Study of B→πlν and B→ρlν decays and determination of |V_{ub}|

P. del Amo Sanchez *et al.* (BABAR Collaboration)

Study of $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ Decays and Determination of $|V_{ub}|$

P. del Amo Sanchez,[1] J. P. Lees,[1] V. Poireau,[1] E. Prencipe,[1] V. Tisserand,[1] J. Garra Tico,[2] E. Grauges,[2]
M. Martinelli[ab],[3] A. Palano[ab],[3] M. Pappagallo[ab],[3] G. Eigen,[4] B. Stugu,[4] L. Sun,[4] M. Battaglia,[5] D. N. Brown,[5]
B. Hooberman,[5] L. T. Kerth,[5] Yu. G. Kolomensky,[5] G. Lynch,[5] I. L. Osipenkov,[5] T. Tanabe,[5] C. M. Hawkes,[6]
N. Soni,[6] A. T. Watson,[6] H. Koch,[7] T. Schroeder,[7] D. J. Asgeirsson,[8] C. Hearty,[8] T. S. Mattison,[8] J. A. McKenna,[8]
A. Khan,[9] A. Randle-Conde,[9] V. E. Blinov,[10] A. R. Buzykaev,[10] V. P. Druzhinin,[10] V. B. Golubev,[10]
A. P. Onuchin,[10] S. I. Serednyakov,[10] Yu. I. Skovpen,[10] E. P. Solodov,[10] K. Yu. Todyshev,[10] A. N. Yushkov,[10]
M. Bondioli,[11] S. Curry,[11] D. Kirkby,[11] A. J. Lankford,[11] M. Mandelkern,[11] E. C. Martin,[11] D. P. Stoker,[11]
H. Atmacan,[12] J. W. Gary,[12] F. Liu,[12] O. Long,[12] G. M. Vitug,[12] Z. Yasin,[12] V. Sharma,[13] C. Campagnari,[14]
T. M. Hong,[14] D. Kovalskyi,[14] J. D. Richman,[14] A. M. Eisner,[15] C. A. Heusch,[15] J. Kroseberg,[15] W. S. Lockman,[15]
A. J. Martinez,[15] T. Schalk,[15] B. A. Schumm,[15] A. Seiden,[15] L. O. Winstrom,[15] C. H. Cheng,[16] D. A. Doll,[16]
B. Echenard,[16] D. G. Hitlin,[16] P. Ongmongkolkul,[16] F. C. Porter,[16] A. Y. Rakitin,[16] R. Andreassen,[17]
M. S. Dubrovin,[17] G. Mancinelli,[17] B. T. Meadows,[17] M. D. Sokoloff,[17] P. C. Bloom,[18] W. T. Ford,[18] A. Gaz,[18]
J. F. Hirschauer,[18] M. Nagel,[18] U. Nauenberg,[18] J. G. Smith,[18] S. R. Wagner,[18] R. Ayad,[19, *] W. H. Toki,[19]
A. Hauke,[20] H. Jasper,[20] T. M. Karbach,[20] J. Merkel,[20] A. Petzold,[20] B. Spaan,[20] K. Wacker,[20] M. J. Kobel,[21]
K. R. Schubert,[21] R. Schwierz,[21] D. Bernard,[22] M. Verderi,[22] P. J. Clark,[23] S. Playfer,[23] J. E. Watson,[23]
M. Andreotti[ab],[24] D. Bettoni[a],[24] C. Bozzi[a],[24] R. Calabrese[ab],[24] A. Cecchi[ab],[24] G. Cibinetto[ab],[24] E. Fioravanti[ab],[24]
P. Franchini[ab],[24] E. Luppi[ab],[24] M. Munerato[ab],[24] M. Negrini[ab],[24] A. Petrella[ab],[24] L. Piemontese[a],[24]
R. Baldini-Ferroli,[25] A. Calcaterra,[25] R. de Sangro,[25] G. Finocchiaro,[25] M. Nicolaci,[25] S. Pacetti,[25] P. Patteri,[25]
I. M. Peruzzi,[25, †] M. Piccolo,[25] M. Rama,[25] A. Zallo,[25] R. Contri[ab],[26] E. Guido[ab],[26] M. Lo Vetere[ab],[26]
M. R. Monge[ab],[26] S. Passaggio[a],[26] C. Patrignani[ab],[26] E. Robutti[a],[26] S. Tosi[ab],[26] B. Bhuyan,[27] M. Morii,[28]
A. Adametz,[29] J. Marks,[29] S. Schenk,[29] U. Uwer,[29] F. U. Bernlochner,[30] H. M. Lacker,[30] T. Lueck,[30] A. Volk,[30]
P. D. Dauncey,[31] M. Tibbetts,[31] P. K. Behera,[32] U. Mallik,[32] C. Chen,[33] J. Cochran,[33] H. B. Crawley,[33] L. Dong,[33]
W. T. Meyer,[33] S. Prell,[33] E. I. Rosenberg,[33] A. E. Rubin,[33] Y. Y. Gao,[34] A. V. Gritsan,[34] Z. J. Guo,[34]
N. Arnaud,[35] M. Davier,[35] D. Derkach,[35] J. Firmino da Costa,[35] G. Grosdidier,[35] F. Le Diberder,[35] A. M. Lutz,[35]
B. Malaescu,[35] A. Perez,[35] P. Roudeau,[35] M. H. Schune,[35] J. Serrano,[35] V. Sordini,[35, ‡] A. Stocchi,[35] L. Wang,[35]
G. Wormser,[35] D. J. Lange,[36] D. M. Wright,[36] I. Bingham,[37] J. P. Burke,[37] C. A. Chavez,[37] J. P. Coleman,[37]
J. R. Fry,[37] E. Gabathuler,[37] R. Gamet,[37] D. E. Hutchcroft,[37] D. J. Payne,[37] C. Touramanis,[37] A. J. Bevan,[38]
F. Di Lodovico,[38] R. Sacco,[38] M. Sigamani,[38] G. Cowan,[39] S. Paramesvaran,[39] A. C. Wren,[39] D. N. Brown,[40]
C. L. Davis,[40] A. G. Denig,[41] M. Fritsch,[41] W. Gradl,[41] A. Hafner,[41] K. E. Alwyn,[42] D. Bailey,[42] R. J. Barlow,[42]
G. Jackson,[42] G. D. Lafferty,[42] T. J. West,[42] J. Anderson,[43] R. Cenci,[43] A. Jawahery,[43] D. A. Roberts,[43]
G. Simi,[43] J. M. Tuggle,[43] C. Dallapiccola,[44] E. Salvati,[44] R. Cowan,[45] D. Dujmic,[45] P. H. Fisher,[45] G. Sciolla,[45]
R. K. Yamamoto,[45] M. Zhao,[45] P. M. Patel,[46] S. H. Robertson,[46] M. Schram,[46] P. Biassoni[ab],[47] A. Lazzaro[ab],[47]
V. Lombardo[a],[47] F. Palombo[ab],[47] S. Stracka[ab],[47] L. Cremaldi,[48] R. Godang,[48, §] R. Kroeger,[48] P. Sonnek,[48]
D. J. Summers,[48] H. W. Zhao,[48] X. Nguyen,[49] M. Simard,[49] P. Taras,[49] G. De Nardo[ab],[50] D. Monorchio[ab],[50]
G. Onorato[ab],[50] C. Sciacca[ab],[50] G. Raven,[51] H. L. Snoek,[51] C. P. Jessop,[52] K. J. Knoepfel,[52] J. M. LoSecco,[52]
W. F. Wang,[52] L. A. Corwin,[53] K. Honscheid,[53] R. Kass,[53] J. P. Morris,[53] A. M. Rahimi,[53] N. L. Blount,[54] J. Brau,[54]
R. Frey,[54] O. Igonkina,[54] J. A. Kolb,[54] R. Rahmat,[54] N. B. Sinev,[54] D. Strom,[54] J. Strube,[54] E. Torrence,[54]
G. Castelli[ab],[55] E. Feltresi[ab],[55] N. Gagliardi[ab],[55] M. Margoni[ab],[55] M. Morandin[a],[55] M. Posocco[a],[55] M. Rotondo[a],[55]
F. Simonetto[ab],[55] R. Stroili[ab],[55] E. Ben-Haim,[56] G. R. Bonneaud,[56] H. Briand,[56] J. Chauveau,[56] O. Hamon,[56]
Ph. Leruste,[56] G. Marchiori,[56] J. Ocariz,[56] J. Prendki,[56] S. Sitt,[56] M. Biasini[ab],[57] E. Manoni[ab],[57] C. Angelini[ab],[58]
G. Batignani[ab],[58] S. Bettarini[ab],[58] G. Calderini[ab],[58, ¶] M. Carpinelli[ab],[58, **] A. Cervelli[ab],[58] F. Forti[ab],[58]
M. A. Giorgi[ab],[58] A. Lusiani[ac],[58] N. Neri[ab],[58] E. Paoloni[ab],[58] G. Rizzo[ab],[58] J. J. Walsh[a],[58] D. Lopes Pegna,[59]
C. Lu,[59] J. Olsen,[59] A. J. S. Smith,[59] A. V. Telnov,[59] F. Anulli[a],[60] E. Baracchini[ab],[60] G. Cavoto[a],[60] R. Faccini[ab],[60]
F. Ferrarotto[a],[60] F. Ferroni[ab],[60] M. Gaspero[ab],[60] L. Li Gioi[a],[60] M. A. Mazzoni[a],[60] G. Piredda[a],[60] F. Renga[ab],[60]
M. Ebert,[61] T. Hartmann,[61] T. Leddig,[61] H. Schröder,[61] R. Waldi,[61] T. Adye,[62] B. Franek,[62] E. O. Olaiya,[62]
F. F. Wilson,[62] S. Emery,[63] G. Hamel de Monchenault,[63] G. Vasseur,[63] Ch. Yèche,[63] M. Zito,[63] M. T. Allen,[64]
D. Aston,[64] D. J. Bard,[64] R. Bartoldus,[64] J. F. Benitez,[64] C. Cartaro,[64] M. R. Convery,[64] J. C. Dingfelder,[64, ††]
J. Dorfan,[64] G. P. Dubois-Felsmann,[64] W. Dunwoodie,[64] R. C. Field,[64] M. Franco Sevilla,[64] B. G. Fulsom,[64]

