV. The sizeable resolution on the reconstruction of the X system will result in a non-negligible amount of these backgrounds to also be present in the low- and high-M_X region. The determined signal yields are used to measure the partial branching fraction of $M_X < 1.7$ GeV and $E_\ell^B > 1$ GeV. We thus use two signal templates and split events according to generator-level $M_X < 1.7$ and $M_X > 1.7$ GeV.
2. The four-momentum-transfer squared, q^2: Signal will on average have a higher q^2 than $B \to X_c\ell^+\nu_\ell$ background, whose kinematic end point is $q^2 = (m_B - m_D)^2 \approx 11.6$ GeV^2. However, the reconstructed q^2 of $B \to X_c\ell^+\nu_\ell$ events is smeared over the entire kinematic range due to the sizeable resolution in the reconstruction of the inclusive X system and the B_{tag} reconstruction. To reduce background from $B \to X_c\ell^+\nu_\ell$ events, we apply a cut on the reconstructed M_X and require a value smaller than 1.7 GeV. The determined signal yields are used to measure the partial branching fraction of $M_X < 1.7$ GeV, $q^2 > 8$ GeV^2, and $E_\ell^B > 1$ GeV. We use two signal templates: template (a) is defined as signal events with generator-level values of $M_X < 1.7$ GeV and $q^2 > 8$ GeV^2 and template (b) contains all other signal events.
3. The lepton energy in the B meson rest frame, E_ℓ^B: Signal and $B \to X_c\ell^+\nu_\ell$ can be separated beyond the kinematic end point of the $B \to X_c\ell^+\nu_\ell$ background, which is $\frac{1}{2m_B}(m_B^2 - m_D^2 + m_\ell^2) \approx 2.3$ GeV. The lepton energy is reconstructed using its momentum ($E_\ell^B = |\mathbf{p}_\ell^B|$), which has excellent resolution. This makes the measurement more sensitive to the exact composition of the $B \to X_c\ell^+\nu_\ell$ background and $B \to X_u\ell^+\nu_\ell$ signal. To minimize the dependence on the signal modeling, the end point of the lepton

spectrum, ranging from $E_\ell^B \in [2.5, 2.7]$ GeV, is treated as a single coarse bin in the fit. To reduce the dependence on the exact modeling of $B \to X_c \ell^+ \nu_\ell$ we require $M_X < 1.7$ GeV. The determined signal yields are used to measure the partial branching fraction with $M_X < 1.7$ GeV and the signal templates are split accordingly into a matching generator-level template and all other signal events.

4. The next fit also analyzes E_ℓ^B, but uses the determined signal yields to measure the partial branching fraction with $E_\ell^B > 1$ GeV. Thus no separation of signal events in different categories is used.
5. The final fit uses M_X and q^2 simultaneously in a two-dimensional fit. This fit also measure the partial branching fraction with $E_\ell^B > 1$ GeV and no separation into different categories of signal events is used.

A summary of the binning choices of the kinematic variables is provided in Table III. The likelihood (20) is numerically maximized to fit the value of the different components η_k from the observed events and by using the sequential least squares programming method implementation of Ref. [65]. Confidence intervals (CI) are constructed using the profile likelihood ratio method. For a given component η_k the ratio is

$$\Lambda(\eta_k) = -2 \ln \frac{\mathcal{L}(\eta_k, \hat{\hat{\boldsymbol{\eta}}}_{\eta_k}, \hat{\hat{\boldsymbol{\theta}}}_{\eta_k})}{\mathcal{L}(\hat{\eta}_k, \hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\theta}})}, \tag{22}$$

where $\hat{\eta}_k, \hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\theta}}$ are the values of the component of interest, the remaining components, and a vector of NPs, respectively, that maximize the likelihood function, whereas the remaining components $\hat{\hat{\boldsymbol{\eta}}}_{\eta_k}$ and nuisance parameters $\hat{\hat{\boldsymbol{\theta}}}_{\eta_k}$ maximize the likelihood for the specific value η_k. In the asymptotic limit, the test statistic (22) can be used to construct approximate confidence intervals through

$$1 - \text{C.L.} = \int_{\Lambda(\eta_k)}^{\infty} f_{\chi^2}(x; 1\,\text{d.o.f.})\mathrm{d}x, \tag{23}$$

with $f_{\chi^2}(x; 1\,\text{d.o.f.})$ denoting the χ^2 distribution of the variable x with a single degree of freedom. Further, C.L. denotes the desired confidence level. The determined signal yields $\hat{\eta}_k = \hat{\eta}_{\text{sig}}$ are translated into partial branching fractions via

$$\Delta\mathcal{B}(B \to X_u \ell^+ \nu_\ell; \text{Reg}) = \frac{\hat{\eta}_{\text{sig}} \cdot \epsilon_{\Delta\mathcal{B}(\text{Reg})}}{4(\epsilon_{\text{tag}} \cdot \epsilon_{\text{sel}}) \cdot N_{BB}}. \tag{24}$$

Here ϵ_{tag} denotes the tagging efficiency, as determined after applying the calibration factor introduced in Sec. III D. Further, ϵ_{sel} and $\epsilon_{\Delta\mathcal{B}(\text{Reg})}$ denote the signal side selection efficiency and a correction to the efficiency to account for the fraction of $B \to X_u \ell^+ \nu_\ell$ phase-space region that is measured. The factor of 4 in the denominator is due to the factor $N_{BB} = (771.58 \pm 9.78) \times 10^6$ B meson pairs and our averaging over electron and muon final states.

To validate the fit procedure we generated ensembles of pseudoexperiments for different input branching fractions for $B \to X_u \ell^+ \nu_\ell$ signal and $B \to X_c \ell^+ \nu_\ell$ background. Fits to these ensembles show no biases in central values and no under- or overcoverage of CI. Using the current world average of $\mathcal{B}(B \to X_u \ell^+ \nu_\ell) = (2.13 \pm 0.30) \times 10^{-3}$, we expect approximately between 930 and 2070 $B \to X_u \ell^+ \nu_\ell$ signal events with significances $s = \hat{\eta}_{\text{sig}}/\epsilon$ ranging from about 9 to 15 standard deviations, depending on the signal region under study, and with ϵ being the expected fit error determined from Asimov datasets [66].

TABLE III. The binning choices of the five fit scenarios are given.

Fit variable	Bins
M_X	[0, 1.5, 1.9, 2.5, 3.1, 5.0] GeV
q^2	[0, 2, 4, 6, 8, 10, 12, 14, 26] GeV^2
E_ℓ^B	15 equidist. bins in [1, 2.5] & [2.5, 2.7] GeV
$M_X : q^2$	[0, 1.5] GeV × [0, 2, 4, 6, 8, 10, 12, 14, 26] GeV^2
	[1.5, 1.9] GeV × [0, 2, 4, 6, 26] GeV^2
	[1.9, 2.5] GeV × [0, 2, 4, 26] GeV^2
	[2.5, 4.0] GeV × [0, 2, 26] GeV^2

V. SYSTEMATIC UNCERTAINTIES

Several systematic uncertainties affect the determination of the reported partial branching fractions. The most important uncertainties arise from the modeling of the $B \to X_u \ell^+ \nu_\ell$ signal component and from the tagging calibration correction. This is followed by uncertainties on particle identification of kaons and leptons, the uncertainty on the number of B meson pairs, the statistical uncertainty on the used MC samples, and uncertainties related to the efficiency of the track reconstructions. Table IV summarizes the systematic uncertainties for the five measured partial branching fractions probing three phase-space regions. The table separates uncertainties that originate from the background subtraction ("additive uncertainties") and uncertainties related to the translation of the fitted signal yields into partial branching fractions ("multiplicative uncertainties").

The tagging calibration uncertainties are evaluated by producing different sets of calibration factors. These sets take into account the correlation structure from common systematic uncertainties (cf. Sec. III D) and that individual channels and ranges of the output classifier are statistically independent. When applying the different sets of calibration factors, we notice only negligible shape changes on the signal and background template shapes, but the overall tagging efficiency is affected. The associated uncertainty on the calibration factors is found to be 3.6% and is identical

TABLE IV. The relative uncertainty on the extracted $B \to X_u \ell^+ \nu_\ell$ partial branching fractions are shown. For definitions of additive and multiplicative errors, see text.

	Relative uncertainties (%)				
Phase-space region	$M_X < 1.7$ GeV, $E_\ell^B > 1$ GeV	$M_X < 1.7$ GeV, $E_\ell^B > 1$ GeV	$M_X < 1.7$ GeV, $q^2 > 8$ GeV^2, $E_\ell^B > 1$ GeV	$E_\ell^B > 1$ GeV	$E_\ell^B > 1$ GeV
Fit variable(s)	(M_X fit)	(E_ℓ^B fit)	(q^2 fit)	(E_ℓ^B fit)	(M_X:q^2 fit)
Additive uncertainties					
$B \to X_u \ell^+ \nu_\ell$ modeling					
$B \to \pi \ell^+ \nu_\ell$ FFs	0.1	0.7	1.4	0.6	0.4
$B \to \rho \ell^+ \nu_\ell$ FFs	0.2	1.9	4.3	1.9	0.7
$B \to \omega \ell^+ \nu_\ell$ FFs	0.5	3.2	5.2	3.1	0.8
$B \to \eta \ell^+ \nu_\ell$ FFs	0.1	1.6	3.0	1.6	0.3
$B \to \eta' \ell^+ \nu_\ell$ FFs	0.1	1.6	3.0	1.6	1.6
$\mathcal{B}(B \to \pi \ell^+ \nu_\ell)$	0.2	0.1	0.2	0.1	0.2
$\mathcal{B}(B \to \rho \ell^+ \nu_\ell)$	0.3	0.7	0.8	0.5	0.4
$\mathcal{B}(B \to \omega \ell^+ \nu_\ell)$	<0.1	0.1	0.8	0.1	0.1
$\mathcal{B}(B \to \eta \ell^+ \nu_\ell)$	<0.1	0.1	<0.1	0.1	<0.1
$\mathcal{B}(B \to \eta' \ell^+ \nu_\ell)$	<0.1	<0.1	0.1	0.1	<0.1
$\mathcal{B}(B \to X_u \ell^+ \nu)$	0.7	2.0	2.1	2.1	2.1
DFN parameters	2.3	3.5	1.1	3.5	5.0
Hybrid model	2.7	8.7	4.6	8.7	3.1
$B \to X_c \ell^+ \nu_\ell$ modeling					
$B \to D \ell^+ \nu_\ell$ FFs	0.1	0.1	0.9	0.1	<0.1
$B \to D^* \ell^+ \nu_\ell$ FFs	1.4	1.2	3.0	1.3	1.1
$B \to D^{**} \ell^+ \nu_\ell$ FFs	0.4	0.5	0.3	0.5	0.4
$\mathcal{B}(B \to D \ell^+ \nu_\ell)$	0.1	<0.1	0.2	<0.1	0.2
$\mathcal{B}(B \to D^* \ell^+ \nu_\ell)$	<0.1	<0.1	0.3	<0.1	0.2
$\mathcal{B}(B \to D^{**} \ell^+ \nu_\ell)$	0.6	0.1	0.3	0.1	0.5
Gap modeling	1.1	0.1	0.3	0.1	1.0
MC statistics	1.3	1.6	3.8	1.7	1.6
Tracking efficiency	0.3	⋯	0.8	⋯	0.4
$\mathcal{L}_{\ell\,\mathrm{ID}}$ shape	1.0	0.5	1.3	0.6	1.2
$\mathcal{L}_{\mathrm{K}/\pi\mathrm{ID}}$ shape	1.2	-	1.3	-	1.0
$D \to X \ell \nu_\ell$	0.1	0.1	0.1	0.1	0.1
π_s efficiency	<0.1	⋯	0.1	⋯	0.1
Multiplicative uncertainties					
$B \to X_u \ell^+ \nu_\ell$ modeling					
$B \to \pi \ell^+ \nu_\ell$ FFs	0.2	0.2	1.9	0.2	0.2
$B \to \rho \ell^+ \nu_\ell$ FFs	0.7	0.8	3.7	0.8	0.6
$B \to \omega \ell^+ \nu_\ell$ FFs	1.3	1.6	6.1	1.6	1.1
$B \to \eta \ell^+ \nu_\ell$ FFs	0.3	0.3	1.8	0.3	0.2
$B \to \eta' \ell^+ \nu_\ell$ FFs	0.2	0.3	1.8	0.3	0.2
$\mathcal{B}(B \to \pi \ell^+ \nu_\ell)$	0.3	0.4	0.4	0.4	0.3
$\mathcal{B}(B \to \rho \ell^+ \nu_\ell)$	0.4	0.6	0.6	0.6	0.4
$\mathcal{B}(B \to \omega \ell^+ \nu_\ell)$	<0.1	<0.1	0.1	<0.1	<0.1
$\mathcal{B}(B \to \eta \ell^+ \nu_\ell)$	0.1	0.1	<0.1	0.1	<0.1
$\mathcal{B}(B \to \eta' \ell^+ \nu_\ell)$	0.1	0.1	0.1	0.1	0.1
$\mathcal{B}(B \to X_u \ell^+ \nu)$	3.0	3.2	2.9	4.8	3.8
DFN parameters	2.5	2.5	2.7	6.8	3.6
Hybrid model	0.2	0.8	1.4	4.7	2.8
π^+ multiplicity	1.7	2.5	2.3	3.1	1.7
γ_s ($s\bar{s}$ fragmentation)	0.5	0.8	1.1	1.1	0.8
$\mathcal{L}_{\ell\mathrm{ID}}$ efficiency	1.5	1.6	1.6	1.6	1.5
$\mathcal{L}_{\mathrm{K}/\pi\mathrm{ID}}$ efficiency	0.7	0.6	0.6	0.6	0.7
$N_{B\bar{B}}$	1.3	1.3	1.3	1.3	1.3
Tracking efficiency	0.8	0.3	0.8	0.3	0.9
Tagging calibration	3.6	3.6	3.6	3.6	3.6
Total systematic uncertainty	7.8	12.6	14.6	15.4	10.4

for the five measured partial branching fractions. The $B \to X_u \ell^+ \nu_\ell$ and $B \to X_c \ell^+ \nu_\ell$ modeling uncertainties do directly affect the shapes of M_X, q^2, and E_ℓ^B signal and background distributions. Further, the $B \to X_u \ell^+ \nu_\ell$ modeling affects the overall reconstruction efficiencies and migrations of events inside and outside of the phase-space regions we measure. We evaluate the uncertainties on the composition of the hybrid $B \to X_u \ell^+ \nu_\ell$ MC sample by variations of the $B \to \pi \ell^+ \nu_\ell$, $B \to \rho \ell^+ \nu_\ell$, $B \to \omega \ell^+ \nu_\ell$, $B \to \eta \ell^+ \nu_\ell$, and $B \to \eta' \ell^+ \nu_\ell$ branching fractions and form factors (FFs). The uncertainty on nonresonant $B \to X_u \ell^+ \nu_\ell$ contributions in the hybrid model is estimated by changing the underlying model from that of DFN [54] to that of BLNP [17]. In addition, the uncertainty on the used DFN parameters m_b^{1S} and a (cf. Sec. II) are incorporated. For each of these variations, new hybrid weights are calculated to propagate the uncertainties into shapes and efficiencies. We estimate the uncertainties of X_u fragmentation into $s\bar{s}$ quark pairs by variations of the corresponding JETSET parameter γ_s (cf. Ref. [57]). As our BDT is trained to reject final states with kaon candidates, a change in this fraction will directly impact the signal efficiency. The $s\bar{s}$ production probability has been measured by Refs. [67,68] at center-of-mass energies of 12 and 36 GeV with values of $\gamma_s = 0.35 \pm 0.05$ and $\gamma_s = 0.27 \pm 0.06$, respectively. We adopt the value and error of $\gamma_s = 0.30 \pm 0.09$, which spans the range of both measurements including their uncertainties. The X_u system of the nonresonant signal component is hadronized by JETSET into final states with two or more pions. We test the impact on the signal efficiency by changing the postfit charged pion multiplicity of nonresonant $B \to X_u \ell^+ \nu_\ell$ to the distribution observed in data in the signal-enriched region of $M_X < 1.7$ GeV (cf. Sec. VIII E and Appendix C). The $B \to X_c \ell^+ \nu_\ell$ background after the BDT selection is dominated by $B \to D \ell^+ \nu_\ell$ and $B \to D^* \ell^+ \nu_\ell$ decays. We evaluate the uncertainties on the modeling of $B \to D \ell^+ \nu_\ell$, $B \to D^* \ell^+ \nu_\ell$, and $B \to D^{**} \ell^+ \nu_\ell$ by variations of the BGL parameters and heavy-quark form factors within their uncertainties. In addition, we propagate the branching fraction uncertainties. The uncertainties on the $B \to X_c \ell^+ \nu_\ell$ gap branching fractions are taken to be large enough to account for the difference between the sum of all exclusive branching fractions measured and the inclusive branching fraction measured. We also evaluate the impact on the efficiency of the lepton- and hadron-identification uncertainties and the overall tracking efficiency uncertainty. The statistical uncertainty on all generated MC samples is also evaluated and propagated into the systematic errors.

