

Measurements of q^2 moments of inclusive $B \rightarrow X_c \ell^+ \nu_\ell$ decays with hadronic tagging

R. van Tonder^{1,*}, L. Cao,¹ W. Sutcliffe,¹ M. Welsch,¹ F. U. Bernlochner,^{1,†} I. Adachi,^{16,13} H. Aihara,⁸² D. M. Asner,³ T. Aushev,¹⁸ R. Ayad,⁷⁷ V. Babu,⁸ P. Behera,²⁴ K. Belous,²⁸ J. Bennett,⁴⁹ M. Bessner,¹⁵ V. Bhardwaj,²¹ B. Bhuyan,²² T. Bilka,⁵ J. Biswal,³³ A. Bobrov,^{4,62} D. Bodrov,^{18,40} J. Borah,²² A. Bozek,⁵⁹ M. Bračko,^{46,33} P. Branchini,³⁰ A. Budano,³⁰ M. Campajola,^{29,54} D. Červenkov,⁵ M.-C. Chang,⁹ P. Chang,⁵⁸ V. Chekelian,⁴⁷ A. Chen,⁵⁶ B. G. Cheon,¹⁴ K. Chilikin,⁴⁰ H. E. Cho,¹⁴ K. Cho,³⁶ S.-J. Cho,⁸⁹ Y. Choi,⁷⁵ S. Choudhury,²³ D. Cinabro,⁸⁷ S. Cunliffe,⁸ S. Das,⁴⁵ G. De Nardo,^{29,54} G. De Pietro,³⁰ R. Dhamija,²³ F. Di Capua,^{29,54} J. Dingfelder,¹ Z. Doležal,⁵ T. V. Dong,¹⁰ S. Dubey,¹⁵ D. Epifanov,^{4,62} T. Ferber,⁸ B. G. Fulsom,⁶⁴ R. Garg,⁶⁵ V. Gaur,⁸⁶ N. Gabyshev,^{4,62} A. Giri,²³ P. Goldenzweig,³⁴ E. Graziani,³⁰ T. Gu,⁶⁷ Y. Guan,⁷ K. Gudkova,^{4,62} C. Hadjivasiliou,⁶⁴ S. Halder,⁷⁸ O. Hartbrich,¹⁵ K. Hayasaka,⁶¹ H. Hayashii,⁵⁵ W.-S. Hou,⁵⁸ C.-L. Hsu,⁷⁶ T. Iijima,^{53,52} K. Inami,⁵² A. Ishikawa,^{16,13} R. Itoh,^{16,13} M. Iwasaki,⁶³ W. W. Jacobs,²⁵ S. Jia,¹⁰ Y. Jin,⁸² K. K. Joo,⁶ A. B. Kaliyar,⁷⁸ K. H. Kang,³⁸ G. Karyan,⁸ H. Kichimi,¹⁶ C. Kiesling,⁴⁷ C. H. Kim,¹⁴ D. Y. Kim,⁷⁴ K.-H. Kim,⁸⁹ Y.-K. Kim,⁸⁹ K. Kinoshita,⁷ P. Kodyš,⁵ T. Konno,³⁵ S. Korpar,^{46,33} E. Kovalenko,^{4,62} P. Križan,^{42,33} R. Kroeger,⁴⁹ P. Krokovny,^{4,62} T. Kuhr,⁴³ M. Kumar,⁴⁵ R. Kumar,⁶⁸ K. Kumara,⁸⁷ A. Kuzmin,^{4,62} Y.-J. Kwon,⁸⁹ J. S. Lange,¹¹ M. Laurenza,^{30,71} S. C. Lee,³⁸ P. Lewis,¹ C. H. Li,⁴¹ J. Li,³⁸ L. K. Li,⁷ Y. B. Li,⁶⁶ L. Li Gioi,⁴⁷ J. Libby,²⁴ K. Lieret,⁴³ D. Liventsev,^{87,16} C. MacQueen,⁴⁸ M. Masuda,^{81,69} T. Matsuda,⁵⁰ D. Matvienko,^{4,62,40} M. Merola,^{29,54} F. Metzner,³⁴ K. Miyabayashi,⁵⁵ R. Mizuk,^{40,18} G. B. Mohanty,⁷⁸ S. Mohanty,^{78,85} M. Mrvar,²⁷ M. Nakao,^{16,13} Z. Natkaniec,⁵⁹ A. Natochii,¹⁵ L. Nayak,²³ N. K. Nisar,³ S. Nishida,^{16,13} K. Nishimura,¹⁵ K. Ogawa,⁶¹ S. Ogawa,⁸⁰ H. Ono,^{60,61} P. Oskin,⁴⁰ P. Pakhlov,^{40,51} G. Pakhlova,^{18,40} T. Pang,⁶⁷ S. Pardi,²⁹ H. Park,³⁸ S.-H. Park,¹⁶ A. Passeri,³⁰ S. Patra,²¹ S. Paul,^{79,47} T. K. Pedlar,⁴⁴ R. Pestotnik,³³ L. E. Piilonen,⁸⁶ T. Podobnik,^{42,33} V. Popov,¹⁸ E. Prencipe,¹⁹ M. T. Prim,¹ A. Rabusov,⁷⁹ M. Röhrken,⁸ A. Rostomyan,⁸ N. Rout,²⁴ G. Russo,⁵⁴ D. Sahoo,⁷⁸ L. Santelj,^{42,33} V. Savinov,⁶⁷ G. Schnell,^{2,20} C. Schwanda,²⁷ Y. Seino,⁶¹ K. Senyo,⁸⁸ M. E. Sevier,⁴⁸ M. Shapkin,²⁸ C. Sharma,⁴⁵ C. P. Shen,¹⁰ J.-G. Shiu,⁵⁸ A. Sokolov,²⁸ E. Solovieva,⁴⁰ M. Starič,³³ Z. S. Stottler,⁸⁶ J. F. Strube,⁶⁴ M. Sumihama,¹² T. Sumiyoshi,⁸⁴ M. Takizawa,^{73,17,70} U. Tamponi,³¹ K. Tanida,³² F. Tenchini,⁸ K. Trabelsi,³⁹ M. Uchida,⁸³ T. Uglov,^{40,18} Y. Unno,¹⁴ S. Uno,^{16,13} P. Urquijo,⁴⁸ S. E. Vahsen,¹⁵ G. Varner,¹⁵ K. E. Varvell,⁷⁶ E. Waheed,¹⁶ C. H. Wang,⁵⁷ E. Wang,⁶⁷ M.-Z. Wang,⁵⁸ P. Wang,²⁶ X. L. Wang,¹⁰ M. Watanabe,⁶¹ S. Watanuki,³⁹ B. D. Yabsley,⁷⁶ W. Yan,⁷² S. B. Yang,³⁷ H. Ye,⁸ J. H. Yin,³⁷ Z. P. Zhang,⁷² V. Zhilich,^{4,62} and V. Zhukova⁴⁰

(Belle Collaboration)

¹University of Bonn, 53115 Bonn

²Department of Physics, University of the Basque Country, UPV/EHU, 48080 Bilbao

³Brookhaven National Laboratory, Upton, New York 11973

⁴Budker Institute of Nuclear Physics SB RAS, Novosibirsk 630090

⁵Faculty of Mathematics and Physics, Charles University, 121 16 Prague

⁶Chonnam National University, Gwangju 61186

⁷University of Cincinnati, Cincinnati, Ohio 45221

⁸Deutsches Elektronen-Synchrotron, 22607 Hamburg

⁹Department of Physics, Fu Jen Catholic University, Taipei 24205

¹⁰Key Laboratory of Nuclear Physics and Ion-beam Application (MOE) and Institute of Modern Physics, Fudan University, Shanghai 200443

¹¹Justus-Liebig-Universität Gießen, 35392 Gießen

¹²Gifu University, Gifu 501-1193

¹³SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193

¹⁴Department of Physics and Institute of Natural Sciences, Hanyang University, Seoul 04763

¹⁵University of Hawaii, Honolulu, Hawaii 96822

¹⁶High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801

¹⁷J-PARC Branch, KEK Theory Center, High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801

¹⁸National Research University Higher School of Economics, Moscow 101000

¹⁹Forschungszentrum Jülich, 52425 Jülich

²⁰IKERBASQUE, Basque Foundation for Science, 48013 Bilbao

²¹Indian Institute of Science Education and Research Mohali, SAS Nagar, 140306

²²Indian Institute of Technology Guwahati, Assam 781039

- ²³*Indian Institute of Technology Hyderabad, Telangana 502285*
²⁴*Indian Institute of Technology Madras, Chennai 600036*
²⁵*Indiana University, Bloomington, Indiana 47408*
²⁶*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049*
²⁷*Institute of High Energy Physics, Vienna 1050*
²⁸*Institute for High Energy Physics, Protvino 142281*
²⁹*INFN–Sezione di Napoli, I-80126 Napoli*
³⁰*INFN–Sezione di Roma Tre, I-00146 Roma*
³¹*INFN–Sezione di Torino, I-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*Kitasato University, Sagami-hara 252-0373*
³⁶*Korea Institute of Science and Technology Information, Daejeon 34141*
³⁷*Korea University, Seoul 02841*
³⁸*Kyungpook National University, Daegu 41566*
³⁹*Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay*
⁴⁰*P.N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow 119991*
⁴¹*Liaoning Normal University, Dalian 116029*
⁴²*Faculty of Mathematics and Physics, University of Ljubljana, 1000 Ljubljana*
⁴³*Ludwig Maximilians University, 80539 Munich*
⁴⁴*Luther College, Decorah, Iowa 52101*
⁴⁵*Malaviya National Institute of Technology Jaipur, Jaipur 302017*
⁴⁶*Faculty of Chemistry and Chemical Engineering, University of Maribor, 2000 Maribor*
⁴⁷*Max-Planck-Institut für Physik, 80805 München*
⁴⁸*School of Physics, University of Melbourne, Victoria 3010*
⁴⁹*University of Mississippi, University, Mississippi 38677*
⁵⁰*University of Miyazaki, Miyazaki 889-2192*
⁵¹*Moscow Physical Engineering Institute, Moscow 115409*
⁵²*Graduate School of Science, Nagoya University, Nagoya 464-8602*
⁵³*Kobayashi-Maskawa Institute, Nagoya University, Nagoya 464-8602*
⁵⁴*Università di Napoli Federico II, I-80126 Napoli*
⁵⁵*Nara Women's University, Nara 630-8506*
⁵⁶*National Central University, Chung-li 32054*
⁵⁷*National United University, Miao Li 36003*
⁵⁸*Department of Physics, National Taiwan University, Taipei 10617*
⁵⁹*H. Niewodniczanski Institute of Nuclear Physics, Krakow 31-342*
⁶⁰*Nippon Dental University, Niigata 951-8580*
⁶¹*Niigata University, Niigata 950-2181*
⁶²*Novosibirsk State University, Novosibirsk 630090*
⁶³*Osaka City University, Osaka 558-8585*
⁶⁴*Pacific Northwest National Laboratory, Richland, Washington 99352*
⁶⁵*Panjab University, Chandigarh 160014*
⁶⁶*Peking University, Beijing 100871*
⁶⁷*University of Pittsburgh, Pittsburgh, Pennsylvania 15260*
⁶⁸*Punjab Agricultural University, Ludhiana 141004*
⁶⁹*Research Center for Nuclear Physics, Osaka University, Osaka 567-0047*
⁷⁰*Meson Science Laboratory, Cluster for Pioneering Research, RIKEN, Saitama 351-0198*
⁷¹*Dipartimento di Matematica e Fisica, Università di Roma Tre, I-00146 Roma*
⁷²*Department of Modern Physics and State Key Laboratory of Particle Detection and Electronics, University of Science and Technology of China, Hefei 230026*
⁷³*Showa Pharmaceutical University, Tokyo 194-8543*
⁷⁴*Soongsil University, Seoul 06978*
⁷⁵*Sungkyunkwan University, Suwon 16419*
⁷⁶*School of Physics, University of Sydney, New South Wales 2006*
⁷⁷*Department of Physics, Faculty of Science, University of Tabuk, Tabuk 71451*
⁷⁸*Tata Institute of Fundamental Research, Mumbai 400005*
⁷⁹*Department of Physics, Technische Universität München, 85748 Garching*
⁸⁰*Toho University, Funabashi 274-8510*
⁸¹*Earthquake Research Institute, University of Tokyo, Tokyo 113-0032*

⁸²*Department of Physics, University of Tokyo, Tokyo 113-0033*⁸³*Tokyo Institute of Technology, Tokyo 152-8550*⁸⁴*Tokyo Metropolitan University, Tokyo 192-0397*⁸⁵*Utkal University, Bhubaneswar 751004*⁸⁶*Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061*⁸⁷*Wayne State University, Detroit, Michigan 48202*⁸⁸*Yamagata University, Yamagata 990-8560*⁸⁹*Yonsei University, Seoul 03722*

(Received 7 September 2021; accepted 6 December 2021; published 30 December 2021)

We present the measurement of the first to fourth order moments of the four-momentum transfer squared, q^2 , of inclusive $B \rightarrow X_c \ell^+ \nu_\ell$ decays using the full Belle dataset of 711 fb^{-1} of integrated luminosity at the $\Upsilon(4S)$ resonance where $\ell = e, \mu$. The determination of these moments and their systematic uncertainties open new pathways to determine the absolute value of the Cabibbo-Kobayashi-Maskawa matrix element V_{cb} using a reduced set of matrix elements of the heavy quark expansion. In order to identify and reconstruct the X_c system, we reconstruct one of the two B -mesons using machine learning techniques in fully hadronic decay modes. The moments are measured with progressively increasing threshold selections on q^2 starting with a lower value of 3.0 GeV^2 in steps of 0.5 GeV^2 up to a value of 10.0 GeV^2 . The measured moments are further unfolded, correcting for reconstruction and selection effects as well as QED final state radiation. We report the moments separately for electron and muon final states and observe no lepton flavor universality violating effects.

DOI: [10.1103/PhysRevD.104.112011](https://doi.org/10.1103/PhysRevD.104.112011)

I. INTRODUCTION

Precise measurements of the absolute value of the Cabibbo-Kobayashi-Maskawa (CKM) matrix element V_{cb} are important to deepen our understanding of the Standard Model of Particle Physics (SM) [1,2]. The CKM matrix is a 3×3 unitary matrix, whose complex phase is responsible for all known charge-parity (CP) violating effects in the quark sector [3] through the KM mechanism [2]. There exists evidence for conceptually similar CP -violating processes to be present in the neutrino sector [4]. The quark sector associated CP violation is not sufficient to explain the matter dominance present in the Universe today, and it is also unclear if the CP violation in the neutrino sector can produce a baryon asymmetry of the required size. This motivates searches for new CP -violating phenomena e.g. in the form of processes involving heavy exotic particles. If such new exotic states interact with quarks in some notable form, their existence might alter the properties of measurements constraining the unitarity property of the CKM matrix [5]. Precise measurements of $|V_{ub}|$, $|V_{cb}|$, and the CKM angle $\gamma = \phi_3$ are imperative to isolate such effects. These quantities can be measured

using tree-level dominated processes, which are expected to remain unaffected by new physics contributions in most models. Thus combining their results one can obtain an unbiased measure for the amount of CP violation in the SM quark sector.

Semileptonic $B \rightarrow X_c \ell^+ \nu_\ell$ decays offer a theoretically clean avenue to determine $|V_{cb}|$ and Fig. 1 depicts the processes involving B^+ and B^0 -meson decays.¹ Due to the factorization (up to small known electroweak corrections, cf. [6]) of the hadronic and lepton-neutrino final states, these processes are theoretically better understood than hadronic transitions involving V_{cb} . Furthermore, until future lepton colliders will provide clean samples of B_c decays, measurements involving the theoretically better understood purely leptonic decays are not feasible [7] even with the large samples of B_c mesons recorded at the LHC. The existing $|V_{cb}|$ determinations either focus on exclusive final states, with $B \rightarrow D^* \ell^+ \nu_\ell$ and $B \rightarrow D \ell^+ \nu_\ell$ transitions providing the most precise values to date [8], or on the study of inclusive $B \rightarrow X_c \ell^+ \nu_\ell$ decay. First measurements using $B_s \rightarrow D_s^* \ell^+ \nu_\ell$ and $B_s \rightarrow D_s \ell^+ \nu_\ell$ decays were recently reported in Ref. [9], indicating that in the future such channels will play a prominent role in the precision determination of $|V_{cb}|$ from exclusive final states. To translate measured (differential or partial) branching fractions of exclusive decays into values of $|V_{cb}|$, information

*vantonder@physik.uni-bonn.de

†florian.bernlochner@uni-bonn.de

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

¹Charge conjugation is implied and $B \rightarrow X_c \ell^+ \nu_\ell$ is defined as the average branching fraction of B^+ and B^0 meson decays and $\ell = e$ or μ .

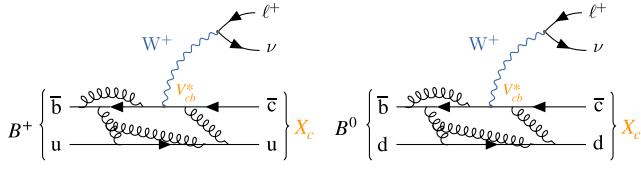


FIG. 1. The inclusive $B \rightarrow X_c \ell^+ \nu_\ell$ semileptonic processes for a B^+ (left) or a B^0 (right) meson decay.

regarding the normalization and dynamics of the non-perturbative form factors is needed. Recent studies indicate that the underlying assumptions employed to simplify form factor parametrizations should be reevaluated, since the use of more generalized and model-independent forms hint at higher values of $|V_{cb}|$, see e.g., Refs. [10–13]. Inclusive determinations of $|V_{cb}|$ exploit the fact that the total decay rate can be expanded into a manageable number of non-perturbative matrix elements using the heavy quark expansion (HQE). For instance, the total rate of the semileptonic transition can be represented in the HQE in a series of terms proportional to the inverse bottom quark mass times the QCD scale parameter, Λ_{QCD}/m_b , of increasing powers and corrections proportional to the strong coupling constant, $\alpha_s(m_b)$, can be systematically included. This approach is independent from the considerations that enter exclusive determinations involving form factors and therefore provide complementary information. Spectral moments, such as the moments of the lepton energy, hadronic mass or hadronic energy spectra, can be expressed in the same way as the total rate using the HQE. The expressions describing these observables have been calculated to next-to-next-to-leading order precision in α_s and at leading order in the HQE, at next-to-leading order in α_s and for HQE up to $\mathcal{O}(1/m_b^2)$, and up to the HQE of order $\mathcal{O}(1/m_b^5)$ at tree level with respect to the strong interaction [14–21]. With measurements of the total $B \rightarrow X_c \ell^+ \nu_\ell$ branching fraction, the energy and the hadronic mass moments [22–28], the nonperturbative parameters and $|V_{cb}|$ can be determined using combined fits of all the relevant inputs [29,30]. A comprehensive review of this approach can be found in Refs. [29,31].

The world averages of $|V_{cb}|$ from exclusive and inclusive determinations are [8]

$$|V_{cb}^{\text{excl}}| = (39.25 \pm 0.56) \times 10^{-3}, \quad (1)$$

$$|V_{cb}^{\text{incl}}| = (42.19 \pm 0.78) \times 10^{-3}. \quad (2)$$

Here the uncertainties are the sum from experiment and theory. Both world averages exhibit a disagreement of about 3 standard deviations. It is noteworthy that newer measurements of exclusive $|V_{cb}|$ tend to agree better with the inclusive value, but these also have larger uncertainties. The total uncertainty in the inclusive determination is dominated by theory errors, mainly to cover uncertainties

from missing higher-order contributions [32,33]. This disagreement between the two measurements is limiting the reach of present-day searches for loop-level new physics in the CKM sector of the SM, see e.g., Ref. [34] for a recent analysis.

