Belle Preprint 2021-18, KEK Preprint 2021-22

Measurements of q^2 Moments of Inclusive $B \to X_c\, \ell^+\, \nu_\ell$ 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. Sevior,[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. Zhukova[40]

(The Belle Collaboration)

[1] *University of Bonn, 53115 Bonn*
[2] *Department of Physics, University of the Basque Country UPV/EHU, 48080 Bilbao*
[3] *Brookhaven National Laboratory, Upton, New York 11973*
[4] *Budker Institute of Nuclear Physics SB RAS, Novosibirsk 630090*
[5] *Faculty of Mathematics and Physics, Charles University, 121 16 Prague*
[6] *Chonnam National University, Gwangju 61186*
[7] *University of Cincinnati, Cincinnati, Ohio 45221*
[8] *Deutsches Elektronen–Synchrotron, 22607 Hamburg*
[9] *Department of Physics, Fu Jen Catholic University, Taipei 24205*
[10] *Key Laboratory of Nuclear Physics and Ion-beam Application (MOE) and Institute of Modern Physics, Fudan University, Shanghai 200443*
[11] *Justus-Liebig-Universität Gießen, 35392 Gießen*
[12] *Gifu University, Gifu 501-1193*
[13] *SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193*
[14] *Department of Physics and Institute of Natural Sciences, Hanyang University, Seoul 04763*
[15] *University of Hawaii, Honolulu, Hawaii 96822*
[16] *High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801*
[17] *J-PARC Branch, KEK Theory Center, High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801*
[18] *National Research University Higher School of Economics, Moscow 101000*
[19] *Forschungszentrum Jülich, 52425 Jülich*
[20] *IKERBASQUE, Basque Foundation for Science, 48013 Bilbao*
[21] *Indian Institute of Science Education and Research Mohali, SAS Nagar, 140306*
[22] *Indian 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, Sagamihara 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

(Dated: September 7, 2021)

We present the measurement of the first to fourth order moments of the four-momentum transfer squared, q^2, of inclusive $B \to X_c \ell^+ \nu_\ell$ decays using the full Belle data set of $711\,\mathrm{fb}^{-1}$ 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 CKM 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 $3.0\,\mathrm{GeV}^2$ in steps of $0.5\ \mathrm{GeV}^2$ up to a value of $10.0\,\mathrm{GeV}^2$. 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.

PACS numbers: 12.15.Hh, 13.20.-v, 14.40.Nd

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 $B \to X_c \ell^+ \nu_\ell$ 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 $B \to D^* \ell^+ \nu_\ell$ and $B \to D \ell^+ \nu_\ell$ transitions providing the most precise values to date [8], or on the study of inclusive $B \to X_c \ell^+ \nu_\ell$ decay. First measurements using $B_s \to D_s^* \ell^+ \nu_\ell$ and $B_s \to D_s \ell^+ \nu_\ell$ decays were recently reported by 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 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 re-evaluated, since the use of more generalised 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, $\Lambda_{\rm QCD}/m_b$, 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

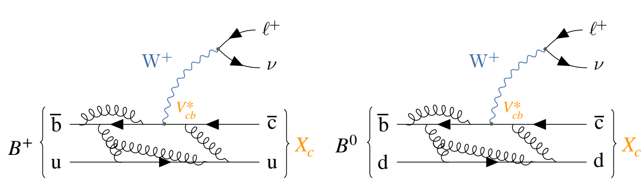

FIG. 1. The inclusive $B \to X_c \ell^+ \nu_\ell$ semileptonic processes for a B^+ (left) or a B^0 (right) meson decay.

* vantonder@physik.uni-bonn.de
† florian.bernlochner@uni-bonn.de

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

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 $B \to X_c\,\ell^+\,\nu_\ell$ branching fraction, the energy and the hadronic mass moments [22–28], the non-perturbative 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]:

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

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

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 ca. 2% uncertainty in the inclusive determination is dominated by theory errors, mainly to cover uncertainties from missing higher-order contributions [32, 33]. This disagreement from both 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 non-perturbative matrix elements present in the total rate can be reduced to a set of only 8 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 8 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 selections[2] $q^2_{\rm sel} \in [3.0, 10.0]\,{\rm GeV}^2$. 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. In this paper a first systematic study is reported with progressively increasing threshold selections on q^2. Furthermore, 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 $3\,{\rm GeV}^2$, to ensure that the Belle detector can still reliably reconstruct and identify the lepton from the $B \to X_c\,\ell^+\,\nu_\ell$ decay.

The remainder of this paper is organized as follows: Section II introduces the data set and simulated event samples, as well as the applied corrections to the simulated samples. In Section 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 discusses the calibration procedure used to reverse detector resolution, selection, and acceptance effects of the measured moments of q^2. In Section V the systematic uncertainties affecting the moment measurements are discussed, while Section VI reports the measured moments and compares them to the expectation from simulated $B \to X_c\,\ell^+\,\nu_\ell$ decays. Finally, Section VII summarizes our findings and presents our conclusions.

II. DATA SET AND SIMULATED SAMPLES

The results are based on an analysis of the full Belle data set 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 $\sqrt{s} = 10.58\,{\rm GeV}$ corresponding to the $\Upsilon(4S)$ resonance. An additional $79\,{\rm fb}^{-1}$ 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 sub-detectors: 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, $\mathcal{L}_{\rm LID}$.

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

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 MeV and 600 MeV, respectively, in the laboratory frame of Belle. We further identify charged tracks as pions or kaons using a likelihood ratio $\mathcal{L}_{\mathrm{K}/\pi\,\mathrm{ID}} = \mathcal{L}_{\mathrm{K\,ID}}/\left(\mathcal{L}_{\mathrm{K\,ID}} + \mathcal{L}_{\pi\,\mathrm{ID}}\right)$. The most important identifying features of the kaon ($\mathcal{L}_{\mathrm{K\,ID}}$) and pion ($\mathcal{L}_{\pi\,\mathrm{ID}}$) likelihoods for low momentum particles with transverse momentum below 1 GeV in the laboratory frame are the recorded energy loss by ionization, $\mathrm{d}E/\mathrm{d}x$, in the CDC, and the time of flight information from the TOF. Higher-momentum kaon and pion classification relies on the Cherenkov light recorded in the ACC. In order to avoid the difficulties in understanding the efficiencies of reconstructing K_L^0 mesons, they are not explicitly reconstructed or used in this 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 $E_\gamma > 100\,\mathrm{MeV}$, $150\,\mathrm{MeV}$, and $50\,\mathrm{MeV}$ in the forward endcap, and backward endcap 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 $m_{\gamma\gamma} \in [0.12, 0.15]\,\mathrm{GeV}$, corresponding to about 2.5 times the π^0 mass resolution.

Monte Carlo (MC) samples of B meson decays and continuum processes ($e^+e^- \to q\bar{q}$ 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 three 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. The efficiencies in the MC are corrected using data-driven methods to account for, e.g., differences in identification and reconstruction efficiencies.

Inclusive semileptonic $B \to X_c\,\ell^+\,\nu_\ell$ decays are dominated by $B \to D\,\ell^+\,\nu_\ell$ and $B \to D^*\,\ell^+\,\nu_\ell$ transitions. We model the $B \to D\,\ell^+\,\nu_\ell$ decays using the BGL parametrization [43] with form factor central values and uncertainties taken from the fit in Ref. [44]. For $B \to D^*\,\ell^+\,\nu_\ell$ 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 $B \to D^{**}\,\ell^+\,\nu_\ell$ decays with $D^{**} = \{D_0^*, D_1^*, D_1, D_2^*\}$ denoting the four orbitally excited charmed mesons are modeled using the heavy-quark-symmetry-based form factors proposed in Ref. [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]. To correct for the fact that all measurements were targeting the $D^{**\,0} \to D^{(*)+}\,\pi^-$ decay modes, we correct for the missing isospin modes with a factor of

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

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

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

is applied. The world average of $B \to D_1^*\,\ell\bar{\nu}_\ell$ 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 $\mathcal{B}(B \to D_1^*\,\ell\bar{\nu}_\ell)$ being of similar size than $\mathcal{B}(B \to D_0\,\ell\bar{\nu}_\ell)$ [51, 52]. We perform our own average excluding the conflicting measurement and use

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

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

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

We subtract the contribution of $D_1 \to D\pi\pi$ from the measured non-resonant plus resonant $B \to D\pi\pi\ell\bar{\nu}_\ell$ branching fraction of Ref. [54]. 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

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

The uncertainty quoted here takes into account the full spread of final states ($f_0(500) \to \pi\pi$ or $\rho \to \pi\pi$ result in

$f_{\pi\pi} = 2/3$ and $1/3$, respectively) and the non-resonant three-body decays ($f_{\pi\pi} = 3/7$). We further make the implicit assumption that

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

For the remaining $B \to D^{(*)}\,\pi\,\pi\,\ell^+\,\nu_\ell$ contributions we use the measured value of Ref. [54]. The remaining 'gap' between the sum of all considered exclusive modes and the inclusive $B \to X_c\,\ell^+\,\nu_\ell$ branching fraction is filled in equal parts with $B \to D\,\eta\,\ell^+\,\nu_\ell$ and $B \to D^*\,\eta\,\ell^+\,\nu_\ell$ and for both we assume a 100% uncertainty. We simulate $B \to D^{(*)}\,\pi\,\pi\,\ell^+\nu_\ell$ and $B \to D^{(*)}\,\eta\,\ell^+\nu_\ell$ final states assuming that they are produced by the decay of two broad resonant states $D^{**}_{\rm gap}$ with masses and widths identical to D_1^* and D_0^*. Although there is currently no experimental evidence for decays of charm $1P$ states into these final states or the existence of such an additional broad state (e.g. a $2S$) in semileptonic transitions, this description provides a better kinematic description of the initial three-body decay, $B \to D^{**}_{\rm gap}\,\ell\bar{\nu}_\ell$, 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]. In what follows, we will associate this component with a 100% uncertainty.

Semileptonic $B \to X_u\,\ell^+\,\nu_\ell$ decays are modeled as a mixture of specific exclusive modes and non-resonant 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 $B \to X_c\,\ell^+\,\nu_\ell$ processes that we use.

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

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

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 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 system of the $B \to X_c\,\ell^+\,\nu_\ell$ process of interest and we label the reconstructed B mesons reconstructed in hadronic modes in the following as $B_{\rm tag}$. 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 ($e^+, \mu^+, K^+, \pi^+, \gamma$), neutral π^0 candidates, or K_S^0 candidates. These candidate particles are then combined during a second stage into heavier meson candidates ($J/\psi, D^0, D^+, D_s$). 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 $B_{\rm tag}$ 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 $\mathcal{O}_{\rm FR}$. 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 $\mathcal{O}_{\rm Cont}$. 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 $B_{\rm tag}$ candidates that show at least moderate agreement based on these two outputs, and require that $\mathcal{O}_{\rm FR} > 10^{-4}$ and $\mathcal{O}_{\rm Cont} > 10^{-4}$.

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 $B_{\rm tag}$ candidate. We require the $B_{\rm tag}$ candidates to have at least a beam-constrained mass of

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

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

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

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, $E_{\rm tag}$ denotes the energy of the $B_{\rm tag}$ candidate in the center-of-mass frame of the colliding e^+e^--pair. In each event a single $B_{\rm tag}$ candidate is then selected according to the highest $\mathcal{O}_{\rm FR}$ score value. After the reconstruction of the $B_{\rm tag}$ 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 Section II. The signal B rest frame is calculated with the momentum of the $B_{\rm tag}$ candidate via

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

with $p_{e^+e^-}$ denoting the four-momentum of the colliding electron-positron pair.

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 $b \to c \to s$ cascade. For charged $B_{\rm tag}$ candidates, we demand that the charge assignment of the signal-side lepton be opposite to that of the $B_{\rm tag}$ charge. The hadronic X system is reconstructed from the remaining unassigned charged particles and neutral energy depositions. Its four-momentum is calculated as

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

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,

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

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

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

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

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

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

$$q^2 = \left(p_{\rm sig} - p_X\right)^2\,. \quad (16)$$

The resolution of both variables for $B \to X_c\,\ell^+\,\nu_\ell$ 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 small peak at 0 GeV is from $B^+ \to \overline{D}^0\,\ell^+\,\nu_\ell$ with $\overline{D}^0 \to K^+\pi^-$ and other low-multiplicity final states comprising only charged pions. The four-momentum transfer squared q^2 exhibits a large resolution, which is caused by a combination of the tag-side B and the X reconstruction. The RMS deviation for q^2 is $3.35\,{\rm GeV}^2$. 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 $q^2 > 3.0\,{\rm GeV}^2$, 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 $3.0\,{\rm GeV}^2$ is dominated by processes other than $B \to X_c\,\ell^+\,\nu_\ell$: secondary leptons from cascade decays and fake lepton candidates make up a prominent fraction of the selected event candidates.

Fig. 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 $\Upsilon(4S) \to B\bar{B}$. In the following, we separate the electron and muon candidates and analyze them separately.