We incorporate the effect of additive systematic uncertainties directly into the likelihood function. This can be done by introducing a vector of NPs $\boldsymbol{\theta}_k$ for each fit template of a process k (e.g., signal or background). Each element of this vector represents one bin of the fitted observables of interest (e.g., M_X,q^2, E_ℓ^B, or a 2D bin of $M_X : q^2$). These NPs are constrained parameters in the likelihood (20) using multivariate Gaussian distributions, $\mathcal{G}_k = \mathcal{G}_k(\mathbf{0}; \boldsymbol{\theta}_k, \Sigma_k)$. Here Σ_k denotes the systematic covariance matrix for a given template k and $\boldsymbol{\theta}_k$ is a vector of NPs. The covariance Σ_k is the sum over all possible uncertainty sources for a given template k,

$$\Sigma_k = \sum_s^{\text{error sources}} \Sigma_{ks}, \tag{25}$$

with Σ_{ks} denoting the covariance matrix of error source s. The covariance matrices Σ_{ks} depend on uncertainty vectors $\boldsymbol{\sigma_{ks}}$, which represent the absolute error in bins of the fit variable of template k. Uncertainties from the same error source are either fully correlated, or for the case of MC events or other statistical uncertainties, are treated as uncorrelated. Both cases can be expressed as $\Sigma_{ks} = \boldsymbol{\sigma_{ks}} \otimes \boldsymbol{\sigma_{ks}}$ or $\Sigma_{ks} = \mathrm{Diag}(\boldsymbol{\sigma_{ks}}^2)$, respectively. For particle-identification uncertainties, we estimate Σ_{ks} using sets of correction tables, sampled according to their statistical and systematic uncertainties. The systematic NPs are incorporated in Eq. (21) by rewriting the fractions f_{ik} for all templates as

$$f_{ik} = \frac{\eta_{ik}^{\mathrm{MC}}}{\sum_j \eta_{jk}^{\mathrm{MC}}} \to \frac{\eta_{ik}^{\mathrm{MC}}(1+\theta_{ik})}{\sum_j \eta_{jk}^{\mathrm{MC}}(1+\theta_{jk})}, \tag{26}$$

to take into account changes in the signal or background shape. Here η_{ik}^{MC} denotes the predicted number of MC events of a given bin i and a process k, and θ_{ik} is the associated nuisance parameter constrained by $\mathcal{G}_k$.

VI. $B \to X_c \ell \bar{\nu}_\ell$ CONTROL REGION

Figure 5 compares the reconstructed M_X, q^2, and E_ℓ^B distributions with the expectation from MC before applying the background suppression BDT. All corrections are applied and the MC uncertainty contains all systematic uncertainties discussed in Sec. V. The agreement of M_X and q^2 is excellent, but some differences in the shape of the lepton momentum spectrum are seen. This is likely due to imperfections of the modeling of the inclusive $B \to X_c \ell^+ \nu_\ell$ background. The discrepancy reduces in the $M_X < 1.7$ GeV region. The main results of this paper will be produced by fitting q^2 and M_X in two dimensions. We use the lepton spectrum to measure the same regions of phase space, to validate the obtained results.

VII. $B \to X_u \ell^+ \nu_\ell$ SIGNAL REGION

Figure 6 shows the reconstructed M_X, q^2, and E_ℓ^B distributions after the BDT selection is applied. The $B \to X_u \ell^+ \nu_\ell$ contribution is now clearly visible at low M_X and high E_ℓ^B, while the reconstructed events and the MC expectation show good agreement. The $B \to X_c \ell^+ \nu_\ell$ background is dominated by contributions from $B \to D \ell^+ \nu_\ell$ and $B \to D^* \ell^+ \nu_\ell$ decays, and the remaining background is

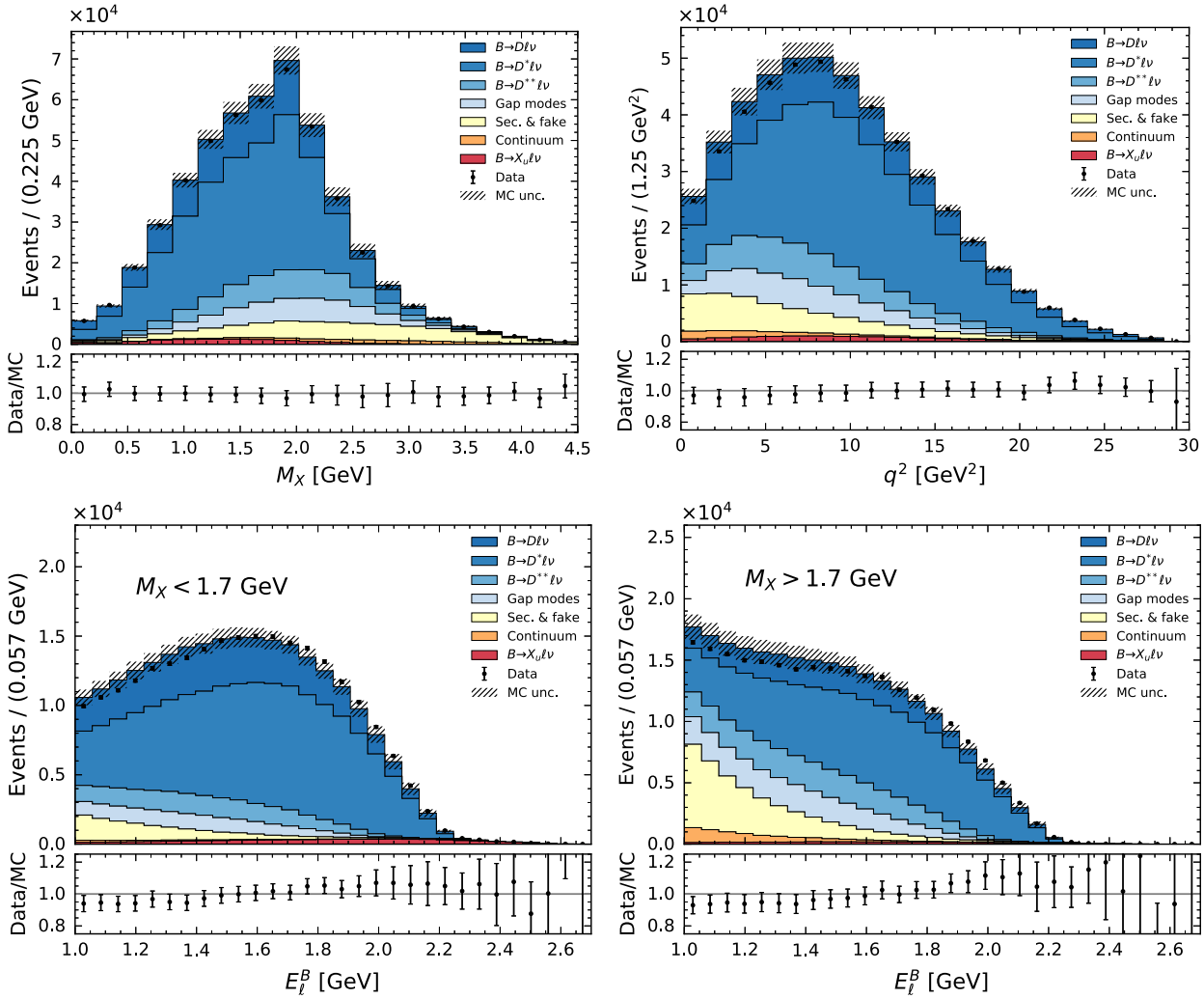


FIG. 5. Top: the M_X and q^2 spectra of the selected candidates prior to applying the background BDT are shown. Bottom: the E_ℓ^B spectrum of the selected candidates prior to applying the background BDT are shown for events with $M_X < 1.7$ and $M_X > 1.7$ GeV.

predominantly from secondary leptons and misidentified lepton candidates.

VIII. RESULTS

We report partial branching fractions for three phase-space regions from five fits to the reconstructed variables introduced in Sec. IV. All partial branching fractions correspond to a selection with $E_\ell^B > 1$ GeV, also reverting the effect of final-state radiation photons and possible additional phase-space restrictions. The resulting fit yields are listed in Table V.

A. Partial branching fraction results

For the partial branching fraction with $M_X < 1.7$ GeV from the fit to M_X we find

$$\Delta\mathcal{B}(B \to X_u \ell^+ \nu_\ell) = (1.09 \pm 0.05 \pm 0.08) \times 10^{-3}, \quad (27)$$

with the first and second error denoting the statistical and systematic uncertainty, respectively. The resulting postfit distribution is shown in the top panel of Fig. 7. With this selection about 56% of the available $B \to X_u \ell^+ \nu_\ell$ phase space is probed. The partial branching fraction is in good agreement with the value obtained by fitting E_ℓ^B and corrected to the same phase space. The fit is shown in Fig. 8 and we measure

$$\Delta\mathcal{B}(B \to X_u \ell^+ \nu_\ell) = (1.11 \pm 0.06 \pm 0.14) \times 10^{-3}, \quad (28)$$

with a larger systematic and statistical uncertainty than Eq. (27). To further probe the $B \to X_u \ell^+ \nu_\ell$ -enriched region, we carry out a measurement for $M_X < 1.7$ GeV and $q^2 > 8$ GeV^2 from a fit to the q^2 spectrum. This selection only probes about 31% of the available $B \to X_u \ell^+ \nu_\ell$ phase space. We find

$$\Delta\mathcal{B}(B \to X_u \ell^+ \nu_\ell) = (0.67 \pm 0.07 \pm 0.10) \times 10^{-3}. \quad (29)$$

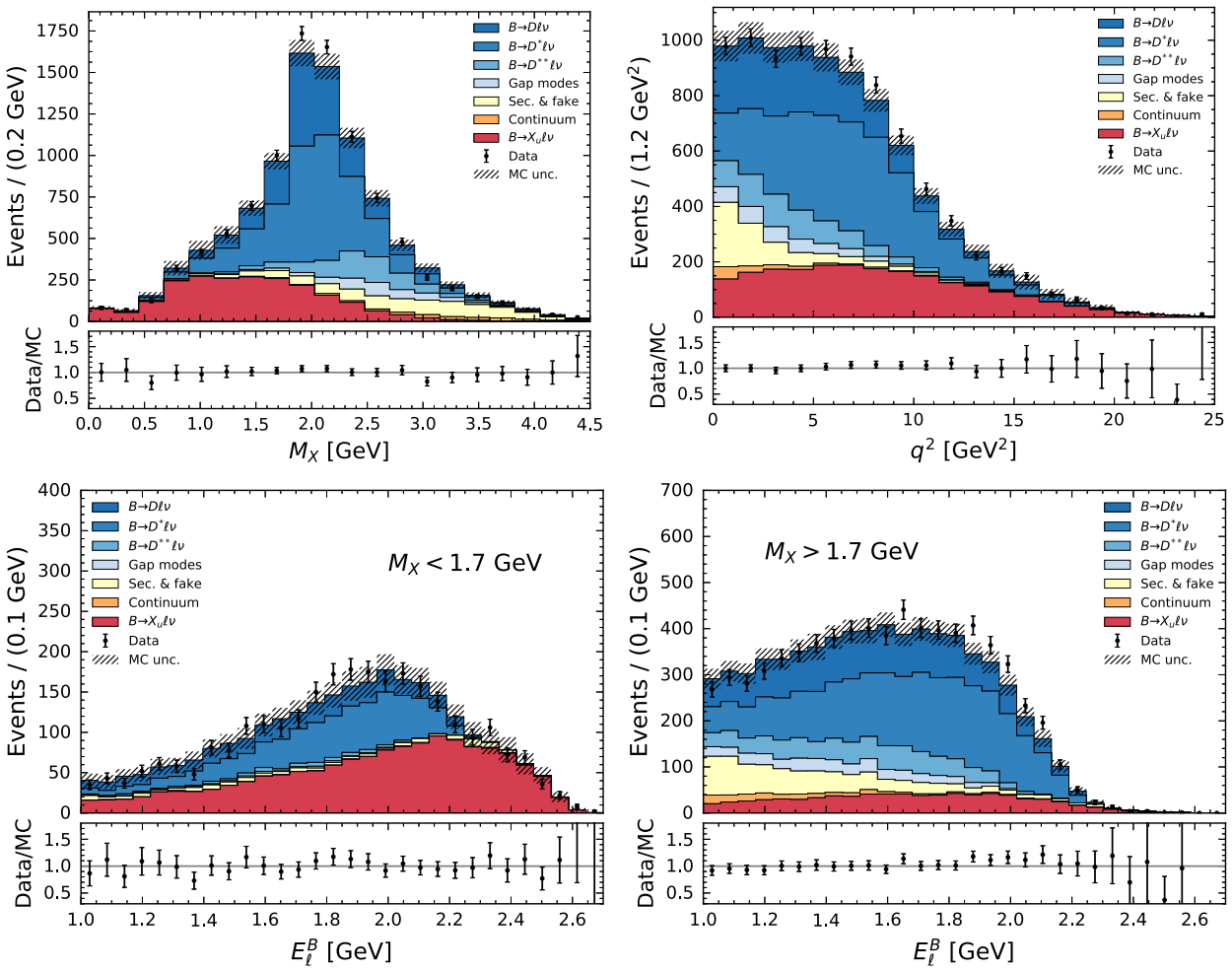
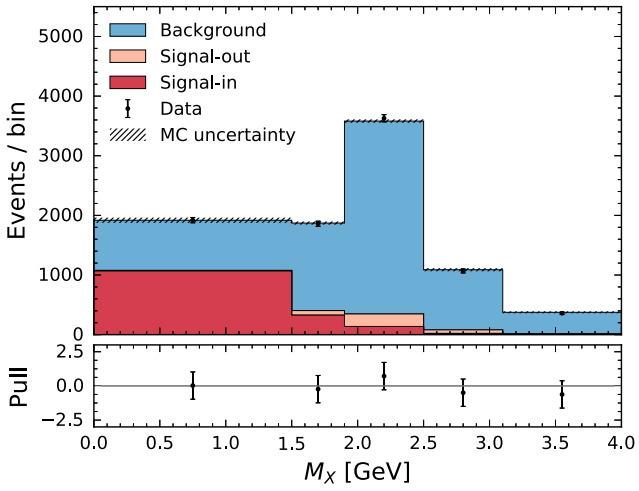


FIG. 6. The M_X, q^2, and E_ℓ^B spectra after applying the background BDT but before the fit are shown. The $B \to X_u \ell^+ \nu_\ell$ contribution is shown in red and scaled to the world average of $\mathcal{B}(B \to X_u \ell^+ \nu_\ell) = (2.13 \pm 0.30) \times 10^{-3}$. The data and MC agreement is reasonable in all variables. The E_ℓ^B spectra is shown with selections of $M_X < 1.7$ and $M_X > 1.7$ GeV. The cut of $M_X < 1.7$ GeV is later used in the fit to reduce the dependence on the $B \to X_c \ell^+ \nu_\ell$ modeling of higher charmed states.