In Ref. [35] a novel and alternative approach to determine $|V_{cb}|$ from inclusive decays was outlined: by exploiting reparametrization invariance the authors could demonstrate that the full set of 13 nonperturbative matrix elements present in the total rate can be reduced to a set of only eight parameters at the order of $\mathcal{O}(1/m_b^4)$. This reduction eases the proliferation of new parameters when considering higher orders of the HQE. New measurements are needed to determine this reduced set of eight parameters, as the key prerequisite that gives rise to the reparametrization invariance in the total rate is violated in measurements of moments of lepton momentum, hadronic mass, and hadronic energy spectra. The authors of Ref. [35] thus propose a systematic measurement of moments of the four-momentum transfer squared, q^2 , of the B meson system to the X_c system with progressively increasing requirements on q^2 itself. These measurements either require the identification and explicit reconstruction of the X_c system or the reconstruction of the missing neutrino three momentum. In this paper the former approach is used in combination with the full reconstruction of the second B meson produced in the collision event. This is achieved efficiently with the use of neural networks. We report measurements of the first to fourth moments, $\langle q^2 \rangle$, $\langle q^4 \rangle$, $\langle q^6 \rangle$, $\langle q^8 \rangle$ using a set of threshold selections² $q_{\text{sel}}^2 \in [3.0, 10.0] \text{ GeV}^2$. The first measurement of the first moment of the q^2 spectrum was reported in Ref. [23] with a lepton energy requirement of 1 GeV. However, this requirement reintroduces the full set of nonperturbative matrix elements, since the lepton energy is not a reparametrization invariant quantity. Furthermore, the moments of higher order exhibit greater sensitivity to the higher order terms of the HQE. In this paper a first systematic study is reported with progressively increasing threshold selections on q^2 . Additionally, the third and fourth order q^2 moments are reported for the first time. Due to the relationship of q^2 with the lepton momentum, we will restrict our measurement to moments with a minimum threshold selection on q^2 of 3 GeV^2 , to ensure that the Belle detector can still reliably reconstruct and identify the lepton from the $B \rightarrow X_c \ell^+ \nu_\ell$ decay.

The remainder of this paper is organized as follows: Section II introduces the dataset and simulated event samples, as well as the applied corrections to the simulated samples. In Sec. III the employed analysis strategy is outlined and the reconstruction of the X_c system and q^2 , and the background subtraction are discussed. Section IV

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

discusses the calibration procedure used to reverse detector resolution, selection, and acceptance effects of the measured moments of q^2 . In Sec. V the systematic uncertainties affecting the moment measurements are discussed, while Sec. VI reports the measured moments and compares them to the expectation from simulated $B \rightarrow X_c \ell^+ \nu_\ell$ decays. Finally, Sec. VII summarizes our findings and presents our conclusions.

II. DATASET AND SIMULATED SAMPLES

The results are based on an analysis of the full Belle dataset of $(772 \pm 10) \times 10^6$ B meson pairs, which were produced at the KEKB accelerator complex [36] with a center-of-mass energy of $\sqrt{s} = 10.58$ GeV corresponding to the $\Upsilon(4S)$ resonance. An additional 79 fb^{-1} of collision events recorded 60 MeV below the $\Upsilon(4S)$ resonance peak is used to derive corrections and for cross-checks.

The Belle detector is a large-solid-angle magnetic spectrometer. The detector consists of several subdetectors: a silicon vertex detector, a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), a barrel-like arrangement of time-of-flight 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 and to identify muons (KLM). More details about the detector layout and performance can be found in Ref. [37] and in references therein.

We identify charged tracks as electron or muon candidates by combining the information of multiple subdetectors into a lepton identification likelihood ratio, $\mathcal{L}_{\text{LID}}$. The most important identifying features for electrons 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 [38]. We identify muon candidates 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 [39]. Electron and muon candidates are required to have a minimum transverse momentum of 300 and 600 MeV, respectively, in the laboratory frame of Belle. We further identify charged tracks as pions or kaons using a likelihood ratio $\mathcal{L}_{K/\pi \text{ ID}} = \mathcal{L}_{K \text{ ID}} / (\mathcal{L}_{K \text{ ID}} + \mathcal{L}_{\pi \text{ ID}})$. The most important identifying features of the kaon ($\mathcal{L}_{K \text{ ID}}$) and pion ($\mathcal{L}_{\pi \text{ ID}}$) likelihoods for low momentum particles with transverse momentum below 1 GeV in the laboratory frame are the recorded energy loss by ionization, dE/dx , 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 measurement.

We identify photons 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 selected. We reconstruct π^0 candidates from photon candidates and require the invariant mass to fall into a window of $m_{\gamma\gamma} \in [0.12, 0.15]$ GeV, corresponding to about 2.5 times the π^0 mass resolution.

Monte Carlo (MC) samples of B meson decays and continuum processes ($e^+e^- \rightarrow q\bar{q}$ with $q = u, d, s, c$) are simulated using the EvtGen generator [40]. These samples are used to evaluate reconstruction efficiencies and acceptance, and to estimate background contaminations. The sample sizes correspond to approximately 3 times the Belle collision data for B meson and continuum decays. The interactions of particles traversing the detector are simulated using GEANT3 [41]. Electromagnetic final-state radiation is simulated using the PHOTOS [42] software package for all charged final-state particles. To account for, e.g. differences in identification and reconstruction efficiencies in the MC, we employ data-driven methods to derive efficiency corrections. For the particle-identification likelihood ratios these corrections are parametrized as a function of the polar angle and laboratory frame momentum of the particle candidates.

Inclusive semileptonic $B \rightarrow X_c \ell^+ \nu_\ell$ decays are dominated by $B \rightarrow D \ell^+ \nu_\ell$ and $B \rightarrow D^* \ell^+ \nu_\ell$ transitions. We model the $B \rightarrow D \ell^+ \nu_\ell$ decays using the Boyd-Grinstein-Lebed (BGL) parametrization [43] with form factor central values and uncertainties taken from the fit in Ref. [44]. For $B \rightarrow D^* \ell^+ \nu_\ell$ we use the BGL implementation proposed by Refs. [11, 45] with form factor central values and uncertainties from the fit to the measurement of Ref. [46]. Both processes are normalized to the average branching fraction of Ref. [8] assuming isospin symmetry. Semileptonic $B \rightarrow D^{**} \ell^+ \nu_\ell$ 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. [47]. The D^{**} decays are simulated using masses and widths from Ref. [48]. For the branching fractions the values of Ref. [8] are adopted and we correct them to account for missing isospin-conjugated and other established decay modes, following the prescription given in Ref. [47]. Since all measurements target the $D^{**0} \rightarrow D^{(*)+} \pi^-$ decay modes, where D^{**0} refers to all neutral D^{**} states, we correct for the missing isospin modes with a factor of

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

The measurement of the $B \rightarrow D_2^* \ell \bar{\nu}_\ell$ in Ref. [8] are converted to be respective to the $\bar{D}_2^{*0} \rightarrow D^{*-} \pi^+$ final states only. To also account for $\bar{D}_2^{*0} \rightarrow D^- \pi^+$ contributions, a factor of [48]

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

is applied. The world average of $B \rightarrow D_1^* \ell \bar{\nu}_\ell$ given in Ref. [8] combines measurements, which show very poor agreement, and the resulting probability of the combination is below 0.01%. Notably, the measurement of Ref. [49] is in conflict with the measured branching fractions of Refs. [24,50] and with the expectation of $\mathcal{B}(B \rightarrow D_1^* \ell \bar{\nu}_\ell)$ being of similar size to $\mathcal{B}(B \rightarrow D_0^* \ell \bar{\nu}_\ell)$ [51,52]. We perform our own average excluding the conflicting measurement and use

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

This slightly deviates from the treatment in Ref. [48] that omits the measurements of Refs. [24,49]. The world average of $B \rightarrow D_1 \ell \bar{\nu}_\ell$ does not account for contributions from three-body decays of the form $D_1 \rightarrow D \pi \pi$. We account for these using a factor [53]

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

To account for missing isospin-conjugated modes of the three-hadron final states we adopt the prescription from Ref. [54], which quotes an average isospin correction factor of

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

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

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

For the remaining $B \rightarrow D^{(*)} \pi \pi \ell^+ \nu_\ell$ contributions we use the measured value of Ref. [54], while subtracting the contribution of $D_1 \rightarrow D \pi \pi$ from the measured nonresonant plus resonant $B \rightarrow D \pi \pi \ell \bar{\nu}_\ell$ branching fraction. The remaining “gap” between the sum of all considered

exclusive modes and the inclusive $B \rightarrow X_c \ell^+ \nu_\ell$ branching fraction is filled in equal parts with $B \rightarrow D \eta \ell^+ \nu_\ell$ and $B \rightarrow D^* \eta \ell^+ \nu_\ell$ and for both we assume a 100% uncertainty. We simulate $B \rightarrow D^{(*)} \pi \pi \ell^+ \nu_\ell$ and $B \rightarrow D^{(*)} \eta \ell^+ \nu_\ell$ 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, $B \rightarrow D_{\text{gap}}^{**} \ell \bar{\nu}_\ell$, than e.g., a model based on the equidistribution of all final-state particles in phase space. For the form factors we adapt Ref. [47]. Comparisons of kinematic distributions for the different $B \rightarrow D_{\text{gap}}^{**} \ell \bar{\nu}_\ell$ models are found in Appendix A. In what follows, we will associate this component with a 100% uncertainty.

Semileptonic $B \rightarrow X_u \ell^+ \nu_\ell$ decays are modeled as a mixture of specific exclusive modes and nonresonant contributions. They are mixed using a “hybrid” approach proposed by Ref. [55] and our modeling is identical to the approach detailed in Ref. [56].

In Table I we summarize the branching fractions and uncertainties for the signal $B \rightarrow X_c \ell^+ \nu_\ell$ processes that we use.

TABLE I. Branching fractions for $B \rightarrow X_c \ell^+ \nu_\ell$ signal and $B \rightarrow X_u \ell^+ \nu_\ell$ signal processes that were used are listed. More details on the applied corrections can be found in the text.

B	Value B^+	Value B^0
$B \rightarrow X_c \ell^+ \nu_\ell$		
$B \rightarrow D \ell^+ \nu_\ell$	$(2.5 \pm 0.1) \times 10^{-2}$	$(2.3 \pm 0.1) \times 10^{-2}$
$B \rightarrow D^* \ell^+ \nu_\ell$	$(5.4 \pm 0.1) \times 10^{-2}$	$(5.1 \pm 0.1) \times 10^{-2}$
$B \rightarrow D_0^* \ell^+ \nu_\ell$	$(4.2 \pm 0.8) \times 10^{-3}$	$(3.9 \pm 0.7) \times 10^{-3}$
($\hookrightarrow D \pi$)		
$B \rightarrow D_1^* \ell^+ \nu_\ell$	$(4.2 \pm 0.8) \times 10^{-3}$	$(3.9 \pm 0.8) \times 10^{-3}$
($\hookrightarrow D^* \pi$)		
$B \rightarrow D_1 \ell^+ \nu_\ell$	$(4.2 \pm 0.3) \times 10^{-3}$	$(3.9 \pm 0.3) \times 10^{-3}$
($\hookrightarrow D^* \pi$)		
$B \rightarrow D_2^* \ell^+ \nu_\ell$	$(1.2 \pm 0.1) \times 10^{-3}$	$(1.1 \pm 0.1) \times 10^{-3}$
($\hookrightarrow D^* \pi$)		
$B \rightarrow D_2 \ell^+ \nu_\ell$	$(1.8 \pm 0.2) \times 10^{-3}$	$(1.7 \pm 0.2) \times 10^{-3}$
($\hookrightarrow D \pi$)		
$B \rightarrow D_1 \ell^+ \nu_\ell$	$(2.4 \pm 1.0) \times 10^{-3}$	$(2.3 \pm 0.9) \times 10^{-3}$
($\hookrightarrow D \pi \pi$)		
$B \rightarrow D \pi \pi \ell^+ \nu_\ell$	$(0.6 \pm 0.6) \times 10^{-3}$	$(0.6 \pm 0.6) \times 10^{-3}$
$B \rightarrow D^* \pi \pi \ell^+ \nu_\ell$	$(2.2 \pm 1.0) \times 10^{-3}$	$(2.0 \pm 1.0) \times 10^{-3}$
$B \rightarrow D \eta \ell^+ \nu_\ell$	$(4.0 \pm 4.0) \times 10^{-3}$	$(4.0 \pm 4.0) \times 10^{-3}$
$B \rightarrow D^* \eta \ell^+ \nu_\ell$	$(4.0 \pm 4.0) \times 10^{-3}$	$(4.0 \pm 4.0) \times 10^{-3}$
$B \rightarrow X_c \ell^+ \nu_\ell$	$(10.8 \pm 0.4) \times 10^{-2}$	$(10.1 \pm 0.4) \times 10^{-2}$
$B \rightarrow X_u \ell^+ \nu_\ell$	$(2.2 \pm 0.3) \times 10^{-3}$	$(2.0 \pm 0.3) \times 10^{-3}$

III. ANALYSIS STRATEGY, X RECONSTRUCTION, AND BACKGROUND SUBTRACTION

A. Neutral network based tag side reconstruction

The collision events are reconstructed using the hadronic Full Reconstruction (FR) algorithm detailed in Ref. [57]. The algorithm reconstructs one of the two B mesons produced in the collision event fully using hadronic decay channels. This allows for the explicit identification of the constituents of the hadronic X_c system of the $B \rightarrow X_c \ell^+ \nu_\ell$ process of interest and we label the B mesons reconstructed in hadronic modes in the following as B_{tag} . Instead of attempting to reconstruct as many B meson decay cascades as possible, the Full Reconstruction algorithm employs a hierarchical approach in four sequential stages: at the first stage, neural networks are trained to identify charged tracks and neutral energy depositions as detector stable particles (e^+ , μ^+ , K^+ , π^+ , γ), neutral π^0 candidates, or K_S^0 candidates. These candidate particles are then combined during a second stage into heavier meson candidates (J/ψ , D^0 , D^+ , D_s). For each target final state a neural network is trained to identify probable candidate combinations. The input variables for these neural networks are the output classifiers from the first reconstruction stage, vertex fit probabilities of the candidate combinations, and the four-momenta of all the particles used to reconstruct the composite particle in question. The third stage forms candidates for D^{*0} , D^{*+} , and D_s^* mesons, and for each a separate neural network is trained to identify viable combinations. The input layer aggregates the output classifiers from all previous reconstruction stages and additional information such as four-momenta of the combined particles. In the final stage the information from all previous stages is used to form B_{tag} candidates. The viability of such combinations is again assessed by a dedicated neural network that was trained to recognize correctly reconstructed candidates from incorrect combinations and whose output classifier score we denote by $\mathcal{O}_{\text{FR}}$. In total 1104 decay cascades are reconstructed in this way. This results in an efficiency of 0.28% and 0.18% for charged and neutral B meson pairs [58], respectively. As a final step, 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 $\mathcal{O}_{\text{Cont}}$. Both classifier scores are mapped to a range of [0, 1] signifying the estimated reconstruction quality from poor to excellent candidates. For the analysis we select B_{tag} candidates that show at least moderate agreement based on these two outputs, and require that $\mathcal{O}_{\text{FR}} > 10^{-4}$ and $\mathcal{O}_{\text{Cont}} > 10^{-4}$.

We use the charges and four-momenta of the decay constituents in combination with the known beam-energy to infer the flavor and four-momentum of the B_{tag} candidate.

We require the B_{tag} candidates to have at least a beam-constrained mass of

$$M_{\text{bc}} = \sqrt{E_{\text{beam}}^2 - |\mathbf{p}_{\text{tag}}|^2} > 5.27 \text{ GeV}. \quad (9)$$

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

$$\Delta E = E_{\text{tag}} - E_{\text{beam}}, \quad (10)$$

is already used in the input layer of the neural network trained in the final stage of the reconstruction, so no further requirements are imposed. Additionally, E_{tag} denotes the energy of the B_{tag} candidate in the center-of-mass frame of the colliding e^+e^- pair. In each event a single B_{tag} candidate is then selected according to the highest $\mathcal{O}_{\text{FR}}$ score value. After the reconstruction of the B_{tag} candidate, all remaining tracks and clusters are used to define and reconstruct the signal side.

B. Signal side reconstruction

The signal side of the event is reconstructed by identifying a well-reconstructed lepton using the likelihood described in Sec. II. The signal B rest frame is calculated with the momentum of the B_{tag} candidate via

$$p_{\text{sig}} = p_{e^+e^-} - \left(\sqrt{m_B^2 + |\mathbf{p}_{\text{tag}}|^2}, \mathbf{p}_{\text{tag}} \right), \quad (11)$$

with $p_{e^+e^-} = (\sqrt{s}, \mathbf{0})$ denoting the four-momentum of the colliding electron-positron pair and $m_B = 5.279 \text{ GeV}$ the nominal B meson mass.

If multiple lepton candidates are present on the signal side, the event is discarded to avoid additional neutrinos on the signal side, since multiple leptons are likely to originate from a double semileptonic $b \rightarrow c \rightarrow s$ cascade. For charged B_{tag} candidates, we demand that the charge assignment of the signal-side lepton be opposite to that of the B_{tag} charge. The hadronic X system is reconstructed from the remaining unassigned charged particles and neutral energy depositions. Its four-momentum is calculated as

$$p_X = \sum_i \left(\sqrt{m_\pi^2 + |\mathbf{p}_i|^2}, \mathbf{p}_i \right) + \sum_j (E_j, \mathbf{k}_j), \quad (12)$$

with $E_j = |\mathbf{k}_j|$ the energy of the neutral energy depositions and all charged particles with momentum $\mathbf{p}_i$ assumed to be pions. With the X system reconstructed, we can also reconstruct the missing four-momentum,

$$P_{\text{miss}} = (E_{\text{miss}}, \mathbf{p}_{\text{miss}}) = p_{\text{sig}} - p_X - p_\ell, \quad (13)$$

which estimates the four-momentum of the neutrino in the event. Here E_{miss} and $\mathbf{p}_{\text{miss}}$ denote the missing energy and momentum, respectively. For correctly reconstructed semi-leptonic $B \rightarrow X_c \ell^+ \nu_\ell$ decays $E_{\text{miss}} \approx |\mathbf{p}_{\text{miss}}|$ and we require

$$-0.5 \text{ GeV} < E_{\text{miss}} - |\mathbf{p}_{\text{miss}}| < 0.5 \text{ GeV}. \quad (14)$$

The hadronic mass of the X system is later used to discriminate $B \rightarrow X_c \ell^+ \nu_\ell$ decays from backgrounds. It is reconstructed using

$$M_X = \sqrt{(p_X)^\mu (p_X)_\mu}, \quad (15)$$

and exhibits a distinct peak at $\approx 2 \text{ GeV}$. We require the total observed charge of the event to be $|Q_{\text{tot}}| = |Q_{B_{\text{tag}}} + Q_X + Q_\ell| \leq 1$, allowing for charge imbalance in events with low-momentum tracks. Finally, we reconstruct the four-momentum transfer squared q^2 as

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

The resolution of both variables for $B \rightarrow X_c \ell^+ \nu_\ell$ is shown in Fig. 2 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 0.45 GeV and exhibits a tail towards negative values of the residuals from not reconstructed constituents of the hadronic X system. 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 3.35 GeV^2 . The core resolution is dominated by the tagging resolution, whereas the large positive tail is dominated from the resolution of the reconstruction of the X system.

In the following we analyze only the events with $q^2 > 3.0 \text{ GeV}^2$, corresponding to a lepton momentum of at least 300 MeV in the B rest frame. This corresponds also to the region of phase space in the laboratory frame at which the Belle detector operates efficiently in the

identification and reconstruction of leptons. The region of phase space below 3.0 GeV^2 is dominated by processes other than $B \rightarrow X_c \ell^+ \nu_\ell$: secondary leptons from cascade decays and fake lepton candidates make up a prominent fraction of the selected event candidates.