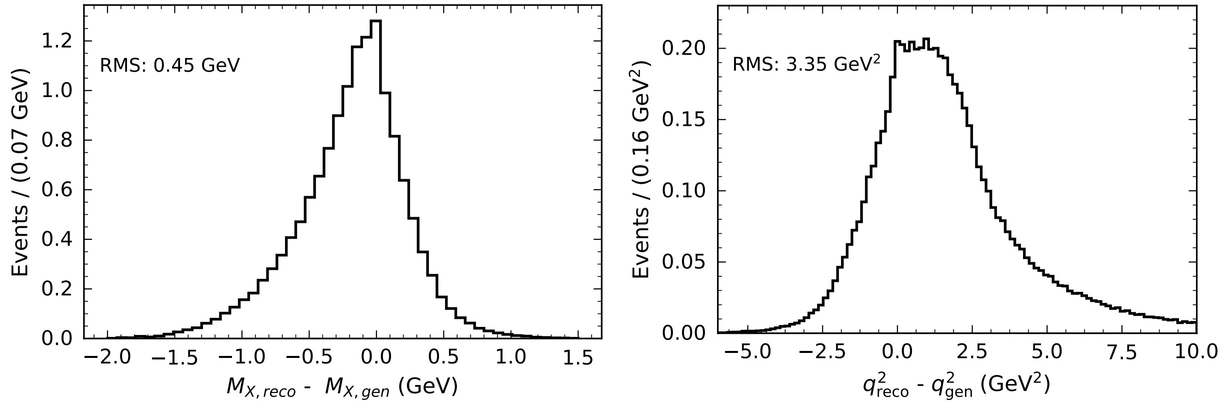

FIG. 2. The resolution of the reconstructed M_X and q^2 values for $B \to X_c\,\ell^+\,\nu_\ell$ signal is shown as a residual with respect to the generated values.

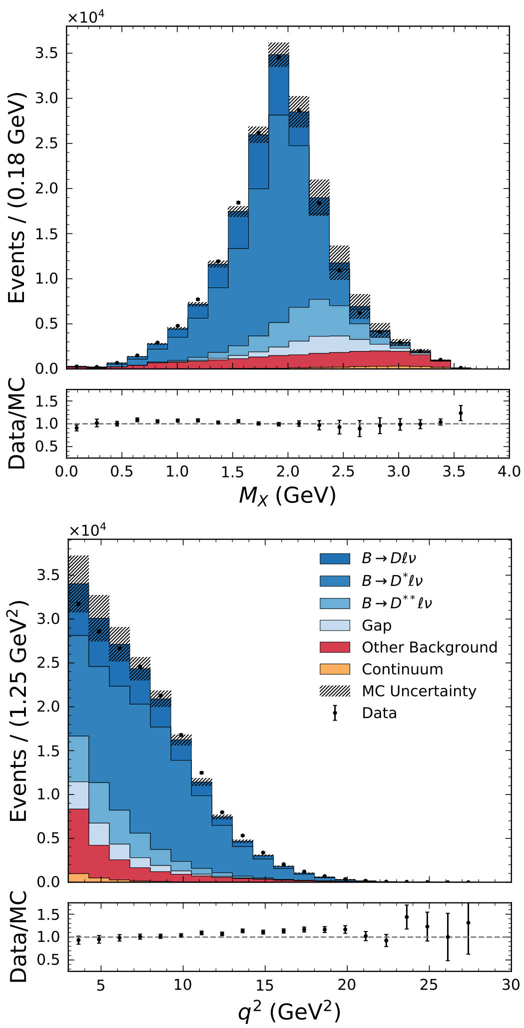

FIG. 3. The reconstructed M_X and q^2 distributions are shown. The error band of the simulated samples incorporates the full set of systematic uncertainties discussed in Section V.

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 $\mathcal{P}$,

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

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 labelling a given uncertainty source. More details of this procedure will be given in Section V. The Poisson term $\mathcal{P}_{\rm cont}$ 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

$$\nu_i = \eta^{\rm sig} f_i^{\rm sig} + \eta^{B\,\rm bkg} f_i^{B\,\rm bkg} + \eta^{\rm cont} f_i^{\rm cont}\,. \tag{18}$$

Here, $\eta^{\rm sig}$ is the total number of $B \to X_c\,\ell^+\,\nu_\ell$ signal events. Furthermore, $\eta^{B\,\rm bkg}$ denotes the background events stemming from double semileptonic cascades, $B \to X_u\,\ell^+\,\nu_\ell$ decays, and from hadrons misidentified as leptons originating from B meson decays. The number of continuum events is denoted as $\eta^{\rm cont}$. 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. Eq. 17 is numerically maximized to determine both the total number of $B \to X_c\,\ell^+\,\nu_\ell$ and background events from the

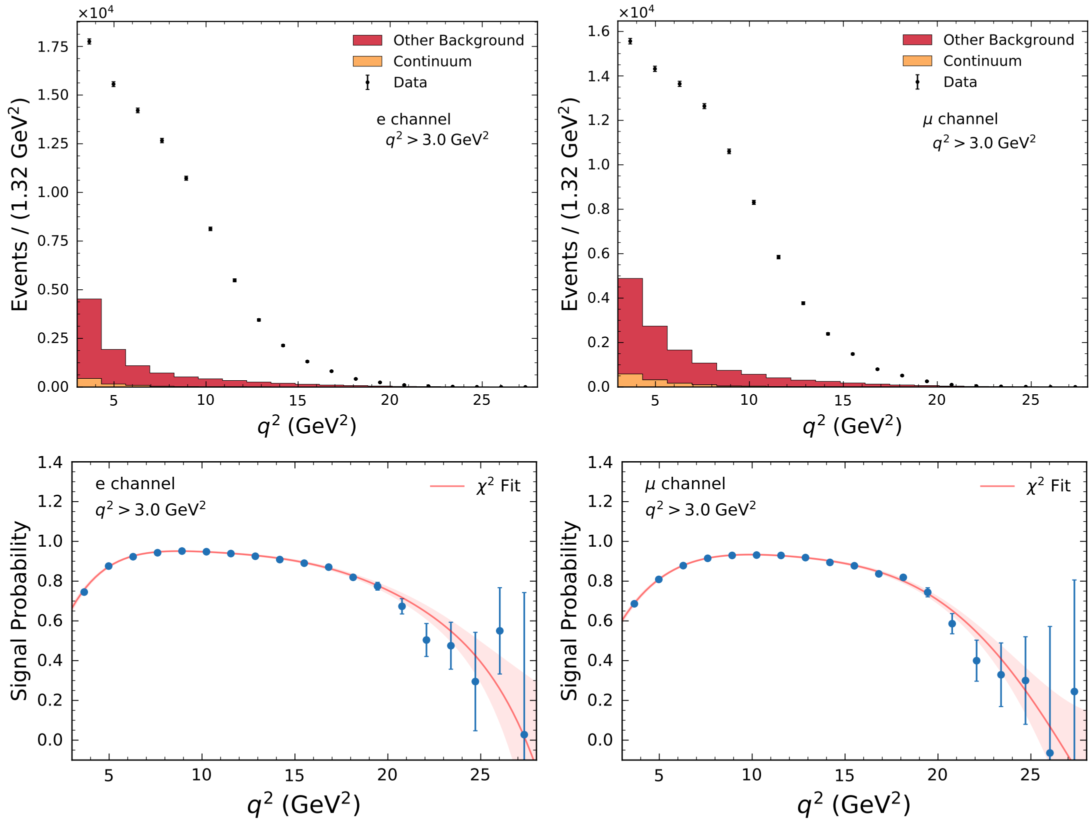

FIG. 4. The reconstructed q^2 distributions with an example q^2 threshold selection of $3.0\,\mathrm{GeV}^2$ (top) for electron (left) and muon (right) candidates and the determined binned signal probabilities (bottom) are shown. The background contributions are scaled to their estimated values using 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).

observed event yields. This is done using the sequential least squares programming method implementation of Ref. [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 $B \to X_c\, \ell^+\, \nu_\ell$ and background templates (more details about these will be discussed in Section V). The continuum constraint $\mathcal{P}_{\mathrm{cont}}$ is adjusted to reflect the number of continuum events for a given q^2 selection value.

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

$$w_i = 1 - \frac{\widehat{\eta}^{\mathrm{bkg}} \tilde{f}_i^{\,\mathrm{bkg}} + \widehat{\eta}^{\mathrm{cont}} \tilde{f}_i^{\,\mathrm{cont}}}{n_i}\,. \tag{19}$$

Here, $\tilde{f}_i$ denotes the estimated fractions of events being reconstructed in a bin i of q^2 for a given background category. Fig. 4 shows the q^2 spectrum for electron and muon candidates and the w_i distribution for the threshold selection with $q^2 > 3.0\,\mathrm{GeV}^2$. 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 A.

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 $B \to X_c\, \ell^+\, \nu_\ell$ q^2 spectrum, we need to correct for these effects. This is done in three sequential steps:

- First, a calibration function is applied event-by-

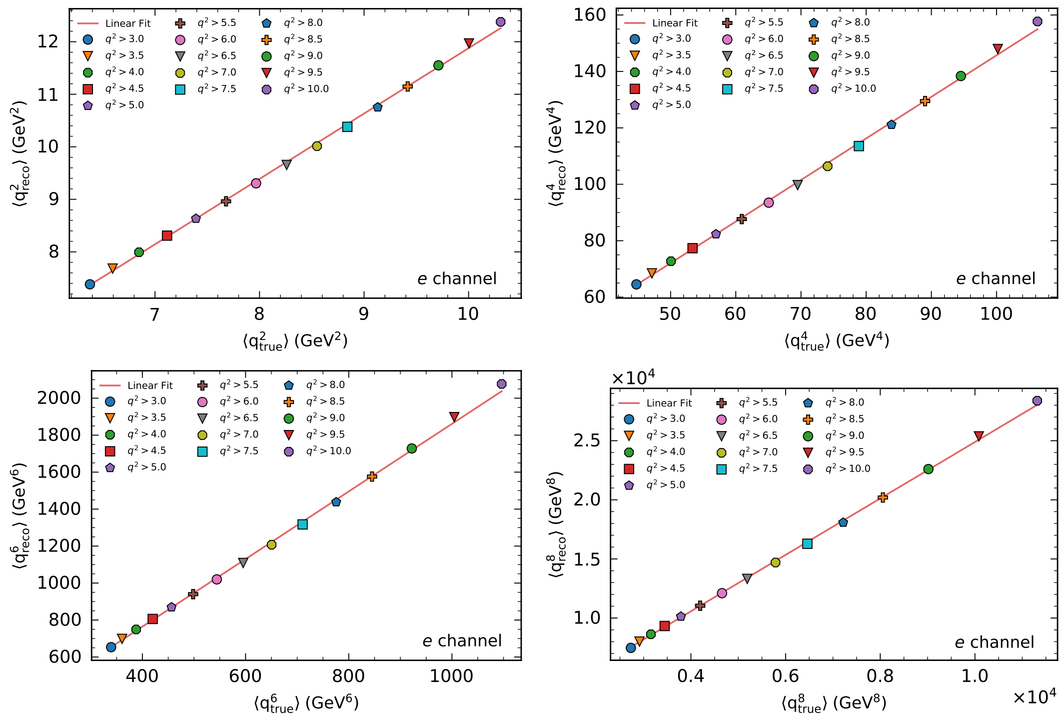

FIG. 5. The generator-level, $\langle q^{2m}_{\rm true} \rangle$, and reconstructed, $\langle q^{2m}_{\rm reco} \rangle$, $B \to X_c\, \ell^+\, \nu_\ell$ moments are shown for the first to fourth q^2 moment for electrons. The different markers represent different threshold selections on the generator-level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

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

- In a second step we correct for any residual bias not captured by the linear calibration function for a given moment using a global correction factor, which we denote as $C_{\rm cal}$.

- Thirdly, we correct for acceptance and selection effects by applying a global correction factor, $C_{\rm acc}$. 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. Fig. 5 shows the first to fourth generator-level and reconstructed $B \to X_c\, \ell^+\, \nu_\ell$ 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,

$$q^{2m}_{{\rm cal}\, i} = \frac{q_i^{2m} - a_m}{b_m} \,. \qquad (20)$$

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 B and show the same linear behavior. The residual bias correction factor $C_{\rm cal}$ 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 $B \to X_c\, \ell^+\, \nu_\ell$ processes, we determine an additional factor accounting for selection and acceptance effects. The correction factor, $C_{\rm acc}$, is calculated by comparing the moments of the generator-level simulated events with a sample without any selection criteria applied. Fig. 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, selection and acceptance efficiencies for generator-level simulated $B \to X_c\, \ell^+\, \nu_\ell$ samples are shown in Appendix C.