A. M. Gabareen,[64] M. T. Graham,[64] P. Grenier,[64] C. Hast,[64] W. R. Innes,[64] M. H. Kelsey,[64] H. Kim,[64]
P. Kim,[64] M. L. Kocian,[64] D. W. G. S. Leith,[64] S. Li,[64] B. Lindquist,[64] S. Luitz,[64] V. Luth,[64] H. L. Lynch,[64]
D. B. MacFarlane,[64] H. Marsiske,[64] D. R. Muller,[64] H. Neal,[64] S. Nelson,[64] C. P. O'Grady,[64] I. Ofte,[64] M. Perl,[64]
B. N. Ratcliff,[64] A. Roodman,[64] A. A. Salnikov,[64] R. H. Schindler,[64] J. Schwiening,[64] A. Snyder,[64] D. Su,[64]
M. K. Sullivan,[64] K. Suzuki,[64] J. M. Thompson,[64] J. Va'vra,[64] A. P. Wagner,[64] M. Weaver,[64] C. A. West,[64]
W. J. Wisniewski,[64] M. Wittgen,[64] D. H. Wright,[64] H. W. Wulsin,[64] A. K. Yarritu,[64] V. Santoro,[64] C. C. Young,[64]
V. Ziegler,[64] X. R. Chen,[65] W. Park,[65] M. V. Purohit,[65] R. M. White,[65] J. R. Wilson,[65] S. J. Sekula,[66] M. Bellis,[67]
P. R. Burchat,[67] A. J. Edwards,[67] T. S. Miyashita,[67] S. Ahmed,[68] M. S. Alam,[68] J. A. Ernst,[68] B. Pan,[68]
M. A. Saeed,[68] S. B. Zain,[68] N. Guttman,[69] A. Soffer,[69] P. Lund,[70] S. M. Spanier,[70] R. Eckmann,[71] J. L. Ritchie,[71]
A. M. Ruland,[71] C. J. Schilling,[71] R. F. Schwitters,[71] B. C. Wray,[71] J. M. Izen,[72] X. C. Lou,[72] F. Bianchi[ab,73]
D. Gamba[ab,73] M. Pelliccioni[ab,73] M. Bomben[ab,74] G. Della Ricca[ab,74] L. Lanceri[ab,74] L. Vitale[ab,74] V. Azzolini,[75]
N. Lopez-March,[75] F. Martinez-Vidal,[75] D. A. Milanes,[75] A. Oyanguren,[75] J. Albert,[76] Sw. Banerjee,[76]
H. H. F. Choi,[76] K. Hamano,[76] G. J. King,[76] R. Kowalewski,[76] M. J. Lewczuk,[76] I. M. Nugent,[76] J. M. Roney,[76]
R. J. Sobie,[76] T. J. Gershon,[77] P. F. Harrison,[77] J. Ilic,[77] T. E. Latham,[77] G. B. Mohanty,[77] E. M. T. Puccio,[77]
H. R. Band,[78] X. Chen,[78] S. Dasu,[78] K. T. Flood,[78] Y. Pan,[78] R. Prepost,[78] C. O. Vuosalo,[78] and S. L. Wu[78]

(The *BABAR* Collaboration)