Figure 3 compares the selected events with the expectation from simulation: the small continuum contribution is normalized using the off-resonance event sample, while the remaining simulated events are normalized to the number of reconstructed events from $\Upsilon(4S) \rightarrow B\bar{B}$. In the following, we separate the electron and muon candidates and analyze them separately.

C. Background subtraction

In order to subtract background events, we carry out a two-step procedure. First a binned likelihood fit of the M_X distribution determines the number of expected signal and background events. For this fit we construct a likelihood function $\mathcal{L}$ as the product of individual Poisson distributions $\mathcal{P}$,

$$\mathcal{L} = \prod_i^{\text{bins}} \mathcal{P}(n_i; \nu_i) \times \prod_k \mathcal{G}_k \times \mathcal{P}_{\text{cont}}, \quad (17)$$

where n_i denotes the number of observed data events and ν_i is the total number of expected signal and background events in a given bin i . Furthermore, the $\mathcal{G}_k$ denote nuisance-parameter (NP) constraints, whose role is to incorporate systematic uncertainties on e.g., signal and background shapes directly into the fit procedure, with the index k labeling a given uncertainty source. More details of this procedure will be given in Sec. V. The Poisson term $\mathcal{P}_{\text{cont}}$ constrains the normalization of the continuum contribution to its expectation as determined from off-resonance collision events. The number of expected signal and background events in a given bin, ν_i , is estimated using simulated collision events and is given by

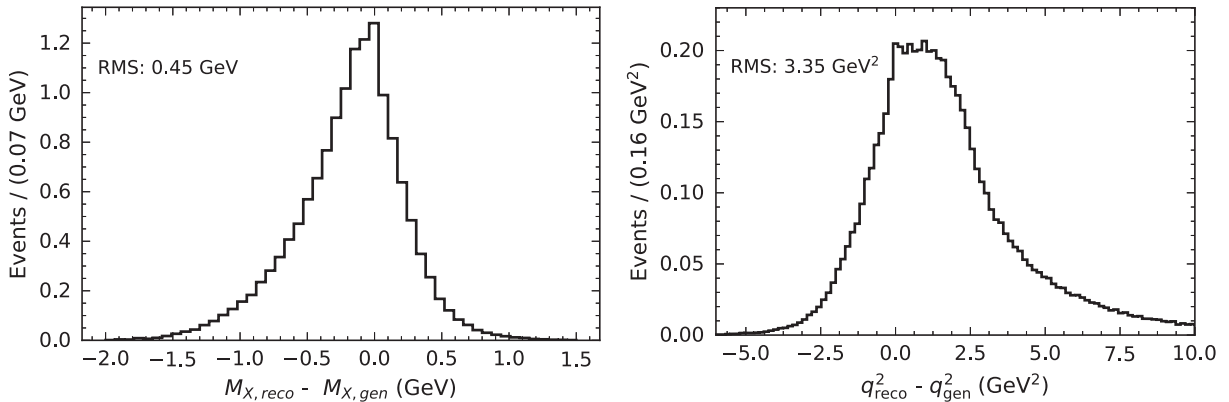


FIG. 2. The resolution of the reconstructed M_X and q^2 values for the $B \rightarrow X_c \ell^+ \nu_\ell$ signal is shown as a residual with respect to the generated values.

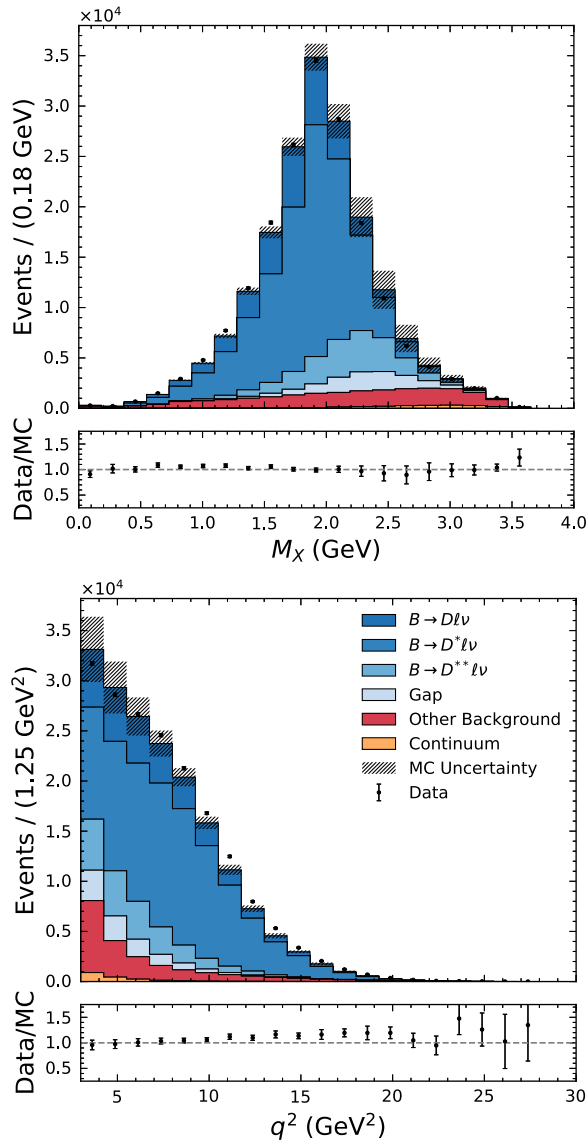


FIG. 3. The reconstructed M_X and q^2 distributions are shown. The error band of the simulated samples incorporates the full set of systematic uncertainties discussed in Sec. V.

$$\nu_i = \eta^{\text{sig}} f_i^{\text{sig}} + \eta^{B \text{ bkg}} f_i^{B \text{ bkg}} + \eta^{\text{cont}} f_i^{\text{cont}}. \quad (18)$$

Here, η^{sig} is the total number of $B \rightarrow X_c \ell^+ \nu_\ell$ signal events. Furthermore, $\eta^{B \text{ bkg}}$ denotes the background events stemming from double semileptonic cascades, $B \rightarrow X_u \ell^+ \nu_\ell$ decays, and from hadrons misidentified as leptons originating from B meson decays. The number of continuum events is denoted as η^{cont} . Furthermore, f_i denotes the fraction of events being reconstructed in a bin i with shapes as determined by the MC simulation for a given event category. Equation (17) is numerically maximized to determine both the total number of $B \rightarrow X_c \ell^+ \nu_\ell$ and background events from the observed event yields. This is

done using the sequential least squares programming method implementation of Refs. [59,60]. The fit is carried out in 20 equidistant bins of M_X ranging from 0 to 3.5 GeV to determine the number of background events for each studied threshold selection on q^2 , taking into account systematic uncertainties on the composition of $B \rightarrow X_c \ell^+ \nu_\ell$ and background templates (more details about these will be discussed in Sec. V). The continuum constraint $\mathcal{P}_{\text{cont}}$ is adjusted to reflect the number of continuum events for a given q^2 selection value as determined by the off-resonance sample, for which the M_{bc} selection was adjusted to account for the difference in center-of-mass energies.

In a second step, the determined number of background ($\hat{\eta}^{\text{bkg}}$) and continuum ($\hat{\eta}^{\text{cont}}$) events are used to construct binned signal probabilities as a function of q^2 , which is defined as

$$w_i = 1 - \frac{\hat{\eta}^{\text{bkg}} \tilde{f}_i^{\text{bkg}} + \hat{\eta}^{\text{cont}} \tilde{f}_i^{\text{cont}}}{n_i}. \quad (19)$$

Here, $\tilde{f}_i$ denotes the estimated fractions of events being reconstructed in a bin i of q^2 for a given background category as determined by the MC simulation. Figure 4 shows the q^2 spectrum for electron and muon candidates and the w_i distribution for the threshold selection with $q^2 > 3.0$ GeV². To avoid dependence on binning effects, we fit the binned signal probabilities for each q^2 selection with a polynomial function of a given order n to determine event-by-event weights, $w(q^2)$, by performing a χ^2 minimization. The order of the polynomial is determined using a nested hypothesis test and we accept a polynomial of order n over $n-1$ if the improvement in χ^2 is larger than one. Furthermore, the χ^2 takes into account the full experimental covariance of the background shapes. The resulting polynomial, $w(q^2)$, allows for an unbinned background subtraction and is required to have positive or zero event weights. We set all weights with negative values to zero. The matching figures for other q^2 threshold selections can be found in Appendix B.

IV. MOMENT CALIBRATION MASTER FORMULA

The reconstructed q^2 values are distorted by the reconstruction of the X_c system and selection criteria. To measure the first to the fourth moment of the $B \rightarrow X_c \ell^+ \nu_\ell$ q^2 spectrum, we need to correct for these effects. This is done in three sequential steps:

- (a) First, a calibration function is applied event-by-event as a function of the reconstructed q^2 value to correct for resolution distortions. This linear correction is determined separately for each order of the moment we wish to measure. For a given event i , we denote the

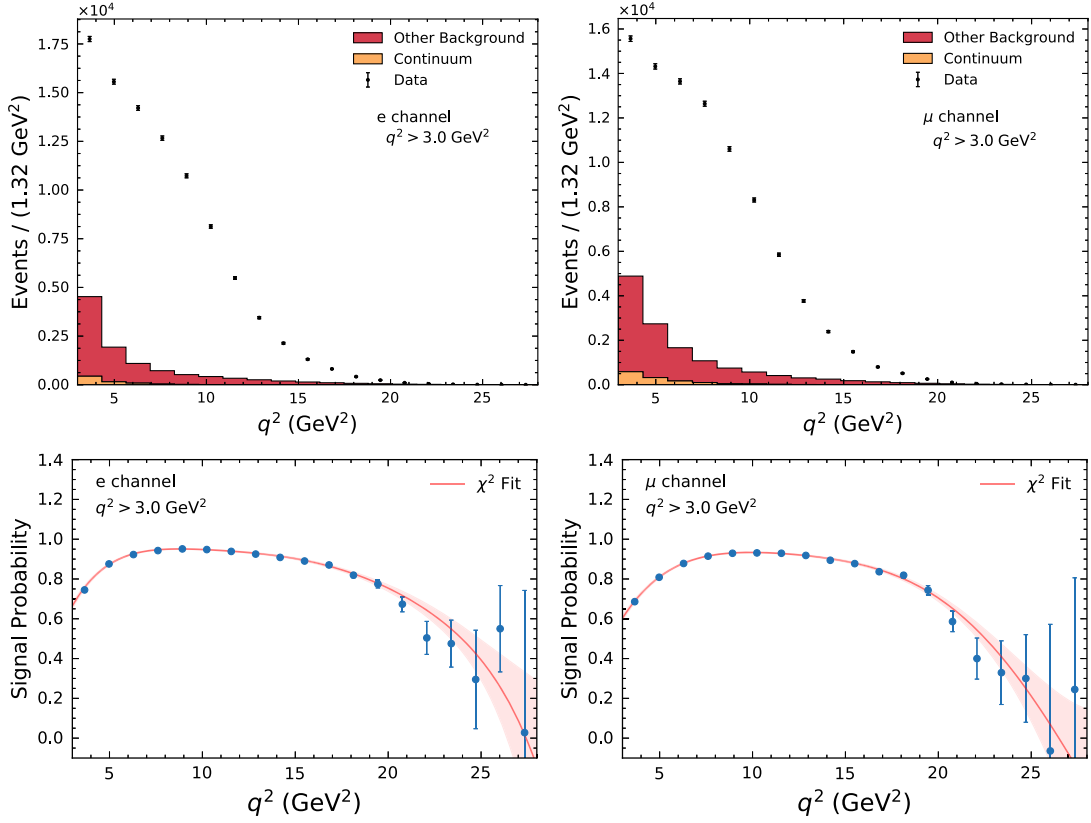


FIG. 4. The reconstructed q^2 distributions with an example q^2 threshold selection of 3.0 GeV^2 (top) for electron (left) and muon (right) candidates and the determined binned signal probabilities (bottom) are shown. The background contributions are scaled to their estimated values using the fit described in the text. The binned signal probabilities are obtained by a fit of a polynomial of a given order n (red curve).

calibrated q^2 value in the following as $q_{\text{cal } i}^{2m}$ with m being the order of the moment.

- (b) In a second step we correct for any residual bias not captured by the linear calibration function for a given moment using a q^2 threshold and moment-order dependent correction factor, which we denote as C_{cal} .
- (c) Third, we correct for acceptance and selection effects by applying a correction factor, C_{acc} . This correction is calculated for each q^2 threshold selection as well as each order of moment.

The linear calibration functions and the two correction factors are determined using simulated event samples, which are statistically independent from the simulated samples used in the background subtraction procedure. The calibration function is determined by comparing the reconstructed and generator-level moments. Figure 5 shows the first to fourth generator-level and reconstructed $B \rightarrow X_c \ell^+ \nu_\ell$ moments with different selections on the generator-level and reconstructed q^2 value for electrons. There exists a strong linear relationship for the studied order of moments, and we determine a linear calibration function of the form

$$q_{\text{cal } i}^{2m} = \frac{q_i^{2m} - a_m}{b_m}, \quad (20)$$

with a_m and b_m denoting the offset and slope for moments of the order m . The corresponding figures for muons can be found in Appendix C and show the same linear behavior. In addition, a summary of the fitted parameters of the determined calibration curves is also given in Appendix C. The residual bias correction factor C_{cal} is determined by comparing simulated samples of the calibrated and generator-level moments for each q^2 threshold selection and order of moment under study. Lastly, since different selection efficiencies are observed for different $B \rightarrow X_c \ell^+ \nu_\ell$ processes, we determine an additional factor accounting for selection and acceptance effects. The correction factor, C_{acc} , is calculated by comparing the moments of the generator-level simulated events with a sample without any selection criteria applied. Figure 6 shows the size of both calibration factors for the first to fourth q^2 moment for electrons. Both factors are close to unity and the corresponding factors for muons, displaying a similar behavior, are found in Appendix B. In addition,

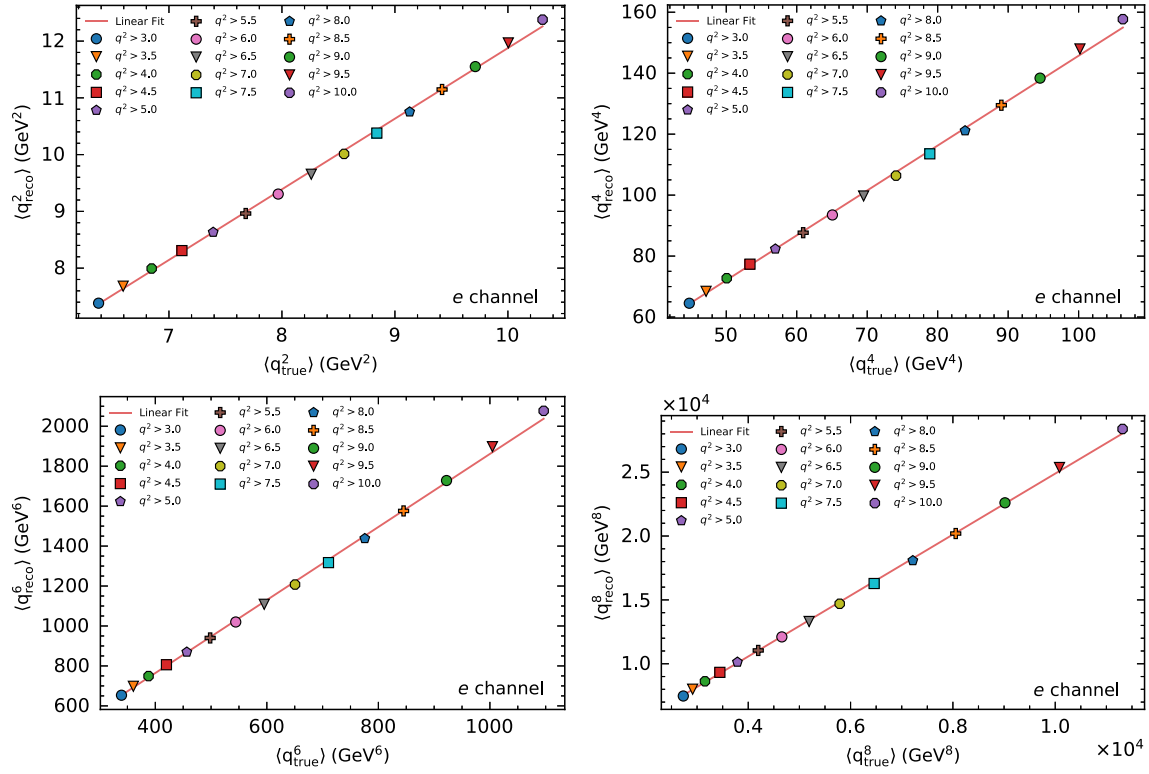


FIG. 5. The generator-level, $\langle q^2_{\text{true}} \rangle$, and reconstructed, $\langle q^2_{\text{reco}} \rangle$, $B \rightarrow X_c \ell^+ \nu_\ell$ moments are shown for the first to fourth q^2 moment for electrons. The different markers represent different threshold selections on the generator-level and reconstructed q^2 and the relation between the moments can be described with a linear parametrization of the form $a_m + \langle q^2_{\text{true}} \rangle \cdot b_m = \langle q^2_{\text{reco}} \rangle$. The corresponding plots for muons can be found in Appendix B.

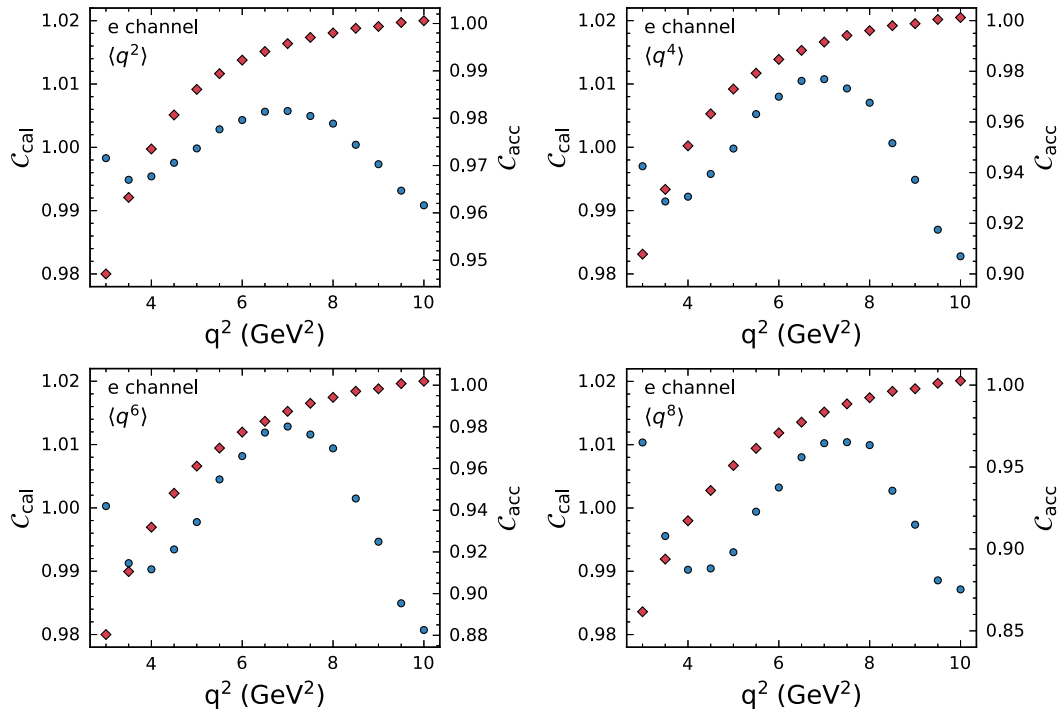


FIG. 6. The residual bias and acceptance correction factors, denoted as C_{cal} (circles) and C_{acc} (diamonds), respectively, are shown for the first to fourth q^2 moment for electrons. The corresponding plots for muons can be found in Appendix B.

selection and acceptance efficiencies for generator-level simulated $B \rightarrow X_c \ell^+ \nu_\ell$ samples are shown in Appendix D.