The q^2 moments are then given by

$$\langle q^{2m} \rangle = \frac{C_{\rm cal} \cdot C_{\rm acc}}{\sum_i^{\rm events} w(q_i^2)} \times \sum_i^{\rm events} w(q_i^2) \cdot q^{2m}_{{\rm cal}\, i} \,. \qquad (21)$$

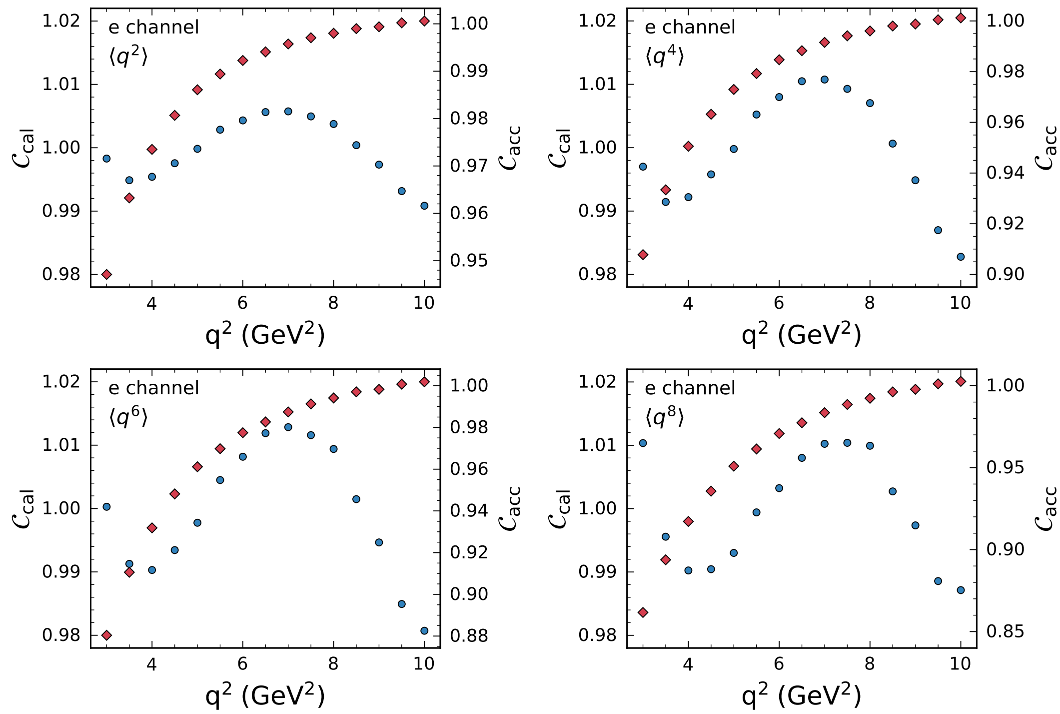

FIG. 6. The residual bias and acceptance correction factors, denoted as $C_{\rm cal}$ (circles) and C_{acc} (diamonds), respectively, are shown for the first to fourth q^2 moment for electrons. The corresponding plots for muons can be found in Appendix B.

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 $q^{2m}_{{\rm cal}\,i}$ 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

A. Overview

Several systematic uncertainties affect the measured q^2 moments. 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. The most important sources of systematic uncertainty are associated with the assumed composition of the $B \to X_c\,\ell^+\,\nu_\ell$ 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. We evaluate the uncertainties on the background subtraction by considering various sources of systematic uncertainty included as NP constraints in the binned likelihood fits (labeled "Bkg. subtraction" in Tables II and III). We consider signal modeling uncertainties by variations of the BGL parameters and heavy quark form factors of $B \to D\,\ell^+\,\nu_\ell$, $B \to D^*\,\ell^+\,\nu_\ell$, and $B \to D^{**}\,\ell^+\,\nu_\ell$ within their uncertainties. In addition, we propagate the branching fraction uncertainties, cf. Table I. The uncertainties on the $B \to X_c\,\ell^+\,\nu_\ell$ gap branching fractions are taken to be large enough to account for the difference between the sum of all exclusive branching fractions measured and the inclusive branching fraction measured. We also evaluate the impact on the efficiency of the lepton- and hadron-identification uncertainties, and the overall tracking efficiency uncertainty. The statistical uncertainty on all generated MC samples is also evaluated and propagated into the systematic errors. The $B \to X_u\,\ell^+\,\nu_\ell$ background component is varied within its branching fraction uncertainty ("$B \to X_u \ell\nu$ BF").

For the calibration we change the composition of the assumed $B \to X_c\,\ell^+\,\nu_\ell$ process and redetermine the calibration function as well as the bias and acceptance correction factors ("$B \to X_c \ell\nu$ BF"). For the $B \to D^{**}\,\ell^+\,\nu_\ell$ 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 $B \to D\,\ell^+\,\nu_\ell$ and $B \to D^*\,\ell^+\,\nu_\ell$ within their respective branching fractions. In addition, we evaluate the impact of the modeling of non-resonant decays by replacing the model described in Section II with a model based on the equidistribution of all final-state particles in phase-space ("Non-resonant model"). The signal modeling is evaluated in a similar manner

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)	1.416	1.360	1.312	1.271	1.235	1.206	1.185	1.170	1.163	1.163	1.172	1.189	1.215	1.249	1.290
Bkg. subtraction	1.057	0.764	0.568	0.404	0.372	0.405	0.478	0.551	0.598	0.635	0.670	0.705	0.775	0.829	0.908
$B \to X_u \ell \nu$ BF	1.805	1.663	1.517	1.095	0.783	0.566	0.415	0.291	0.233	0.186	0.156	0.132	0.122	0.147	0.103
$B \to X_c \ell \nu$ BF	4.942	4.719	4.898	4.546	4.469	4.295	3.773	3.467	3.281	2.988	2.491	1.963	1.413	1.273	1.164
Non-resonant model	13.198	11.539	10.140	8.328	6.932	5.780	4.498	3.542	2.764	2.214	1.626	1.241	0.812	0.772	0.663
$B \to X_c \ell \nu$ FF	1.632	1.466	1.271	1.101	0.936	0.839	0.788	0.746	0.685	0.632	0.558	0.493	0.424	0.402	0.367
N_{tracks} res.	4.665	4.397	4.167	3.910	3.717	3.500	3.314	3.134	2.955	2.825	2.659	2.500	2.400	2.243	2.132
N_γ res.	0.423	0.414	0.367	0.361	0.335	0.290	0.277	0.273	0.273	0.258	0.245	0.232	0.235	0.238	0.226
MC non-closure	0.064	0.092	0.029	0.014	0.014	0.026	0.025	0.012	0.012	0.011	0.000	0.032	0.000	0.010	0.000
Cal. function	0.135	0.079	0.026	0.025	0.076	0.124	0.171	0.216	0.259	0.299	0.341	0.380	0.418	0.455	0.490
Stat. bias corr.	1.223	1.182	1.145	1.112	1.085	1.061	1.043	1.029	1.021	1.019	1.022	1.032	1.049	1.072	1.102
PID eff.	0.174	0.156	0.144	0.133	0.127	0.108	0.101	0.088	0.090	0.085	0.081	0.077	0.067	0.055	0.047
Track eff.	0.416	0.386	0.361	0.335	0.313	0.290	0.271	0.253	0.233	0.219	0.204	0.188	0.177	0.161	0.151
Sys. error (total)	15.143	13.491	12.248	10.458	9.213	8.162	6.910	6.048	5.409	4.888	4.239	3.702	3.252	3.092	2.973
Total rel. error in ‰	15.209	13.559	12.318	10.535	9.295	8.250	7.011	6.161	5.533	5.025	4.398	3.889	3.472	3.335	3.241
$\langle q^4 \rangle$ in GeV^4	43.01	46.28	49.74	53.42	57.42	61.64	65.96	70.45	75.16	79.92	85.18	90.50	96.05	101.71	108.10
Stat. error (data)	3.229	3.133	3.043	2.959	2.883	2.817	2.771	2.739	2.722	2.722	2.738	2.778	2.840	2.918	3.007
Bkg. subtraction	1.775	1.346	1.138	0.986	1.061	1.168	1.334	1.492	1.580	1.646	1.713	1.776	1.929	2.048	2.226
$B \to X_u \ell \nu$ BF	3.841	3.524	3.156	2.287	1.642	1.212	0.895	0.634	0.525	0.431	0.366	0.314	0.290	0.341	0.240
$B \to X_c \ell \nu$ BF	9.447	9.452	9.934	9.479	9.409	9.102	8.134	7.537	7.106	6.449	5.404	4.307	3.183	2.875	2.622
Non-resonant model	25.190	22.358	19.837	16.571	13.952	11.720	9.247	7.364	5.805	4.678	3.487	2.693	1.825	1.725	1.492
$B \to X_c \ell \nu$ FF	3.132	2.840	2.506	2.218	1.953	1.805	1.719	1.637	1.518	1.409	1.255	1.121	0.979	0.925	0.846
N_{tracks} res.	10.437	9.946	9.493	8.980	8.560	8.088	7.670	7.256	6.842	6.521	6.126	5.747	5.484	5.114	4.837
N_γ res.	0.970	0.949	0.868	0.850	0.798	0.712	0.684	0.672	0.668	0.633	0.605	0.575	0.579	0.580	0.556
MC non-closure	0.028	0.013	0.014	0.011	0.009	0.011	0.014	0.009	0.009	0.010	0.004	0.011	0.002	0.003	0.004
Cal. function	0.269	0.149	0.030	0.084	0.201	0.312	0.417	0.517	0.613	0.702	0.792	0.876	0.955	1.028	1.097
Stat. bias corr.	2.684	2.626	2.571	2.519	2.472	2.431	2.396	2.370	2.355	2.347	2.351	2.370	2.401	2.447	2.504
PID eff.	0.364	0.335	0.312	0.290	0.278	0.241	0.225	0.199	0.200	0.187	0.178	0.166	0.144	0.119	0.101
Track eff.	0.904	0.850	0.803	0.751	0.706	0.657	0.616	0.575	0.531	0.498	0.464	0.427	0.400	0.364	0.340
Sys. error (total)	29.489	26.816	24.657	21.539	19.273	17.284	14.928	13.256	11.949	10.842	9.503	8.382	7.450	7.077	6.793
Total rel. error in ‰	29.665	26.999	24.844	21.742	19.487	17.512	15.183	13.536	12.255	11.179	9.890	8.831	7.972	7.655	7.428
$\langle q^6 \rangle$ in GeV^6	326.35	355.59	387.60	423.96	465.18	510.57	558.68	610.83	668.01	728.17	797.42	870.56	948.71	1031.66	1128.53
Stat. error (data)	5.835	5.695	5.555	5.403	5.258	5.125	5.031	4.959	4.908	4.889	4.891	4.938	5.027	5.145	5.270
Bkg. subtraction	2.552	2.175	2.199	2.137	2.380	2.544	2.808	3.048	3.138	3.217	3.298	3.374	3.616	3.806	4.108
$B \to X_u \ell \nu$ BF	6.434	5.824	5.072	3.656	2.622	1.968	1.454	1.039	0.884	0.744	0.646	0.554	0.511	0.589	0.417
$B \to X_c \ell \nu$ BF	13.833	14.413	15.245	14.805	14.753	14.317	12.966	12.087	11.373	10.312	8.699	7.013	5.315	4.811	4.378
Non-resonant model	35.926	32.230	28.839	24.431	20.788	17.602	14.080	11.342	9.036	7.329	5.546	4.331	3.032	2.853	2.482
$B \to X_c \ell \nu$ FF	4.422	4.101	3.733	3.410	3.119	2.956	2.841	2.717	2.536	2.359	2.119	1.905	1.682	1.579	1.443
N_{tracks} res.	17.448	16.759	16.094	15.315	14.632	13.866	13.166	12.460	11.747	11.158	10.462	9.786	9.284	8.632	8.122
N_γ res.	1.698	1.665	1.556	1.522	1.443	1.317	1.269	1.242	1.224	1.165	1.114	1.062	1.060	1.054	1.012
MC non-closure	0.007	0.000	0.004	0.003	0.003	0.002	0.003	0.001	0.001	0.002	0.001	0.002	0.001	0.000	0.001
Cal. function	0.421	0.220	0.022	0.178	0.379	0.572	0.753	0.924	1.087	1.239	1.388	1.527	1.654	1.768	1.874
Stat. bias corr.	4.681	4.600	4.519	4.440	4.364	4.293	4.229	4.177	4.139	4.113	4.102	4.115	4.146	4.200	4.273
PID eff.	0.571	0.532	0.500	0.468	0.448	0.395	0.369	0.330	0.328	0.305	0.288	0.266	0.229	0.190	0.160
Track eff.	1.468	1.395	1.327	1.251	1.181	1.105	1.038	0.971	0.899	0.841	0.781	0.718	0.670	0.609	0.565
Sys. error (total)	43.374	40.113	37.316	33.227	30.150	27.350	24.063	21.648	19.666	17.935	15.885	14.142	12.702	12.058	11.557
Total rel. error in ‰	43.765	40.515	37.727	33.664	30.605	27.826	24.583	22.209	20.270	18.589	16.621	14.980	13.660	13.110	12.702
$\langle q^8 \rangle$ in GeV^8	2664.70	2914.93	3192.31	3531.36	3925.00	4374.90	4862.59	5411.82	6037.81	6721.98	7539.01	8441.07	9424.26	10507.43	11816.35
Stat. error (data)	9.768	9.579	9.382	9.110	8.845	8.583	8.395	8.230	8.093	8.000	7.932	7.933	8.012	8.133	8.249
Bkg. subtraction	4.024	3.885	4.315	4.242	4.674	4.850	5.204	5.519	5.514	5.582	5.647	5.705	6.046	6.313	6.768
$B \to X_u \ell \nu$ BF	9.951	8.850	7.469	5.319	3.786	2.872	2.111	1.517	1.321	1.135	1.009	0.866	0.797	0.905	0.638
$B \to X_c \ell \nu$ BF	19.150	20.111	21.080	20.543	20.401	19.775	18.077	16.927	15.922	14.460	12.294	10.031	7.782	7.060	6.419
Non-resonant model	45.701	41.230	37.105	31.734	27.222	23.213	18.805	15.314	12.331	10.077	7.733	6.107	4.404	4.130	3.613
$B \to X_c \ell \nu$ FF	5.568	5.325	5.032	4.751	4.488	4.329	4.188	4.012	3.759	3.502	3.162	2.854	2.538	2.369	2.158
N_{tracks} res.	25.809	24.892	23.999	22.905	21.906	20.796	19.767	18.715	17.642	16.715	15.643	14.598	13.779	12.781	11.972
N_γ res.	2.675	2.625	2.488	2.430	2.317	2.148	2.073	2.022	1.980	1.887	1.806	1.722	1.704	1.680	1.612
MC non-closure	0.001	0.000	0.001	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Cal. function	0.621	0.311	0.012	0.308	0.624	0.926	1.205	1.469	1.720	1.951	2.178	2.387	2.572	2.736	2.886
Stat. bias corr.	7.594	7.452	7.307	7.164	7.024	6.890	6.761	6.650	6.560	6.482	6.426	6.403	6.405	6.442	6.507
PID eff.	0.792	0.744	0.704	0.661	0.632	0.564	0.527	0.475	0.467	0.432	0.405	0.371	0.320	0.265	0.224
Track eff.	2.113	2.020	1.932	1.829	1.730	1.624	1.529	1.433	1.329	1.242	1.152	1.058	0.982	0.892	0.825
Sys. error (total)	57.775	53.971	50.603	45.647	41.857	38.325	34.268	31.191	28.548	26.189	23.438	21.056	19.106	18.145	17.391
Total rel. error in ‰	58.595	54.815	51.465	46.547	42.781	39.274	35.281	32.259	29.673	27.383	24.744	22.501	20.718	19.884	19.248