The corresponding postfit distribution of q^2 is shown in the bottom panel of Fig. 7. The most precise determinations of $B \to X_u \ell^+ \nu_\ell$ are obtained from a two-dimensional fit, exploiting the full combined discriminatory power of M_X and q^2. The resulting partial branching fraction probes about 86% of the available $B \to X_u \ell^+ \nu_\ell$ phase space. We measure

TABLE V. The fitted signal yields in ($\hat{\eta}_{\rm sig}$) and outside ($\hat{\eta}_{\rm sig-out}$) the measured phase-space regions, the background yields ($\hat{\eta}_{\rm bkg}$), and the product of tagging and selection efficiency are listed.

Phase-space region	Additional selection (GeV)	Fit variable(s)	$\hat{\eta}_{\rm sig}$	$\hat{\eta}_{\rm sig-out}$	$\hat{\eta}_{\rm bkg}$	$n_{\rm data}$	$10^3(\epsilon_{\rm tag} \cdot \epsilon_{\rm sel})$
$M_X < 1.7$ GeV, $E_\ell^B > 1$ GeV	⋯	M_X fit	$1558 \pm 69 \pm 71$	364 ± 51	6912 ± 138	8833 ± 94	0.26 ± 0.07
$M_X < 1.7$ GeV, $E_\ell^B > 1$ GeV	$M_X < 1.7$	E_ℓ^B fit	$1285 \pm 68 \pm 139$	22 ± 3	1362 ± 155	2669 ± 52	0.21 ± 0.07
$M_X < 1.7$ GeV, $q^2 > 8$ GeV^2, $E_\ell^B > 1$ GeV	$M_X < 1.7$	q^2 fit	$938 \pm 99 \pm 100$	474 ± 57	1253 ± 192	2665 ± 52	0.14 ± 0.07
$E_\ell^B > 1$ GeV	$M_X < 1.7$	E_ℓ^B fit	$1303 \pm 69 \pm 138$	⋯	1366 ± 154	2669 ± 52	0.21 ± 0.19
$E_\ell^B > 1$ GeV		M_X:q^2 fit	$1801 \pm 81 \pm 123$	⋯	7031 ± 164	8833 ± 94	0.31 ± 0.12



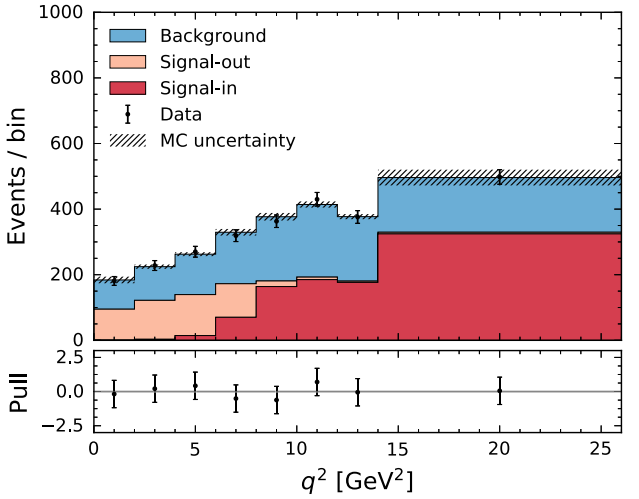


FIG. 7. The postfit distributions of the one-dimensional fits to M_X and q^2 are shown, corresponding to the measured partial branching fractions for $E_\ell^B > 1$ GeV with additional requirements of $M_X < 1.7$ and $M_X < 1.7$ GeV, and $q^2 > 8$ GeV^2, respectively.

$$\Delta\mathcal{B}(B \to X_u \ell^+ \nu_\ell) = (1.59 \pm 0.07 \pm 0.16) \times 10^{-3}. \quad (30)$$

The projection of the 2D fit onto M_X and the q^2 distribution for the signal-enriched region of $M_X < 1.5$ GeV are shown in Fig. 9. The remaining q^2 distributions are given in Appendix D. The partial branching fraction is also in good agreement from the measurement obtained by fitting E_ℓ^B, covering the same phase space (cf. Fig. 8),

$$\Delta\mathcal{B}(B \to X_u \ell^+ \nu_\ell) = (1.69 \pm 0.09 \pm 0.26) \times 10^{-3}. \quad (31)$$

The uncertainties are larger, but both results are compatible. The nuisance parameter pulls of all fits are provided in Appendix D. The result of Eq. (30) can be further compared with the most precise measurement to date of this region of Ref. [69], where $\Delta\mathcal{B}(B \to X_u \ell \nu_\ell) = (1.55 \pm 0.12) \times 10^{-3}$, and shows good agreement. The measurement can also be compared to Ref. [15] using a similar experimental approach.

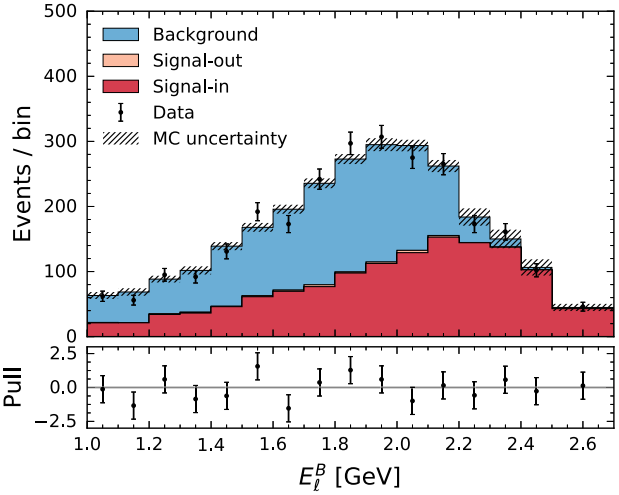


FIG. 8. The postfit distributions of the fit to E_ℓ^B with $M_X < 1.7$ GeV is shown. The resulting yields were corrected to correspond to the partial branching fraction with $E_\ell^B > 1$ GeV with and without an additional requirement of $M_X < 1.7$ GeV, respectively.

The measured partial branching fraction of $E_\ell^B > 1$ GeV is $\Delta\mathcal{B}(B \to X_u \ell \nu_\ell) = (1.82 \pm 0.19) \times 10^{-3}$, which is compatible with Eq. (30) within 0.9 standard deviations. Belle previously reported in Ref. [16] using also a similar approach for the same phase space a higher value of $\Delta\mathcal{B}(B \to X_u \ell \nu_\ell) = (1.96 \pm 0.19) \times 10^{-3}$. We cannot quantify the statistical overlap between both results, but by comparing the number of determined signal events one can estimate it to be below 55%. The dominant systematic uncertainties of Ref. [16] were evaluated using different approaches, but fully correlating the dominant systematic uncertainties and assuming a statistical correlation of 55% we obtain a compatibility of 1.7 standard deviations. The main difference of this analysis with Ref. [16] lies in the modeling of signal and background processes: since its publication our understanding improved and more precise measurements of branching fractions and form factors were made available. Further, for the $B \to X_u \ell^+ \nu_\ell$ signal process in this paper a hybrid approach was adopted (see Sec. II and Appendix A), whereas Ref. [16] used an alternative approach to model signal as a mix of inclusive and exclusive decay modes. Note that this work supersedes Ref. [16].

B. $|V_{ub}|$ determination

We determine $|V_{ub}|$ from the measured partial branching fractions using a range of theoretical rate predictions. In principle, the total $B \to X_u \ell^+ \nu_\ell$ decay rate can be calculated using the same approach as $B \to X_c \ell^+ \nu_\ell$ using the heavy-quark expansion in inverse powers of m_b. Unfortunately, the measurement requirements necessary to separate $B \to X_u \ell^+ \nu_\ell$ from the dominant $B \to X_c \ell^+ \nu_\ell$ background spoil the convergence of this approach. In the predictions for the partial rates corresponding to our measurements, perturbative and nonperturbative uncertainties are largely enhanced

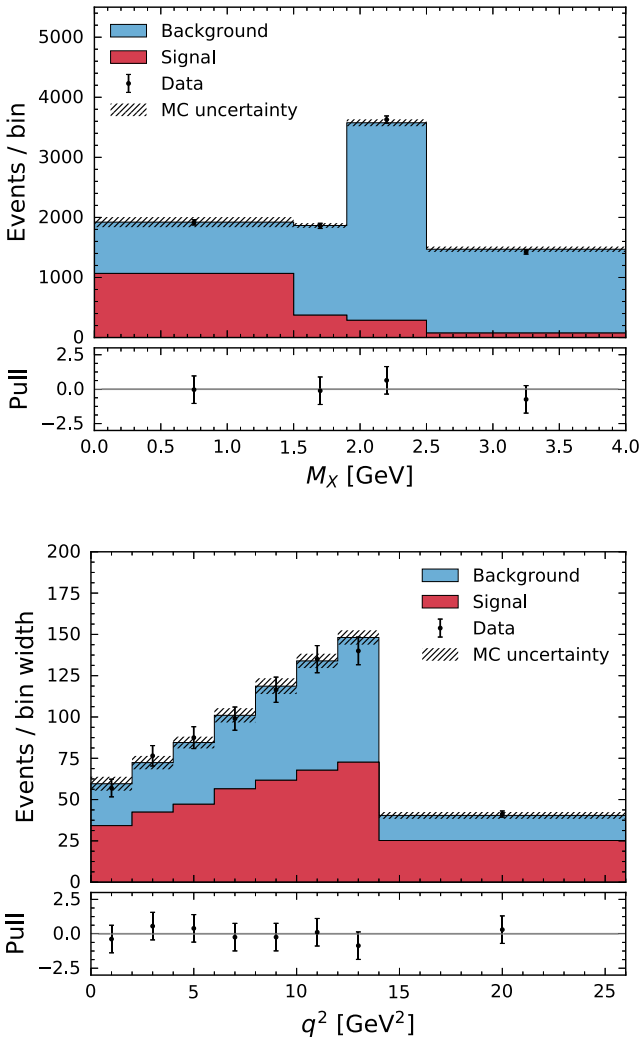


FIG. 9. The postfit projection of M_X of the two-dimensional fit to M_X:q^2 on M_X and the q^2 distribution in the range of $M_X \in [0, 1.5]$ GeV are shown. The resulting yields are corrected to correspond to a partial branching fraction with $E_\ell^B > 1$ GeV. The remaining q^2 distributions are given in Fig. 22 (Appendix D).

and, as outlined in the Introduction, the predictions are sensitive to the shape-function modeling.

The relationship between measured partial branching fractions, predictions of the rate (omitting CKM factors) $\Delta\Gamma(B \to X_u \ell^+ \nu_\ell)$, and $|V_{ub}|$ is

$$|V_{ub}| = \sqrt{\frac{\Delta\mathcal{B}(B \to X_u \ell^+ \nu_\ell)}{\tau_B \cdot \Delta\Gamma(B \to X_u \ell^+ \nu_\ell)}}, \tag{32}$$

with $\tau_B = (1.579 \pm 0.004)$ ps denoting the average of the charged and neutral B meson lifetime [39]. We use four predictions for the theoretical partial rates. All predictions use the same input values as Ref. [6] chooses for their world averages. The four predictions are

(i) BLNP: The prediction of Bosch, Lange, Neubert, and Paz of Ref. [17,70,71] provides a prediction at next-to-leading-order accuracy in terms of the strong coupling constant α_s and incorporates all known corrections. Predictions are interpolated between the shape-function dominated region (end point of the lepton spectrum, small hadronic mass) to the region of phase space, that can be described via the operator product expansion. As input we use $m_b^{\mathrm{SF}} = 4.58 \pm 0.03$ GeV and $\mu_\pi^{2\,\mathrm{SF}} = 0.20^{+0.09}_{-0.10}$ GeV^2.

(ii) DGE: The dressed gluon approximation from Andersen and Gardi [19,20] makes predictions by avoiding the direct use of shape functions, but produces predictions for hadronic observables using the on shell b-quark mass. The calculation is carried out in the $\overline{\mathrm{MS}}$ scheme and we use $m_b(\overline{\mathrm{MS}}) = 4.19 \pm 0.04$ GeV.

(iii) GGOU: The prediction from Gambino, Giordano, Ossola, and Uraltsev [18] (GGOU) incorporates all known perturbative and nonperturbative effects up to the order $\mathcal{O}(\alpha_s^2\beta_0)$ and $\mathcal{O}(1/m_b^3)$, respectively. The shape-function dependence is incorporated by parametrizing its effects in each structure function with a single light-cone function. The calculation is carried out in the kinetic scheme and we use as inputs $m_b^{\mathrm{kin}} = 4.55 \pm 0.02$ GeV and $\mu_\pi^{2\,\mathrm{kin}} = 0.46 \pm 0.08$ GeV^2.