The q^2 moments are then given by

$$\langle q^{2m} \rangle = \frac{C_{\text{cal}} \cdot C_{\text{acc}}}{\sum_i^{\text{events}} w(q_i^2)} \times \sum_i^{\text{events}} w(q_i^2) \cdot q_{\text{cal } i}^{2m}. \quad (21)$$

Here the sums run over all selected events and q_i^2 denotes the measured four-momentum transfer squared of a given event i with a corresponding calibrated four-momentum transfer squared $q_{\text{cal } i}^{2m}$ to the power of m . The continuous signal probability $w(q_i^2)$ is calculated for each event, while the calculated moments are normalized by the sum of signal probabilities. The full background subtraction and calibration procedure was tested on ensembles of statistically independent simulated samples and no statistical significant biases in the unfolded moments are observed.

V. SYSTEMATIC UNCERTAINTIES

Several systematic uncertainties affect the measured q^2 moments and their impact on the background subtraction and calibration procedure are discussed in the following sections. The most important sources of systematic uncertainty are associated with the assumed composition of the $B \rightarrow X_c \ell^+ \nu_\ell$ process: the decays involving the higher mass states beyond the $1S$ D and D^* meson are poorly known and the composition affects the background subtraction as well as the calibration of the measured moments. In addition, we observe systematic shifts in the energy and momentum of the X system, whose size we use to estimate a q^2 scale uncertainty. Tables II and III summarize the relative statistical and systematic uncertainties on the measured q^2 moments for electron and muon final states, given in permille.

A. Background subtraction

We evaluate the uncertainties on the background subtraction by considering various sources of systematic uncertainty included as NP constraints in the binned likelihood fits. By taking into account the full experimental covariance of the background shapes when performing the χ^2 minimization, we directly propagate these sources of uncertainty into the signal probability functions, $w(q^2)$. Subsequently, we determine orthogonal variations for the parameters of the fitted polynomial curves and extract varied sets of moments to evaluate the total uncertainty (labeled “Bkg. subtraction” in Tables II and III). Specifically, we consider signal modeling uncertainties by variations of the BGL parameters and heavy quark form factors of $B \rightarrow D \ell^+ \nu_\ell$, $B \rightarrow D^* \ell^+ \nu_\ell$, and $B \rightarrow D^{**} \ell^+ \nu_\ell$ within their uncertainties. In addition, we propagate the branching fraction uncertainties, cf. Table I. The uncertainties on the $B \rightarrow X_c \ell^+ \nu_\ell$ gap branching fractions are

TABLE II. Summary of statistical and systematic uncertainties for the moments $\langle q^{2,4,6,8} \rangle$ for the electron channel. The values are given as the relative error in permille.

q^2 selection in GeV ²	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
$\langle q^2 \rangle$ in GeV ²	6.21	6.51	6.81	7.10	7.41	7.71	8.01	8.30	8.60	8.88	9.19	9.48	9.78	10.07	10.38
Stat. error (data)	1.42	1.36	1.31	1.27	1.24	1.21	1.18	1.17	1.16	1.16	1.17	1.19	1.22	1.25	1.29
Bkg. subtraction	1.06	0.76	0.57	0.40	0.37	0.41	0.48	0.55	0.60	0.63	0.67	0.70	0.77	0.83	0.91
$B \rightarrow X_u \ell^+ \nu_\ell$ BF	1.80	1.66	1.52	1.10	0.78	0.57	0.42	0.29	0.23	0.19	0.16	0.13	0.12	0.15	0.10
$B \rightarrow X_c \ell^+ \nu_\ell$ BF	4.94	4.72	4.90	4.55	4.47	4.30	3.77	3.47	3.28	2.99	2.49	1.96	1.41	1.27	1.16
Nonresonant model	13.20	11.54	10.14	8.33	6.93	5.78	4.50	3.54	2.76	2.21	1.63	1.24	0.81	0.77	0.66
$B \rightarrow X_c \ell^+ \nu_\ell$ FF	1.63	1.47	1.27	1.10	0.94	0.84	0.79	0.75	0.69	0.63	0.56	0.49	0.42	0.40	0.37
N_{tracks} res.	4.67	4.40	4.17	3.91	3.72	3.50	3.31	3.13	2.95	2.83	2.66	2.50	2.40	2.24	2.13
N_γ res.	0.50	0.49	0.44	0.43	0.40	0.37	0.36	0.36	0.36	0.35	0.33	0.31	0.31	0.32	0.32
$F_{\text{miss}} - \mathbf{p}_{\text{miss}} $ shape	0.71	0.69	0.73	0.68	0.77	0.86	0.89	0.94	0.97	1.06	1.14	1.20	1.30	1.29	1.31
q^2 scale	8.98	6.77	6.12	5.77	5.70	5.50	5.12	4.86	4.72	4.62	4.25	4.33	3.91	3.90	3.94
MC nonclosure	0.06	0.09	0.03	0.01	0.01	0.03	0.02	0.01	0.01	0.01	0.00	0.03	0.00	0.01	0.00
Cal. function	0.13	0.08	0.03	0.02	0.08	0.12	0.17	0.22	0.26	0.30	0.34	0.38	0.42	0.45	0.49
Stat. bias corr.	1.22	1.18	1.14	1.11	1.08	1.06	1.04	1.03	1.02	1.02	1.02	1.03	1.05	1.07	1.10
PID eff.	0.17	0.16	0.14	0.13	0.13	0.11	0.10	0.09	0.09	0.08	0.08	0.08	0.07	0.06	0.05
Track eff.	0.42	0.39	0.36	0.33	0.31	0.29	0.27	0.25	0.23	0.22	0.20	0.19	0.18	0.16	0.15
$B^0/B^\pm$ tag eff.	0.21	0.25	0.35	0.41	0.52	0.44	0.45	0.37	0.55	0.74	0.72	0.73	0.60	0.47	0.45
Sys. error (total)	17.62	15.11	13.72	11.98	10.88	9.89	8.66	7.83	7.27	6.85	6.16	5.87	5.29	5.17	5.13

(Table continued)

TABLE II. (Continued)

q^2 selection in GeV ²	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
Total rel. error in % _o	17.68	15.17	13.78	12.04	10.95	9.96	8.74	7.92	7.36	6.95	6.27	5.99	5.43	5.32	5.29
$\langle q^4 \rangle$ in GeV ⁴	42.99	46.27	49.74	53.41	57.42	61.64	65.96	70.46	75.15	79.92	85.20	90.50	96.07	101.73	108.10
Stat. error (data)	3.23	3.13	3.04	2.96	2.88	2.82	2.77	2.74	2.72	2.72	2.74	2.78	2.84	2.92	3.01
Bkg. subtraction	1.78	1.35	1.14	0.99	1.06	1.17	1.33	1.49	1.58	1.65	1.71	1.78	1.93	2.05	2.23
$B \rightarrow X_u \ell^+ \nu$ BF	3.84	3.52	3.16	2.29	1.64	1.21	0.90	0.63	0.52	0.43	0.37	0.31	0.29	0.34	0.24
$B \rightarrow X_c \ell^+ \nu$ BF	9.45	9.45	9.93	9.48	9.41	9.10	8.13	7.54	7.11	6.45	5.40	4.31	3.18	2.88	2.62
Nonresonant model	25.19	22.36	19.84	16.57	13.95	11.72	9.25	7.36	5.81	4.68	3.49	2.69	1.82	1.73	1.49
$B \rightarrow X_c \ell^+ \nu$ FF	3.13	2.84	2.51	2.22	1.95	1.81	1.72	1.64	1.52	1.41	1.26	1.12	0.98	0.92	0.85
N_{tracks} res.	10.44	9.95	9.49	8.98	8.56	8.09	7.67	7.26	6.84	6.52	6.13	5.75	5.48	5.11	4.84
N_γ res.	1.18	1.16	1.07	1.05	1.00	0.93	0.91	0.90	0.90	0.86	0.82	0.79	0.78	0.81	0.80
$E_{\text{miss}} - \mathbf{p}_{\text{miss}} $ shape	2.17	2.11	2.15	2.05	2.20	2.32	2.36	2.43	2.47	2.62	2.75	2.83	3.00	2.95	2.93
q^2 scale	18.61	15.14	14.05	13.28	13.10	12.57	11.79	11.15	10.79	10.59	9.81	9.95	9.04	9.03	9.15
MC nonclosure	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.00	0.00	0.00
Cal. function	0.27	0.15	0.03	0.08	0.20	0.31	0.42	0.52	0.61	0.70	0.79	0.88	0.95	1.03	1.10
Stat. bias corr.	2.68	2.63	2.57	2.52	2.47	2.43	2.40	2.37	2.35	2.35	2.35	2.37	2.40	2.45	2.50
PID eff.	0.36	0.33	0.31	0.29	0.28	0.24	0.22	0.20	0.20	0.19	0.18	0.17	0.14	0.12	0.10
Track eff.	0.90	0.85	0.80	0.75	0.71	0.66	0.62	0.58	0.53	0.50	0.46	0.43	0.40	0.36	0.34
$B^0/B^\pm$ tag eff.	0.81	0.89	1.03	1.12	1.30	1.14	1.15	1.00	1.32	1.65	1.60	1.58	1.30	1.03	0.97
Sys. error (total)	34.95	30.89	28.49	25.42	23.45	21.54	19.21	17.53	16.35	15.48	14.04	13.42	12.18	11.90	11.82
Total rel. error in % _o	35.10	31.04	28.65	25.59	23.63	21.72	19.41	17.74	16.58	15.72	14.30	13.71	12.50	12.25	12.20
$\langle q^6 \rangle$ in GeV ⁶	326.23	355.51	387.59	423.92	465.17	510.57	558.65	610.98	667.93	728.29	797.82	870.48	949.14	1031.95	1128.69
Stat. error (data)	5.84	5.70	5.56	5.40	5.26	5.13	5.03	4.96	4.91	4.89	4.89	4.94	5.03	5.14	5.27
Bkg. subtraction	2.55	2.18	2.20	2.14	2.38	2.54	2.81	3.05	3.14	3.22	3.30	3.37	3.62	3.81	4.11
$B \rightarrow X_u \ell^+ \nu$ BF	6.43	5.82	5.07	3.66	2.62	1.97	1.45	1.04	0.88	0.74	0.65	0.55	0.51	0.59	0.42
$B \rightarrow X_c \ell^+ \nu$ BF	13.83	14.41	15.25	14.81	14.75	14.32	12.97	12.09	11.37	10.31	8.70	7.01	5.31	4.81	4.38
Nonresonant model	35.93	32.23	28.84	24.43	20.79	17.60	14.08	11.34	9.04	7.33	5.55	4.33	3.03	2.85	2.48
$B \rightarrow X_c \ell^+ \nu$ FF	4.42	4.10	3.73	3.41	3.12	2.96	2.84	2.72	2.54	2.36	2.12	1.90	1.68	1.58	1.44
N_{tracks} res.	17.45	16.76	16.09	15.31	14.63	13.87	13.17	12.46	11.75	11.16	10.46	9.79	9.28	8.63	8.12
N_γ res.	2.15	2.11	1.99	1.95	1.87	1.76	1.72	1.69	1.68	1.61	1.54	1.48	1.47	1.49	1.47
$E_{\text{miss}} - \mathbf{p}_{\text{miss}} $ shape	4.56	4.44	4.43	4.27	4.41	4.53	4.53	4.58	4.59	4.75	4.88	4.93	5.10	4.96	4.86
q^2 scale	29.62	25.28	23.94	22.61	22.35	21.35	20.15	19.01	18.35	18.07	16.85	17.06	15.55	15.56	15.83
MC nonclosure	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cal. function	0.42	0.22	0.02	0.18	0.38	0.57	0.75	0.92	1.09	1.24	1.39	1.53	1.65	1.77	1.87
Stat. bias corr.	4.68	4.60	4.52	4.44	4.36	4.29	4.23	4.18	4.14	4.11	4.10	4.12	4.15	4.20	4.27
PID eff.	0.57	0.53	0.50	0.47	0.45	0.40	0.37	0.33	0.33	0.30	0.29	0.27	0.23	0.19	0.16
Track eff.	1.47	1.40	1.33	1.25	1.18	1.10	1.04	0.97	0.90	0.84	0.78	0.72	0.67	0.61	0.57
$B^0/B^\pm$ tag eff.	1.79	1.87	2.03	2.13	2.34	2.11	2.09	1.86	2.29	2.71	2.58	2.50	2.06	1.64	1.52
Sys. error (total)	52.77	47.68	44.62	40.49	37.88	35.07	31.80	29.25	27.40	26.06	23.83	22.86	20.84	20.39	20.28
Total rel. error in % _o	53.09	48.01	44.96	40.85	38.24	35.44	32.20	29.67	27.84	26.52	24.33	23.39	21.44	21.03	20.95
$\langle q^8 \rangle$ in GeV ⁸	2663.52	2914.20	3192.80	3531.01	3925.20	4375.31	4862.26	5414.29	6036.81	6723.88	7545.39	8439.78	9431.26	10512.44	11818.99

(Table continued)

TABLE II. (Continued)

q^2 selection in GeV^2	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
Stat. error (data)	9.77	9.58	9.38	9.11	8.84	8.58	8.40	8.23	8.09	8.00	7.93	7.93	8.01	8.13	8.25
Bkg. subtraction	4.02	3.88	4.31	4.24	4.67	4.85	5.20	5.52	5.51	5.58	5.65	5.71	6.05	6.31	6.77
$B \rightarrow X_u \ell \nu$ BF	9.95	8.85	7.47	5.32	3.79	2.87	2.11	1.52	1.32	1.13	1.01	0.87	0.80	0.91	0.64
$B \rightarrow X_c \ell \nu$ BF	19.15	20.11	21.08	20.54	20.40	19.78	18.08	16.93	15.92	14.46	12.29	10.03	7.78	7.06	6.42
Nonresonant model	45.70	41.23	37.11	31.73	27.22	23.21	18.80	15.31	12.33	10.08	7.73	6.11	4.40	4.13	3.61
$B \rightarrow X_c \ell \nu$ FF	5.57	5.32	5.03	4.75	4.49	4.33	4.19	4.01	3.76	3.50	3.16	2.85	2.54	2.37	2.16
N_{tracks} res.	25.81	24.89	24.00	22.90	21.91	20.80	19.77	18.71	17.64	16.71	15.64	14.60	13.78	12.78	11.97
N_γ res.	3.51	3.45	3.30	3.22	3.10	2.94	2.87	2.81	2.77	2.66	2.55	2.44	2.41	2.42	2.37
$E_{\text{miss}} - \mathbf{p}_{\text{miss}} $ shape	7.95	7.73	7.63	7.37	7.46	7.52	7.44	7.42	7.34	7.44	7.51	7.47	7.59	7.30	7.07
q^2 scale	42.51	37.31	35.84	33.74	33.46	31.82	30.22	28.43	27.42	27.12	25.46	25.75	23.55	23.65	24.18
MC nonclosure	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cal. function	0.62	0.31	0.01	0.31	0.62	0.93	1.20	1.47	1.72	1.95	2.18	2.39	2.57	2.74	2.89
Stat. bias corr.	7.59	7.45	7.31	7.16	7.02	6.89	6.76	6.65	6.56	6.48	6.43	6.40	6.40	6.44	6.51
PID eff.	0.79	0.74	0.70	0.66	0.63	0.56	0.53	0.48	0.47	0.43	0.41	0.37	0.32	0.27	0.22
Track eff.	2.11	2.02	1.93	1.83	1.73	1.62	1.53	1.43	1.33	1.24	1.15	1.06	0.98	0.89	0.83
$B^0/B^\pm$ tag eff.	3.03	3.10	3.26	3.34	3.56	3.24	3.18	2.86	3.35	3.81	3.59	3.43	2.81	2.24	2.05
Sys. error (total)	72.27	66.18	62.60	57.37	54.26	50.52	46.44	42.99	40.44	38.66	35.64	34.31	31.44	30.82	30.73
Total rel. error in %	72.93	66.86	63.30	58.09	54.97	51.25	47.19	43.77	41.24	39.48	36.51	35.21	32.44	31.87	31.81

TABLE III. Summary of statistical and systematic uncertainties for the moments $\langle q^{2,4,6,8} \rangle$ for the muon channel. The values are given as the relative error in permille.

q^2 selection in GeV^2	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
$\langle q^2 \rangle$ in GeV^2	6.25	6.54	6.83	7.13	7.42	7.72	8.02	8.32	8.61	8.91	9.21	9.51	9.80	10.09	10.40
Stat. error (data)	1.51	1.45	1.39	1.34	1.30	1.27	1.24	1.22	1.21	1.20	1.20	1.22	1.23	1.26	1.29
Bkg. subtraction	1.34	1.12	0.90	0.71	0.59	0.53	0.49	0.49	0.57	0.63	0.65	0.70	0.76	0.77	0.82
$B \rightarrow X_u \ell \nu$ BF	2.18	2.04	1.75	1.48	1.35	1.04	0.76	0.54	0.38	0.29	0.19	0.16	0.12	0.05	0.05
$B \rightarrow X_c \ell \nu$ BF	4.82	5.02	5.14	5.14	5.05	5.00	4.67	4.05	3.51	3.11	2.66	2.21	1.75	1.36	1.16
Nonresonant model	14.25	12.72	11.04	9.28	7.83	6.62	5.42	4.00	3.02	2.28	1.65	1.43	1.04	0.86	0.78
$B \rightarrow X_c \ell \nu$ FF	1.43	1.30	1.16	1.03	0.91	0.85	0.82	0.74	0.69	0.62	0.54	0.48	0.42	0.39	0.35
N_{tracks} res.	5.66	5.31	4.96	4.65	4.36	4.06	3.78	3.52	3.29	3.06	2.85	2.66	2.51	2.38	2.20
N_γ res.	0.39	0.38	0.34	0.31	0.30	0.28	0.30	0.31	0.32	0.28	0.27	0.26	0.25	0.27	0.29
$E_{\text{miss}} - \mathbf{p}_{\text{miss}} $ shape	1.29	1.26	1.21	1.17	1.15	1.11	1.04	1.05	1.06	1.09	1.16	1.20	1.30	1.33	1.29
q^2 scale	9.48	7.15	6.65	6.65	6.12	5.91	5.83	5.48	5.26	4.69	4.27	4.42	3.91	3.94	4.38
MC nonclosure	0.19	0.11	0.12	0.11	0.11	0.05	0.05	0.06	0.08	0.07	0.11	0.04	0.04	0.06	0.02
Cal. function	0.13	0.08	0.03	0.02	0.07	0.12	0.17	0.22	0.26	0.31	0.35	0.39	0.43	0.47	0.51
Stat. bias corr.	1.32	1.27	1.23	1.19	1.16	1.13	1.10	1.08	1.07	1.06	1.06	1.06	1.07	1.09	1.11
PID eff.	0.16	0.14	0.14	0.13	0.13	0.12	0.11	0.10	0.10	0.10	0.09	0.08	0.08	0.07	0.06
Track eff.	0.44	0.42	0.39	0.36	0.34	0.31	0.29	0.27	0.25	0.23	0.21	0.20	0.18	0.17	0.15
$B^0/B^\pm$ tag eff.	0.46	0.58	0.50	0.44	0.51	0.40	0.28	0.34	0.36	0.38	0.29	0.23	0.20	0.12	0.47
Sys. error (total)	18.99	16.65	15.03	13.62	12.22	11.19	10.17	8.86	7.97	7.06	6.30	6.09	5.44	5.27	5.50