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	6.26	6.54	6.84	7.13	7.42	7.72	8.02	8.32	8.61	8.91	9.21	9.51	9.80	10.09	10.41
Stat. error (data)	1.511	1.447	1.390	1.341	1.300	1.267	1.239	1.219	1.207	1.202	1.205	1.215	1.234	1.257	1.289
Bkg. subtraction	1.337	1.121	0.900	0.714	0.594	0.529	0.486	0.491	0.570	0.626	0.649	0.704	0.758	0.768	0.818
$B \to X_u \ell \nu$ BF	2.182	2.044	1.753	1.483	1.352	1.036	0.755	0.540	0.383	0.290	0.189	0.162	0.119	0.053	0.048
$B \to X_c \ell \nu$ BF	4.817	5.017	5.142	5.141	5.048	4.999	4.672	4.049	3.511	3.111	2.656	2.206	1.755	1.363	1.157
Non-resonant model	14.247	12.717	11.040	9.284	7.825	6.624	5.416	3.997	3.019	2.281	1.646	1.425	1.041	0.858	0.778
$B \to X_c \ell \nu$ FF	1.433	1.300	1.158	1.029	0.914	0.846	0.820	0.743	0.687	0.622	0.540	0.481	0.425	0.388	0.355
N_{tracks} res.	5.658	5.313	4.958	4.654	4.361	4.061	3.777	3.519	3.288	3.055	2.853	2.660	2.508	2.381	2.197
N_γ res.	0.293	0.277	0.258	0.234	0.227	0.214	0.213	0.209	0.205	0.186	0.184	0.157	0.151	0.150	0.161
MC non-closure	0.192	0.107	0.117	0.112	0.108	0.052	0.050	0.060	0.081	0.067	0.109	0.042	0.041	0.059	0.019
Cal. function	0.133	0.082	0.029	0.022	0.072	0.121	0.169	0.217	0.262	0.306	0.352	0.394	0.435	0.474	0.511
Stat. bias corr.	1.321	1.274	1.229	1.190	1.156	1.125	1.100	1.081	1.067	1.059	1.057	1.061	1.072	1.089	1.112
PID eff.	0.157	0.141	0.135	0.126	0.128	0.116	0.108	0.103	0.095	0.099	0.091	0.082	0.081	0.072	0.062
Track eff.	0.444	0.415	0.387	0.363	0.338	0.312	0.289	0.268	0.248	0.230	0.212	0.195	0.181	0.169	0.154
Sys. error (total)	16.399	14.972	13.412	11.819	10.501	9.427	8.264	6.869	5.876	5.138	4.471	4.010	3.553	3.237	3.020
Total rel. error in ‰	16.468	15.042	13.484	11.895	10.582	9.512	8.356	6.976	5.998	5.277	4.631	4.190	3.762	3.473	3.283
$\langle q^4 \rangle$ in GeV^4	43.58	46.75	50.20	53.86	57.70	61.77	66.14	70.71	75.42	80.37	85.59	91.00	96.40	102.29	108.60
Stat. error (data)	3.453	3.331	3.220	3.117	3.027	2.950	2.885	2.832	2.803	2.788	2.792	2.815	2.858	2.908	2.975
Bkg. subtraction	2.410	2.097	1.765	1.479	1.358	1.338	1.302	1.339	1.527	1.646	1.673	1.804	1.897	1.899	2.001
$B \to X_u \ell \nu$ BF	4.657	4.389	3.803	3.198	2.898	2.225	1.609	1.144	0.811	0.610	0.390	0.337	0.256	0.106	0.096
$B \to X_c \ell \nu$ BF	9.548	10.302	10.678	10.776	10.674	10.504	9.839	8.632	7.523	6.669	5.690	4.730	3.784	2.968	2.522
Non-resonant model	27.340	24.667	21.669	18.477	15.721	13.364	10.997	8.235	6.299	4.817	3.533	3.049	2.267	1.877	1.691
$B \to X_c \ell \nu$ FF	2.729	2.514	2.287	2.077	1.897	1.796	1.751	1.606	1.495	1.362	1.197	1.071	0.952	0.869	0.792
N_{tracks} res.	12.449	11.812	11.129	10.508	9.892	9.247	8.627	8.044	7.513	6.978	6.495	6.038	5.664	5.337	4.903
N_γ res.	0.679	0.649	0.613	0.568	0.550	0.523	0.515	0.503	0.490	0.450	0.440	0.384	0.371	0.367	0.387
MC non-closure	0.055	0.036	0.038	0.033	0.031	0.019	0.018	0.018	0.027	0.020	0.029	0.013	0.009	0.012	0.007
Cal. function	0.268	0.156	0.040	0.075	0.187	0.300	0.409	0.516	0.615	0.713	0.811	0.900	0.983	1.062	1.135
Stat. bias corr.	2.920	2.845	2.771	2.702	2.638	2.577	2.526	2.483	2.450	2.428	2.416	2.417	2.431	2.457	2.495
PID eff.	0.346	0.318	0.306	0.289	0.289	0.264	0.246	0.233	0.217	0.220	0.201	0.181	0.176	0.156	0.134
Track eff.	0.966	0.914	0.860	0.810	0.758	0.705	0.654	0.607	0.563	0.520	0.479	0.440	0.406	0.376	0.342
Sys. error (total)	32.229	29.893	27.187	24.353	21.926	19.799	17.516	14.826	12.843	11.324	9.919	8.926	7.957	7.256	6.760
Total rel. error in ‰	32.413	30.078	27.377	24.551	22.134	20.017	17.752	15.095	13.145	11.662	10.304	9.359	8.455	7.817	7.386
$\langle q^6 \rangle$ in GeV^6	332.46	360.82	393.33	429.39	468.40	512.19	561.28	614.68	671.93	734.16	803.00	876.94	953.08	1041.01	1137.25
Stat. error (data)	6.223	6.029	5.845	5.660	5.494	5.340	5.204	5.083	5.011	4.961	4.939	4.960	5.016	5.075	5.156
Bkg. subtraction	3.615	3.280	2.944	2.667	2.667	2.767	2.710	2.769	3.084	3.256	3.244	3.484	3.566	3.524	3.678
$B \to X_u \ell \nu$ BF	7.775	7.301	6.340	5.260	4.735	3.623	2.571	1.807	1.273	0.952	0.594	0.515	0.413	0.154	0.136
$B \to X_c \ell \nu$ BF	14.817	16.142	16.677	16.806	16.674	16.299	15.287	13.537	11.864	10.535	8.996	7.496	6.036	4.776	4.063
Non-resonant model	39.151	35.583	31.544	27.197	23.346	19.947	16.529	12.551	9.719	7.518	5.593	4.818	3.640	3.023	2.708
$B \to X_c \ell \nu$ FF	3.864	3.649	3.418	3.189	2.996	2.883	2.820	2.608	2.438	2.232	1.978	1.775	1.582	1.439	1.305
N_{tracks} res.	20.450	19.543	18.546	17.594	16.629	15.592	14.582	13.603	12.703	11.789	10.934	10.136	9.458	8.849	8.097
N_γ res.	1.194	1.150	1.097	1.031	0.999	0.953	0.932	0.904	0.878	0.812	0.787	0.700	0.675	0.663	0.686
MC non-closure	0.008	0.006	0.007	0.006	0.006	0.003	0.003	0.003	0.006	0.005	0.007	0.004	0.003	0.003	0.002
Cal. function	0.424	0.236	0.040	0.155	0.346	0.541	0.727	0.911	1.078	1.241	1.404	1.547	1.679	1.803	1.915
Stat. bias corr.	5.096	4.981	4.863	4.748	4.638	4.528	4.429	4.344	4.270	4.212	4.169	4.146	4.144	4.158	4.192
PID eff.	0.566	0.528	0.510	0.484	0.480	0.442	0.413	0.389	0.362	0.361	0.329	0.294	0.283	0.249	0.213
Track eff.	1.569	1.496	1.417	1.341	1.262	1.176	1.094	1.016	0.942	0.870	0.799	0.732	0.673	0.620	0.561
Sys. error (total)	47.847	44.886	41.289	37.447	34.096	30.968	27.638	23.776	20.853	18.542	16.349	14.778	13.248	12.091	11.257
Total rel. error in ‰	48.250	45.289	41.701	37.872	34.536	31.425	28.124	24.313	21.446	19.194	17.078	15.588	14.166	13.113	12.382
$\langle q^8 \rangle$ in GeV^8	2726.97	2969.47	3260.51	3593.17	3958.00	4390.97	4895.15	5460.63	6089.21	6791.52	7606.31	8511.15	9470.48	10648.37	11956.19
Stat. error (data)	10.350	10.068	9.785	9.475	9.189	8.893	8.626	8.360	8.191	8.052	7.936	7.910	7.939	7.946	7.988
Bkg. subtraction	5.566	5.251	4.985	4.802	4.991	5.227	5.022	5.056	5.510	5.707	5.583	5.998	5.959	5.813	6.017
$B \to X_u \ell \nu$ BF	11.940	11.101	9.606	7.821	6.997	5.308	3.661	2.528	1.757	1.305	0.787	0.694	0.591	0.196	0.157
$B \to X_c \ell \nu$ BF	21.509	22.912	23.239	23.142	22.841	22.135	20.758	18.498	16.308	14.526	12.432	10.401	8.438	6.736	5.742
Non-resonant model	49.931	45.521	40.561	35.221	30.447	26.127	21.796	16.749	13.122	10.264	7.729	6.656	5.104	4.252	3.791
$B \to X_c \ell \nu$ FF	4.908	4.764	4.596	4.403	4.235	4.120	4.033	3.755	3.520	3.234	2.883	2.592	2.312	2.093	1.889
N_{tracks} res.	29.723	28.505	27.155	25.823	24.475	22.985	21.536	20.089	18.761	17.405	16.092	14.887	13.834	12.863	11.732
N_γ res.	1.877	1.814	1.740	1.649	1.598	1.526	1.485	1.434	1.386	1.289	1.241	1.118	1.075	1.048	1.067
MC non-closure	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.000	0.000	0.000
Cal. function	0.631	0.342	0.042	0.261	0.552	0.856	1.140	1.424	1.676	1.920	2.166	2.374	2.564	2.742	2.898
Stat. bias corr.	8.211	8.005	7.793	7.585	7.385	7.180	6.994	6.825	6.671	6.541	6.428	6.347	6.294	6.265	6.268
PID eff.	0.814	0.766	0.740	0.704	0.693	0.642	0.600	0.564	0.523	0.515	0.467	0.417	0.396	0.347	0.296
Track eff.	2.253	2.156	2.050	1.946	1.836	1.716	1.600	1.486	1.378	1.272	1.164	1.065	0.977	0.894	0.807
Sys. error (total)	64.139	60.464	55.941	51.099	46.912	42.794	38.471	33.549	29.788	26.721	23.716	21.581	19.450	17.774	16.563
Total rel. error in ‰	64.969	61.296	56.790	51.970	47.804	43.708	39.426	34.575	30.894	27.908	25.009	22.985	21.008	19.469	18.389