[1]*Laboratoire d'Annecy-le-Vieux de Physique des Particules (LAPP),
Université de Savoie, CNRS/IN2P3, F-74941 Annecy-Le-Vieux, France*
[2]*Universitat de Barcelona, Facultat de Fisica, Departament ECM, E-08028 Barcelona, Spain*
[3]*INFN Sezione di Bari[a]; Dipartimento di Fisica, Università di Bari[b], I-70126 Bari, Italy*
[4]*University of Bergen, Institute of Physics, N-5007 Bergen, Norway*
[5]*Lawrence Berkeley National Laboratory and University of California, Berkeley, California 94720, USA*
[6]*University of Birmingham, Birmingham, B15 2TT, United Kingdom*
[7]*Ruhr Universität Bochum, Institut für Experimentalphysik 1, D-44780 Bochum, Germany*
[8]*University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1*
[9]*Brunel University, Uxbridge, Middlesex UB8 3PH, United Kingdom*
[10]*Budker Institute of Nuclear Physics, Novosibirsk 630090, Russia*
[11]*University of California at Irvine, Irvine, California 92697, USA*
[12]*University of California at Riverside, Riverside, California 92521, USA*
[13]*University of California at San Diego, La Jolla, California 92093, USA*
[14]*University of California at Santa Barbara, Santa Barbara, California 93106, USA*
[15]*University of California at Santa Cruz, Institute for Particle Physics, Santa Cruz, California 95064, USA*
[16]*California Institute of Technology, Pasadena, California 91125, USA*
[17]*University of Cincinnati, Cincinnati, Ohio 45221, USA*
[18]*University of Colorado, Boulder, Colorado 80309, USA*
[19]*Colorado State University, Fort Collins, Colorado 80523, USA*
[20]*Technische Universität Dortmund, Fakultät Physik, D-44221 Dortmund, Germany*
[21]*Technische Universität Dresden, Institut für Kern- und Teilchenphysik, D-01062 Dresden, Germany*
[22]*Laboratoire Leprince-Ringuet, CNRS/IN2P3, Ecole Polytechnique, F-91128 Palaiseau, France*
[23]*University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom*
[24]*INFN Sezione di Ferrara[a]; Dipartimento di Fisica, Università di Ferrara[b], I-44100 Ferrara, Italy*
[25]*INFN Laboratori Nazionali di Frascati, I-00044 Frascati, Italy*
[26]*INFN Sezione di Genova[a]; Dipartimento di Fisica, Università di Genova[b], I-16146 Genova, Italy*
[27]*Indian Institute of Technology Guwahati, Guwahati, Assam, 781 039, India*
[28]*Harvard University, Cambridge, Massachusetts 02138, USA*
[29]*Universität Heidelberg, Physikalisches Institut, Philosophenweg 12, D-69120 Heidelberg, Germany*
[30]*Humboldt-Universität zu Berlin, Institut für Physik, Newtonstr. 15, D-12489 Berlin, Germany*
[31]*Imperial College London, London, SW7 2AZ, United Kingdom*
[32]*University of Iowa, Iowa City, Iowa 52242, USA*
[33]*Iowa State University, Ames, Iowa 50011-3160, USA*
[34]*Johns Hopkins University, Baltimore, Maryland 21218, USA*
[35]*Laboratoire de l'Accélérateur Linéaire, IN2P3/CNRS et Université Paris-Sud 11,
Centre Scientifique d'Orsay, B. P. 34, F-91898 Orsay Cedex, France*
[36]*Lawrence Livermore National Laboratory, Livermore, California 94550, USA*
[37]*University of Liverpool, Liverpool L69 7ZE, United Kingdom*
[38]*Queen Mary, University of London, London, E1 4NS, United Kingdom*
[39]*University of London, Royal Holloway and Bedford New College, Egham, Surrey TW20 0EX, United Kingdom*
[40]*University of Louisville, Louisville, Kentucky 40292, USA*
[41]*Johannes Gutenberg-Universität Mainz, Institut für Kernphysik, D-55099 Mainz, Germany*
[42]*University of Manchester, Manchester M13 9PL, United Kingdom*

^{43}University of Maryland, College Park, Maryland 20742, USA

^{44}University of Massachusetts, Amherst, Massachusetts 01003, USA

^{45}Massachusetts Institute of Technology, Laboratory for Nuclear Science, Cambridge, Massachusetts 02139, USA

^{46}McGill University, Montréal, Québec, Canada H3A 2T8

^{47}INFN Sezione di Milano^a; Dipartimento di Fisica, Università di Milano^b, I-20133 Milano, Italy

^{48}University of Mississippi, University, Mississippi 38677, USA

^{49}Université de Montréal, Physique des Particules, Montréal, Québec, Canada H3C 3J7

^{50}INFN Sezione di Napoli^a; Dipartimento di Scienze Fisiche,
Università di Napoli Federico II^b, I-80126 Napoli, Italy

^{51}NIKHEF, National Institute for Nuclear Physics and High Energy Physics, NL-1009 DB Amsterdam, The Netherlands

^{52}University of Notre Dame, Notre Dame, Indiana 46556, USA

^{53}Ohio State University, Columbus, Ohio 43210, USA

^{54}University of Oregon, Eugene, Oregon 97403, USA

^{55}INFN Sezione di Padova^a; Dipartimento di Fisica, Università di Padova^b, I-35131 Padova, Italy

^{56}Laboratoire de Physique Nucléaire et de Hautes Energies,
IN2P3/CNRS, Université Pierre et Marie Curie-Paris6,
Université Denis Diderot-Paris7, F-75252 Paris, France

^{57}INFN Sezione di Perugia^a; Dipartimento di Fisica, Università di Perugia^b, I-06100 Perugia, Italy

^{58}INFN Sezione di Pisa^a; Dipartimento di Fisica,
Università di Pisa^b; Scuola Normale Superiore di Pisa^c, I-56127 Pisa, Italy

^{59}Princeton University, Princeton, New Jersey 08544, USA

^{60}INFN Sezione di Roma^a; Dipartimento di Fisica,
Università di Roma La Sapienza^b, I-00185 Roma, Italy

^{61}Universität Rostock, D-18051 Rostock, Germany

^{62}Rutherford Appleton Laboratory, Chilton, Didcot, Oxon, OX11 0QX, United Kingdom

^{63}CEA, Irfu, SPP, Centre de Saclay, F-91191 Gif-sur-Yvette, France

^{64}SLAC National Accelerator Laboratory, Stanford, California 94309 USA

^{65}University of South Carolina, Columbia, South Carolina 29208, USA

^{66}Southern Methodist University, Dallas, Texas 75275, USA

^{67}Stanford University, Stanford, California 94305-4060, USA

^{68}State University of New York, Albany, New York 12222, USA

^{69}Tel Aviv University, School of Physics and Astronomy, Tel Aviv, 69978, Israel

^{70}University of Tennessee, Knoxville, Tennessee 37996, USA

^{71}University of Texas at Austin, Austin, Texas 78712, USA

^{72}University of Texas at Dallas, Richardson, Texas 75083, USA

^{73}INFN Sezione di Torino^a; Dipartimento di Fisica Sperimentale, Università di Torino^b, I-10125 Torino, Italy