(Table continued)

TABLE III. (Continued)

q^2 selection in GeV ²	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
Total rel. error in % _o	19.05	16.71	15.09	13.68	12.29	11.26	10.25	8.94	8.06	7.16	6.41	6.21	5.58	5.42	5.65
$\langle q^4 \rangle$ in GeV ⁴	43.52	46.73	50.16	53.81	57.65	61.75	66.09	70.71	75.36	80.34	85.63	90.96	96.39	102.18	108.50
Stat. error (data)	3.45	3.33	3.22	3.12	3.03	2.95	2.88	2.83	2.80	2.79	2.79	2.81	2.86	2.91	2.97
Bkg. subtraction	2.41	2.10	1.76	1.48	1.36	1.34	1.30	1.34	1.53	1.65	1.67	1.80	1.90	1.90	2.00
$B \rightarrow X_u \ell^+ \nu$ BF	4.66	4.39	3.80	3.20	2.90	2.23	1.61	1.14	0.81	0.61	0.39	0.34	0.26	0.11	0.10
$B \rightarrow X_c \ell^+ \nu$ BF	9.55	10.30	10.68	10.78	10.67	10.50	9.84	8.63	7.52	6.67	5.69	4.73	3.78	2.97	2.52
Nonresonant model	27.34	24.67	21.67	18.48	15.72	13.36	11.00	8.24	6.30	4.82	3.53	3.05	2.27	1.88	1.69
$B \rightarrow X_c \ell^+ \nu$ FF	2.73	2.51	2.29	2.08	1.90	1.80	1.75	1.61	1.49	1.36	1.20	1.07	0.95	0.87	0.79
N_{tracks} res.	12.45	11.81	11.13	10.51	9.89	9.25	8.63	8.04	7.51	6.98	6.49	6.04	5.66	5.34	4.90
N_γ res.	0.95	0.92	0.85	0.79	0.78	0.74	0.77	0.78	0.78	0.71	0.69	0.66	0.64	0.67	0.71
$E_{\text{miss}} - \mathbf{p}_{\text{miss}} $ shape	3.35	3.26	3.12	3.01	2.92	2.81	2.66	2.65	2.64	2.67	2.76	2.81	2.95	2.98	2.85
q^2 scale	19.97	16.30	15.38	15.26	14.19	13.68	13.33	12.59	11.97	10.81	9.86	10.03	8.97	9.17	10.05
MC nonclosure	0.06	0.04	0.04	0.03	0.03	0.02	0.02	0.02	0.03	0.02	0.03	0.01	0.01	0.01	0.01
Cal. function	0.27	0.16	0.04	0.07	0.19	0.30	0.41	0.52	0.62	0.71	0.81	0.90	0.98	1.06	1.14
Stat. bias corr.	2.92	2.85	2.77	2.70	2.64	2.58	2.53	2.48	2.45	2.43	2.42	2.42	2.43	2.46	2.49
PID eff.	0.35	0.32	0.31	0.29	0.29	0.26	0.25	0.23	0.22	0.22	0.20	0.18	0.18	0.16	0.13
Track eff.	0.97	0.91	0.86	0.81	0.76	0.70	0.65	0.61	0.56	0.52	0.48	0.44	0.41	0.38	0.34
$B^0/B^\pm$ tag eff.	0.92	1.09	0.95	0.84	0.94	0.74	0.51	0.59	0.62	0.63	0.43	0.29	0.22	0.42	1.12
Sys. error (total)	38.08	34.23	31.41	28.91	26.30	24.24	22.18	19.65	17.77	15.90	14.27	13.73	12.36	12.08	12.51
Total rel. error in % _o	38.23	34.39	31.58	29.08	26.48	24.42	22.37	19.85	17.99	16.14	14.54	14.02	12.69	12.43	12.86
$\langle q^6 \rangle$ in GeV ⁶	331.70	360.44	392.58	428.49	467.70	511.77	560.28	614.62	670.90	733.57	803.64	876.13	952.91	1038.72	1135.12
Stat. error (data)	6.22	6.03	5.84	5.66	5.49	5.34	5.20	5.08	5.01	4.96	4.94	4.96	5.02	5.07	5.16
Bkg. subtraction	3.62	3.28	2.94	2.67	2.67	2.77	2.71	2.77	3.08	3.26	3.24	3.48	3.57	3.52	3.68
$B \rightarrow X_u \ell^+ \nu$ BF	7.78	7.30	6.34	5.26	4.74	3.62	2.57	1.81	1.27	0.95	0.59	0.52	0.41	0.15	0.14
$B \rightarrow X_c \ell^+ \nu$ BF	14.82	16.14	16.68	16.81	16.67	16.30	15.29	13.54	11.86	10.53	9.00	7.50	6.04	4.78	4.06
Nonresonant model	39.15	35.58	31.54	27.20	23.35	19.95	16.53	12.55	9.72	7.52	5.59	4.82	3.64	3.02	2.71
$B \rightarrow X_c \ell^+ \nu$ FF	3.86	3.65	3.42	3.19	3.00	2.88	2.82	2.61	2.44	2.23	1.98	1.78	1.58	1.44	1.30
N_{tracks} res.	20.45	19.54	18.55	17.59	16.63	15.59	14.58	13.60	12.70	11.79	10.93	10.14	9.46	8.85	8.10
N_γ res.	1.77	1.73	1.63	1.54	1.52	1.45	1.46	1.47	1.46	1.34	1.30	1.25	1.21	1.24	1.29
$E_{\text{miss}} - \mathbf{p}_{\text{miss}} $ shape	6.30	6.10	5.85	5.63	5.43	5.22	4.94	4.87	4.79	4.76	4.83	4.82	4.95	4.92	4.65
q^2 shift	32.17	27.63	26.32	26.02	24.37	23.45	22.66	21.49	20.24	18.50	16.91	16.97	15.32	15.86	17.17
MC nonclosure	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00
Cal. function	0.42	0.24	0.04	0.16	0.35	0.54	0.73	0.91	1.08	1.24	1.40	1.55	1.68	1.80	1.92
Stat. bias corr.	5.10	4.98	4.86	4.75	4.64	4.53	4.43	4.34	4.27	4.21	4.17	4.15	4.14	4.16	4.19
PID eff.	0.57	0.53	0.51	0.48	0.48	0.44	0.41	0.39	0.36	0.36	0.33	0.29	0.28	0.25	0.21
Track eff.	1.57	1.50	1.42	1.34	1.26	1.18	1.09	1.02	0.94	0.87	0.80	0.73	0.67	0.62	0.56
$B^0/B^\pm$ tag eff.	1.16	1.36	1.18	1.03	1.15	0.85	0.53	0.63	0.65	0.63	0.32	0.10	0.02	0.97	1.98
Sys. error (total)	58.03	53.09	49.34	45.97	42.29	39.22	36.10	32.45	29.49	26.65	24.04	23.03	20.88	20.59	21.17
Total rel. error in % _o	58.36	53.43	49.68	46.32	42.65	39.58	36.47	32.84	29.91	27.11	24.54	23.56	21.47	21.21	21.79
$\langle q^8 \rangle$ in GeV ⁸	2717.22	2963.88	3248.31	3578.45	3947.44	4384.73	4878.23	5458.95	6072.92	6780.95	7616.67	8497.60	9466.03	10603.31	11917.23

(Table continued)

TABLE III. (Continued)

q^2 selection in GeV^2	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
Stat. error (data)	10.35	10.07	9.78	9.47	9.19	8.89	8.63	8.36	8.19	8.05	7.94	7.91	7.94	7.95	7.99
Bkg. subtraction	5.57	5.25	4.98	4.80	4.99	5.23	5.02	5.06	5.51	5.71	5.58	6.00	5.96	5.81	6.02
$B \rightarrow X_u \ell \nu$ BF	11.94	11.10	9.61	7.82	7.00	5.31	3.66	2.53	1.76	1.30	0.79	0.69	0.59	0.20	0.16
$B \rightarrow X_c \ell \nu$ BF	21.51	22.91	23.24	23.14	22.84	22.14	20.76	18.50	16.31	14.53	12.43	10.40	8.44	6.74	5.74
Nonresonant model	49.93	45.52	40.56	35.22	30.45	26.13	21.80	16.75	13.12	10.26	7.73	6.66	5.10	4.25	3.79
$B \rightarrow X_c \ell \nu$ FF	4.91	4.76	4.60	4.40	4.23	4.12	4.03	3.75	3.52	3.23	2.88	2.59	2.31	2.09	1.89
N_{tracks} res.	29.72	28.51	27.15	25.82	24.47	22.99	21.54	20.09	18.76	17.40	16.09	14.89	13.83	12.86	11.73
N_{γ} res.	2.95	2.89	2.75	2.62	2.58	2.46	2.46	2.44	2.39	2.22	2.16	2.07	2.00	2.01	2.06
$E_{\text{miss}} - \mathbf{p}_{\text{miss}} $ shape	10.18	9.83	9.42	9.05	8.69	8.33	7.89	7.70	7.50	7.35	7.33	7.21	7.26	7.11	6.66
q^2 scale	46.61	41.26	39.53	39.00	36.70	35.23	33.82	32.22	30.11	27.83	25.47	25.28	23.04	24.16	25.90
MC nonclosure	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cal. function	0.63	0.34	0.04	0.26	0.55	0.86	1.14	1.42	1.68	1.92	2.17	2.37	2.56	2.74	2.90
Stat. bias corr.	8.21	8.00	7.79	7.59	7.38	7.18	6.99	6.83	6.67	6.54	6.43	6.35	6.29	6.27	6.27
PID eff.	0.81	0.77	0.74	0.70	0.69	0.64	0.60	0.56	0.52	0.51	0.47	0.42	0.40	0.35	0.30
Track eff.	2.25	2.16	2.05	1.95	1.84	1.72	1.60	1.49	1.38	1.27	1.16	1.06	0.98	0.89	0.81
$B^0/B^\pm$ tag eff.	0.97	1.19	0.98	0.80	0.94	0.58	0.18	0.29	0.30	0.26	0.14	0.42	0.59	1.79	3.06
Sys. error (total)	79.98	73.90	69.18	64.95	60.23	56.09	51.86	47.19	43.06	39.32	35.61	34.06	31.06	30.93	31.65
Total rel. error in %	80.64	74.58	69.87	65.64	60.93	56.79	52.58	47.93	43.83	40.13	36.48	34.97	32.06	31.93	32.64

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 particle-identification uncertainties on the background shapes, and the overall tracking efficiency uncertainty. The statistical uncertainty on all generated MC samples is also evaluated and propagated into the systematic errors. The $B \rightarrow X_u \ell^+ \nu_\ell$ background component is varied within its branching fraction uncertainty (“ $B \rightarrow X_u \ell \nu$ BF”).

B. Calibration

For the calibration we change the composition of the assumed $B \rightarrow X_c \ell^+ \nu_\ell$ process and redetermine the calibration functions as well as the bias and acceptance correction factors (“ $B \rightarrow X_c \ell \nu$ BF”). Specifically, for the $B \rightarrow D^{**} \ell^+ \nu_\ell$ and gap contributions, we redetermine these factors by completely removing their respective contributions, to account for the poor knowledge of their precise decay processes. We vary the $B \rightarrow D \ell^+ \nu_\ell$ and $B \rightarrow D^* \ell^+ \nu_\ell$ within their respective branching fractions. In addition, we evaluate the impact of the modeling of nonresonant decays by replacing the model described in Sec. II with a model based on the equidistribution of all final-state particles in phase space (“Nonresonant model”). The signal modeling is evaluated in a similar manner as described above by using the variations of the BGL parameters and heavy quark form factors (“ $B \rightarrow X_c \ell \nu$ FF”). To estimate an uncertainty associated with the modeling of the resolution, we calculate binwise weight functions in a signal enriched region by comparing distributions of the number of charged particles and neutral clusters in the X system to the number observed in the recorded collision events. The weight functions are subsequently applied to the simulated MC samples and we redetermine the calibration functions and correction factors (“ N_{tracks} res.” and “ N_γ res.”). Additionally, we employ a similar strategy to take into account a potential mismodeling of the shape of the missing energy and momentum distribution (“ $E_{\text{miss}} - |\mathbf{p}_{\text{miss}}|$ shape”). To estimate an uncertainty on the modeling of the scale of q^2 (“ q^2 scale”), we shift the reconstructed value in the simulation by 1%, and redetermine the background subtraction weights and the calibration functions. The shift in q^2 corresponds to the observed differences in the mean values of the energy and momentum of the reconstructed X system, and we take the full difference to the moments determined with no correction as an uncertainty. The additional remaining bias due to the moment extraction method is estimated by making use of the ensemble tests described in Sec. IV (“MC nonclosure”). We propagate the statistical uncertainty due to the determination of the linear calibration functions by determining orthogonal variations for each of the parameters of the fitted curves and extracting new, varied sets of moments (“Cal. function”). Additionally, we

estimate the statistical uncertainty due to the bias and acceptance correction factors by varying the correction factors within one standard deviation and repeating the moment calculation (“Stat. bias corr.”). The impact of the efficiency of the particle-identification uncertainties (“PID eff.”), and the overall track finding efficiency uncertainty (“Track eff.”), on the calibration curves are estimated by deriving varied calibrations for each efficiency uncertainty under consideration. Lastly, we account for the different reconstruction efficiencies of neutral and charged B mesons in data and simulated samples. We correct for the difference

in MC and repeat the calibration procedure to extract varied sets of q^2 moments (“ $B^0/B^\pm$ tag eff.”) and take the full difference as the systematic uncertainty.

C. Statistical correlations

The statistical correlation of the measured moments is determined using a bootstrapping approach [61,62]: replicas of the measured collision datasets are created using sampling by replacement and the entire analysis procedure is repeated to estimate the statistical correlations between

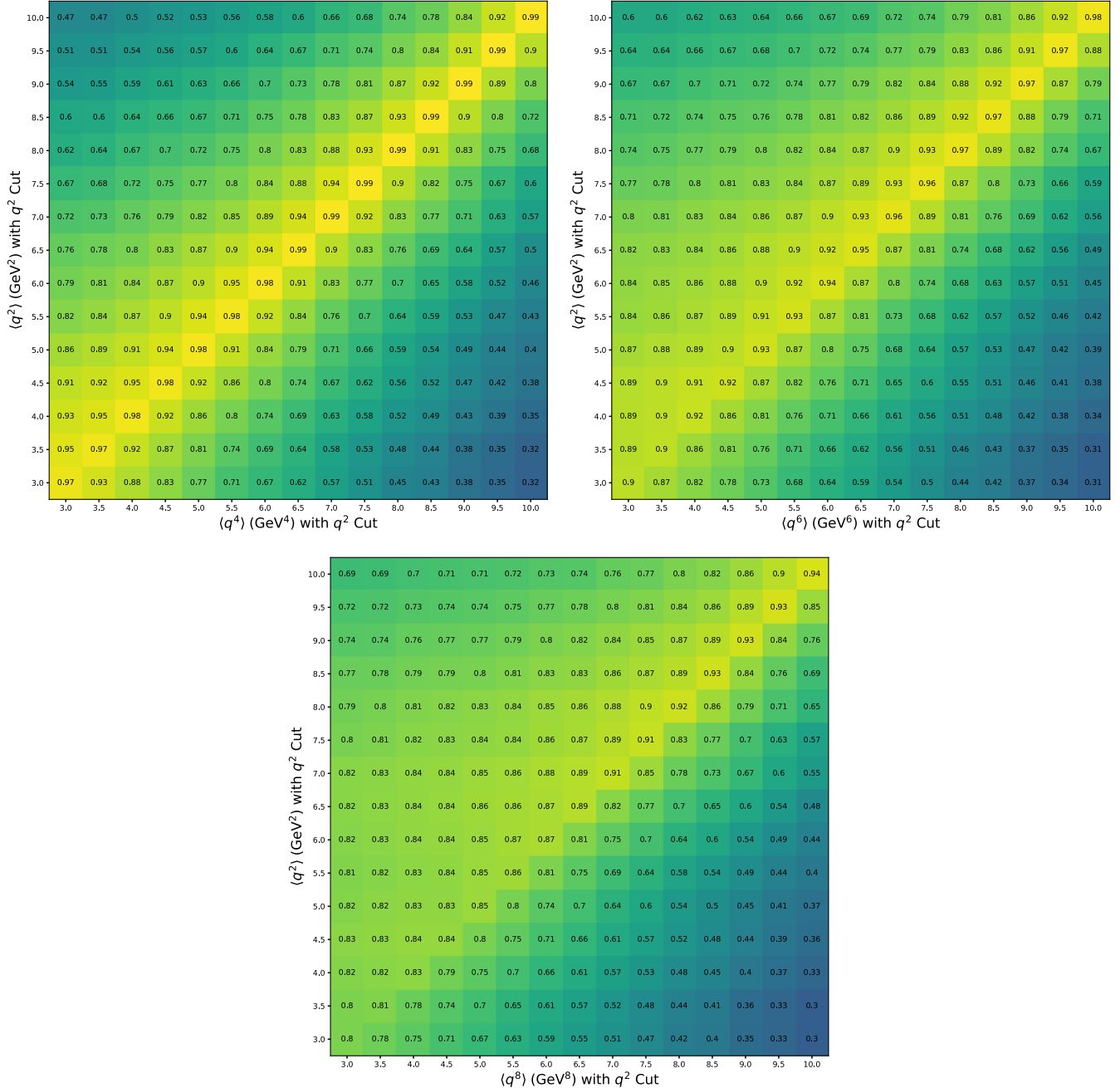


FIG. 7. Statistical correlations between the first and the second, third, and fourth moment for the electron final state.

the different moments. An example for the electron final state is shown in Fig. 7. Typically moments of the same order with similar q^2 threshold selections are highly correlated. The moments of higher order with identical q^2 threshold selections contain more independent information the higher the difference in the order.

D. Systematic correlations

We estimate the systematic covariance from the efficiency of particle identification by creating replicas of the data-driven efficiency corrections, while taking into account correlated and uncorrelated systematic uncertainty sources. We fully correlate or anticorrelate all other identical sources of systematic uncertainty across all q^2 threshold selections and order of moments. For the comparison between the measured electron and muon q^2 moments, we assume that all uncertainties are fully correlated between both sets of measurements with the exception of the contributions that are derived independently for the two different measurements. These sources of systematic uncertainty include the uncertainty on the background subtraction, the MC nonclosure error, the uncertainty on the calibration function, the statistical errors on the bias and acceptance correction factors, the particle-identification efficiency uncertainty, the modeling uncertainty of the resolution, and the modeling of the missing energy and momentum distribution. The full systematic covariance matrix is then determined by combining the statistical and all the individual systematic covariance matrices. These correlation matrices for the moments for both the electron and muon final states are given in Appendix E.

E. Summary

The dominant systematic uncertainties stem from the uncertainty associated with the modeling of the $B \rightarrow X_c \ell^+ \nu_\ell$ composition, especially the nonresonant contributions. The estimated uncertainties associated with these sources for the first moment, with a minimum threshold selection of $q^2 > 3.0 \text{ GeV}^2$ for the electron final state, are found to be 0.49% and 1.32%, respectively. This is followed by the uncertainty associated with the modeling of the number of charged particles in the X system, which remains a leading systematic uncertainty across all q^2 selections for both the muon and electron final states. On the other hand, the uncertainty due to the modeling of the $B \rightarrow X_u \ell^+ \nu_\ell$ background component is found to be a leading systematic uncertainty mainly for low q^2 selections. While the statistical uncertainties due to the additional correction factors and the determination of the linear calibration functions are small contributions to the overall systematic error for low q^2 selections, these sources increase as the selection criteria progress to higher values of q^2 and the number of events gradually decrease. A

similar trend is observed for the uncertainty due to the modeling of the missing energy and momentum, as well as the uncertainty due to the different reconstruction efficiencies of neutral and charged B mesons. Conversely, the uncertainties involving the efficiency of the particle identification, the overall tracking efficiency and the modeling of the number of neutral clusters in the X system are relatively small for the first selections and gradually decrease as tighter q^2 threshold selections are imposed. The smallest source of systematic error across all q^2 selections is the estimated residual bias due to the extraction method.