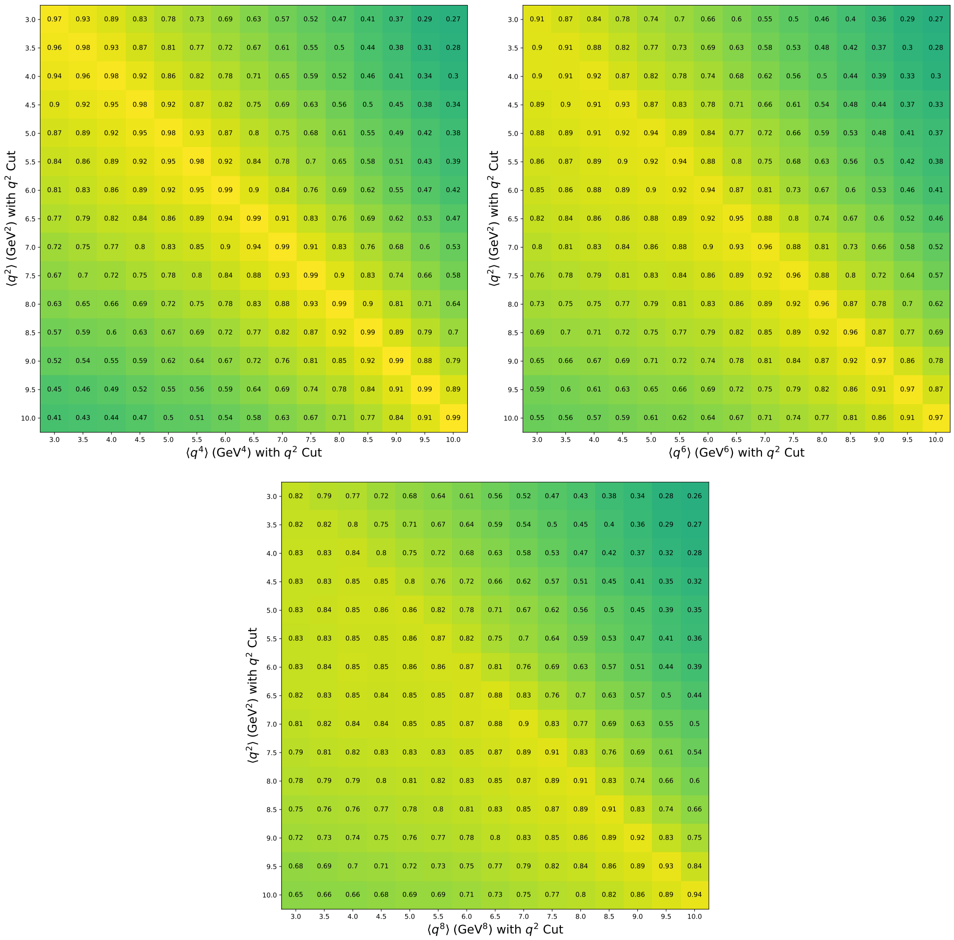

FIG. 7. Statistical correlations between the first and the second, third, and fourth moment for the electron final state.

as described above by using the variations of the BGL parameters and heavy quark form factors ("$B \to X_c \ell\nu$ FF"). We correct the number of charged particles and neutral clusters in the X system to the number observed in the recorded collision events to estimate an uncertainty associated with the modeling of the resolution ("N_{tracks} res." & "N_γ res."). The additional remaining bias due to the moment extraction method is estimated by making use of the ensemble tests described in Section IV ("MC non-closure"). 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."). Lastly, the impact of the efficiency of the lepton- and hadron-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.

The dominant systematic uncertainties stem from the uncertainty associated with the modeling of the $B \to X_c\,\ell^+\,\nu_\ell$ composition, especially the non-resonant 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 $B \to X_u\,\ell^+\,\nu_\ell$ 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. Conversely, the uncertainties involving the efficiency of the lepton- and hadron-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.

B. Statistical Correlations

The statistical correlation of the measured moments is determined using a bootstrapping approach [61, 62]: replicas of the measured collision data sets are created using sampling by replacement and the entire analysis procedure is repeated to estimate the statistical correlations between 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. We similarly estimate the covariance from the efficiency of lepton- and hadron-identification. For the remaining systematic uncertainties, we fully correlate or anti-correlate identical sources across bins. 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 D.

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 $q^2 > 3.0\,\mathrm{GeV}^2$ is about one standard deviation higher in data than in our simulated samples. Fig. 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 $B \to X_c\,\ell^+\,\nu_\ell$ 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 $B \to X_c\,\ell^+\,\nu_\ell$ 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 $B \to X_c\,\ell^+\,\nu_\ell$ model similar to the one used in this measurement. The data prefer to increase the $B \to X_c\,\ell^+\,\nu_\ell$ 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 $B \to X_c\,\ell^+\,\nu_\ell$. 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 calculating the χ^2 for each order of measured q^2 moments. The calculated values range between 2.53 and 5.23 with corresponding p-values of 0.99 and 0.98, respectively.

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 non-linear transformations are applied:

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

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

$$\mathcal{C}' = \mathcal{J} \mathcal{C} \mathcal{J}^T . \tag{23}$$

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 E 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 $B \to X_c\,\ell^+\,\nu_\ell$ 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

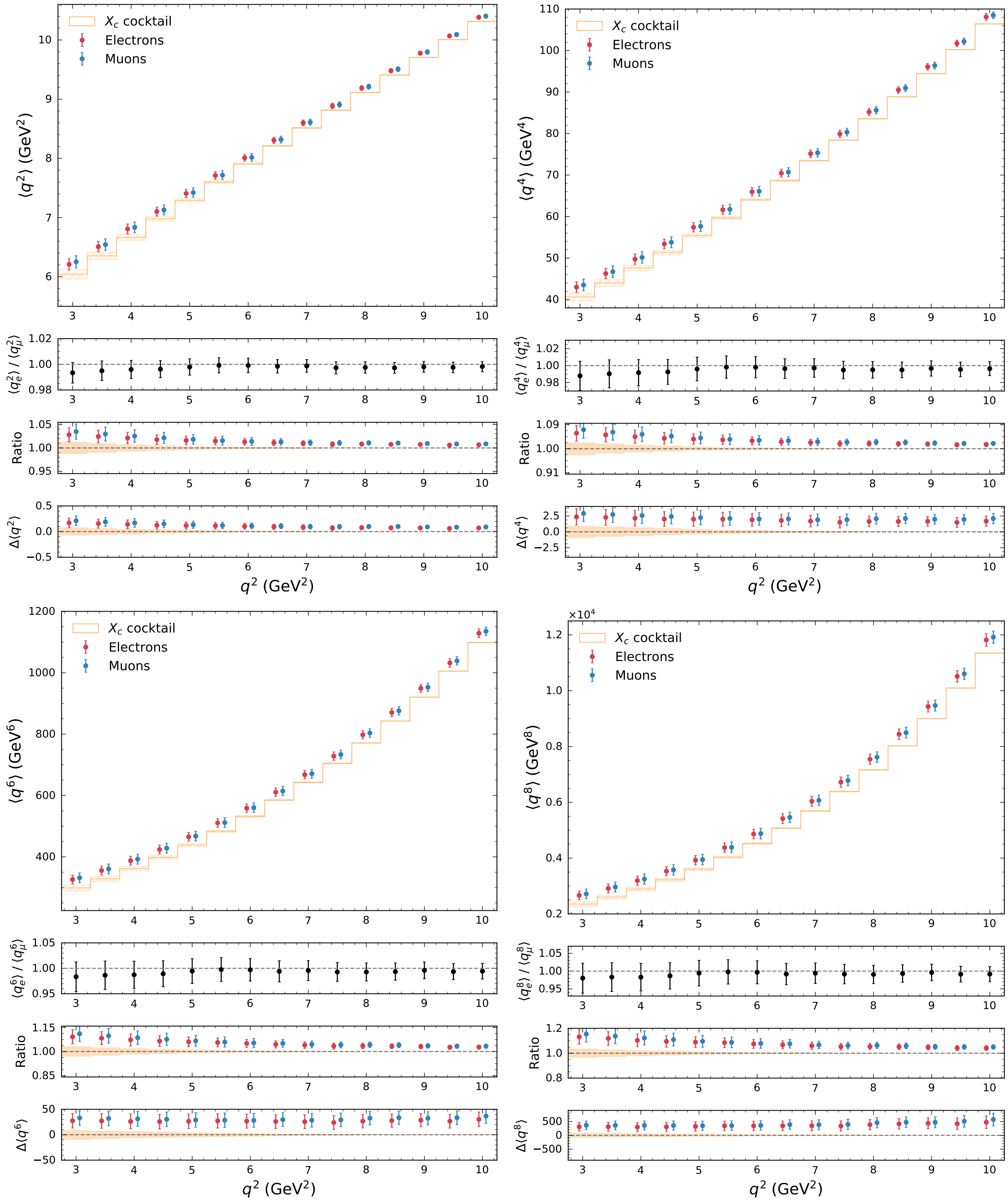

FIG. 8. The expectation of lepton flavor universality of the moments are tested for the first to fourth q^2 moments: in the ratio of electron to muon moments many of the associated systematic uncertainties cancel and all reported moments are compatible with the expectation of lepton flavor universality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

in Ref. [35]: using reparametrization invariance the set of non-perturbative 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 data set 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 $B \to X_c\, \ell^+\, \nu_\ell$ composition. The moments are measured separately for electron and muon $B \to X_c\, \ell^+\, \nu_\ell$ 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 are observed. In addition to measuring the nominal q^2 moments, a non-linear 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 $B \to X_c\, \ell^+\, \nu_\ell$. 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 $B \to X_c\, \ell^+\, \nu_\ell$ model for low q^2 threshold selections. As the $B \to X_c\, \ell^+\, \nu_\ell$ 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 by Refs. [63, 64]. One of the leading sources of systematics for low q^2 threshold selections is observed to be the modeling of the $B \to X_u\, \ell^+\, \nu_\ell$ 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 $B \to X_u\, \ell^+\, \nu_\ell$ component, the authors suggest measuring the full $B \to X\, \ell^+\, \nu_\ell$ spectrum and obtaining the $B \to X_u\, \ell^+\, \nu_\ell$ 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 (https://www.hepdata.net).

ACKNOWLEDGMENTS

We thank Keri Vos, Kevin Olschewsky, and Matteo Fael for useful discussions about the subject matter of this manuscript. RvT, LC, WS, and FB 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 DP180102629, DP170102389, DP170102204, DP150103061, 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, 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 Grant Nos. 2016R1D1A1B-01010135, 2016R1D1A1B02012900, 2018R1A2B3003643, 2018R1A6A1A06024970, 2018R1D1A1B07047294, 2019K1A3A7A09033840, 2019R1I1A3A01058933; Radiation Science Research Institute, Foreign Large-size Research Facility Application Supporting project, the Global Science Experimental Data Hub Center of the Korea Institute of Science and Technology Information and KREONET/GLORIAD; the Polish Ministry of Science and Higher Education and the National Science Center; the Ministry of Science and Higher Education of the Russian Federation, Agreement 14.W03.31.0026, and the HSE University Basic Research Program, Moscow; University of Tabuk research grants S-1440-0321, S-0256-1438, and S-0280-1439 (Saudi Arabia); the Slovenian

Research Agency Grant Nos. J1-9124 and 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.