(iv) ADFR: The calculation of Aglietti, Di Lodovico, Ferrera, and Ricciardi [21,22] makes use of the ratio of $B \to X_u \ell^+ \nu_\ell$ to $B \to X_c \ell^+ \nu_\ell$ rates and soft-gluon resummation at next-to-next-to-leading order and an effective QCD coupling approach. The calculation uses the $\overline{\mathrm{MS}}$ scheme and we use $m_b(\overline{\mathrm{MS}}) = 4.19 \pm 0.04$ GeV.

Table VI lists the decay rates and their associated uncertainties for the probed regions of phase space, which we use to extract $|V_{ub}|$ from the measured partial branching fractions with Eq. (32).

C. $|V_{ub}|$ results

From the partial branching fractions with $E_\ell^B > 1$ and $M_X < 1.7$ GeV determined from fitting M_X we find

$$\begin{aligned}
|V_{ub}|(\mathrm{BLNP}) &= (3.90 \pm 0.09 \pm 0.15 \pm 0.21) \times 10^{-3}, \\
|V_{ub}|(\mathrm{DGE}) &= (4.08 \pm 0.09 \pm 0.16^{+0.20}_{-0.26}) \times 10^{-3}, \\
|V_{ub}|(\mathrm{GGOU}) &= (3.97 \pm 0.09^{+0.15+0.15}_{-0.16-0.16}) \times 10^{-3}, \\
|V_{ub}|(\mathrm{ADFR}) &= (3.63 \pm 0.08 \pm 0.14 \pm 0.17) \times 10^{-3}.
\end{aligned} \tag{33}$$

The uncertainties denote the statistical uncertainty, the systematic uncertainty, and the theory error from the partial rate prediction. For the partial branching fraction with $E_\ell^B > 1$ GeV, $M_X < 1.7$ GeV, and $q^2 > 8$ GeV^2 we find

$$\begin{aligned}
|V_{ub}|(\mathrm{BLNP}) &= (4.24^{+0.22+0.30+0.26}_{-0.23-0.32-0.28}) \times 10^{-3}, \\
|V_{ub}|(\mathrm{DGE}) &= (4.16^{+0.21+0.30+0.18}_{-0.23-0.32-0.21}) \times 10^{-3}, \\
|V_{ub}|(\mathrm{GGOU}) &= (4.25^{+0.22+0.30+0.24}_{-0.23-0.33-0.26}) \times 10^{-3}, \\
|V_{ub}|(\mathrm{ADFR}) &= (3.68^{+0.19+0.26}_{-0.20-0.28} \pm 0.17) \times 10^{-3}.
\end{aligned} \tag{34}$$

TABLE VI. The theory rates $\Delta\Gamma(B \to X_u \ell^+ \nu_\ell)$ from various theory calculations are listed. The rates are given in units of ps^{-1}.

Phase-space region	BLNP [17]	DGE [19,20]	GGOU [18]	ADFR [21,22]
$M_X < 1.7$ GeV	$45.2^{+5.4}_{-4.6}$	$42.3^{+5.8}_{-3.8}$	$43.7^{+3.9}_{-3.2}$	$52.3^{+5.4}_{-4.7}$
$M_X < 1.7$ GeV, $q^2 > 8$ GeV^2	$23.4^{+3.4}_{-2.6}$	$24.3^{+2.6}_{-1.9}$	$23.3^{+3.2}_{-2.4}$	$31.1^{+3.0}_{-2.6}$
$E_\ell^B > 1$ GeV	$61.5^{+6.4}_{-5.1}$	$58.2^{+3.6}_{-3.0}$	$58.5^{+2.7}_{-2.3}$	$61.5^{+5.8}_{-5.1}$

Finally, the most inclusive determination with $E_\ell^B > 1$ GeV from the two-dimensional fit of M_X and q^2 results in

$$
\begin{aligned}
|V_{ub}|(\mathrm{BLNP}) &= (4.05 \pm 0.09^{+0.20+0.18}_{-0.21-0.20}) \times 10^{-3}, \\
|V_{ub}|(\mathrm{DGE}) &= (4.16 \pm 0.09^{+0.21+0.11}_{-0.22-0.12}) \times 10^{-3}, \\
|V_{ub}|(\mathrm{GGOU}) &= (4.15 \pm 0.09^{+0.21+0.08}_{-0.22-0.09}) \times 10^{-3}, \\
|V_{ub}|(\mathrm{ADFR}) &= (4.05 \pm 0.09^{+0.20}_{-0.21} \pm 0.18) \times 10^{-3}. \quad (35)
\end{aligned}
$$

In order to quote a single value for $|V_{ub}|$ we adapt the procedure of Ref. [25] and calculate a simple arithmetic average of the most precise determinations in Eq. (35) to obtain

$$|V_{ub}| = (4.10 \pm 0.09 \pm 0.22 \pm 0.15) \times 10^{-3}. \quad (36)$$

This value is larger, but compatible with the exclusive measurement of $|V_{ub}|$ from $B \to \pi \ell^+ \nu_\ell$ of $|V_{ub}| = (3.67 \pm 0.09 \pm 0.12) \times 10^{-3}$ within 1.3 standard deviations.

D. Stability checks

To check the stability of the result we redetermine the partial branching fractions using two additional working points. We change the BDT selection to increase and decrease the amount of $B \to X_c \ell^+ \nu_\ell$ and other backgrounds, and repeat the full analysis procedure. The resulting values of $\Delta\mathcal{B}(B \to X_u \ell \nu_\ell)$ are determined using the two-dimensional fit of $M_X : q^2$ and are shown in Fig. 10. The background contamination changes by $+37\%$ and -33%, respectively. The small shifts in central value are well contained within the quoted systematic uncertainties. To further estimate the compatibility of the result we determine the full statistical and systematic correlations of the results and recover that the partial branching fraction with looser and tighter BDT selection are in agreement with the nominal result within 1.1 and 1.4 standard deviations, respectively.

E. $B \to X_u \ell^+ \nu_\ell$ charged pion multiplicity

The modeling of the $B \to X_u \ell^+ \nu_\ell$ signal composition is crucial to all presented measurements. One aspect difficult to assess is the X_u fragmentation simulation: the charmless X_u state can decay via many different channels producing a number of charged or neutral pions or kaons. In Sec. V we discussed how we assess the uncertainty on the number of $s\bar{s}$ quark pairs produced in the X_u fragmentation. Due to the BDT removing such events to suppress the dominant $B \to X_c \ell^+ \nu_\ell$ background, no signal-enriched region can be easily obtained. The accuracy of the fragmentation into the number of charged pions can be tested in the signal-enriched region of $M_X < 1.7$ GeV. Figure 11 compares the charged pion multiplicity between simulated signal and background processes and data. The signal and background predictions are scaled to their respective normalizations obtained from the two-dimensional fit in $M_X : q^2$. The uncertainty band shown on the MC sample includes the full systematic uncertainties discussed in Sec. V. The agreement overall lies within the assigned uncertainties, with the data having more events in the zero multiplicity bin and less in the two charged pion multiplicity bin. We use this distribution to correct our simulation to assign an additional uncertainty from the charged pion fragmentation. More details can be found in Sec. V and Appendix C.

F. Lepton flavor universality and weak annihilation contributions

To test the lepton flavor universality in $B \to X_u \ell^+ \nu_\ell$ we also carry out fits to determine the partial branching fraction for electron and muon final states. For this we categorize the

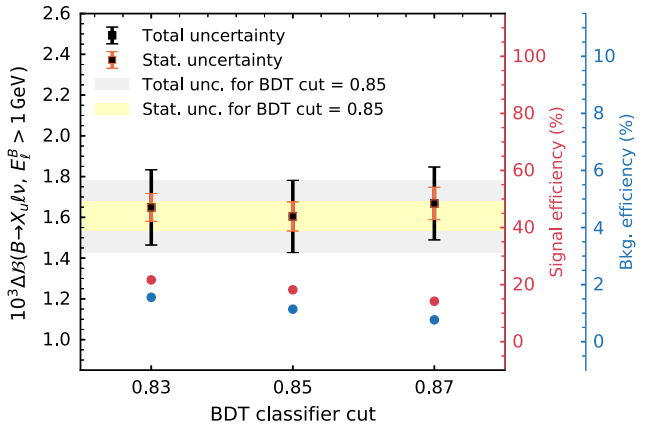


FIG. 10. The stability of the determined partial branching fraction $\Delta\mathcal{B}(B \to X_u \ell \nu_\ell)$ using the $M_X : q^2$ fit is studied as a function of the BDT selection requirement. The classifier output selection of 0.83 and 0.87 correspond to signal efficiencies after the preselection of 22% and 15%, respectively. These selections increase or decrease the background from $B \to X_c \ell^+ \nu_\ell$ and other processes by 37% and 33%, respectively. The gray and yellow bands show the total and statistical error, respectively, with the nominal BDT working point of 0.85.

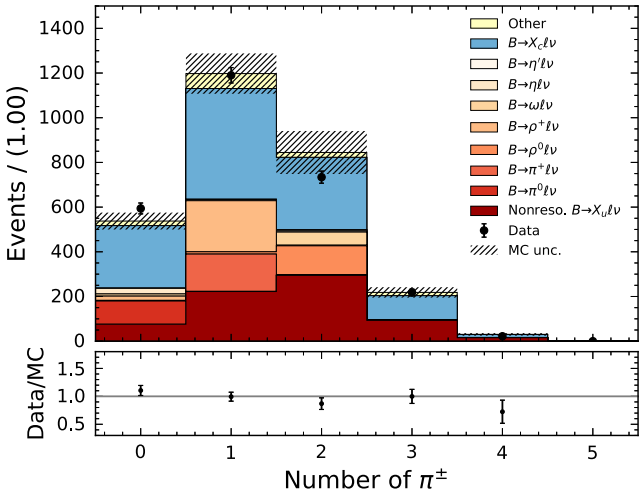


FIG. 11. The postfit charged pion multiplicity is shown for events with $M_X < 1.7$ GeV. The uncertainties on the MC stack include all systematic uncertainties.

selected events accordingly and carry out a fit to the $M_X : q^2$ distributions using the same granularity as the fit described in Sec. VIII A. We carry out a simultaneous analysis of both samples, such that shared NPs for the modeling of the signal or background components can be correctly correlated afterward. The resulting yields are corrected to a partial branching fraction with $E_\ell^B > 1$ GeV and we obtain

$$\Delta\mathcal{B}(B \to X_u e^+ \nu_e) = (1.57 \pm 0.10 \pm 0.16) \times 10^{-3}, \quad (37)$$

$$\Delta\mathcal{B}(B \to X_u \mu^+ \nu_\mu) = (1.62 \pm 0.10 \pm 0.18) \times 10^{-3}, \quad (38)$$

with a total correlation of $\rho = 0.53$. The ratio of the electron to the muon final state is

$$R_{e\mu} = \frac{\Delta\mathcal{B}(B \to X_u e^+ \nu_e)}{\Delta\mathcal{B}(B \to X_u \mu^+ \nu_\mu)} = 0.97 \pm 0.09 \pm 0.04, \quad (39)$$

with the first error denoting the statistical uncertainty and the second the systematic uncertainty. We observe no significant deviation from lepton flavor universality. More details on the fit can be found in Appendix E.

Isospin breaking effects can be studied by separately measuring the partial branching fraction for charged and neutral B meson final states. We determine the ratio

$$R_{\rm iso} = \frac{\tau_{B^0}}{\tau_{B^+}} \times \frac{\Delta\mathcal{B}(B^+ \to X_u \ell^+ \nu_\ell)}{\Delta\mathcal{B}(B^0 \to X_u \ell^+ \nu_\ell)}, \quad (40)$$

by using the information from the composition of the fully reconstructed tag-side B meson decays to separate charged and neutral B candidates. The partial branching fraction is then determined by a simultaneous fit of both samples in $M_X : q^2$ to correctly correlate common systematic uncertainties. To account for the small contamination of wrongly assigned B tag flavors, we use the wrong-tag fractions from our simulation. The measured number of signal events in the reconstructed neutral and charged B candidate categories (denoted in the following as $N^0_{\rm reco}$ and $N^+_{\rm reco}$) are related to the number of neutral and charged B mesons ($N^0_{\rm true}$ and $N^+_{\rm true}$) via

$$N^0_{\rm reco} = \mathcal{P}_{B^0_{\rm true}\to B^0_{\rm reco}} N^0_{\rm true} + \mathcal{P}_{B^+_{\rm true}\to B^0_{\rm reco}} N^+_{\rm true}, \quad (41)$$

$$N^+_{\rm reco} = \mathcal{P}_{B^0_{\rm true}\to B^+_{\rm reco}} N^0_{\rm true} + \mathcal{P}_{B^+_{\rm true}\to B^+_{\rm reco}} N^+_{\rm true}. \quad (42)$$

Here, e.g., $\mathcal{P}_{B^0_{\rm true}\to B^+_{\rm reco}}$ denotes the probability to identify in the reconstruction of the tag-side B meson a true B^0 as a B^+ candidate. In the simulation we find

$$\mathcal{P}_{B^0_{\rm true}\to B^0_{\rm reco}} = 0.985 \qquad \mathcal{P}_{B^0_{\rm true}\to B^+_{\rm reco}} = 0.015, \quad (43)$$

$$\mathcal{P}_{B^+_{\rm true}\to B^+_{\rm reco}} = 0.977 \qquad \mathcal{P}_{B^+_{\rm true}\to B^0_{\rm reco}} = 0.023. \quad (44)$$

Using this procedure we determine for the individual partial branching fractions with $E_\ell^B > 1$ GeV

$$\Delta\mathcal{B}(B^+ \to X_u \ell^+ \nu_\ell) = (1.65 \pm 0.10 \pm 0.18) \times 10^{-3}, \quad (45)$$

$$\Delta\mathcal{B}(B^0 \to X_u \ell^+ \nu_\ell) = (1.51 \pm 0.10 \pm 0.16) \times 10^{-3}, \quad (46)$$

with a total correlation of $\rho = 0.50$ and for the ratio Eq. (40)

$$R_{\rm iso} = 1.01 \pm 0.09 \pm 0.06, \quad (47)$$

compatible with the expectation of equal semileptonic rates for both isospin states. Isospin breaking effects would for instance arise from weak annihilation contributions, which only can contribute to charged B meson final states. Using Eq. (47) the relative contribution from weak annihilation processes to the total semileptonic $B \to X_u \ell^+ \nu_\ell$ rate can be constrained via

$$\frac{\Gamma_{\rm wa}}{\Gamma(B \to X_u \ell^+ \nu_\ell)} = \frac{f_u}{f_{\rm wa}} \times (R_{\rm iso} - 1). \quad (48)$$

Here f_u is a factor that corrects the measured partial branching fraction to the full inclusive phase space. We estimate it using the DFN model [54] (cf. Sec. II for details) and find $f_u = 0.86$. We further assume that $f_{\rm wa} = 1$, as such processes would produce a high-momentum lepton. We recover

$$\frac{\Gamma_{\rm wa}}{\Gamma(B \to X_u \ell^+ \nu_\ell)} = 0.01 \pm 0.09, \quad (49)$$

which translates into a limit of $[-0.14, 0.17]$ at 90% C.L. This result is more stringent than the limit of Ref. [15], but weaker than the result of Ref. [72], that directly used the shape of the q^2 distribution to constrain weak annihilation processes. Our result is also weaker than the estimates of Refs. [73–76] that constrain weak annihilation contributions to be of the order 2%–3%.