VI. RESULTS

The measured q^2 moments for electrons and muons along with a detailed breakdown of the systematic uncertainties are given in Tables II and III. A determination of $|V_{cb}|$ from the measured moments is beyond the scope of this paper and left for future work, as the method proposed in Ref. [35] has not yet been fully implemented and no public code exists. Furthermore, the treatment of theoretical uncertainties on the extraction of $|V_{cb}|$ is central due to the high precision of the measured q^2 moments.

The first moment with the loosest threshold selection of $q^2 > 3.0 \text{ GeV}^2$ is about one standard deviation higher in data than in our simulated samples. Figure 8 shows the comparison for both the electron and muon measurement for this and higher selection criteria on q^2 and also for higher moments. The generator-level moments include uncertainties on the $B \rightarrow X_c \ell^+ \nu_\ell$ composition and form factor variations. The higher the threshold selection on q^2 , the better the agreement becomes. This picture is consistent between the electron and muon channels and also for higher moments. This indicates that the $B \rightarrow X_c \ell^+ \nu_\ell$ composition used tends to overestimate the number of signal events at low q^2 values, which also has been reported by Ref. [63]: there the inclusive lepton spectrum is analyzed using a $B \rightarrow X_c \ell^+ \nu_\ell$ model similar to the one used in this measurement. The data prefer to increase the yield of the $B \rightarrow X_c \ell^+ \nu_\ell$ constituents that produce a more energetic lepton momentum spectrum, resulting in a higher value of the first moment of q^2 . Similar observations have been reported by Ref. [64], in which the moments of the lepton energy, hadronic mass, and hadronic energy spectra were analyzed to determine the exclusive composition of $B \rightarrow X_c \ell^+ \nu_\ell$. The direct measurement of the q^2 moments reported here support this picture.

We test the expectation of lepton flavor universality of identical moments. Figure 8 compares the measured moments for electrons and muons. In the shown ratio many of the systematic uncertainties cancel and we observe no deviation from the expectation of unity. In addition, we estimate a qualitative measure to test the agreement between the electron and muon moment measurement by

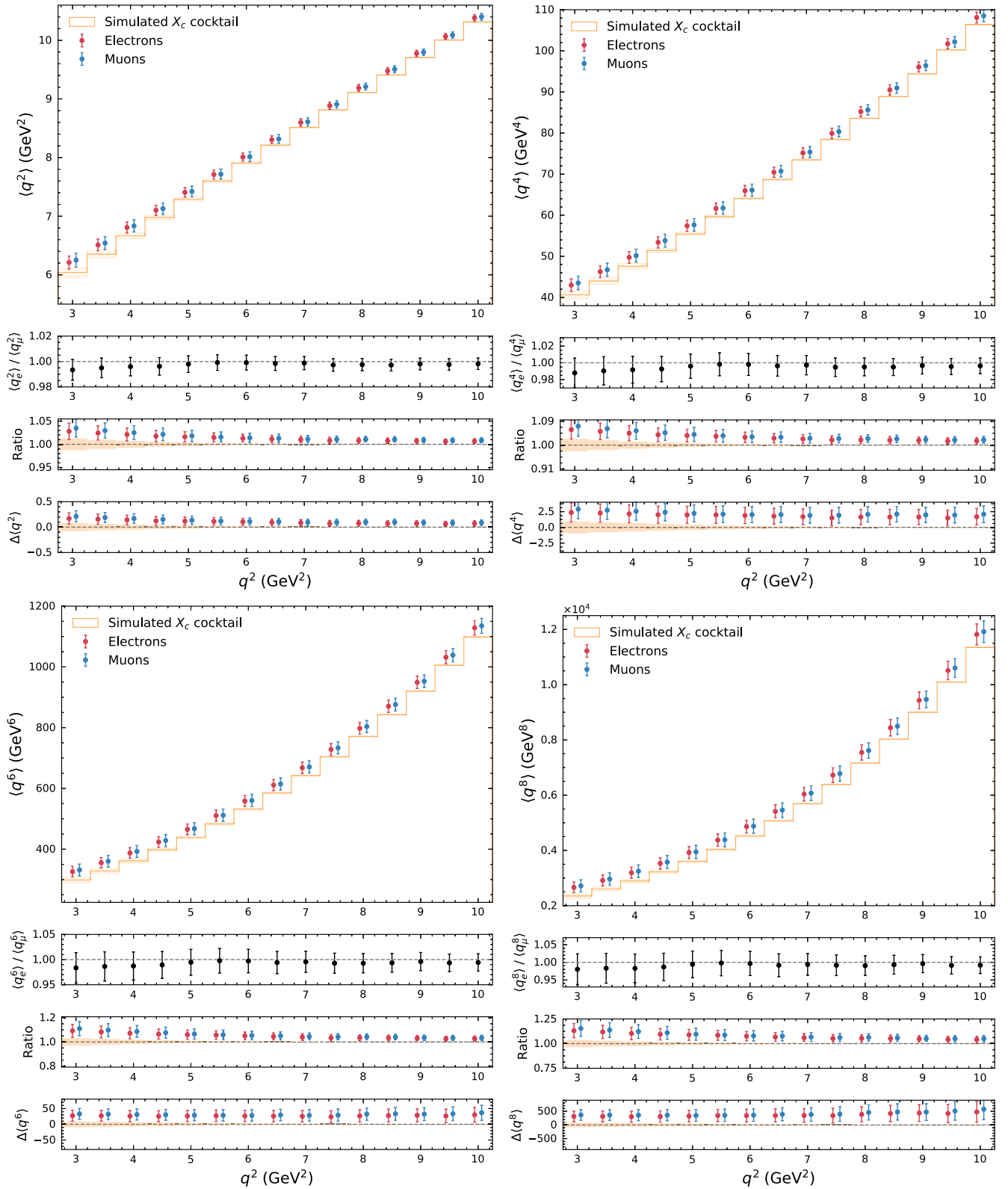


FIG. 8. The expectation of lepton flavor universality of the moments is tested for the first to fourth q^2 moments: in the ratio of electron to muon moments many of the associated systematic uncertainties cancel and all reported moments are compatible with the expectation of lepton flavor universality (bottom top). Note that the individual electron and muon moments are highly correlated. Furthermore, the measured and generated-level moments for all the threshold selections on q^2 are compared as a ratio (bottom middle) and difference (bottom lower) for both electrons and muons.

calculating the χ^2 for each order of measured q^2 moments. The calculated values range between 2.21 and 0.96 with a corresponding p-value of 0.99 for all order of q^2 moments.

We also extract the central, or normalized, moments. The central moments have the advantage of becoming slightly less correlated with respect to the correlations of the nominal moments, especially for the higher moments. To calculate the central moments directly from the nominal measured moments, the following nonlinear transformations are applied:

$$\begin{pmatrix} \langle q^2 \rangle \\ \langle q^4 \rangle \\ \langle q^6 \rangle \\ \langle q^8 \rangle \end{pmatrix} \rightarrow \begin{pmatrix} \langle q^2 \rangle \\ \langle (q^2 - \langle q^2 \rangle)^2 \rangle \\ \langle (q^2 - \langle q^2 \rangle)^3 \rangle \\ \langle (q^2 - \langle q^2 \rangle)^4 \rangle \end{pmatrix}. \quad (22)$$

The new systematic covariance matrix $\mathcal{C}'$ for the vector of central moments is determined by making use of the Jacobian matrix $\mathcal{J}$ for the transformation, together with the initial systematic covariance matrix $\mathcal{C}$ describing the nominal moments, such that

$$\mathcal{C}' = \mathcal{J}\mathcal{C}\mathcal{J}^T. \quad (23)$$

This approximation for the uncertainties of the central moments yields the same results as Gaussian error propagation. Not only are slightly lower correlations between threshold selections observed for central moments of the same order, but also negative correlations between different orders of moments. Appendix F compares the second, third and fourth measured central moments for electrons and muons.

VII. SUMMARY AND CONCLUSIONS

In this paper we report the first systematic study of the first to the fourth moment of the $B \rightarrow X_c \ell^+ \nu_\ell$ q^2 spectra with progressively increasing threshold selections on q^2 . The measured moments are crucial experimental inputs for a novel and alternative approach determining $|V_{cb}|$ from inclusive decays, which was outlined in Ref. [35]: using reparametrization invariance the set of nonperturbative matrix elements can be strongly reduced. As no public code is available to determine $|V_{cb}|$, this is left for future work. The reported measurement uses the full Belle dataset and employs a neural network assisted hadronic tag reconstruction. Explicit reconstruction of the tag B meson allows for the explicit reconstruction of the hadronic X system of the semileptonic decay, and thus the reconstruction of the four-momentum transfer squared q^2 . Background from other processes is subtracted using the reconstructed hadronic mass spectrum M_X in an unbinned approach using event weights. The reconstructed moments are then calibrated in a three step procedure to account for the finite detector resolution and acceptance effects. The background subtraction and calibration procedure

introduces systematic uncertainties, most dominantly from the assumed $B \rightarrow X_c \ell^+ \nu_\ell$ composition. The moments are measured separately for electron and muon $B \rightarrow X_c \ell^+ \nu_\ell$ final states. This allows for testing of lepton flavor universality in inclusive processes involving electrons and muons, and no deviation in the measured moments from the expectation of unity is observed. In addition to measuring the nominal q^2 moments, a nonlinear transformation is applied to extract the central moments, which are slightly less correlated compared to the statistical and systematic correlations of the nominal moments. The measured moments are also compared to the expectation from the exclusive make-up of $B \rightarrow X_c \ell^+ \nu_\ell$. Here, contributions from heavier charmed final states and high multiplicity decays are poorly constrained by current measurements. The measured q^2 spectrum has higher moments than the generator-level $B \rightarrow X_c \ell^+ \nu_\ell$ model for low q^2 threshold selections. As the $B \rightarrow X_c \ell^+ \nu_\ell$ decay gets increasingly dominated at high q^2 by the well measured D and D^* final states, this points towards a problem in the modeling of the other components. Similar observations have been reported in Refs. [63,64]. One of the leading sources of systematics for low q^2 threshold selections is observed to be the modeling of the $B \rightarrow X_u \ell^+ \nu_\ell$ decays. Events originating from this background component are first rejected by imposing selection criteria on decay kinematics, after which the remaining component is further suppressed by making use of the background subtraction procedure. A possible future improvement of the analysis strategy as well as the overall precision of the measurement is outlined in Ref. [65]. Rather than subtracting the $B \rightarrow X_u \ell^+ \nu_\ell$ component, the authors suggest measuring the full $B \rightarrow X \ell^+ \nu_\ell$ spectrum and obtaining the $B \rightarrow X_c \ell^+ \nu_\ell$ contribution precisely from within the HQE. This strategy has the potential to reduce the uncertainty on the measured value of $|V_{cb}|$ determined from the q^2 moments even further. The numerical values and full covariance matrices of the measured moments and central moments will be made available on HEPData [66].

ACKNOWLEDGMENTS

We thank Keri Vos, Kevin Olschewsky, and Matteo Fael for useful discussions about the subject matter of this manuscript. R. v. T., L. C., W. S., and F. U. B. were supported by the DFG Emmy-Noether Grant No. BE 6075/1-1. 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 Volkswagen Stiftung; 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. 2018R1D1A1B-07047294, No. 2019K1A3A7A09033840, and No. 2019-1111A3A01058933; 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 No. 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 United States Department of Energy and the National Science Foundation.

APPENDIX A: COMPARISON OF $B \rightarrow D_{\text{gap}}^{**} \ell \bar{\nu}_\ell$ DECAY MODELS

As discussed in Sec. II, the remaining gap between the sum of all considered exclusive modes and the inclusive $B \rightarrow X_c \ell^+ \nu_\ell$ branching fraction is filled in equal parts with $B \rightarrow D \eta \ell^+ \nu_\ell$ and $B \rightarrow D^* \eta \ell^+ \nu_\ell$ 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^* . This model provides a better kinematic description of the initial three-body decay $B \rightarrow D_{\text{gap}}^{**} \ell \bar{\nu}_\ell$ than a model based on the equidistribution of all final-state particles in phase space. Comparisons of kinematic distributions for the different $B \rightarrow D_{\text{gap}}^{**} \ell \bar{\nu}_\ell$ decay models are shown in Fig. 9.

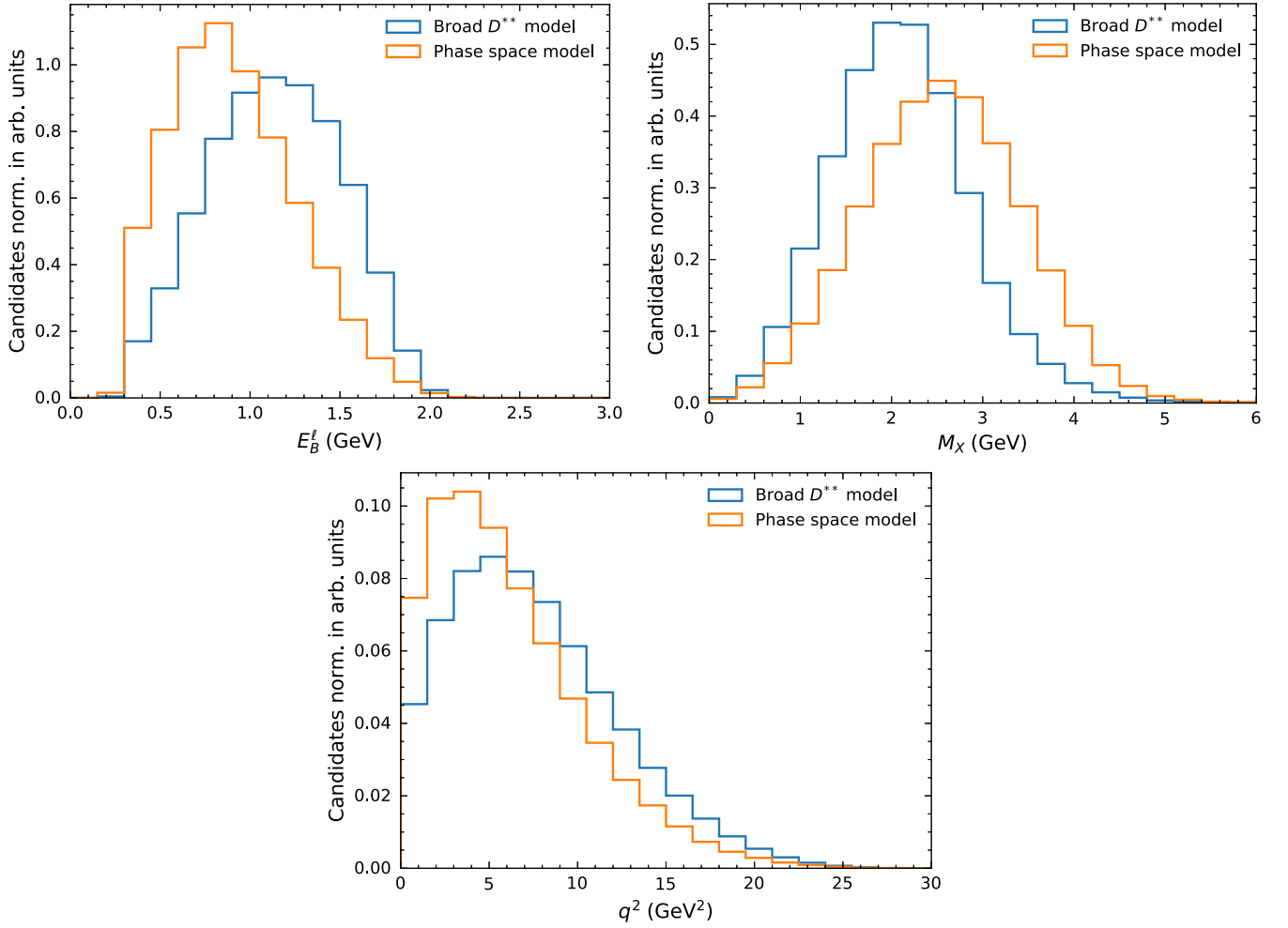


FIG. 9. Comparisons between the different $B \rightarrow D_{\text{gap}}^{**} \ell \bar{\nu}_\ell$ decay models for the reconstructed lepton energy in the signal B rest frame E_ℓ^B (upper left), the M_X (upper right), and the q^2 (bottom) distributions.

APPENDIX B: BACKGROUND SUBTRACTION FITS

The determined background in the q^2 spectrum, after obtaining the expected background yields from the fit to the M_X spectrum, and the signal probability weights are shown in Figs. 10–13.

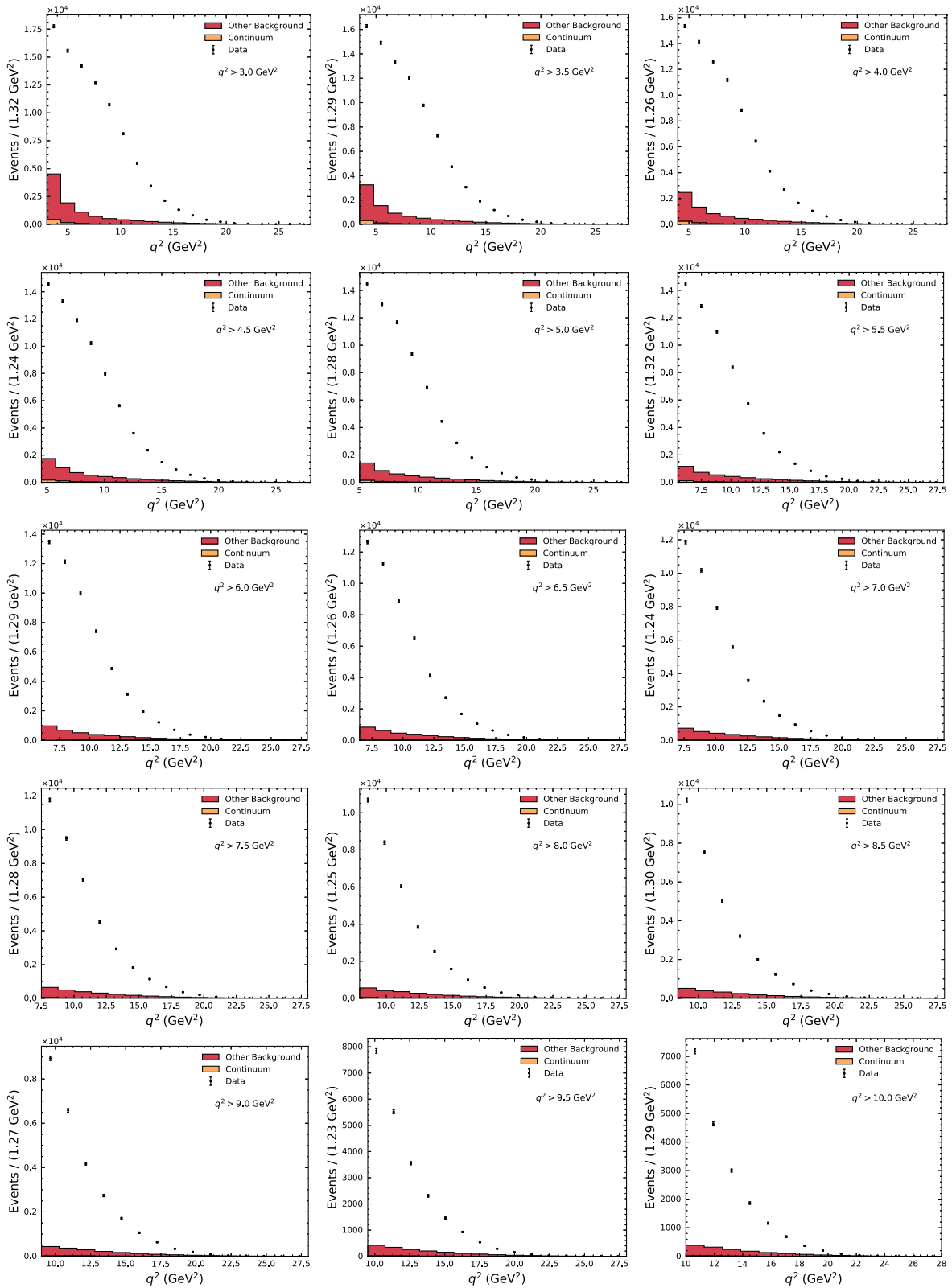


FIG. 10. The determined background in the q^2 spectrum, after obtaining the expected background yields from the fit to the M_X spectrum, for different q^2 threshold selections for electrons are shown.