^{74}INFN Sezione di Trieste^a; Dipartimento di Fisica, Università di Trieste^b, I-34127 Trieste, Italy

^{75}IFIC, Universitat de Valencia-CSIC, E-46071 Valencia, Spain

^{76}University of Victoria, Victoria, British Columbia, Canada V8W 3P6

^{77}Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom

^{78}University of Wisconsin, Madison, Wisconsin 53706, USA

We present an analysis of exclusive charmless semileptonic B-meson decays based on 377 million $B\overline{B}$ pairs recorded with the BABAR detector at the $\Upsilon(4S)$ resonance. We select four event samples corresponding to the decay modes $B^0 \to \pi^-\ell^+\nu$, $B^+ \to \pi^0\ell^+\nu$, $B^0 \to \rho^-\ell^+\nu$, and $B^+ \to \rho^0\ell^+\nu$, and find the measured branching fractions to be consistent with isospin symmetry. Assuming isospin symmetry, we combine the two $B \to \pi\ell\nu$ samples, and similarly the two $B \to \rho\ell\nu$ samples, and measure the branching fractions $\mathcal{B}(B^0 \to \pi^-\ell^+\nu) = (1.41 \pm 0.05 \pm 0.07) \times 10^{-4}$ and $\mathcal{B}(B^0 \to \rho^-\ell^+\nu) = (1.75 \pm 0.15 \pm 0.27) \times 10^{-4}$, where the errors are statistical and systematic. We compare the measured distribution in q^2, the momentum transfer squared, with predictions for the form factors from QCD calculations and determine the CKM matrix element $|V_{ub}|$. Based on the measured partial branching fraction for $B \to \pi\ell\nu$ in the range $q^2 < 12\,\mathrm{GeV}^2$ and the most recent LCSR calculations we obtain $|V_{ub}| = (3.78 \pm 0.13^{+0.55}_{-0.40}) \times 10^{-3}$, where the errors refer to the experimental and theoretical uncertainties. From a simultaneous fit to the data over the full q^2 range and the FNAL/MILC lattice QCD results, we obtain $|V_{ub}| = (2.95 \pm 0.31) \times 10^{-3}$ from $B \to \pi\ell\nu$, where the error is the combined experimental and theoretical uncertainty.

PACS numbers: 13.20.He, 12.15.Hh, 12.38.Qk, 14.40.Nd

*Now at Temple University, Philadelphia, Pennsylvania 19122, USA

I. INTRODUCTION

The elements of the Cabibbo-Kobayashi-Maskawa (CKM) quark-mixing matrix are fundamental parameters of the Standard Model (SM) of electroweak interactions. With the increasingly precise measurements of decay-time-dependent CP asymmetries in B-meson decays, in particular $\sin(2\beta)$ [1, 2], improved measurements of the magnitude of V_{ub} and V_{cb} will allow for more stringent experimental tests of the SM mechanism for CP violation [3]. This is best illustrated in terms of the unitarity triangle, the graphical representation of one of the unitarity conditions for the CKM matrix, for which the length of the side that is opposite to the angle β is proportional to the ratio $|V_{ub}|/|V_{cb}|$. The best method to determine $|V_{ub}|$ and $|V_{cb}|$ is to measure semileptonic decay rates for $B \to X_c\ell\nu$ and $B \to X_u\ell\nu$ (X_c and X_u refer to hadronic states with or without charm), which are proportional to $|V_{cb}|^2$ and $|V_{ub}|^2$, respectively.

There are two methods to extract these two CKM elements from B decays, one based on inclusive and the other on exclusive semileptonic decays. Exclusive decays offer better kinematic constraints and thus more effective background suppression than inclusive decays, but the lower branching fractions result in lower event yields. Since the experimental and theoretical techniques for these two approaches are different and largely independent, they can provide important cross checks of our understanding of the theory and the measurements. An overview of the determination of $|V_{ub}|$ using both approaches can be found in a recent review [4].

In this paper, we present a study of four exclusive charmless semileptonic decay modes, $B^0 \to \pi^-\ell^+\nu$, $B^+ \to \pi^0\ell^+\nu$, $B^0 \to \rho^-\ell^+\nu$, and $B^+ \to \rho^0\ell^+\nu$ [5], and a determination of $|V_{ub}|$. Here ℓ refers to a charged lepton, either e^+ or μ^+, and ν refers to a neutrino, either ν_e or ν_μ. This analysis represents an update of an earlier measurement [6] that was based on a significantly smaller data set. For the current analysis, the signal yields and background suppression have been improved and the systematic uncertainties have been reduced through the use of improved reconstruction and signal extraction methods, combined with more detailed background studies.

The principal experimental challenge is the separation

†Also with Università di Perugia, Dipartimento di Fisica, Perugia, Italy
‡Also with Università di Roma La Sapienza, I-00185 Roma, Italy
§Now at University of South Alabama, Mobile, Alabama 36688, USA
¶Also with Laboratoire de Physique Nucléaire et de Hautes Energies, IN2P3/CNRS, Université Pierre et Marie Curie-Paris6, Université Denis Diderot-Paris7, F-75252 Paris, France
**Also with Università di Sassari, Sassari, Italy
††Now at Physikalisches Institut Freiburg, Hermann-Herder-Strasse 3, 79104 Freiburg, Germany

of the $B \to X_u\ell\nu$ from the dominant $B \to X_c\ell\nu$ decays, for which the inclusive branching fraction is a factor of 50 larger. Furthermore, the isolation of individual exclusive charmless decays from all other $B \to X_u\ell\nu$ decays is difficult, because the exclusive branching ratios are typically only 10% of $\mathcal{B}(B \to X_u\ell\nu) = (2.29\pm0.34)\times10^{-3}$ [7], the inclusive branching fraction for charmless semileptonic B decays.

The reconstruction of signal decays in $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ events requires the identification of three types of particles, the hadronic state X_u producing one or two charged and/or neutral final state pions, the charged lepton, and the neutrino. The presence of the neutrino is inferred from the missing momentum and energy in the whole event.

The event yields for each of the four signal decay modes are extracted from a binned maximum-likelihood fit to the three-dimensional distributions of the variables m_{ES}, the energy-substituted B-meson mass, ΔE, the difference between the reconstructed and the expected B-meson energy, and q^2, the momentum transfer squared from the B meson to the final-state hadron. The measured differential decay rates in combination with recent form-factor calculations are used to determine $|V_{ub}|$. By measuring both $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ decays simultaneously, we reduce the sensitivity to the cross feed between these two decay modes and some of the background contributions.