[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), Prog. Theor. Exp. Phys. **2020 083C01** (2020).
[4] K. Abe *et al.* (T2K Collaboration), Nature **580**, 339 (2020), arXiv:1910.03887 [hep-ex].
[5] E. Kou, P. Urquijo, *et al.* (Belle II Collaboration), Prog. Theor. Exp. Phys. **2019**, 123C01 (2019), [Erratum: PTEP 2020, 029201 (2020)], arXiv:1808.10567 [hep-ex].
[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), arXiv:2007.08234 [hep-ex].
[8] Y. S. Amhis *et al.* (HFLAV Collaboration), Eur. Phys. J. C **81**, 226 (2021), arXiv:1909.12524 [hep-ex].
[9] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. D **101**, 072004 (2020), arXiv:2001.03225 [hep-ex].
[10] D. Bigi and P. Gambino, Phys. Rev. D **94**, 094008 (2016), arXiv:1606.08030 [hep-ph].
[11] D. Bigi, P. Gambino, and S. Schacht, Phys. Lett. B **769**, 441 (2017), arXiv:1703.06124 [hep-ph].
[12] F. U. Bernlochner, Z. Ligeti, M. Papucci, and D. J. Robinson, Phys. Rev. D **96**, 091503 (2017), arXiv:1708.07134 [hep-ph].
[13] D. Ferlewicz, P. Urquijo, and E. Waheed, (2020), arXiv:2008.09341 [hep-ph].
[14] M. Jezabek and J. H. Kuhn, Nucl. Phys. B **314**, 1 (1989).
[15] V. Aquila, P. Gambino, G. Ridolfi, and N. Uraltsev, Nucl. Phys. B **719**, 77 (2005), arXiv:hep-ph/0503083.
[16] A. Pak and A. Czarnecki, Phys. Rev. D **78**, 114015 (2008), arXiv:0808.3509 [hep-ph].
[17] K. Melnikov, Phys. Lett. B **666**, 336 (2008), arXiv:0803.0951 [hep-ph].
[18] T. Becher, H. Boos, and E. Lunghi, JHEP **12**, 062 (2007), arXiv:0708.0855 [hep-ph].
[19] T. Mannel, A. A. Pivovarov, and D. Rosenthal, Phys. Lett. B **741**, 290 (2015), arXiv:1405.5072 [hep-ph].
[20] A. Alberti, P. Gambino, and S. Nandi, JHEP **01**, 147 (2014), arXiv:1311.7381 [hep-ph].
[21] T. Mannel, S. Turczyk, and N. Uraltsev, JHEP **11**, 109 (2010), arXiv:1009.4622 [hep-ph].
[22] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 051103 (2005), arXiv:hep-ex/0502003.
[23] S. E. Csorna *et al.* (CLEO), Phys. Rev. D **70**, 032002 (2004), arXiv:hep-ex/0403052.
[24] J. Abdallah *et al.* (DELPHI Collaboration), Eur. Phys. J. C **45**, 35 (2006), arXiv:hep-ex/0510024.
[25] B. Aubert *et al.* (BaBar Collaboration), Phys. Rev. D **69**, 111104 (2004), arXiv:hep-ex/0403030.
[26] B. Aubert *et al.* (BaBar Collaboration), Phys. Rev. D **81**, 032003 (2010), arXiv:0908.0415 [hep-ex].
[27] P. Urquijo *et al.* (Belle Collaboration), Phys. Rev. D **75**, 032001 (2007), arXiv:hep-ex/0610012.
[28] C. Schwanda *et al.* (Belle Collaboration), Phys. Rev. D **75**, 032005 (2007), arXiv:hep-ex/0611044.
[29] P. Gambino and C. Schwanda, Phys. Rev. D **89**, 014022 (2014), arXiv:1307.4551 [hep-ph].
[30] A. Alberti, P. Gambino, K. J. Healey, and S. Nandi, Phys. Rev. Lett. **114**, 061802 (2015), arXiv:1411.6560 [hep-ph].
[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), arXiv:hep-ph/0302262.
[33] P. Gambino, JHEP **09**, 055 (2011), arXiv:1107.3100 [hep-ph].
[34] M. Bona *et al.* (UTFit Collaboration), Presentation at the ICHEP 2020 Conference (Online) (2020).
[35] M. Fael, T. Mannel, and K. Keri Vos, JHEP **02**, 177 (2019), arXiv:1812.07472 [hep-ph].
[36] S. Kurokawa and E. Kikutani, Nucl. Instr. and. Meth. **A499**, 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. Meth. **A479**, 117 (2002), also see detector section in J. Brodzicka *et al.*, Prog. Theor. Exp. Phys. **2012**, 04D001 (2012).
[38] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima, and T. Tsukamoto, Nucl. Instr. and. Meth. **A485**, 490 (2002).
[39] A. Abashian *et al.*, Nucl. Instr. and. Meth. **A491**, 69 (2002).
[40] D. J. Lange, Nucl. Instr. and. Meth. **A462**, 152 (2001).
[41] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini, CERN-DD-EE-84-1 (1987).
[42] E. Barberio, B. van Eijk, and Z. Wąs, Comput. Phys. Commun. **66**, 115 (1991).
[43] C. G. Boyd, B. Grinstein, and R. F. Lebed, Phys. Rev. Lett. **74**, 4603 (1995), arXiv:hep-ph/9412324 [hep-ph].
[44] R. Glattauer *et al.* (Belle Collaboration), Phys. Rev. D **93**, 032006 (2016), arXiv:1510.03657 [hep-ex].
[45] B. Grinstein and A. Kobach, Phys. Lett. **B 771**, 359 (2017), arXiv:1703.08170 [hep-ph].
[46] E. Waheed *et al.* (Belle Collaboration), Phys. Rev. D **100**, 052007 (2019), arXiv:1809.03290 [hep-ex].
[47] F. U. Bernlochner and Z. Ligeti, Phys. Rev. D **95**, 014022 (2017), arXiv:1606.09300 [hep-ph].
[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), arXiv:0711.3252 [hep-ex].
[50] B. Aubert *et al.* (BaBar Collaboration), Phys. Rev. Lett. **101**, 261802 (2008), arXiv:0808.0528 [hep-ex].
[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), arXiv:0708.1621 [hep-ph].
[52] A. K. Leibovich, Z. Ligeti, I. W. Stewart, and M. B. Wise, Phys. Rev. **D57**, 308 (1998), arXiv:hep-ph/9705467 [hep-ph].

[53] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. D **84**, 092001 (2011), [Erratum: Phys.Rev.D 85, 039904 (2012)], arXiv:1109.6831 [hep-ex].

[54] J. Lees *et al.* (BaBar Collaboration), Phys. Rev. Lett. **116**, 041801 (2016), arXiv:1507.08303 [hep-ex].

[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), arXiv:2102.00020 [hep-ex].

[57] M. Feindt, F. Keller, M. Kreps, T. Kuhr, S. Neubauer, D. Zander, and A. Zupanc, Nucl. Instrum. Meth. A **654**, 432 (2011), arXiv:1102.3876 [hep-ex].

[58] Ed. A.J. Bevan, B. Golob, Th. Mannel, S. Prell, and B.D. Yabsley, Eur. Phys. J. C **74**, 3026 (2014, Page 95), arXiv:1406.6311 [hep-ex].

[59] F. James and M. Roos, Comput. Phys. Commun. **10**, 343 (1975).

[60] H. Dembinski, P. Ongmongkolkul, *et al.*, (2020), 10.5281/zenodo.3949207.

[61] B. Efron, The Annals of Statistics **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), arXiv:1611.05624 [hep-ex].

[64] F. U. Bernlochner, D. Biedermann, H. Lacker, and T. Lück, Eur. Phys. J. C **74**, 2914 (2014), arXiv:1402.2849 [hep-ph].

[65] T. Mannel, M. Rahimi, and K. Keri Vos, (2021), arXiv:2105.02163 [hep-ph].

APPENDIX A. 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 Fig. 9, 10, 11, and 12.

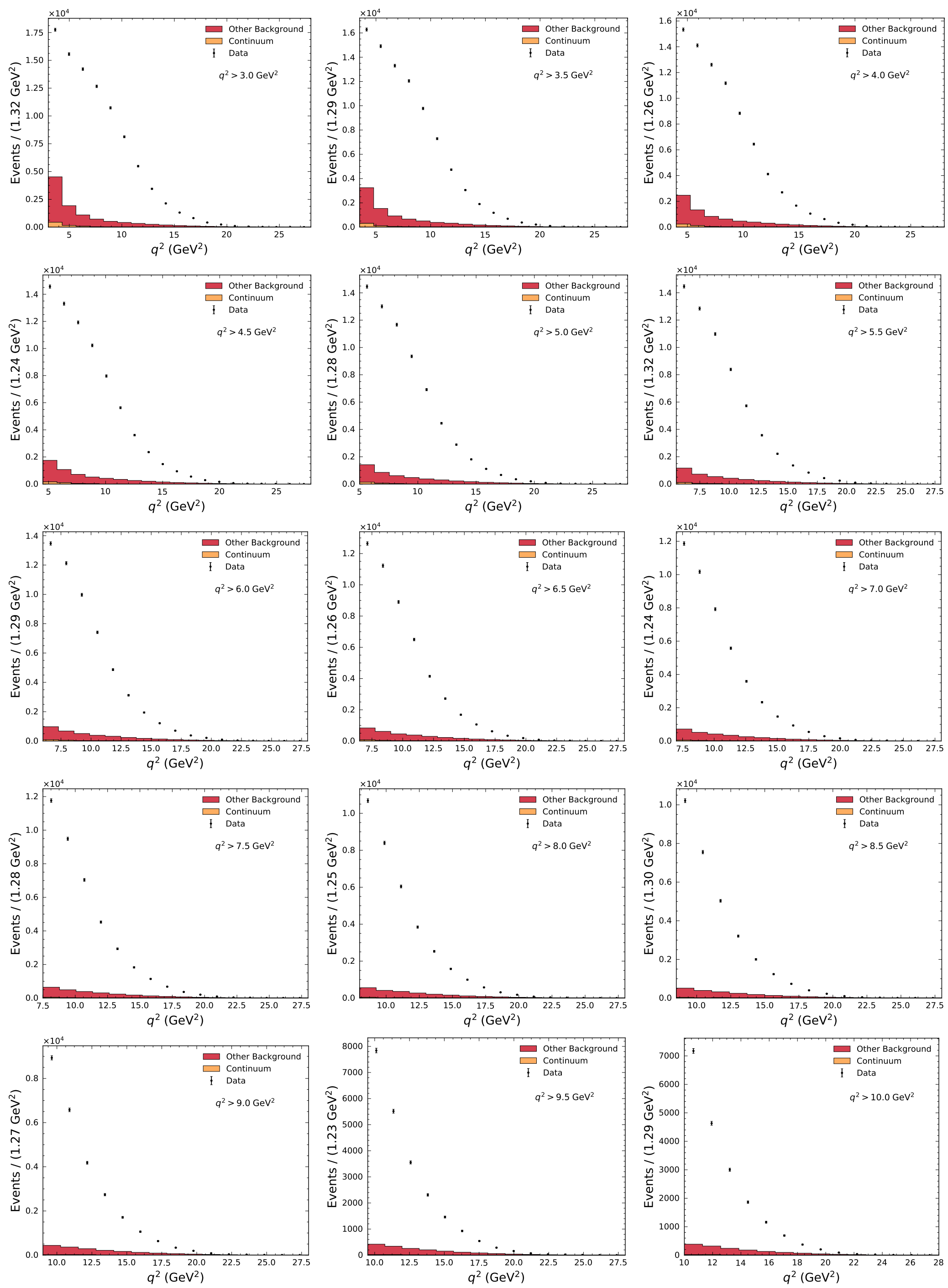

FIG. 9. The determined background in the q^2 spectrum, after obtaining the expected background yields from the fit to the M_X spectrum, for different q^2 threshold selections for electrons are shown.

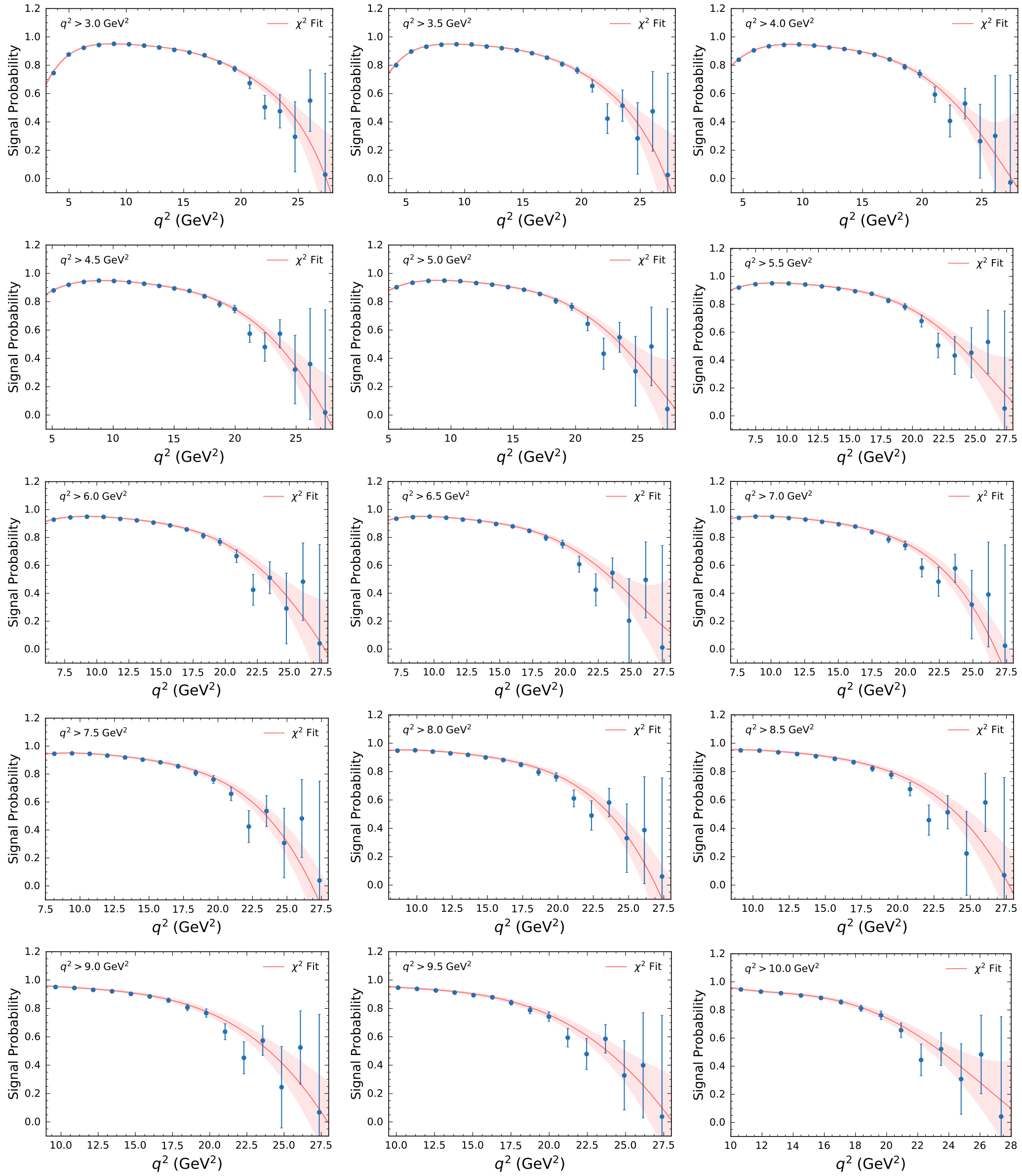

FIG. 10. The determined signal probabilities for different q^2 threshold selections for electrons are shown.