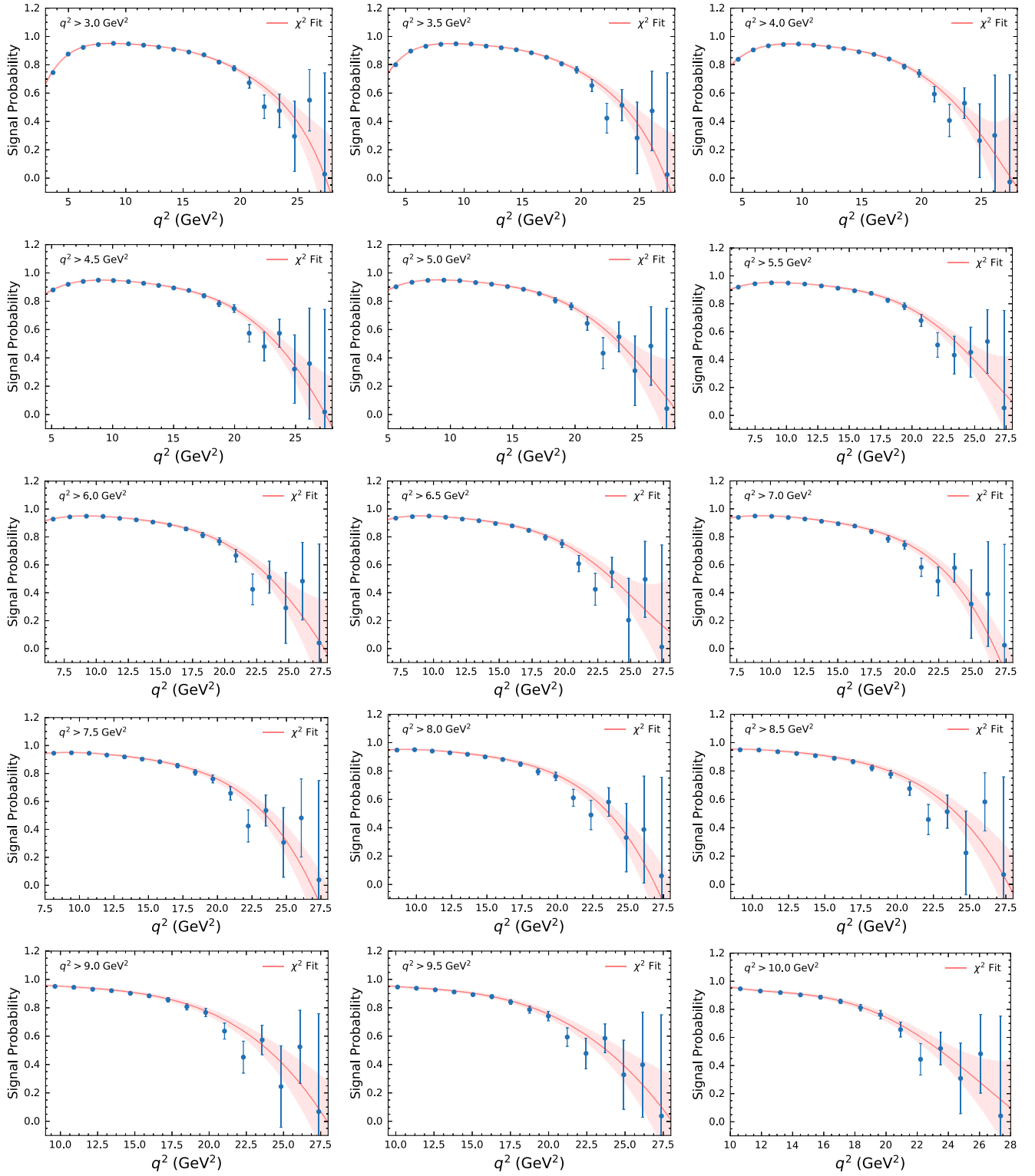


FIG. 11. The determined signal probabilities for different q^2 threshold selections for electrons are shown.

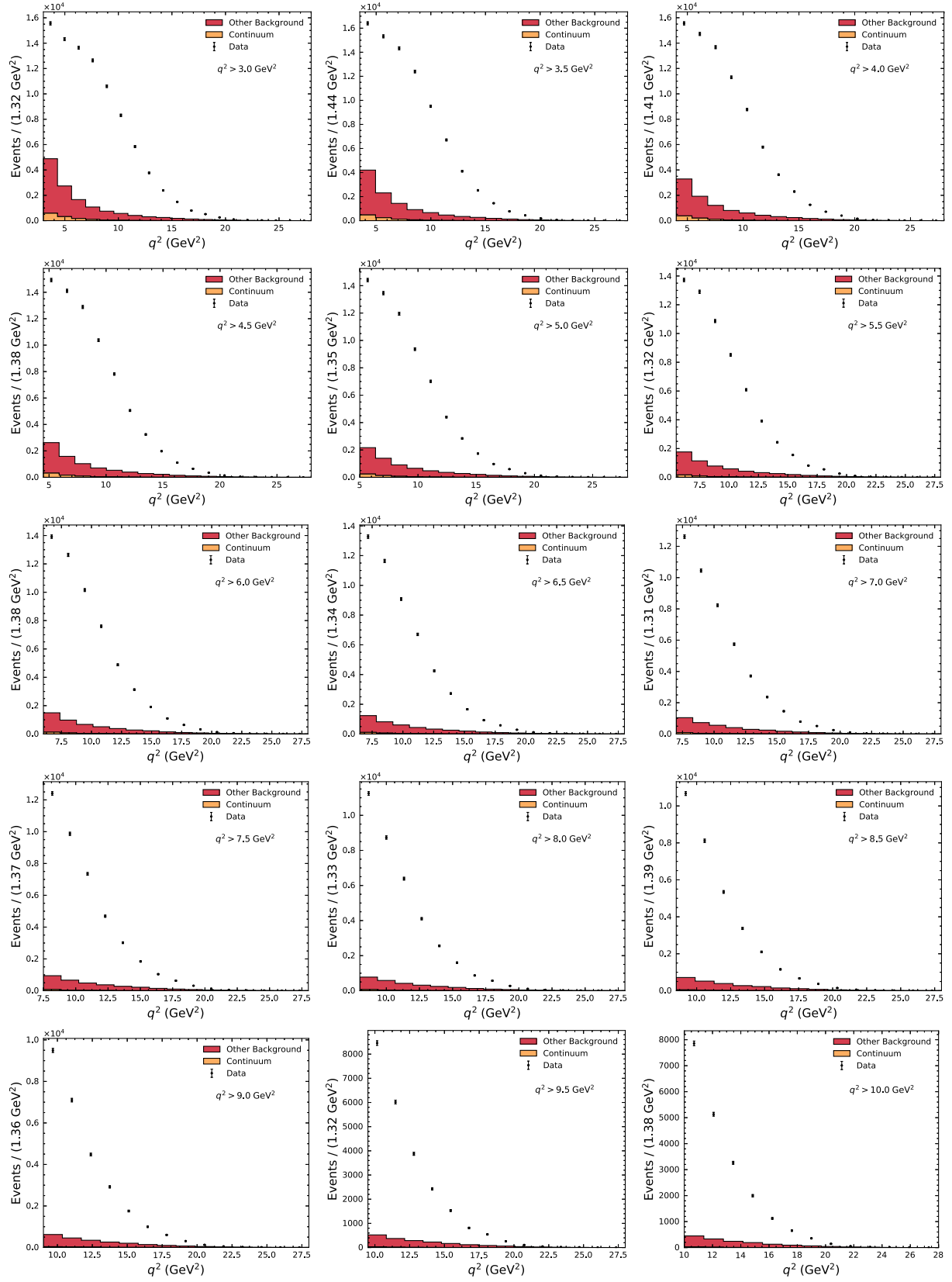
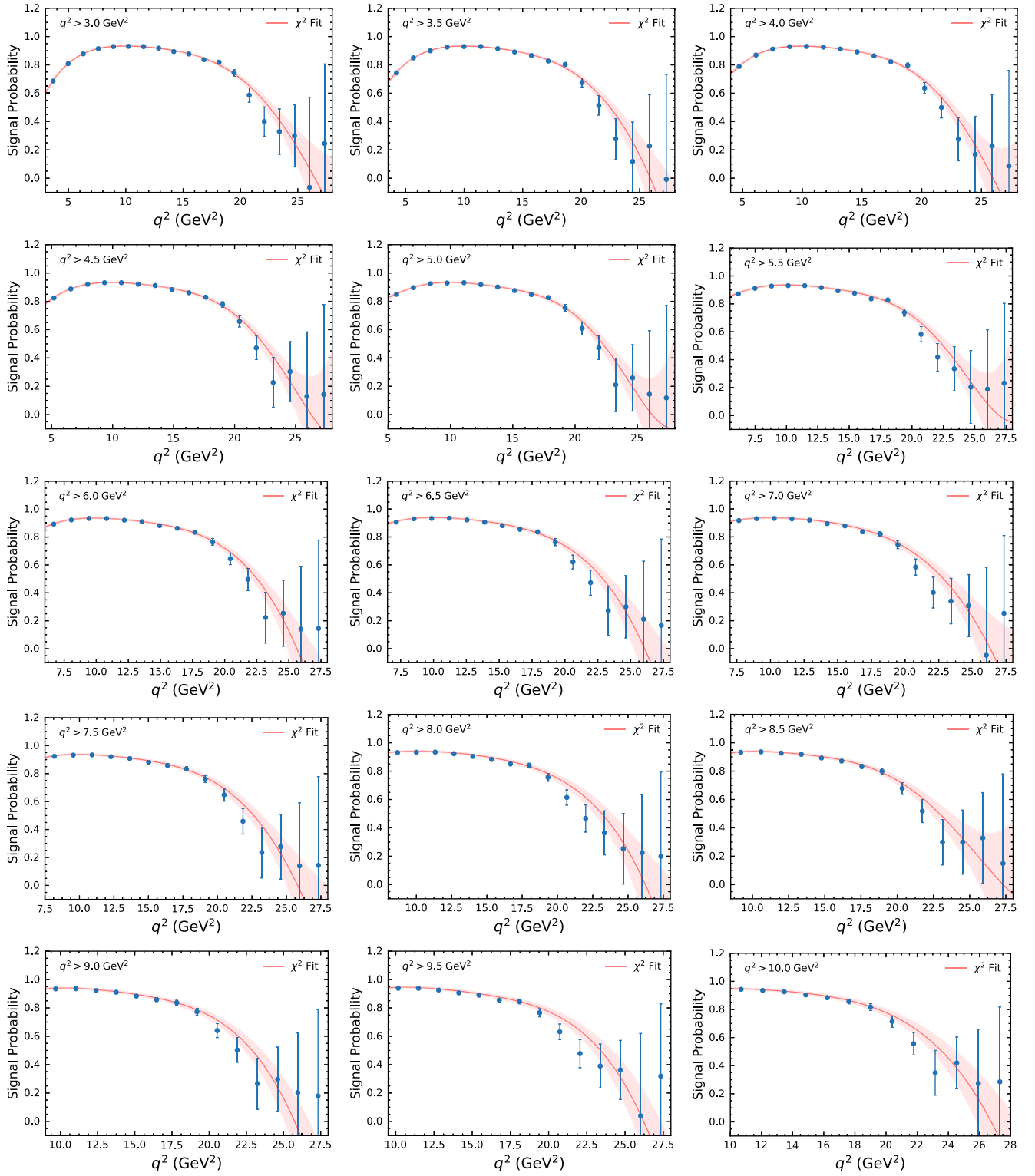


FIG. 12. The determined background in the q^2 spectrum, after obtaining the expected background yields from the fit to the M_X spectrum, for different q^2 threshold selections for muons are shown.

FIG. 13. The determined signal probabilities for different q^2 threshold selections for muons are shown.

APPENDIX C: CALIBRATION PARAMETERS, CALIBRATION CURVES, AND CALIBRATION FACTORS

Figure 14 shows the calibration curves, the bias and acceptance calibration factors, C_{cal} and C_{acc} , for muons for the various selections on q^2 . Furthermore, Table IV summarizes the fitted parameters, a_m and b_m , of the determined linear calibration functions for each order of moment m .

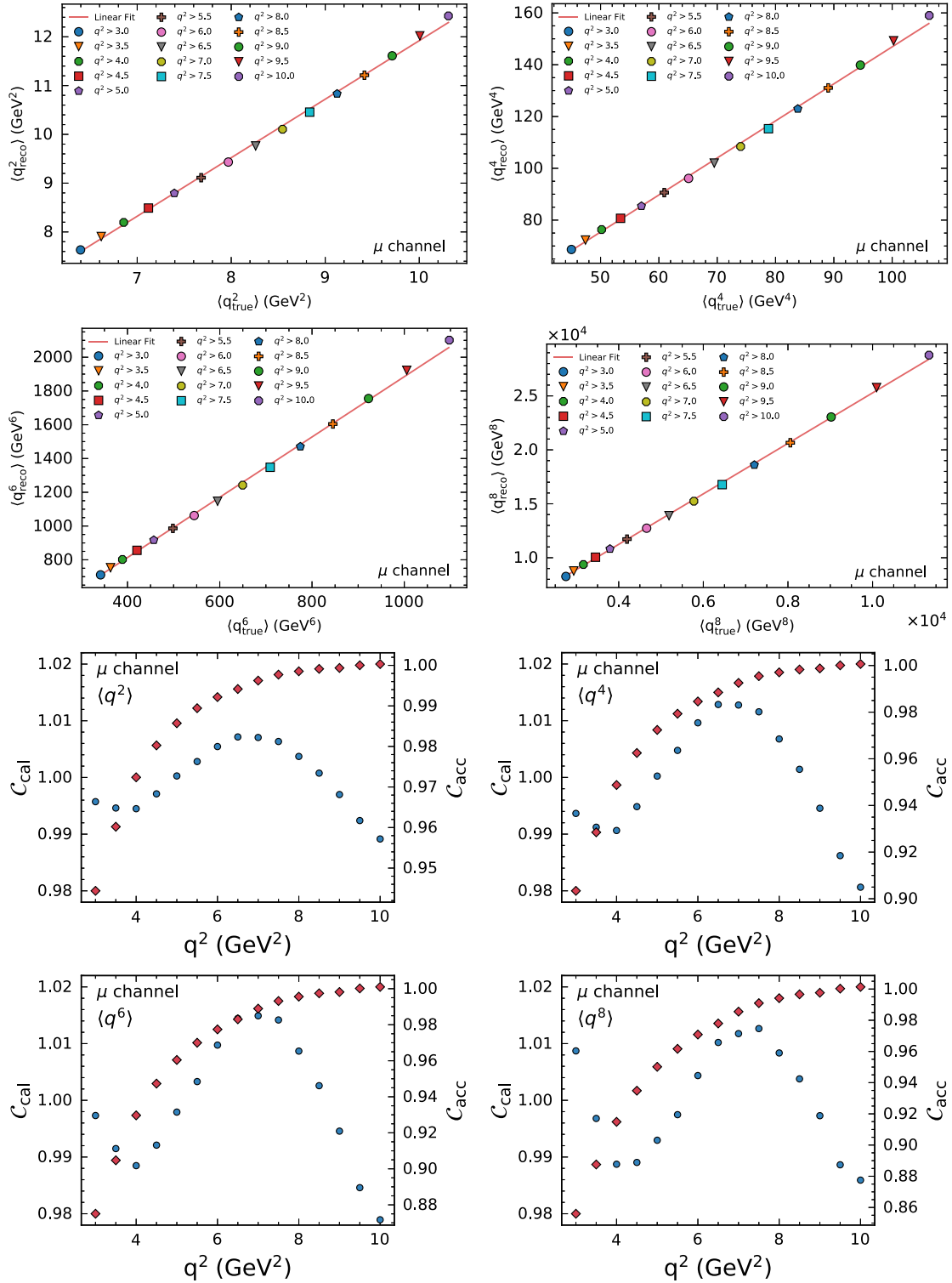


FIG. 14. The calibration curves, bias, and acceptance calibration factors for the first to the fourth moment of q^2 for muons.

TABLE IV. Summary of the fitted parameters of the determined linear calibration functions for electron and muon candidates with a_m and b_m denoting the offset and slope for moments of the order m . The values given in brackets denote the statistical uncertainty on the given parameters estimated from the post-fit covariance.

$\langle q^{2m} \rangle$	Electrons		Muons	
	a_m (GeV 2m)	b_m (GeV 2m)	a_m (GeV 2m)	b_m (GeV 2m)
$\langle q^2 \rangle$	1.245 (0.002)	-0.57 (0.02)	1.200 (0.002)	-0.08 (0.02)
$\langle q^4 \rangle$	1.473 (0.004)	-1.6 (0.2)	1.430 (0.004)	3.9 (0.2)
$\langle q^6 \rangle$	1.828 (0.006)	33.0 (3.0)	1.786 (0.006)	99.0 (3.0)
$\langle q^8 \rangle$	2.386 (0.011)	1027.0 (49.0)	2.340 (0.010)	1882.0 (50.0)

APPENDIX D: SELECTION EFFICIENCY OF $B \rightarrow X_c \ell^+ \nu_\ell$ COMPONENTS

Selection efficiencies for different $B \rightarrow X_c \ell^+ \nu_\ell$ components are shown in Fig. 15 in bins of q^2 . Events are required to pass basic transverse momentum and angular acceptance selection criteria. Different selection and acceptance efficiencies are observed in the low q^2 region for different $B \rightarrow X_c \ell^+ \nu_\ell$ processes, especially for the nonresonant component.

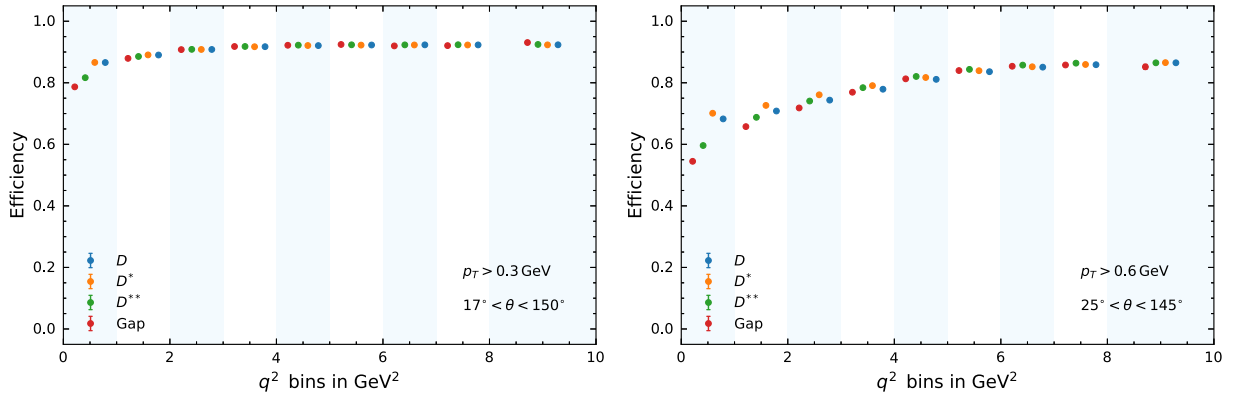


FIG. 15. Selection efficiencies of different $B \rightarrow X_c \ell^+ \nu_\ell$ components for electron (left) and muon (right) final states with basic transverse momentum and angular acceptance requirements in bins of q^2 .