The most promising decay mode for a precise determination of $|V_{ub}|$, both experimentally and theoretically, is the $B \to \pi\ell\nu$ decay for which a number of measurements exist. The first measurement of this type was performed by the CLEO Collaboration [8]. In addition to the earlier BABAR measurement mentioned above [6], there is a more recent BABAR measurement [9] in which somewhat looser criteria on the neutrino selection were applied, resulting in a larger signal sample but also substantially higher backgrounds. These analyses also rely on the measurement of the missing energy and momentum of the whole event to reconstruct the neutrino, without explicitly reconstructing the second B-meson decay in the event, but are based on smaller data sets than the one presented here. Recently a number of measurements of both $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ decays were published, in which the $B\overline{B}$ events were tagged by a fully reconstructed hadronic or semileptonic decay of the second B meson in the event [10, 11]. These analyses have led to a simpler and more precise reconstruction of the neutrino and very low backgrounds. However, this is achieved at the expense of much smaller signal samples, which limit the statistical precision of the form-factor measurement.

II. FORM FACTORS

A. Overview

The advantage of charmless semileptonic decays over charmless hadronic decays of the B meson is that the

leptonic and hadronic components of the matrix element factorize. The hadronic matrix element is difficult to calculate, since it must take into account physical mesons, rather than free quarks. Therefore higher-order perturbative corrections and non-perturbative long-distance hadronization processes cannot be ignored. To overcome these difficulties, a set of Lorentz-invariant form factors has been introduced that give a global description of these QCD processes.

A variety of theoretical predictions for these form factors exist. They are based on QCD calculations, such as lattice QCD and sum rules, in addition to quark models. We will make use of a variety of these calculations to assess their impact on the determination of $|V_{ub}|$ from measurements of the decay rates.

The $V - A$ structure of the hadronic current is invoked, along with the knowledge of the transformation properties of the final-state meson, to formulate these form factors. They are functions of $q^2 = m_W^2$, the mass squared of the virtual W,

$$
\begin{aligned}
q^2 &= (P_\ell + P_\nu)^2 \\
&= (P_B - P_X)^2 = M_B^2 + m_X^2 - 2M_B E_X. \quad (1)
\end{aligned}
$$

Here P_ℓ and P_ν refer to the four-momenta of the charged lepton and the neutrino, M_B and P_B to the mass and the four-momentum of the B meson, and m_X and E_X are the mass and energy (in the B-meson rest frame) of the final-state meson X_u.

We distinguish two main categories of exclusive semileptonic decays: decays to pseudoscalar mesons, $B \to \pi\ell\nu$ or $B^+ \to \eta\ell^+\nu$, and decays to vector mesons, $B \to \rho\ell\nu$ or $B^+ \to \omega\ell^+\nu$.
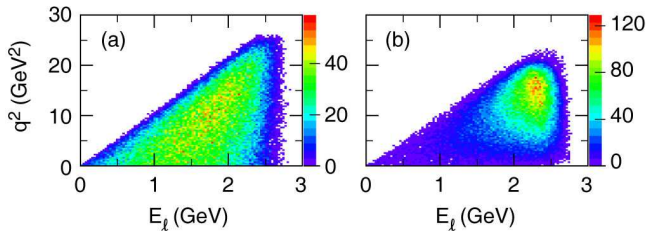
FIG. 1: Simulated distributions of q^2 versus E_ℓ for a) $B \to \pi\ell\nu$ and b) $B \to \rho\ell\nu$ decays. E_ℓ is the lepton energy in the B-meson rest frame.

Figure 1 shows the phase space for $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ decays in terms of q^2 and E_ℓ, the energy of the charged lepton in the B-meson rest frame. The difference between the distributions is due to the different spin structure of the decays.

B. Form Factors

1. B Decays to Pseudoscalar Mesons: $B \to \pi\ell\nu$

For decays to a final-state pseudoscalar meson, the hadronic matrix element is usually written in terms of two form factors, $f_+(q^2)$ and $f_0(q^2)$ [12, 13],

$$
\begin{aligned}
\langle \pi(P_\pi)|\overline{u}\gamma^\mu b|B(P_B)\rangle = \\
f_+(q^2)\left[(P_B + P_\pi)^\mu - \frac{M_B^2 - m_\pi^2}{q^2}q^\mu\right] + \\
f_0(q^2)\frac{M_B^2 - m_\pi^2}{q^2}q^\mu, \quad (2)
\end{aligned}
$$

where P_π and P_B are the four-momenta of the final-state pion and the parent B meson, and m_π and M_B are their masses. This expression can be simplified for leptons with small masses, such as electrons and muons, because in the limit of $m_\ell \ll M_B$ the second term can be neglected. We are left with a single form factor $f_+(q^2)$ and the differential decay rate becomes

$$
\frac{d\Gamma(B^0 \to \pi^-\ell^+\nu)}{dq^2 d\cos\theta_{W\ell}} = |V_{ub}|^2 \frac{G_F^2 p_\pi^3}{32\pi^3}\sin^2\theta_{W\ell}|f_+(q^2)|^2, \quad (3)
$$

where p_π is the momentum of the pion in the rest frame of the B meson, and q^2 varies from zero to $q_{max}^2 = (M_B - m_\pi)^2$.

The decay rate depends on the third power of the pion momentum, suppressing the rate at high q^2. The rate also depends on $\sin^2\theta_{W\ell}$, where $\theta_{W\ell}$ is the angle of the charged-lepton momentum in the W rest frame with respect to direction of the W boost from the B rest frame. The combination of these two factors leads to a lepton-momentum spectrum that is peaked well below the kinematic limit (see Figure 1).

2. B Decays to Vector Mesons: $B \to \rho\ell\nu$

For decays with a vector meson in the final state, the polarization vector ϵ of the vector meson plays an important role. The hadronic current is written in terms of four form factors, of which only three (A_i with $i = 0, 1, 2$) are independent [12, 13],

$$
\begin{aligned}
\langle\rho(P_\rho, \epsilon)|V^\mu - A^\mu|B(P_B)\rangle = \frac{2iV(q^2)}{M_B + m_\rho}\epsilon^{\mu\nu\alpha\beta}\epsilon_\nu^* P_{\rho\alpha}P_{B\beta} \\
-(M_B + m_\rho)A_1(q^2)\epsilon^{*\,\mu} + \frac{A_2(q^2)}{M_B + m_\rho}\epsilon^* \cdot P_B (P_B + P_\rho)^\mu \\
+2m_\rho\frac{\epsilon^* \cdot P_B}{q^2}q^\mu[A_3(q^2) - A_0(q^2)], \quad (4)
\end{aligned}
$$

where m_ρ and P_ρ refer to the vector-meson mass and four-momentum. Again, a simplification can be made for low-mass charged leptons. The term with q^μ can be

neglected, so there are effectively only three form factors for electrons and muons: the axial-vector form factors, $A_1(q^2)$ and $A_2(q^2)$, and the vector form factor, $V(q^2)$. Instead of using these form factors, the full differential decay rate is usually expressed in terms of the helicity amplitudes corresponding to the three helicity states of the ρ meson,

$$
\begin{aligned}
H_\pm(q^2) &= (M_B + m_\rho)\left[A_1(q^2) \mp \frac{2M_B\,p_\rho}{(M_B + m_\rho)^2}V(q^2)\right], \\
H_0(q^2) &= \frac{M_B + m_\rho}{2m_\rho\sqrt{q^2}} \times \left[(M_B^2 - m_\rho^2 - q^2)A_1(q^2)\right. \\
&\left. \quad - \frac{4M_B^2\,p_\rho^2}{(M_B + m_\rho)^2}A_2(q^2)\right],
\end{aligned} \quad (5)
$$

where p_ρ is the momentum of the final-state ρ meson in the B rest frame. While A_1 dominates the three helicity amplitudes, A_2 contributes only to H_0, and V contributes only to $H_\pm$.