IX. SUMMARY AND CONCLUSIONS

We report measurements of partial branching fractions with different requirements on the properties of the

hadronic system of the $B \to X_u \ell^+ \nu_\ell$ decay and with a lepton energy of $E_\ell^B > 1$ GeV in the B rest -frame, covering 31%–86% of the available phase space. The sizeable background from semileptonic $B \to X_c \ell^+ \nu_\ell$ decays is suppressed using multivariate methods in the form of a BDT. This approach allows us to reduce such backgrounds to an acceptable level, while retaining a high signal efficiency. Signal yields are obtained using a binned likelihood fit in either the reconstructed hadronic mass M_X, the four-momentum-transfer squared q^2, or the lepton energy E_ℓ^B. The most precise result is obtained from a two-dimensional fit of M_X and q^2. Translated to a partial branching fraction for $E_\ell^B > 1$ GeV we obtain

$$\Delta\mathcal{B}(B \to X_u \ell^+ \nu_\ell) = (1.59 \pm 0.07 \pm 0.16) \times 10^{-3}, \quad (50)$$

with the errors denoting statistical and systematic uncertainties. The partial branching fraction is compatible with the value obtained by a fit of the lepton energy spectrum E_ℓ^B and with the most precise determination of Ref. [69]. In addition, it is stable under variations of the background suppression BDT. From this partial branching fraction we obtain a value of

$$|V_{ub}| = (4.10 \pm 0.09 \pm 0.22 \pm 0.15) \times 10^{-3} \quad (51)$$

from an average over four theoretical calculations. This value is higher than, but compatible with, the value of $|V_{ub}|$ from exclusive determinations by 1.3 standard deviations. The compatibility with the value expected from CKM unitarity from a fit of Ref. [77] of $|V_{ub}| = (3.62^{+0.11}_{-0.08}) \times 10^{-3}$ is 1.6 standard deviations. Figure 12 summarizes the situation. The result presented here supersedes Ref. [16]: this paper uses a more efficient tagging algorithm, incorporates improvements of the $B \to X_u \ell^+ \nu_\ell$ signal and $B \to X_c \ell^+ \nu_\ell$ background descriptions, and analyzes the full Belle dataset of 711 fb^{-1}. The measurement of kinematic differential shapes of M_X, q^2, and other properties are left for future work. These results will be crucial for future direct measurements with Belle II that will attempt to use data-driven methods to directly constrain the shape function using $B \to X_u \ell^+ \nu_\ell$ information.

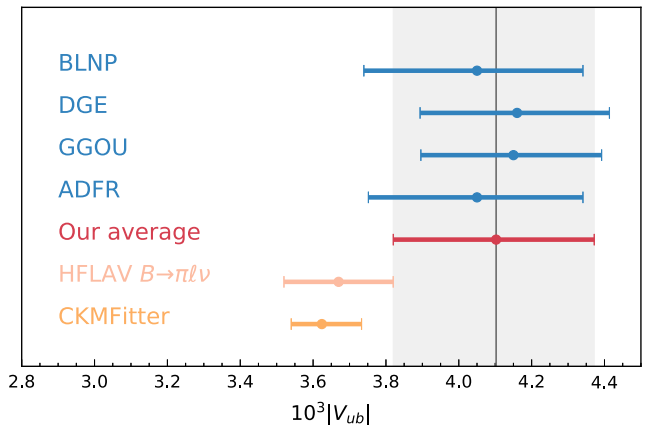


FIG. 12. The obtained values of $|V_{ub}|$ from the four calculations and the arithmetic average is compared to the determination from exclusive $B \to \pi \ell^+ \nu_\ell$ and the expectation from CKM unitarity [77] without the direct constraints from semileptonic and leptonic decays.

ACKNOWLEDGMENTS

We thank Kerstin Tackmann, Frank Tackmann, Zoltan Ligeti, Ian Stewart, Thomas Mannel, and Keri Voss for useful discussions about the subject matter of this manuscript. L. C., W. S., R. V. T., and F. B. were supported by the DFG Emmy-Noether Grant No. BE 6075/1-1. W. S. was supported by the Alexander von Humboldt Foundation. 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 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, and No. FT130100303; Austrian Federal Ministry of Education, Science and Research (FWF) and FWF Austrian Science Fund No. P 31361-N36; the National Natural Science Foundation of China under Contracts No. 11435013, No. 11475187, No. 11521505, No. 11575017, No. 11675166, and 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); the Shanghai Pujiang Program under Grant No. 18PJ1401000; the Shanghai Science and Technology Committee (STCSM) under Grant No. 19ZR1403000; the Ministry of Education, Youth and Sports of the Czech Republic under Contract No. LTT17020; Horizon 2020 ERC Advanced Grant No. 884719 and ERC Starting Grant No. 947006 "InterLeptons" (European Union); the Carl Zeiss Foundation, the Deutsche Forschungsgemeinschaft, the Excellence Cluster Universe, and the VolkswagenStiftung; the Department of Atomic Energy (Project Identification No. RTI 4002) and the Department of Science and Technology of India; the Istituto Nazionale di Fisica Nucleare of Italy; National Research Foundation (NRF) of Korea Grants No. 2016R1D1A1B01010135, No. 2016R1D1A1B02012900, No. 2018R1A2B3003643, No. 2018R1A6A1A06024970, No. 2018R1D1A1B07047294, No. 2019K1A3A7A09033840, and No. 2019R1I1A3A01058933;

Radiation Science Research Institute, Foreign Large-size Research Facility Application Supporting project, the Global Science Experimental Data Hub Center of the Korea Institute of Science and Technology Information and KREONET/GLORIAD; the Polish Ministry of Science and Higher Education and the National Science Center; the Ministry of Science and Higher Education of the Russian Federation, Agreement 14.W03.31.0026, and the HSE University Basic Research Program, Moscow; University of Tabuk research Grants No. S-1440-0321, No. S-0256-1438, and No. S-0280-1439 (Saudi Arabia); the Slovenian Research Agency Grants No. J1-9124 and No. P1-0135; Ikerbasque, Basque Foundation for Science, Spain; the Swiss National Science Foundation; the Ministry of Education and the Ministry of Science and Technology of Taiwan; and the U.S. Department of Energy and the National Science Foundation.

APPENDIX A: $B \to X_u \ell^+ \nu_\ell$ HYBRID MC DETAILS

Figure 13 shows the generator-level hybrid $B \to X_u \ell^+ \nu_\ell$ signal sample for E_ℓ^B, M_X, and q^2 described in Sec. II.

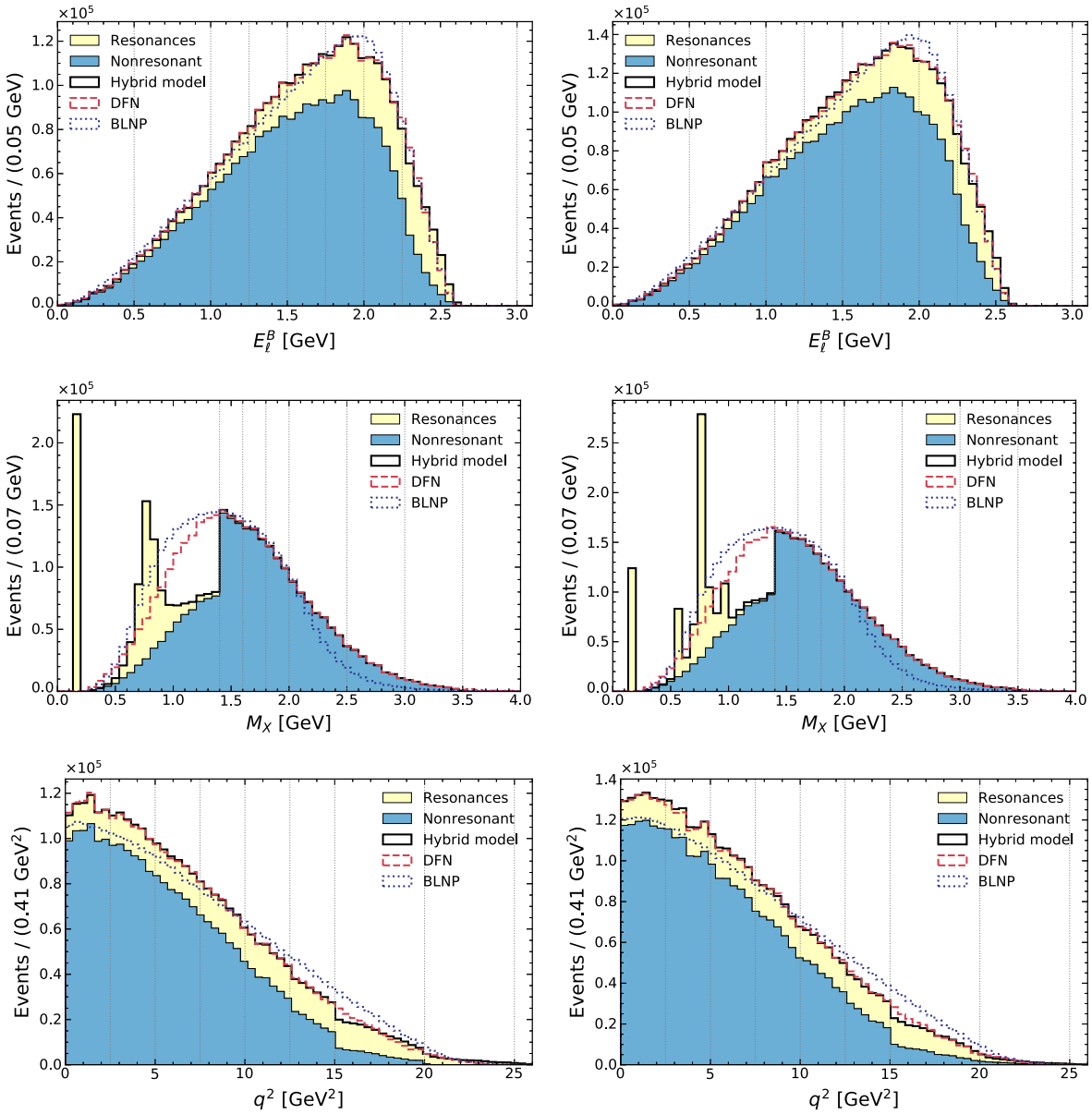


FIG. 13. The generator-level $B \to X_u \ell^+ \nu_\ell$ distributions E_ℓ^B, M_X, and q^2 for neutral (left) and charged (right) B mesons are shown. The black histogram shows the merged hybrid model, composed of resonant and nonresonant contributions. For more details on the used models and how the hybrid $B \to X_u \ell^+ \nu_\ell$ signal sample is constructed, see Sec. II.

APPENDIX B: INPUT VARIABLES OF $B \to X_c \ell \bar{\nu}_\ell$ SUPPRESSION BDT

The shapes of the variables used in the $B \to X_c \ell^+ \nu_\ell$ background suppression BDT are shown in Figs. 14 and 17. The most discriminating variables are M^2_{miss}, the B_{sig} vertex fit probability, and M^2_{miss,D^*}. Figures 15, 16, and 18 show the

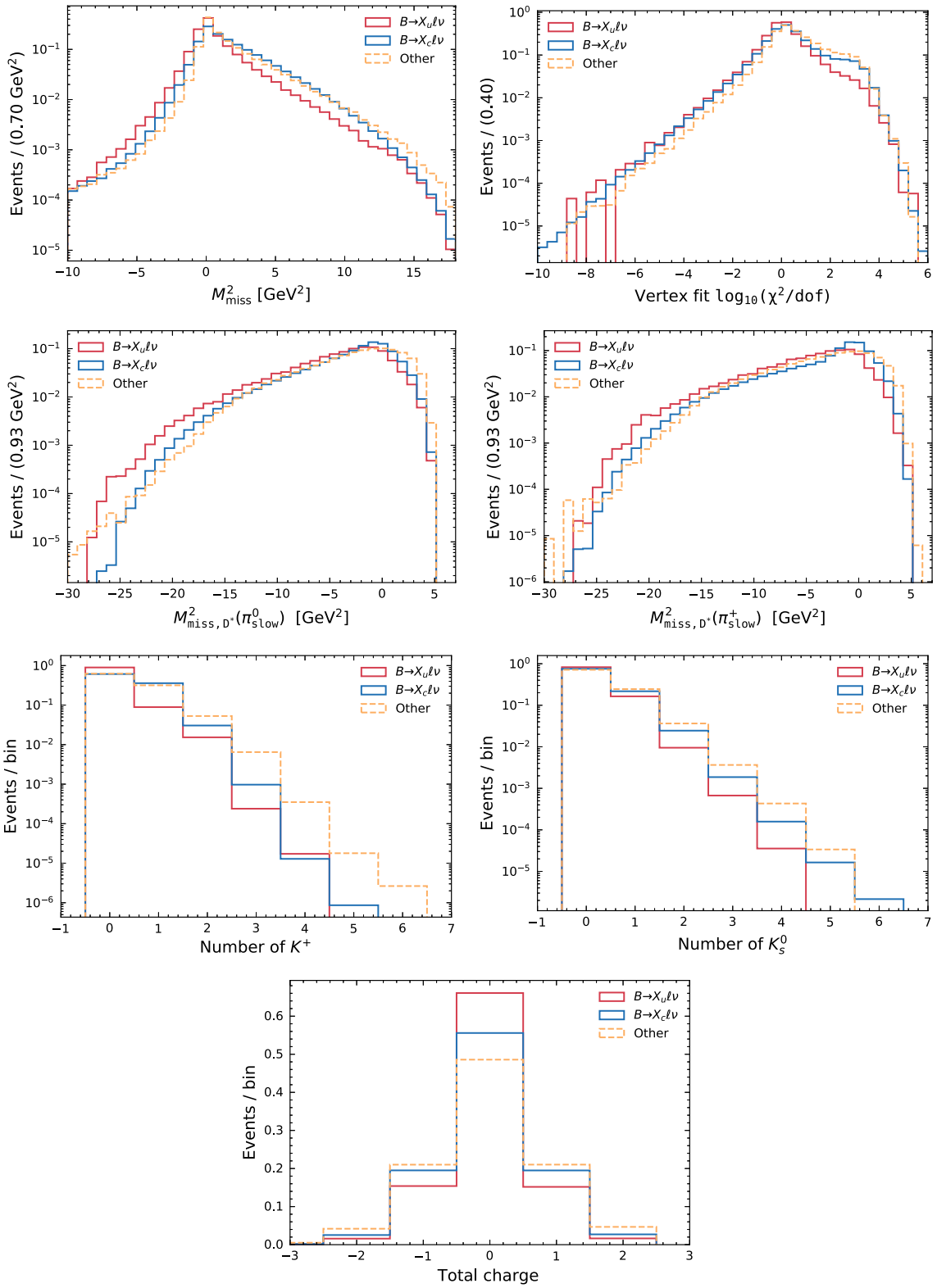


FIG. 14. The shape of the input variables for the $B \to X_c \ell^+ \nu_\ell$ background suppression BDT are shown. For details and definitions, see Sec. III C.