APPENDIX E: EXPERIMENTAL CORRELATION MATRICES

The full systematic correlation matrices that are determined by combining the statistical and systematic covariance matrices are shown in Figs. 16 and 17 for electrons and muons, respectively.

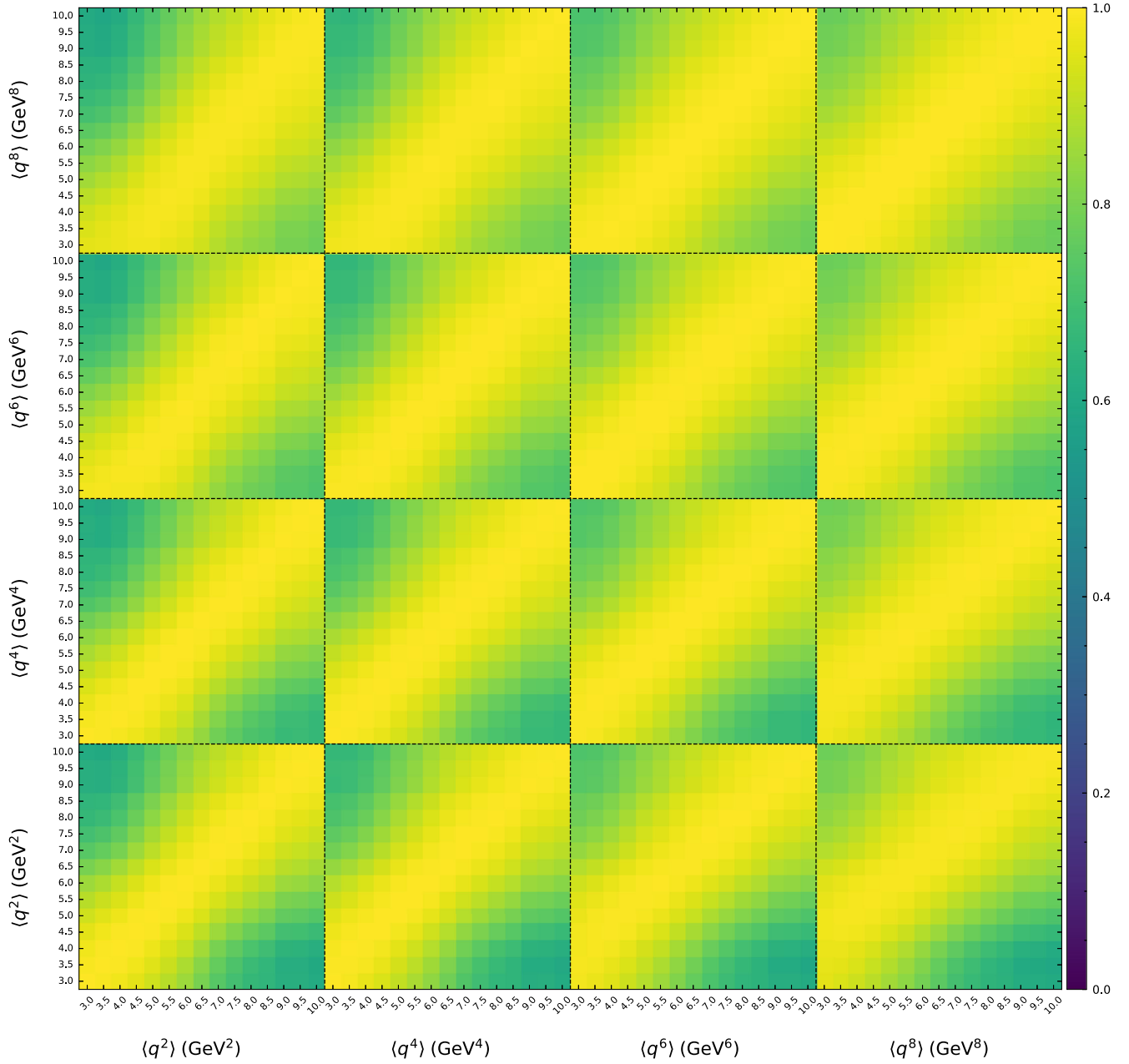


FIG. 16. Color map of the full experimental correlation coefficients determined for the moments for the electron final state.

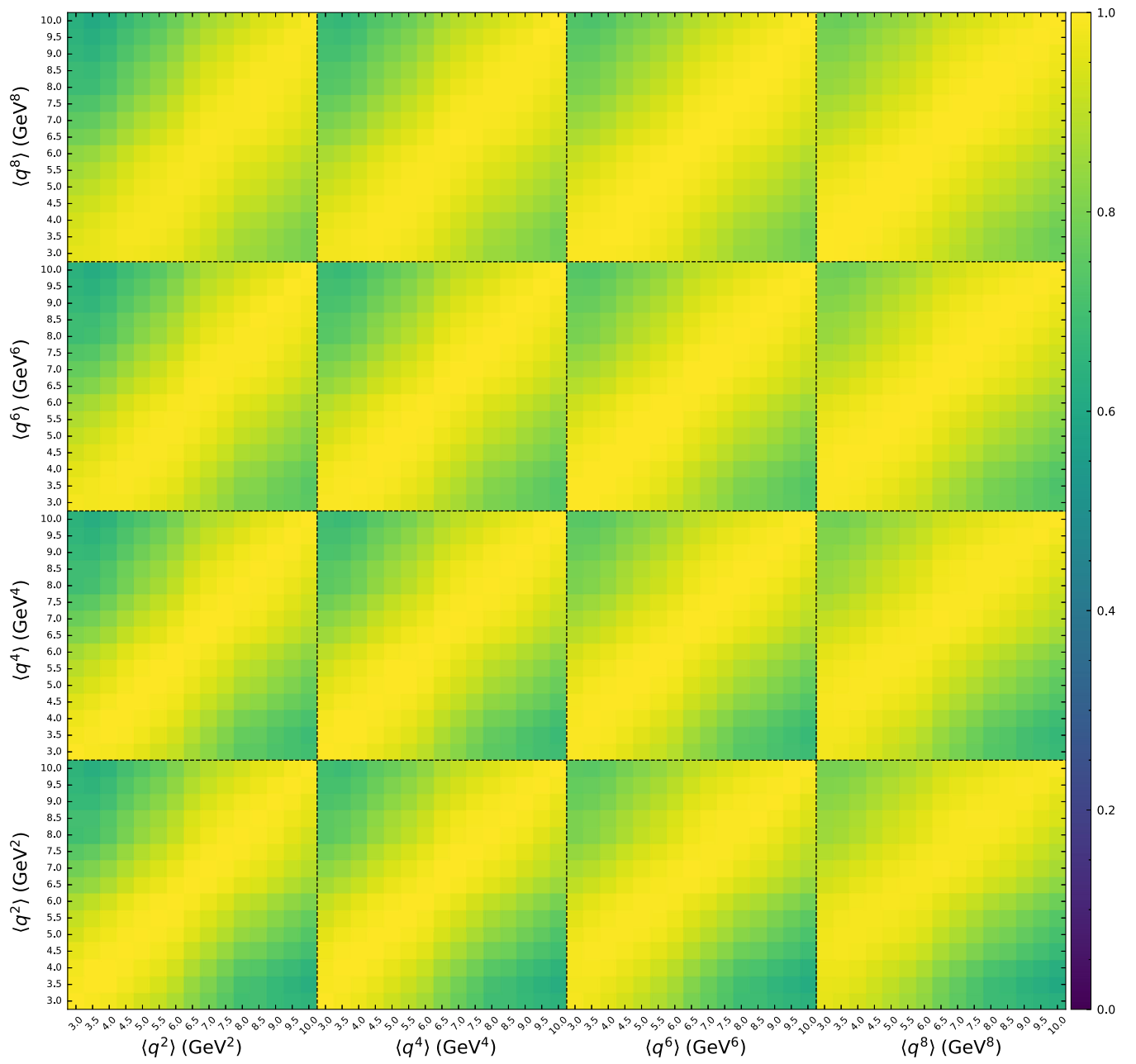


FIG. 17. Color map of the systematic correlation coefficients determined for the moments for the muon final state.

APPENDIX F: CENTRAL MOMENTS

The measured central q^2 moments, discussed in Sec. VI, are shown for both the electron and muon final states in Fig. 18.

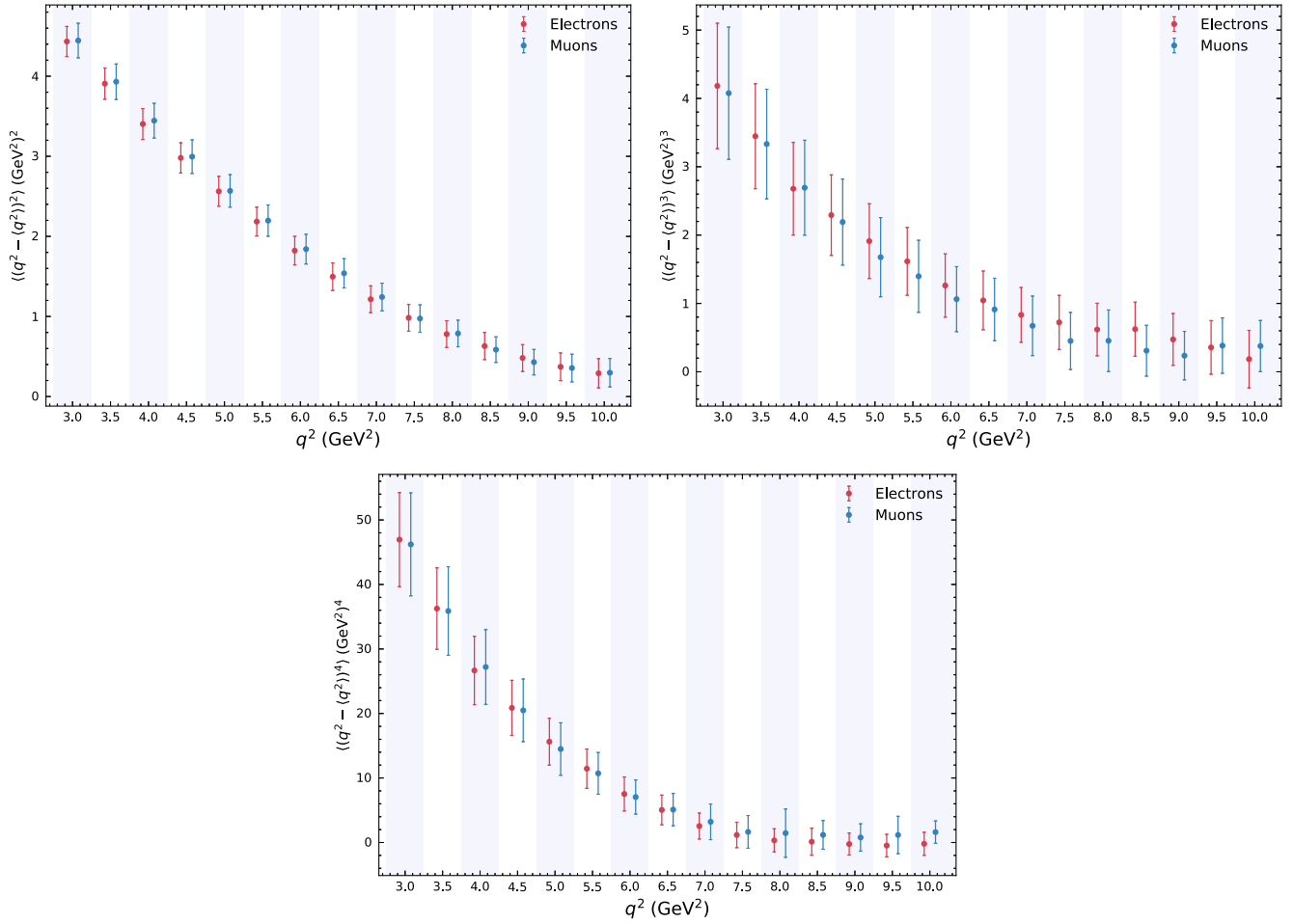


FIG. 18. The measured second (left), third (right), and fourth (bottom) central q^2 moments for both the electron and muon final states.

-
- [1] N. Cabibbo, *Phys. Rev. Lett.* **10**, 531 (1963).
[2] M. Kobayashi and T. Maskawa, *Prog. Theor. Exp. Phys.* **49**, 652 (1973).
[3] P. Zyla *et al.* (Particle Data Group, CKM Quark-Mixing Matrix Review Collaborations), *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] A. Sirlin, *Nucl. Phys.* **B196**, 83 (1982).
[7] T. Zheng, J. Xu, L. Cao, D. Yu, W. Wang, S. Prell, Y.-K. E. Cheung, and M. Ruan, *Chin. Phys. C* **45**, 023001 (2021).
[8] Y. S. Amhis *et al.* (HFLAV Collaboration), *Eur. Phys. J. C* **81**, 226 (2021).
[9] R. Aaij *et al.* (LHCb Collaboration), *Phys. Rev. D* **101**, 072004 (2020).
[10] D. Bigi and P. Gambino, *Phys. Rev. D* **94**, 094008 (2016).
[11] D. Bigi, P. Gambino, and S. Schacht, *Phys. Lett. B* **769**, 441 (2017).
[12] F. U. Bernlochner, Z. Ligeti, M. Papucci, and D. J. Robinson, *Phys. Rev. D* **96**, 091503 (2017).
[13] D. Ferlewicz, P. Urquijo, and E. Waheed, *Phys. Rev. D* **103**, 073005 (2021).
[14] M. Jezabek and J. H. Kuhn, *Nucl. Phys.* **B314**, 1 (1989).
[15] V. Aquila, P. Gambino, G. Ridolfi, and N. Uraltsev, *Nucl. Phys.* **B719**, 77 (2005).
[16] A. Pak and A. Czarnecki, *Phys. Rev. D* **78**, 114015 (2008).
[17] K. Melnikov, *Phys. Lett. B* **666**, 336 (2008).
[18] T. Becher, H. Boos, and E. Lunghi, *J. High Energy Phys.* **12** (2007) 062.
[19] T. Mannel, A. A. Pivovarov, and D. Rosenthal, *Phys. Lett. B* **741**, 290 (2015).
[20] A. Alberti, P. Gambino, and S. Nandi, *J. High Energy Phys.* **01** (2014) 147.
[21] T. Mannel, S. Turczyk, and N. Uraltsev, *J. High Energy Phys.* **11** (2010) 109.

- [22] D. Acosta *et al.* (CDF Collaboration), *Phys. Rev. D* **71**, 051103 (2005).
- [23] S. E. Csorna *et al.* (CLEO Collaboration), *Phys. Rev. D* **70**, 032002 (2004).
- [24] J. Abdallah *et al.* (DELPHI Collaboration), *Eur. Phys. J. C* **45**, 35 (2006).
- [25] B. Aubert *et al.* (BABAR Collaboration), *Phys. Rev. D* **69**, 111104 (2004).
- [26] B. Aubert *et al.* (BABAR Collaboration), *Phys. Rev. D* **81**, 032003 (2010).
- [27] P. Urquijo *et al.* (Belle Collaboration), *Phys. Rev. D* **75**, 032001 (2007).
- [28] C. Schwanda *et al.* (Belle Collaboration), *Phys. Rev. D* **75**, 032005 (2007).
- [29] P. Gambino and C. Schwanda, *Phys. Rev. D* **89**, 014022 (2014).
- [30] A. Alberti, P. Gambino, K. J. Healey, and S. Nandi, *Phys. Rev. Lett.* **114**, 061802 (2015).
- [31] J. Dingfelder and T. Mannel, *Rev. Mod. Phys.* **88**, 035008 (2016).
- [32] D. Benson, I. I. Bigi, T. Mannel, and N. Uraltsev, *Nucl. Phys. B* **665**, 367 (2003).
- [33] P. Gambino, *J. High Energy Phys.* **09** (2011) 055.
- [34] 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>.
- [35] M. Fael, T. Mannel, and K. Keri Vos, *J. High Energy Phys.* **02** (2019) 177.
- [36] 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.
- [37] 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).
- [38] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima, and T. Tsukamoto, *Nucl. Instrum. Methods Phys. Res., Sect. A* **485**, 490 (2002).
- [39] A. Abashian *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **491**, 69 (2002).
- [40] D. J. Lange, *Nucl. Instrum. Methods Phys. Res., Sect. A* **462**, 152 (2001).
- [41] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini, Report No. CERN-DD-EE-84-1, 1987.
- [42] E. Barberio, B. van Eijk, and Z. W̑s, *Comput. Phys. Commun.* **66**, 115 (1991).
- [43] C. G. Boyd, B. Grinstein, and R. F. Lebed, *Phys. Rev. Lett.* **74**, 4603 (1995).
- [44] R. Glattauer *et al.* (Belle Collaboration), *Phys. Rev. D* **93**, 032006 (2016).
- [45] B. Grinstein and A. Kobach, *Phys. Lett. B* **771**, 359 (2017).
- [46] E. Waheed *et al.* (Belle Collaboration), *Phys. Rev. D* **100**, 052007 (2019).
- [47] F. U. Bernlochner and Z. Ligeti, *Phys. Rev. D* **95**, 014022 (2017).
- [48] P. Zyla *et al.* (Particle Data Group), *Prog. Theor. Exp. Phys.* **2020**, 083C01 (2020).
- [49] D. Liventsev *et al.* (Belle Collaboration), *Phys. Rev. D* **77**, 091503 (2008).
- [50] B. Aubert *et al.* (BABAR Collaboration), *Phys. Rev. Lett.* **101**, 261802 (2008).
- [51] I. 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).
- [52] A. K. Leibovich, Z. Ligeti, I. W. Stewart, and M. B. Wise, *Phys. Rev. D* **57**, 308 (1998).
- [53] R. Aaij *et al.* (LHCb Collaboration), *Phys. Rev. D* **84**, 092001 (2011); **85**, 039904(E) (2012).
- [54] J. Lees *et al.* (BABAR Collaboration), *Phys. Rev. Lett.* **116**, 041801 (2016).
- [55] C. Ramirez, J. F. Donoghue, and G. Burdman, *Phys. Rev. D* **41**, 1496 (1990).
- [56] L. Cao *et al.* (Belle Collaboration), *Phys. Rev. D* **104**, 012008 (2021).
- [57] 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).
- [58] A. J. Bevan, B. Golob, Th. Mannel, S. Prell, and B. D. Yabsley, *Eur. Phys. J. C* **74**, 3026 (2014).
- [59] F. James and M. Roos, *Comput. Phys. Commun.* **10**, 343 (1975).
- [60] H. Dembinski, P. Ongmongkolkul *et al.*, 10.5281/zenodo.3949207 (2021).
- [61] B. Efron, *Ann. Stat.* **7**, 1 (1979).
- [62] K. G. Hayes, M. L. Perl, and B. Efron, *Phys. Rev. D* **39**, 274 (1989).
- [63] J. P. Lees *et al.* (BABAR Collaboration), *Phys. Rev. D* **95**, 072001 (2017).
- [64] F. U. Bernlochner, D. Biedermann, H. Lacker, and T. Lück, *Eur. Phys. J. C* **74**, 2914 (2014).
- [65] T. Mannel, M. Rahimi, and K. Keri Vos, *J. High Energy Phys.* **09** (2021) 051.
- [66] <https://www.hepdata.net>.