Thus the differential decay rate can be written as

$$
\begin{aligned}
\frac{d\Gamma(B \to \rho\ell\nu)}{dq^2 d\cos\theta_{W\ell}} &= |V_{ub}|^2\,\frac{G_F^2\,p_\rho\,q^2}{128\pi^3 M_B^2} \times \left[\sin^2\theta_{W\ell}|H_0|^2\right. \\
&\left. +(1 - \cos\theta_{W\ell})^2\frac{|H_+|^2}{2} + (1 + \cos\theta_{W\ell})^2\frac{|H_-|^2}{2}\right]. \quad (6)
\end{aligned}
$$

The $V-A$ nature of the charged weak current leads to a dominant contribution from H_- and a distribution of events characterized by a forward peak in $\cos\theta_{W\ell}$ and high lepton momenta (see Figure 1).

C. Form-Factor Calculations and Models

The q^2 dependence of the form factors can be extracted from the data. Since the differential decay rates are proportional to the product of $|V_{ub}|^2$ and the form-factor terms, we need at least one point in q^2 at which the form factor is predicted in order to extract $|V_{ub}|$ from the measured branching fractions.

Currently predictions of form factors are based on

- quark-model calculations, (ISGW2) [14],
- QCD light-cone sum rules (LCSR) [15–19],
- lattice QCD calculations (LQCD) [20–23].

These calculations will also be used to simulate the kinematics of the signal decay modes and thus might impact the detection efficiency and thereby the branching-fraction measurement. The two QCD calculations result in predictions for different regions of phase space. The lattice calculations are only available in the high-q^2 region, while LCSR provide information near $q^2 = 0$. Interpolations between these two regions can be constrained by unitarity and analyticity requirements [24, 25].

Figure 2 shows the q^2 distributions for $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ decays for various form-factor calculations. The uncertainties in these predictions are not indicated. For $B \to \pi\ell\nu$ decays they are largest at low q^2 for LQCD predictions and largest at high q^2 for LCSR calculations. Estimates of the uncertainties of the calculations are currently not available for $B \to \rho\ell\nu$ decays.
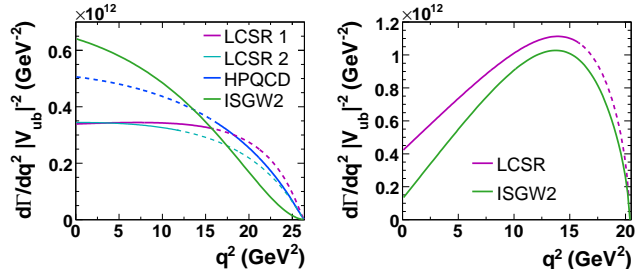
FIG. 2: q^2 distributions for $B \to \pi\ell\nu$ (left) and $B \to \rho\ell\nu$ (right) decays, based on form-factor predictions from the ISGW2 model [14], LCSR calculations (LCSR 1 [15] and LCSR 2 [19] for $B \to \pi\ell\nu$ and LCSR [17] for $B \to \rho\ell\nu$) and the HPQCD [23] lattice calculation. The extrapolations of the QCD predictions to the full q^2 range are marked as dashed lines.

The Isgur-Scora-Grinstein-Wise model (ISGW2) [14] is a constituent quark model with relativistic corrections. Predictions extend over the full q^2 range; they are normalized at $q^2 \approx q_{max}^2$. The form factors are parameterized as

$$
f_+(q^2) = f(q_{max}^2)\left(1 + \frac{1}{6N}\xi^2(q_{max}^2 - q^2)\right)^{-N}, \quad (7)
$$

where ξ is the charge radius of the final-state meson, and $N = 2$ ($N = 3$) for decays to pseudoscalar (vector) mesons. The uncertainties of the predictions by this model are difficult to quantify.

QCD light-cone sum-rule calculations are non-perturbative and combine the idea of QCD sum rules with twist expansions performed to $\mathcal{O}(\alpha_s)$. These calculations provide estimates of various form factors at low to intermediate q^2, for both pseudoscalar and vector decays. The overall normalization is predicted at low q^2 with typical uncertainties of 10-13% [15, 17].

Lattice QCD calculations can potentially provide heavy-to-light-quark form factors from first principles. Unquenched lattice calculations, in which quark-loop effects in the QCD vacuum are incorporated, are now available for the $B \to \pi\ell\nu$ form factors from the Fermilab/MILC [22] and the HPQCD [23] Collaborations. Both calculations account for three dynamical quark flavors, the mass-degenerate u and d quarks and a heavier s quark, but they differ in the way the b quark is simulated. Predictions for $f_0(q^2)$ and $f_+(q^2)$ are shown in Figure 3. The two lattice calculations agree within the stated uncertainties, which are significantly smaller than those of earlier quenched approximations.
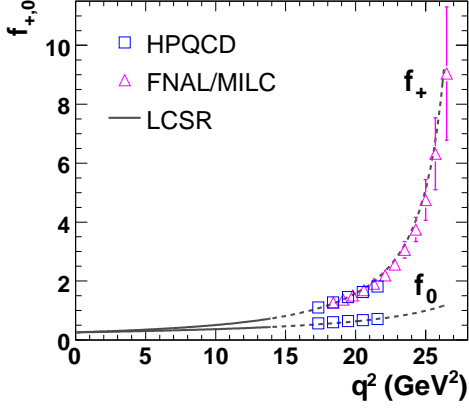
FIG. 3: (color online) Predictions of the form factors f_+ and f_0 for $B \to \pi \ell \nu$ decays based on unquenched LQCD calculations by the FNAL/MILC [22] and HPQCD [23] Collaborations (data points with combined statistical and systematic errors) and LCSR calculations [15] (solid black lines). The dashed lines indicate the extrapolations of the LCSR predictions to $q^2 > 16\,\text{GeV}^2$.

D. Form-Factor Parameterizations

Neither the lattice nor the LCSR QCD calculations predict the form factors over the full q^2 range. Lattice calculations are restricted to small hadron momenta, $i.e.$, to $q^2 \geq q_{max}^2/2$, while LCSR work best at small q^2. If the q^2 spectrum is well measured, the shape of the form factors can be constrained, and the QCD calculations provide the normalization necessary to determine $|V_{ub}|$.