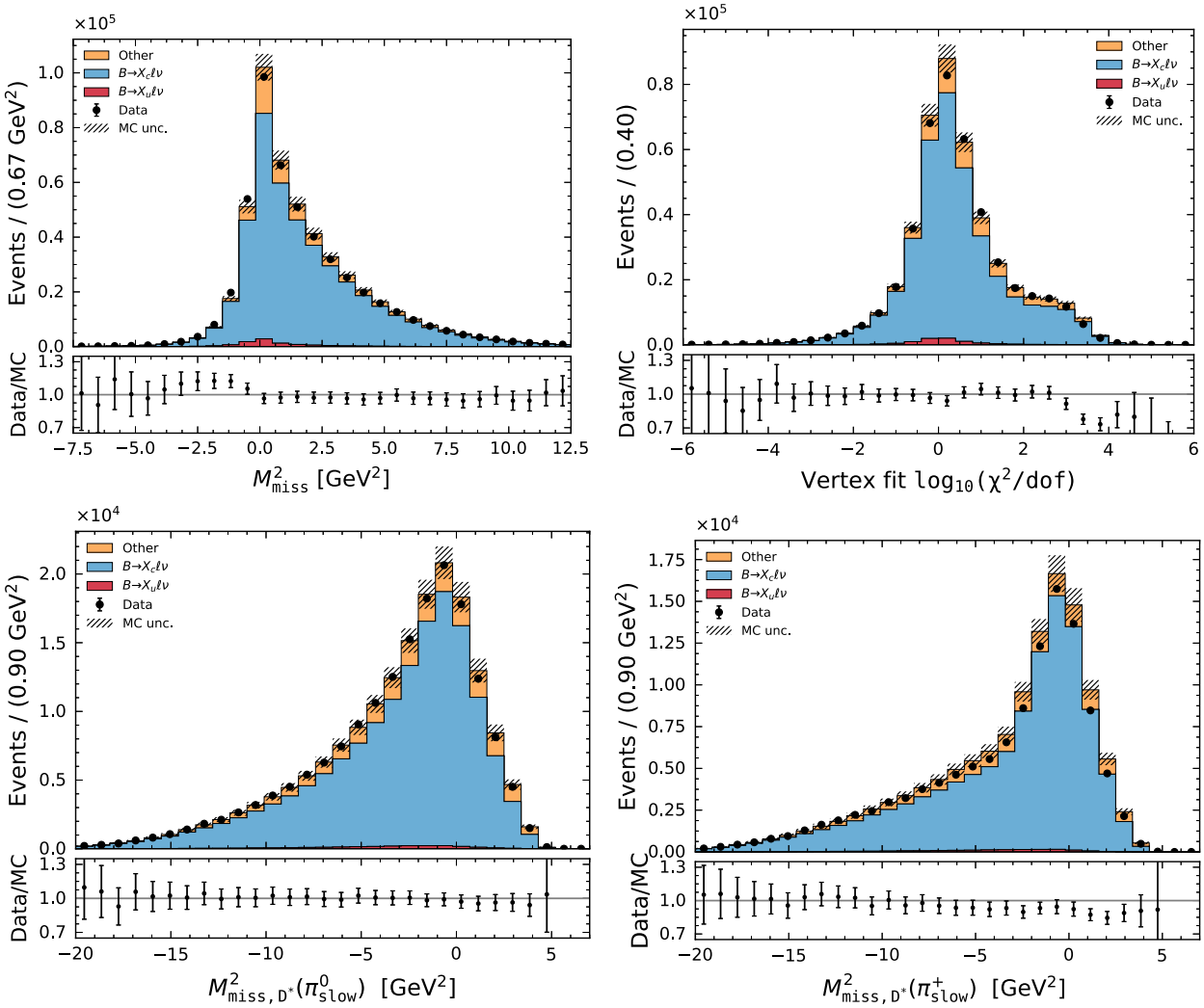


FIG. 15. The input variables for the $B \to X_c \ell^+ \nu_\ell$ background suppression BDT for recorded and simulated events are shown. The uncertainty on the simulated events incorporate the full systematic uncertainties detailed in Sec. V.

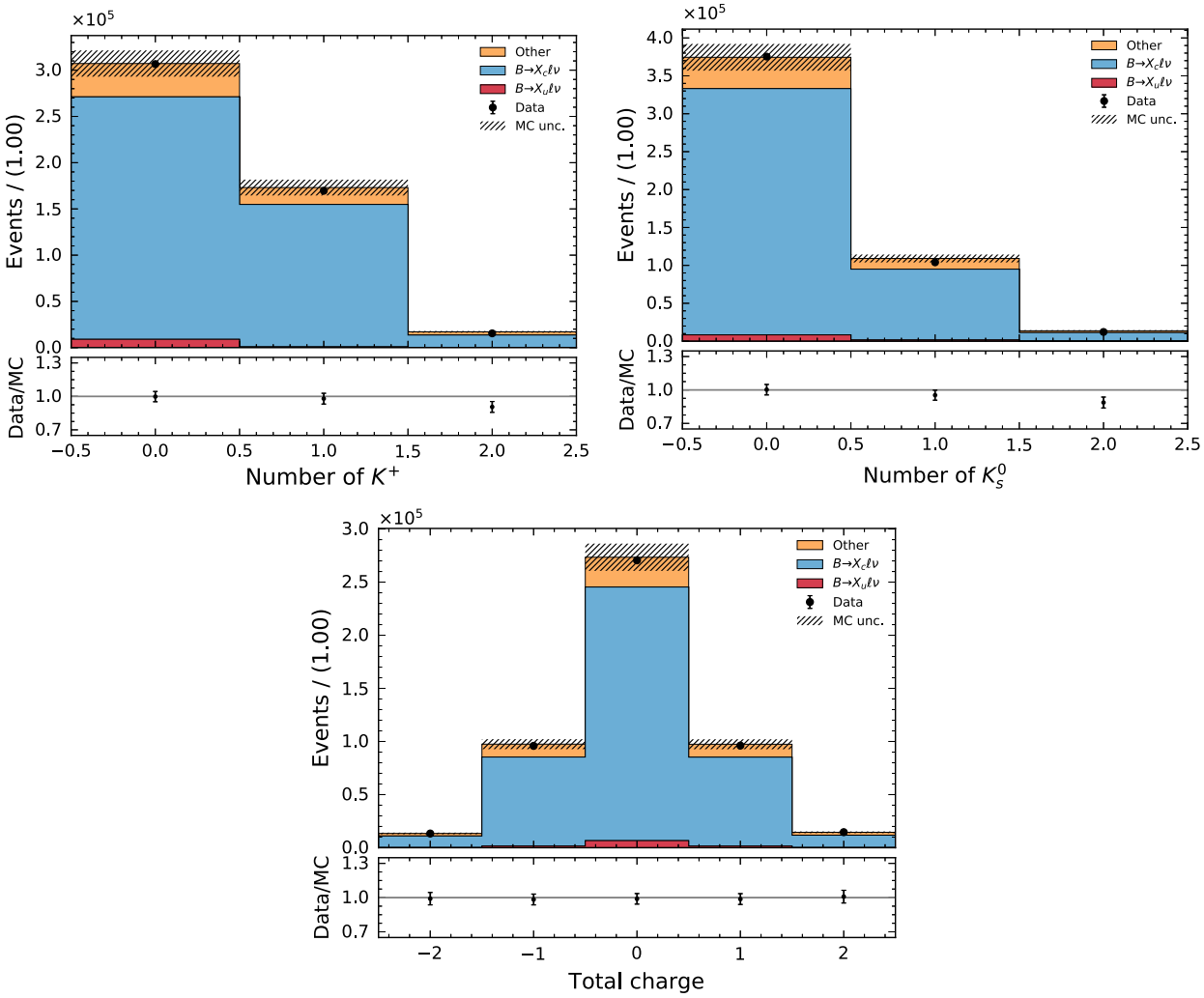


FIG. 16. The input variables for the $B \to X_c \ell^+ \nu_\ell$ background suppression BDT for recorded and simulated events are shown. The uncertainty on the simulated events incorporate the full systematic uncertainties detailed in Sec. V.

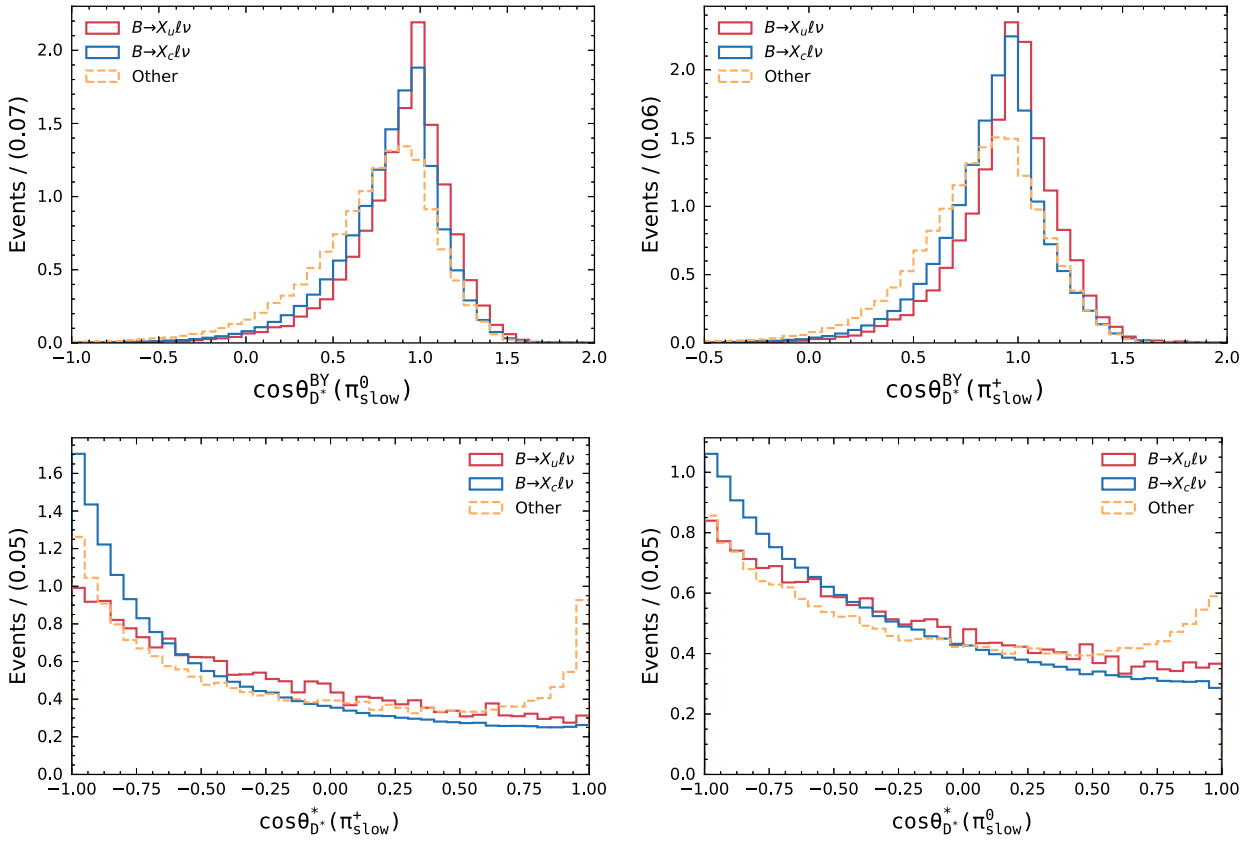


FIG. 17. The shape of the input variables for the $B \to X_c \ell^+ \nu_\ell$ background suppression BDT are shown. For details and definitions, see Sec. III C.

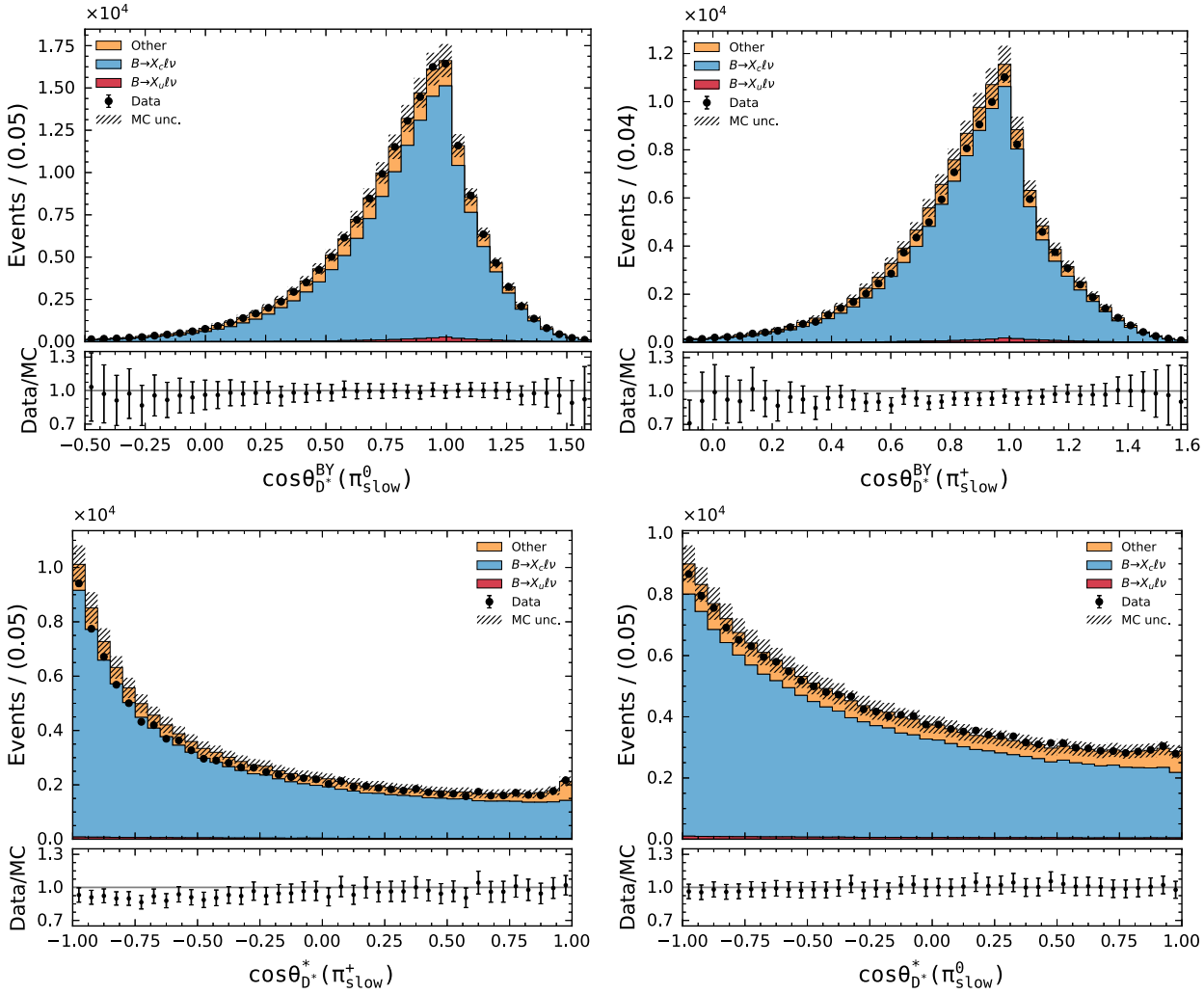


FIG. 18. The input variables for the $B \to X_c \ell^+ \nu_\ell$ background suppression BDT for recorded and simulated events are shown. The uncertainty on the simulated events incorporate the full systematic uncertainties detailed in Sec. V.

agreement between recorded and simulated events, taking into account the full uncertainties detailed in Sec. V. More details about the BDT can be found in Sec. III C.