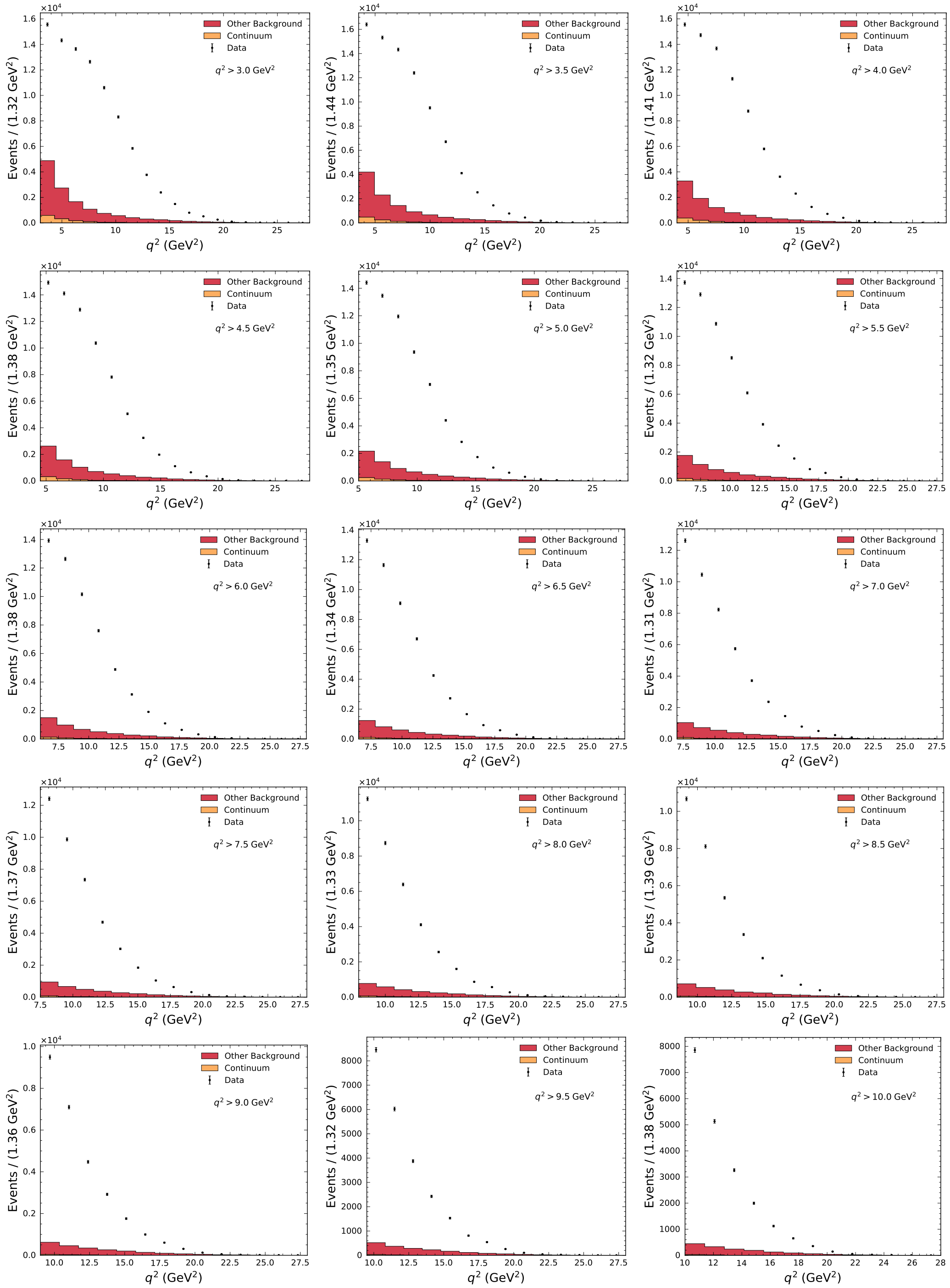

FIG. 11. The determined background in the q^2 spectrum, after obtaining the expected background yields from the fit to the M_X spectrum, for different q^2 threshold selections for muons are shown.

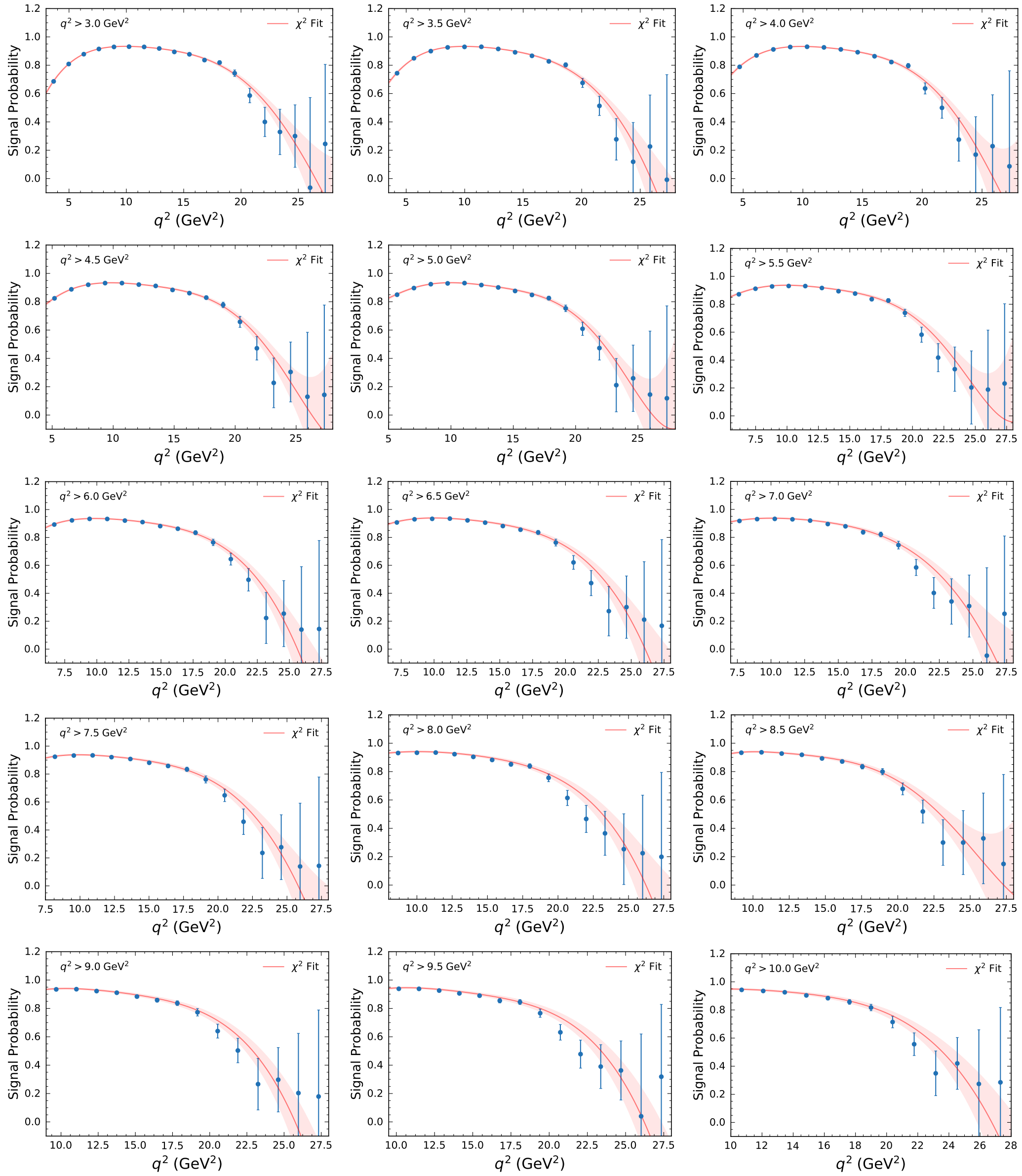

FIG. 12. The determined signal probabilities for different q^2 threshold selections for muons are shown.

APPENDIX B. CALIBRATION CURVES AND CALIBRATION FACTORS FOR MUONS

Fig. 13 shows the calibration curves, the bias and acceptance calibration factors, C_{cal} and C_{acc}, for muons for the various selections on q^2.

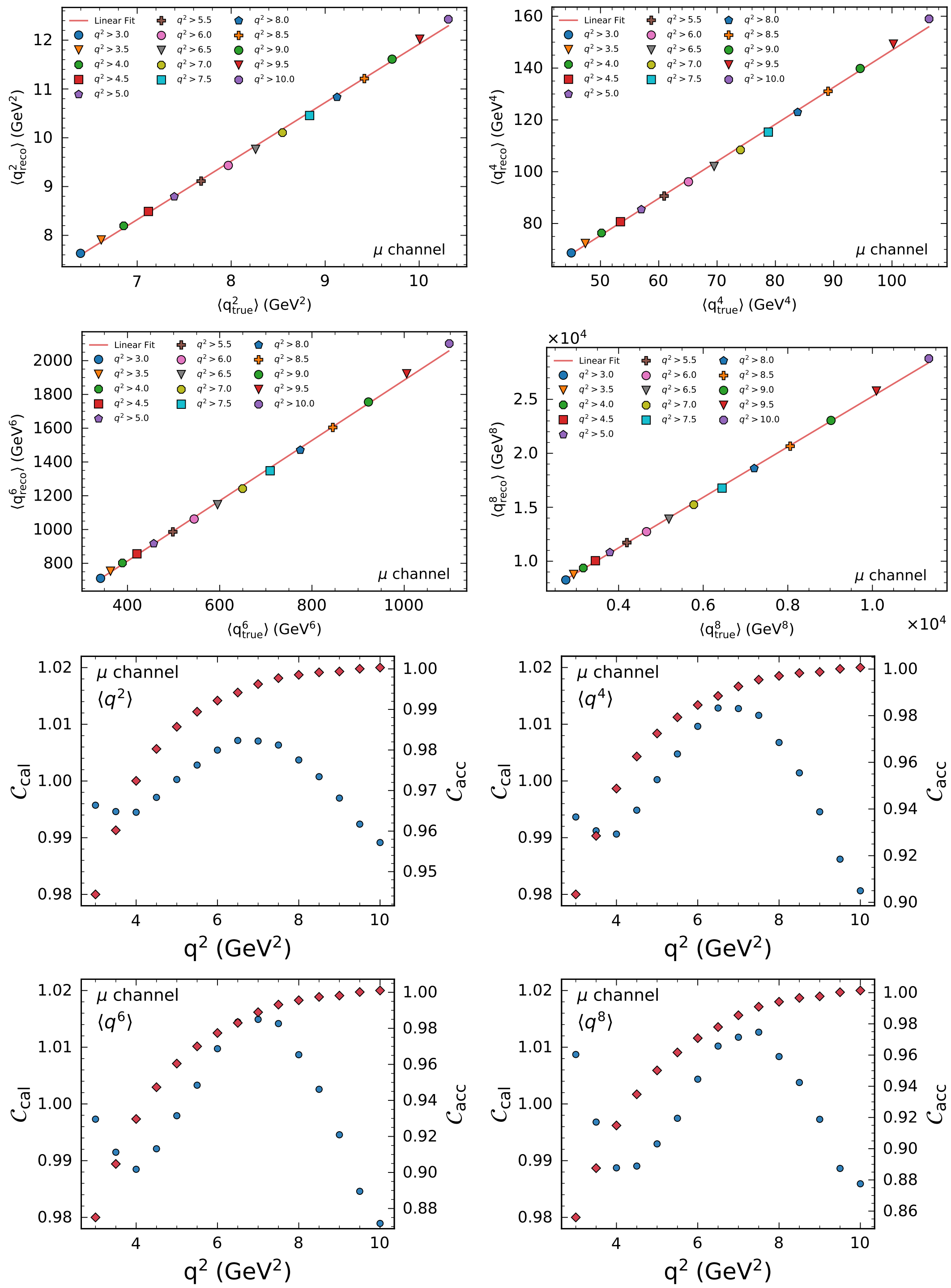

FIG. 13. The calibration curves, bias and acceptance calibration factors for the first to the fourth moment of q^2 for muons.

APPENDIX C. SELECTION EFFICIENCY OF $B \to X_c\,\ell^+\,\nu_\ell$ COMPONENTS

Selection efficiencies for different $B \to X_c\,\ell^+\,\nu_\ell$ components are shown in Fig. 14 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 $B \to X_c\,\ell^+\,\nu_\ell$ processes, especially for the non-resonant component.