A number of parameterizations of the pseudoscalar form factor $f_+(q^2)$ are available in the literature. The following four parameterizations are commonly used. All of them include at least one pole term at $q^2 = m_{B^*}^2$, with $m_{B^*} = 5.325\,\text{GeV} < M_B + m_\pi$.

1. Becirevic-Kaidalov (BK) [26] :

$$f_+(q^2) = \frac{f_+(0)}{(1 - q^2/m_{B^*}^2)(1 - \alpha_{BK} q^2/m_{B^*}^2)}, \quad (8)$$

$$f_0(q^2) = \frac{f_0(0)}{1 - \beta_{BK}^{-1} q^2/m_{B^*}^2} , \quad (9)$$

where $f_+(0)$ and $f_0(0)$ set the normalizations and α_{BK} and β_{BK} define the shapes. The BK parameterization has been applied in fits to the HPQCD lattice predictions for form factors, with the constraint $f_+(0) = f_0(0)$.

2. Ball-Zwicky (BZ) [15, 16] :

$$f_+(q^2) = f_+(0)\left[\frac{1}{1 - q^2/m_{B^*}^2} \right.$$
$$\left. + \frac{r_{BZ} q^2/m_{B^*}^2}{(1 - q^2/m_{B^*}^2)(1 - \alpha_{BZ} q^2/m_{B^*}^2)}\right], \quad (10)$$

where $f_+(0)$ is the normalization, and α_{BZ} and r_{BZ} determine the shape. This is an extension of the BK ansatz, related by the simplification $\alpha_{BK} = \alpha_{BZ} = r_{BZ}$. This ansatz was used to extend the LCSR predictions to higher q^2, as shown in Figure 3.

3. Boyd, Grinstein, Lebed (BGL) [24, 25] :

$$f_+(q^2) = \frac{1}{\mathcal{P}(q^2)\phi(q^2, q_0^2)} \sum_{k=0}^{k_{max}} a_k(q_0^2)[z(q^2, q_0^2)]^k , \quad (11)$$

$$z(q^2, q_0^2) = \frac{\sqrt{m_+^2 - q^2} - \sqrt{m_+^2 - q_0^2}}{\sqrt{m_+^2 - q^2} + \sqrt{m_+^2 - q_0^2}} , \quad (12)$$

where $m_\pm = M_B \pm m_\pi$ and q_0^2 is a free parameter [27]. The so-called Blaschke factor $\mathcal{P}(q^2) = z(q^2, m_{B^*}^2)$ accounts for the pole at $q^2 = m_{B^*}^2$, and $\phi(q^2, q_0^2)$ is an arbitrary analytic function [28] whose choice only affects the particular values of the series coefficients a_k. In this expansion in the variable z, the shape is given by the values of a_k, with truncation at $k_{max} = 2$ or 3. The expansion parameters are constrained by unitarity, $\sum_k a_k^2 \leq 1$. Becher and Hill [25] have pointed out that due to the large b-quark mass, this bound is far from being saturated. Assuming that the ratio Λ/m_b is less than 0.1, the heavy-quark bound is approximately 30 times more constraining than the bound from unitarity alone, $\sum_k a_k^2 \sim (\Lambda/m_b)^3 \approx 0.001$. For more details we refer to the literature [24, 25].

4. Bourrely, Caprini, Lellouch (BCL) [29] :

$$f_+(q^2) = \frac{1}{1 - q^2/m_{B^*}^2} \sum_{k=0}^{k_{max}} b_k(q_0^2)\{[z(q^2, q_0^2)]^k -$$
$$(-1)^{k-k_{max}-1} \frac{k}{k_{max}+1}[z(q^2, q_0^2)]^{k_{max}+1}\} , \quad (13)$$

where the variable z is defined as in Eq. 12 with free parameter q_0^2 [27]. In this expansion the shape is given by the values of b_k, with truncation at $k_{max} = 2$ or 3. The BCL parameterization exhibits the QCD scaling behavior $f_+(q^2) \propto 1/q^2$ at large q^2.

The BK and BZ parameterizations are intuitive and have few free parameters. Fits to the previous $BABAR$ form-factor measurements using these parameterizations have shown that they describe the data quite well [9]. The BGL and BCL parameterizations are based on fundamental theoretical concepts like analyticity and unitarity. The z-expansion avoids ad hoc assumptions about the number of poles and pole masses, and it can be adapted to the precision of the data.

III. DATA SAMPLE, DETECTOR, AND SIMULATION

A. Data Sample

The data used in this analysis were recorded with the *BABAR* detector at the PEP-II energy-asymmetric e^+e^- collider operating at the $\Upsilon(4S)$ resonance. A sample of 377 million $\Upsilon(4S) \to B\overline{B}$ events, corresponding to an integrated luminosity of 349 fb^{-1}, was collected. An additional sample of 35.1 fb^{-1} was recorded at a center-of-mass (c.m.) energy approximately 40 MeV below the $\Upsilon(4S)$ resonance, *i.e.*, just below the threshold for $B\overline{B}$ production. This off-resonance data sample is used to subtract the non-$B\overline{B}$ contributions from the data collected at the $\Upsilon(4S)$ resonance. The principal source of these hadronic non-$B\overline{B}$ events is e^+e^- annihilation in the continuum to $q\overline{q}$ pairs, where $q = u, d, s, c$ refers to quarks. The relative normalization of the off-resonance and on-resonance data samples is derived from luminosity measurements, which are based on the number of detected $\mu^+\mu^-$ pairs and the QED cross section for $e^+e^- \to \mu^+\mu^-$ production, adjusted for the small difference in c.m. energy. The systematic error on the relative normalization is estimated to be 0.25%.

B. BABAR Detector

The *BABAR* detector and event reconstruction are described in detail elsewhere [30, 31]. The momenta and angles of charged particles are measured in a tracking system consisting of a five-layer silicon vertex tracker (SVT) and a 40-layer drift chamber (DCH) filled with a helium-isobutane gas mixture. Charged particles of different masses are distinguished by their ionization energy loss in the tracking devices and by a ring-imaging Cerenkov detector (DIRC). Electromagnetic showers from electrons and photons are measured in a finely segmented CsI(Tl) calorimeter (EMC). These detector components are embedded in the 1.5-T magnetic field of the solenoid. The magnet flux return steel is segmented and instrumented (IFR) with planar resistive plate chambers and limited streamer tubes, which detect particles penetrating the magnet coil and steel.

The efficiency for the reconstruction of charged particles inside the fiducial volume of the tracking system exceeds 96% and is well reproduced by MC simulation. An effort has been made to minimize fake charged tracks, caused by multiple counting of a single low-energy track curling in the DCH, split tracks, or background-generated tracks. The average uncertainty in the track-reconstruction efficiency is estimated to range from 0.25% to 0.5% per track.