APPENDIX C: $B \to X_u \ell^+ \nu_\ell$ CHARGED PION FRAGMENTATION MODELING

Figure 19 compares the charged pion multiplicity at different stages in the selection. This variable is not used in the signal extraction, but its modeling is tested to make sure that the $B \to X_u \ell^+ \nu_\ell$ fragmentation probabilities cannot bias the final result. The agreement in the signal-enriched region with $M_X < 1.7$ GeV after the BDT selection is fair, but shows some deviations. We correct the generator-level charged pion multiplicity to match the $n_{\pi^\pm}$ observed in this selection by assigning the nonresonant $B \to X_u \ell^+ \nu_\ell$ events a correction weight as a function of the true charged pion multiplicity. After this procedure the agreement is perfect and we use the difference in the reconstruction efficiency as an uncertainty on the pion fragmentation on the partial branching fractions and $|V_{ub}|$ (cf. Sec. V).

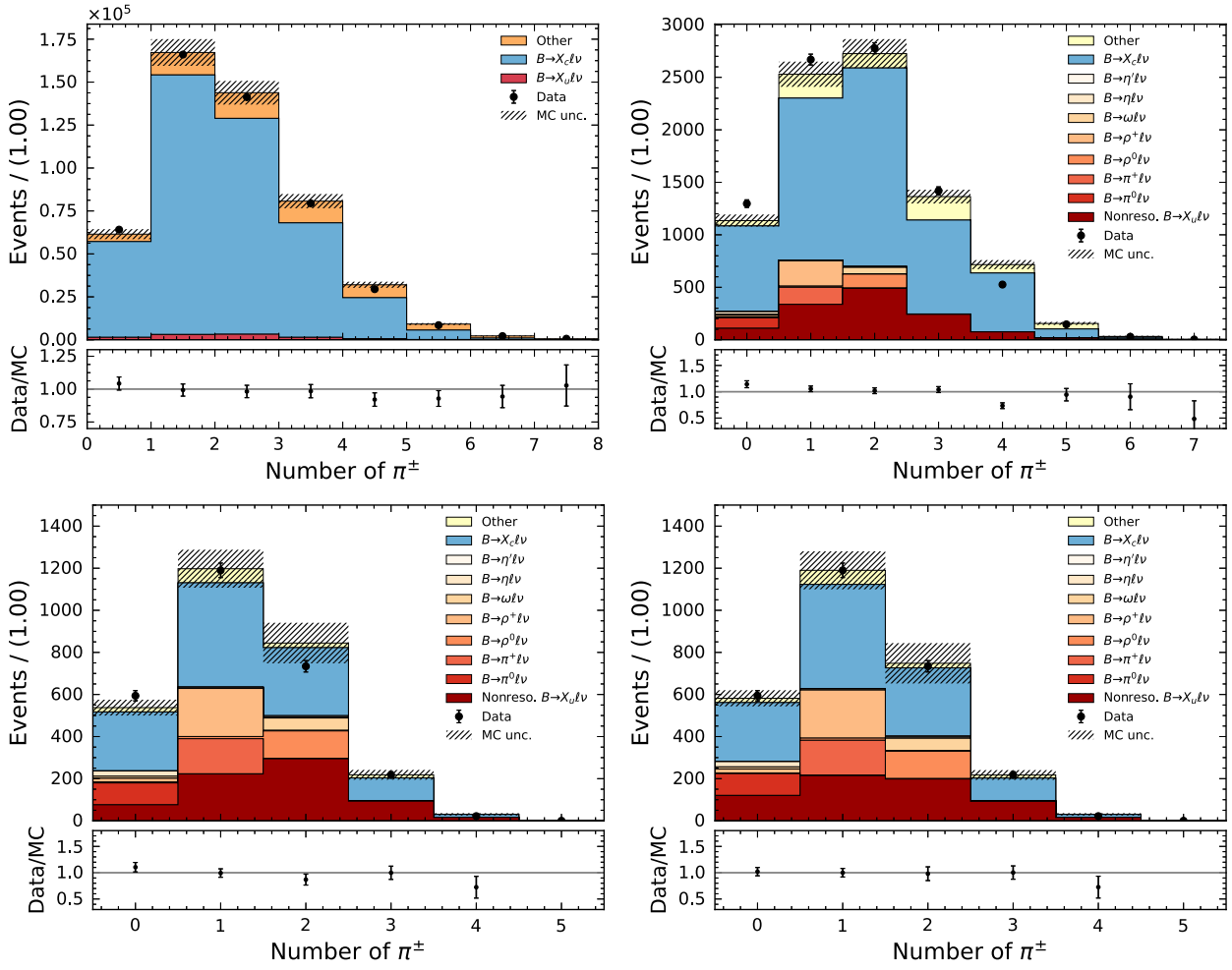


FIG. 19. The charged pion multiplicity ($n_{\pi^{\pm}}$) are compared between data and the simulation: (top left) for all events prior the BDT selection; (top right) for all events after the BDT selection; (bottom left) for the signal-enriched region of $M_X < 1.7$ GeV; (bottom right) for the same region but after rescaling the nonresonant contributions such that the $n_{\pi^{\pm}}$ fragmentation probability to match the one observed in data.

APPENDIX D: NUISANCE PARAMETER PULLS AND ADDITIONAL FIT PLOTS

Figures 20 and 21 show the nuisance parameter pulls for each fit category k and bin i defined as

$$(\hat{\theta}_{ik} - \theta_{ik})/\sqrt{\Sigma_{k,ii}}, \tag{D1}$$

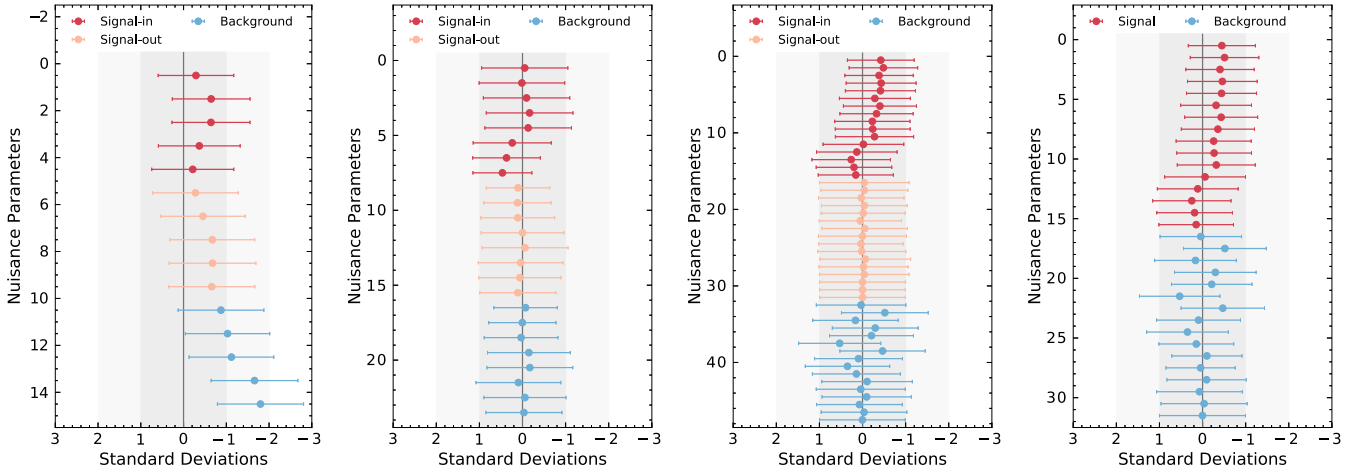


FIG. 20. The nuisance parameter pulls on the 1D fits of M_X, q^2, and E_ℓ^B with and without $M_X < 1.7$ GeV events separated out are shown from left to right.

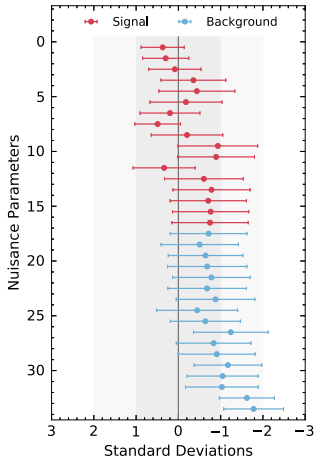


FIG. 21. The nuisance parameter pulls on the 2D fit of $M_X : q^2$ is shown.

of the partial branching fraction fits, with $\hat{\theta}$ (θ) corresponding to the postfit (prefit) value of the nuisance parameter. Note that uncertainties of each pull shows the postfit error

$$\sqrt{\hat{\Sigma}_{k,ii}} \tag{D2}$$

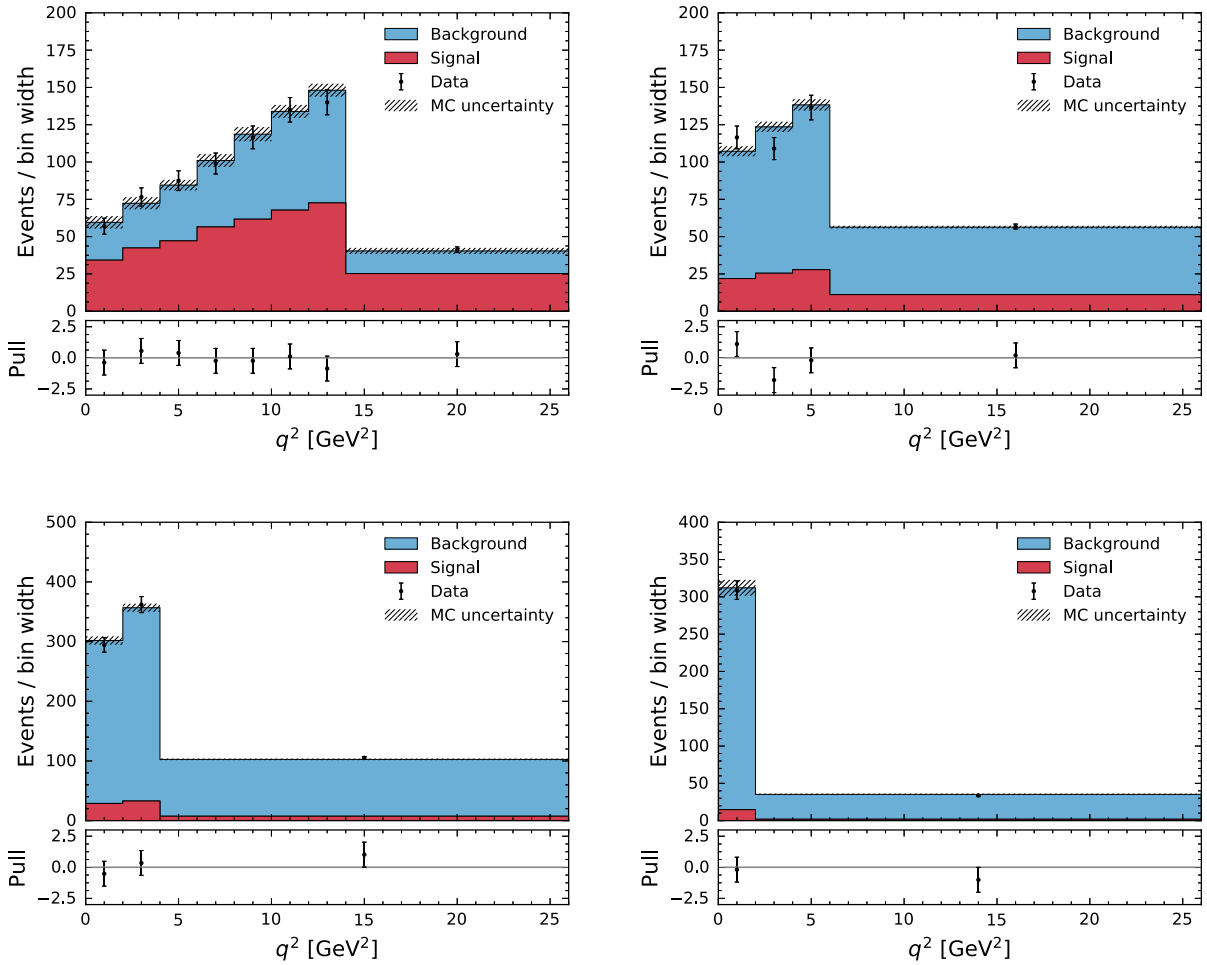


FIG. 22. The postfit q^2 distributions of the two-dimensional fit to $M_X : q^2$ on M_X are shown. The panels correspond to $M_X \in [0, 1.5]$ GeV (top left), $M_X \in [1.5, 1.9]$ GeV (top right), $M_X \in [1.9, 2.4]$ GeV (bottom left), and $M_X \in [2.4, 4]$ GeV (bottom right). The resulting yields are corrected to correspond to a partial branching fraction with $E_\ell^B > 1$ GeV.

TABLE VII. The fitted yields separated in electron and muon candidates, as well as in charged or neutral B mesons.

Decay mode	$\hat{\eta}_{\mathtt{sig}}$	$\hat{\eta}_{\mathtt{bkg}}$	$10^3(\epsilon_{\mathtt{tag}} \cdot \epsilon_{\mathtt{sel}})$	$10^3 \Delta\mathcal{B}$
$B^+ \to X_u \ell^+ \nu$	$914 \pm 56 \pm 64$	$3667 \pm 77 \pm 63$	0.30 ± 0.13	$1.65 \pm 0.10 \pm 0.18$
$B^0 \to X_u \ell^+ \nu$	$879 \pm 58 \pm 65$	$3373 \pm 76 \pm 64$	0.33 ± 0.11	$1.51 \pm 0.10 \pm 0.16$
$B \to X_u e^+ \nu$	$870 \pm 56 \pm 59$	$3311 \pm 75 \pm 60$	0.31 ± 0.12	$1.57 \pm 0.10 \pm 0.16$
$B \to X_u \mu^+ \nu$	$936 \pm 58 \pm 71$	$3716 \pm 78 \pm 71$	0.32 ± 0.13	$1.62 \pm 0.10 \pm 0.18$

normalized to the prefit constraint

$$\sqrt{\Sigma_{k,ii}}. \tag{D3}$$

Figure 22 shows the postfit q^2 distributions of the two-dimensional fit to $M_X : q^2$ on M_X.