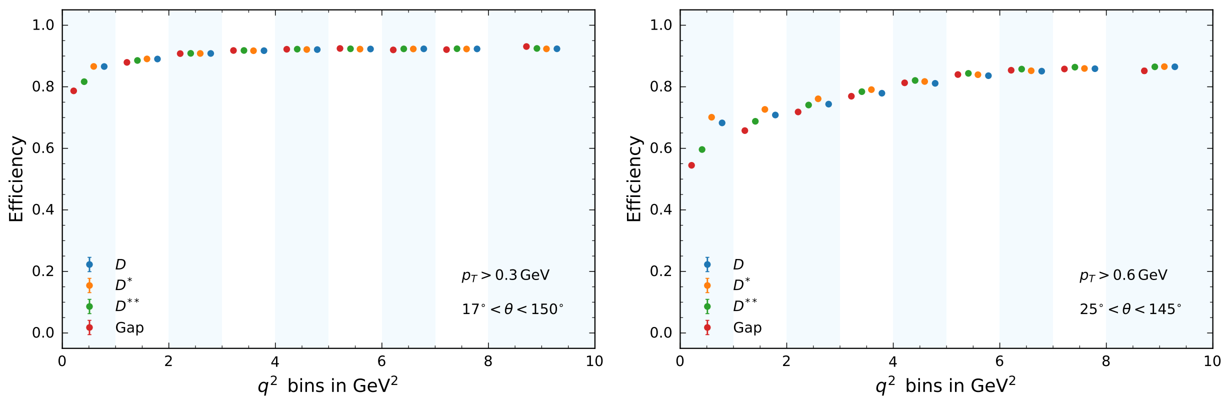

FIG. 14. Selection efficiencies of different $B \to X_c\,\ell^+\,\nu_\ell$ components for electron (left) and muon (right) final states with basic transverse momentum and angular acceptance requirements in bins of q^2.

APPENDIX D. EXPERIMENTAL CORRELATION MATRICES

The full systematic correlation matrices that are determined by combining the statistical and systematic covariance matrices are shown in Fig. 15 and 16 for electrons and muons, respectively.

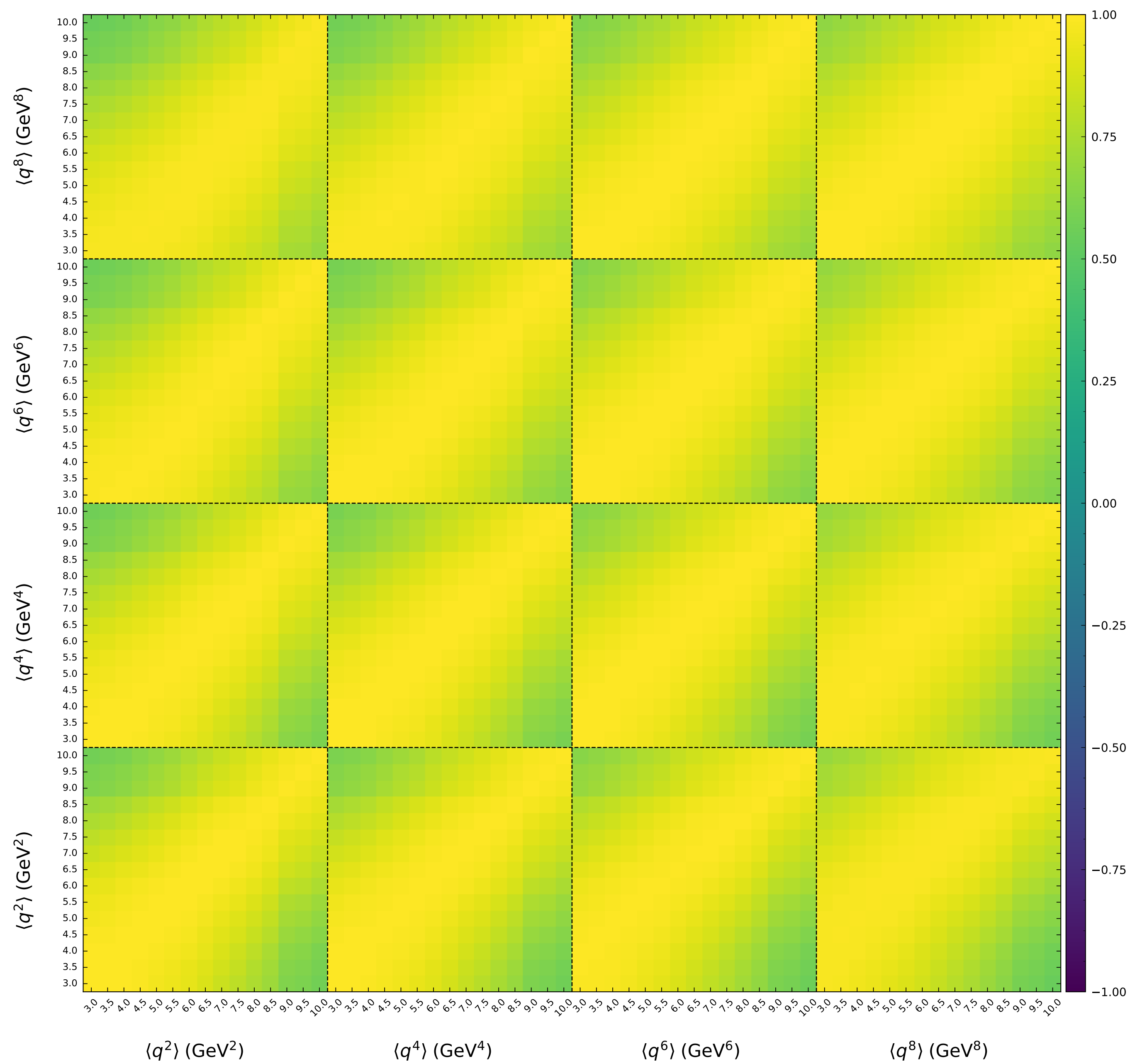

FIG. 15. Color map of the systematic correlation coefficients determined for the moments for the electron final state.

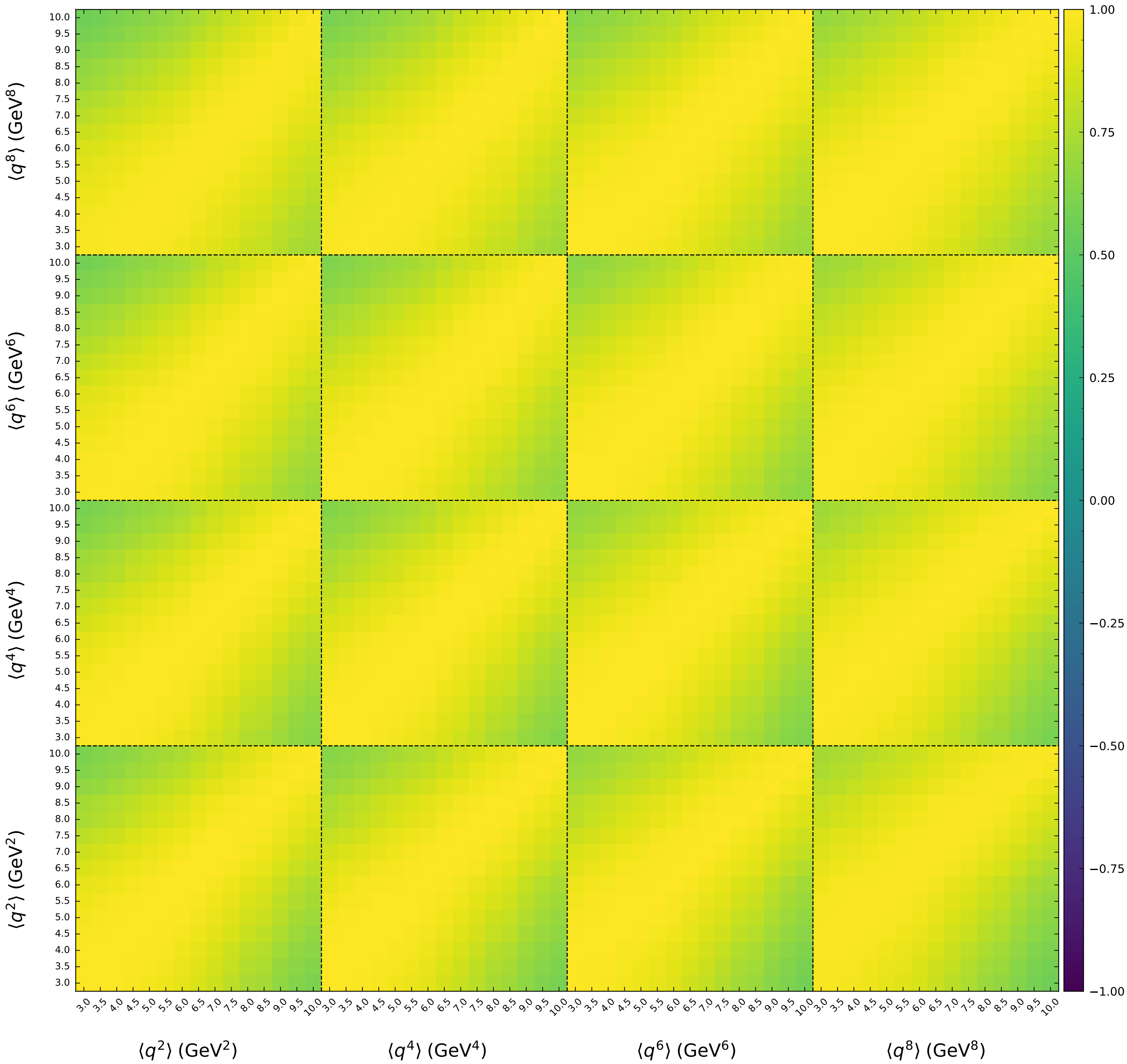

FIG. 16. Color map of the systematic correlation coefficients determined for the moments for the muon final state.

APPENDIX E. CENTRAL MOMENTS

The measured central q^2 moments, discussed in Section VI, are shown for both the electron and muon final states in Fig. 17.

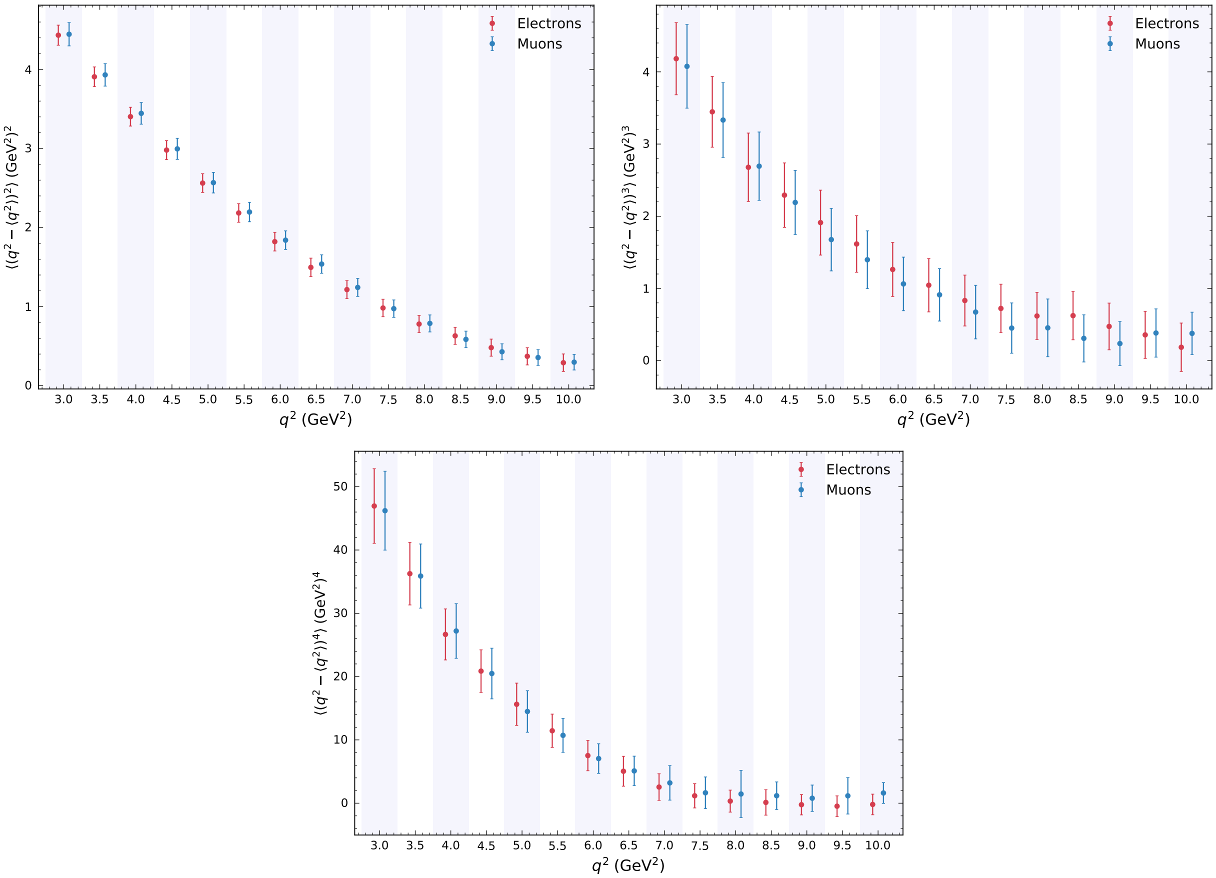

FIG. 17. The measured second (left), third (right), and fourth (bottom) central q^2 moments for both the electron and muon final states.