APPENDIX E: ADDITIONAL FIT DETAILS TO THE LEPTON FLAVOR UNIVERSALITY AND WEAK ANNIHILATION TESTS

The fitted yields of the two-dimensional fit to $M_X : q^2$ separated in electron and muon candidates, as well as in charged or neutral B mesons are listed in Table VII.

APPENDIX F: BDT EFFICIENCIES

Figure 23 shows the efficiency of the BDT selection as a function of the reconstructed variables q^2, M_X, and the lepton energy E_ℓ^B for simulated $B \to X_u \ell^+ \nu_\ell$ events. Although we avoided using these variables in the boosted decision tree, a residual dependence on the kinematic variables is seen. For instance the efficiency increases with an increase in E_ℓ^B and a decrease with respect to high q^2. The efficiency on the hadronic mass M_X is relatively flat. This efficiency dependence is linked to the used variables in the BDT. Although we carefully avoided kinematic variables that would allow the BDT to learn these kinematic properties, there are indirect connections; e.g., high E_ℓ^B

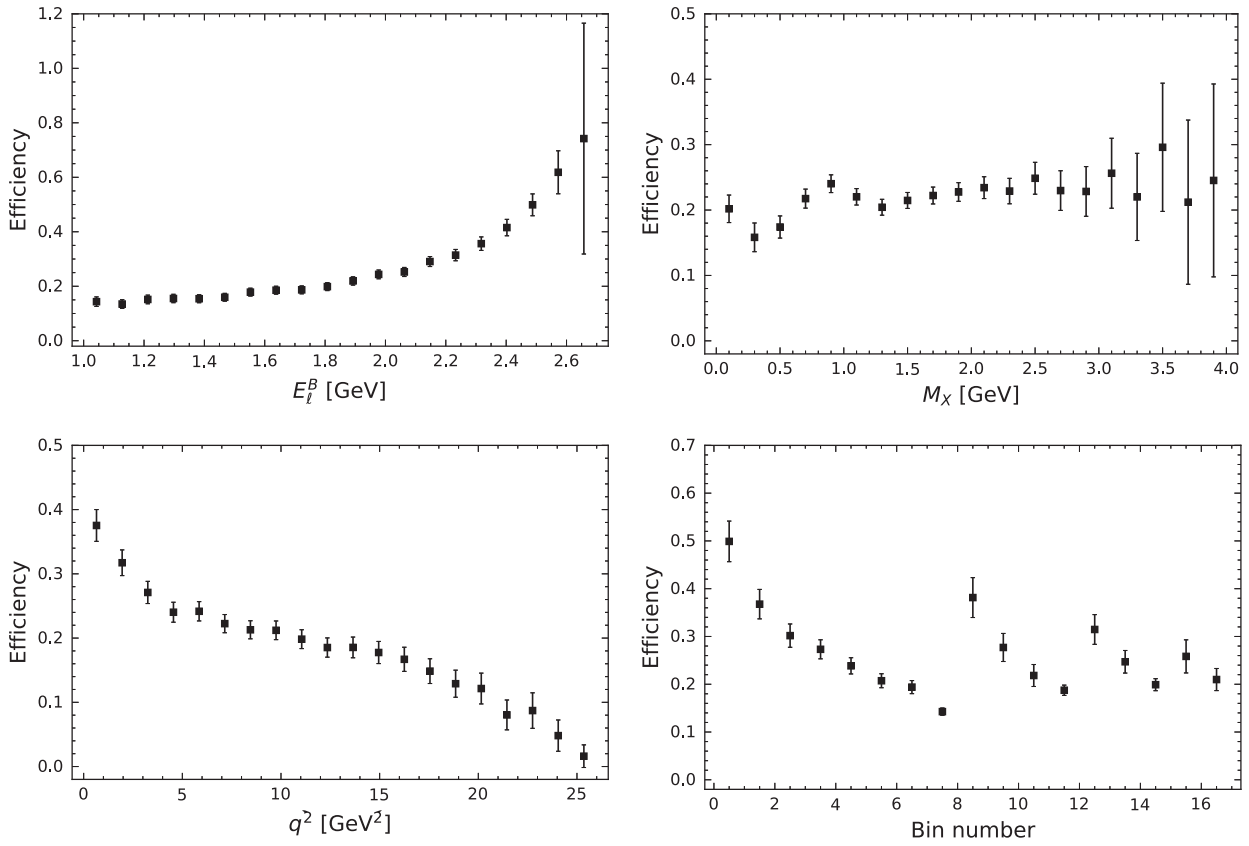


FIG. 23. The $B \to X_u \ell^+ \nu_\ell$ efficiency after the BDT selection is shown as a function of the reconstructed kinematic variables (E_ℓ^B, M_X, q^2) used in the signal extraction. The bottom right plot shows the efficiencies in the bins of $M_X : q^2$ and the binning can be found in the text. The uncertainties are statistical only.

final states have a lower multiplicity as they are dominated by $B \to \pi \ell \bar{\nu}_\ell$ decays. Further, their corresponding hadronic system carries little momentum and on average such decays retain a better resolution in discriminating variables of the background suppression BDT. A concrete example is $M^2_{\rm miss}$ (cf. Fig. 15): high multiplicity $B \to X_u \ell^+ \nu_\ell$ decays will retain a larger tail in this variable and will be selected with a lower efficiency by the BDT.

[1] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).
[2] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
[3] P. Zyla *et al.* (Particle Data Group and CKM Quark-Mixing Matrix Review), Prog. Theor. Exp. Phys. **2020**, 083C01 (2020).
[4] K. Abe *et al.* (T2K Collaboration), Nature (London) **580**, 339 (2020).
[5] E. Kou, P. Urquijo *et al.* (Belle II Collaboration), Prog. Theor. Exp. Phys. **2019**, 123C01 (2019); **2020**, 029201(E) (2020).
[6] Y. S. Amhis *et al.* (Heavy Flavor Averaging Group (HFLAV) Collaboration), Eur. Phys. J. C **81a**, 226 (2021).
[7] R. Aaij *et al.* (LHCb Collaboration), Nat. Phys. **11**, 743 (2015).
[8] S. Aoki *et al.* (Flavour Lattice Averaging Group), Eur. Phys. J. C **80**, 113 (2020).
[9] A. Bharucha, J. High Energy Phys. 05 (2012) 092.
[10] P. Gambino and N. Uraltsev, Eur. Phys. J. C **34**, 181 (2004).
[11] C. W. Bauer, Z. Ligeti, M. Luke, A. V. Manohar, and M. Trott, Phys. Rev. D **70**, 094017 (2004).
[12] D. Benson, I. I. Bigi, and N. Uraltsev, Nucl. Phys. **B710**, 371 (2005).
[13] F. U. Bernlochner, H. Lacker, Z. Ligeti, I. W. Stewart, F. J. Tackmann, and K. Tackmann (SIMBA Collaboration), arXiv:2007.04320.
[14] P. Gambino, K. J. Healey, and C. Mondino, Phys. Rev. D **94**, 014031 (2016).
[15] J. Lees *et al.* (*BABAR* Collaboration), Phys. Rev. D **86**, 032004 (2012).
[16] P. Urquijo *et al.* (Belle Collaboration), Phys. Rev. Lett. **104**, 021801 (2010).
[17] B. O. Lange, M. Neubert, and G. Paz, Phys. Rev. D **72**, 073006 (2005).
[18] P. Gambino, P. Giordano, G. Ossola, and N. Uraltsev, J. High Energy Phys. 10 (**2007**) 058.
[19] J. R. Andersen and E. Gardi, J. High Energy Phys. 01 (2006) 097.
[20] E. Gardi, Frascati Phys. Ser. **47**, 381 (2008), https://inspirehep.net/literature/789304.
[21] U. Aglietti, F. Di Lodovico, G. Ferrera, and G. Ricciardi, Eur. Phys. J. C **59**, 831 (2009).
[22] U. Aglietti, G. Ferrera, and G. Ricciardi, Nucl. Phys. **B768**, 85 (2007).
[23] M. Bona *et al.* (UTFit Collaboration), in *Presentation at the ICHEP 2020 Conference (Online)* (2020), https://indico.cern.ch/event/868940/contributions/3815693/attachments/2082161/3497433/bona-utfit.pdf.
[24] J. Dingfelder and T. Mannel, Rev. Mod. Phys. **88**, 035008 (2016).
[25] P. Zyla *et al.* (Particle Data Group, Semileptonic b-Hadron Decays, Determination of V_{cb} and V_{ub} Review), Prog. Theor. Exp. Phys. **2020**, 083C01 (2020).
[26] 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.
[27] A. Abashian *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 117 (2002); also see detector section in J. Brodzicka *et al.*, Prog. Theor. Exp. Phys. **2012**, 4D001 (2012).
[28] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima, and T. Tsukamoto, Nucl. Instrum. Methods Phys. Res., Sect. A **485**, 490 (2002).
[29] A. Abashian *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **491**, 69 (2002).
[30] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).
[31] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini, Report No. CERN-DD-EE-84-1 (1987).
[32] E. Barberio, B. van Eijk, and Z. Was, Comput. Phys. Commun. **66**, 115 (1991).
[33] C. G. Boyd, B. Grinstein, and R. F. Lebed, Phys. Rev. Lett. **74**, 4603 (1995).
[34] R. Glattauer *et al.* (Belle Collaboration), Phys. Rev. D **93**, 032006 (2016).
[35] B. Grinstein and A. Kobach, Phys. Lett. B **771**, 359 (2017).
[36] D. Bigi, P. Gambino, and S. Schacht, Phys. Lett. B **769**, 441 (2017).
[37] E. Waheed *et al.* (Belle Collaboration), Phys. Rev. D **100**, 052007 (2019).
[38] F. U. Bernlochner and Z. Ligeti, Phys. Rev. D **95**, 014022 (2017).
[39] P. Zyla *et al.* (Particle Data Group), Prog. Theor. Exp. Phys. **2020**, 083C01 (2020).
[40] D. Liventsev *et al.* (Belle Collaboration), Phys. Rev. D **77**, 091503 (2008).
[41] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **101**, 261802 (2008).
[42] J. Abdallah *et al.* (DELPHI Collaboration), Eur. Phys. J. C **45**, 35 (2006).
[43] A. K. Leibovich, Z. Ligeti, I. W. Stewart, and M. B. Wise, Phys. Rev. D **57**, 308 (1998).
[44] I. Bigi, B. Blossier, A. Le Yaouanc, L. Oliver, O. Pene, J.-C. Raynal, A. Oyanguren, and P. Roudeau, Eur. Phys. J. C **52**, 975 (2007).
[45] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. D **84**, 092001 (2011); **85**, 039904 (2012).
[46] J. Lees *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **116**, 041801 (2016).

[47] C. Bourrely, I. Caprini, and L. Lellouch, Phys. Rev. D **79**, 013008 (2009); **82**, 099902(E) (2010).
[48] J. A. Bailey *et al.* (Fermilab Lattice and MILC Collaborations), Phys. Rev. D **92**, 014024 (2015).
[49] F. U. Bernlochner, M. T. Prim, and D. J. Robinson, arXiv:2104.05739.
[50] A. Sibidanov *et al.* (Belle Collaboration), Phys. Rev. D **88**, 032005 (2013).
[51] J. P. Lees *et al.* (*BABAR* Collaboration), Phys. Rev. D **87**, 032004 (2013); **87**, 099904(E) (2013).
[52] P. del Amo Sanchez *et al.* (*BABAR* Collaboration), Phys. Rev. D **83**, 032007 (2011).
[53] G. Duplancic and B. Melic, J. High Energy Phys. 11 (2015) 138.
[54] F. De Fazio and M. Neubert, J. High Energy Phys. 06 (1999) 017.
[55] A. L. Kagan and M. Neubert, Eur. Phys. J. C **7**, 5 (1999).
[56] O. Buchmuller and H. Flacher, Phys. Rev. D **73**, 073008 (2006).
[57] T. Sjöstrand, Comput. Phys. Commun. **82**, 74 (1994).
[58] C. Ramirez, J. F. Donoghue, and G. Burdman, Phys. Rev. D **41**, 1496 (1990).
[59] M. Prim *et al.* (Belle Collaboration), Phys. Rev. D **101**, 032007 (2020).
[60] M. Prim, "b2-hive/effort v0.1.0" (2020), https://doi.org/10.5281/zenodo.3965699.
[61] B. O. Lange, M. Neubert, and G. Paz, Phys. Rev. D **72**, 073006 (2005).
[62] M. Feindt, F. Keller, M. Kreps, T. Kuhr, S. Neubauer, D. Zander, and A. Zupanc, Nucl. Instrum. Methods Phys. Res., Sect. A **654**, 432 (2011).
[63] A. Bevan *et al.*, Eur. Phys. J. C **74**, 3026 (2014).
[64] T. Chen and C. Guestrin, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD '16* (2016), p. 785, https://doi.org/10.1145/2939672.2939785.
[65] P. Ongmongkolkul, C. Deil, H. Dembinski, Dapid, C. Burr, Andrew, F. Rost, A. Pearce, L. Geiger, and O. Zapata, iminuit—minuit fromPYTHON (2012) (Online; accessed March 5, 2018).
[66] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Eur. Phys. J. C **71**, 1554 (2011); **73**, 2501(E) (2013).
[67] M. Althoff *et al.* (TASSO Collaboration), Z. Phys. C **27**, 27 (1985).
[68] W. Bartel *et al.* (JADE Collaboration), Z. Phys. C **20**, 187 (1983).
[69] J. Lees *et al.* (*BABAR* Collaboration), Phys. Rev. D **95**, 072001 (2017).
[70] S. W. Bosch, B. O. Lange, M. Neubert, and G. Paz, Nucl. Phys. **B699**, 335 (2004).
[71] S. W. Bosch, M. Neubert, and G. Paz, J. High Energy Phys. 11 (2004) 073.
[72] J. Rosner *et al.* (CLEO Collaboration), Phys. Rev. Lett. **96**, 121801 (2006).
[73] P. Gambino and J. F. Kamenik, Nucl. Phys. **B840**, 424 (2010).
[74] Z. Ligeti, M. Luke, and A. V. Manohar, Phys. Rev. D **82**, 033003 (2010).
[75] M. Voloshin, Phys. Lett. B **515**, 74 (2001).
[76] I. I. Bigi and N. Uraltsev, Nucl. Phys. **B423**, 33 (1994).
[77] J. Charles, A. Höcker, H. Lacker, S. Laplace, F. R. Le Diberder, J. Malclés, J. Ocariz, M. Pivk, and L. Roos (CKMfitter Group), Eur. Phys. J. C **41**, 1 (2005).