To remove beam generated background and noise in the EMC, photon candidates are required to have an energy of more than 50 MeV and a shower shape that is consistent with an electromagnetic shower. The photon

efficiency and its uncertainty are evaluated by comparing $\tau^\pm \to \pi^\pm \nu$ to $\tau^\pm \to \rho^\pm \nu$ samples and by studying $e^+e^- \to \mu^+\mu^-(\gamma)$ events.

Electron candidates are selected on the basis of the ratio of the energy detected in the EMC and the track momentum, the EMC shower shape, the energy loss in the SVT and DCH, and the angle of the Cerenkov photons reconstructed in the DIRC. The energy of electrons is corrected for bremsstrahlung detected as photons emitted close to the electron direction. Muons are identified by using a neural network that combines the information from the IFR with the measured track momentum and the energy deposition in the EMC.

The electron and muon identification efficiencies and the probabilities to misidentify a pion, kaon, or proton as an electron or muon are measured as a function of the laboratory momentum and angles using high-purity samples of particles selected from data. These measurements are performed separately for positive and negative leptons. For the determination of misidentification probabilities, knowledge of the inclusive momentum spectra of positive and negative hadrons, and the measured fractions of pions, kaons and protons and their misidentification rates is used.

Within the acceptance of the SVT, DCH and EMC defined by the polar angle in the laboratory frame, $-0.72 < \cos\theta_{\text{lab}} < 0.92$, the average electron efficiency for laboratory momenta above 0.5 GeV is 93%, largely independent of momentum. The average hadron misidentification rate is less than 0.2%. Within the same polar-angle acceptance, the average muon efficiency rises with laboratory momentum to reach a plateau of about 70% above 1.4 GeV. The muon efficiency varies between 50% and 80% as a function of the polar angle. The average hadron misidentification rate is 2.5%, varying by about 1% as a function of momentum and polar angle.

Neutral pions are reconstructed from pairs of photon candidates that are detected in the EMC and assumed to originate from the primary vertex. Photon pairs with an invariant mass within 17.5 MeV of the nominal π^0 mass are considered π^0 candidates. The overall detection efficiency, including solid angle restrictions, varies between 55% and 65% for π^0 energies in the range of 0.2 to 2.5 GeV.

C. Monte Carlo Simulation

We assume that the $\Upsilon(4S)$ resonance decays exclusively to $B\overline{B}$ pairs [32] and that the non-resonant cross section for $e^+e^- \to q\overline{q}$ is 3.4 nb, compared to the $\Upsilon(4S)$ peak cross section of 1.05 nb. We use Monte Carlo (MC) techniques [33] to simulate the production and decay of $B\overline{B}$ and $q\overline{q}$ pairs and the detector response [34], to estimate signal and background efficiencies, and to extract the expected signal and background distributions. The size of the simulated sample of generic $B\overline{B}$ events exceeds the $B\overline{B}$ data sample by about a factor of three, while the

MC samples for inclusive and exclusive $B \to X_u\ell\nu$ decays exceed the data samples by factors of 15 or larger. The MC sample for $q\bar{q}$ events is comparable in size to the $q\bar{q}$ data sample recorded at the $\Upsilon(4S)$ resonance.

Information extracted from studies of selected data control samples on efficiencies and resolution is used to improve the accuracy of the simulation. Specifically, comparisons of data with the MC simulations reveal small differences in the tracking efficiencies and calorimeter resolution. We apply corrections to account for these differences. The MC simulations include radiative effects such as bremsstrahlung in the detector material and initial-state and final-state radiation [35]. Adjustments are made to take into account the small variations of the beam energies over time.

For this analysis, no attempt is made to reconstruct K_L^0 interacting in the EMC or IFR. Since a K_L^0 deposits only a small fraction of its energy in the EMC, K_L^0 production can have a significant impact on the energy and momentum balance of the whole event and thereby the neutrino reconstruction. It is therefore important to verify that the production rate of neutral kaons and their interactions in the detector are well reproduced.

From detailed studies of large data and MC samples of $D^0 \to K_L^0\pi^+\pi^-$ and $D^0 \to K_S^0\pi^+\pi^-$ decays, corrections to the simulation of the K_L^0 detection efficiency and energy deposition in the EMC are determined. The MC simulation reproduces the efficiencies well for K_L^0 laboratory momenta above 0.7 GeV. At lower momenta, the difference between MC and data increases significantly; in this range the MC efficiencies are reduced by randomly eliminating a fraction of the associated EMC showers. The energy deposited by K_L^0 in the EMC is significantly underestimated by the simulation for momenta up to 1.5 GeV. At higher momenta the differences decrease. Thus the simulated energies are scaled by factors varying between 1.20 and 1.05 as a function of momentum. Furthermore, assuming equal inclusive production rates for K_L^0 and K_S^0 we verify the production rate as a function of momentum, by comparing data and MC simulated K_S^0 momentum spectra. We observe differences at small momenta; below 0.4 GeV the data rate is lower by as much as $22 \pm 7\%$ compared to the MC simulation. To account for this difference, we reduce the rate of low momentum K_L^0 in the simulation by randomly transforming the excess K_L^0 into a fake photon, $i.e.$, we replace the energy deposited in the EMC by the total K_L^0 energy and set the mass to zero. Thus we correct the overall energy imbalance created by the excess in K_L^0 production.

For reference, the values of the branching fractions, lifetimes, and parameters most relevant to the MC simulation are presented in Tables I and II.

The simulation of inclusive charmless semileptonic decays $B \to X_u\ell\nu$ is based on predictions of a heavy-quark expansion (HQE) (valid to $\mathcal{O}(\alpha_s)$ [36]) for the differential decay rates. This calculation produces a smooth hadronic mass spectrum. The hadronization of X_u with masses above $2m_\pi$ is performed by JETSET [37]. To

TABLE I: Branching fractions and their errors for the semileptonic B decays used in this analysis.

Decay	Unit	B^0	$B^\pm$	Ref.
$B \to \eta\ell\nu$	10^{-4}		0.40 ± 0.09	[41]
$B \to \eta'\ell\nu$	10^{-4}		0.21 ± 0.21	[41]
$B \to \omega\ell\nu$	10^{-4}		1.15 ± 0.16	[41]
$B \to X_u\ell\nu$	10^{-3}	2.25 ± 0.22	2.41 ± 0.22	[7]
$B \to D\ell\nu$	10^{-2}	2.17 ± 0.08	2.32 ± 0.09	[41]
$B \to D^*\ell\nu$	10^{-2}	5.11 ± 0.19	5.48 ± 0.27	[41]
$B \to D_1^*\ell\nu$	10^{-2}	0.69 ± 0.14	0.77 ± 0.15	[41, 45]
$B \to D_2^*\ell\nu$	10^{-2}	0.56 ± 0.11	0.59 ± 0.12	[41, 45]
$B \to D_0^*\ell\nu$	10^{-2}	0.81 ± 0.24	0.88 ± 0.26	[41, 45]
$B \to D_1'\ell\nu$	10^{-2}	0.76 ± 0.22	0.82 ± 0.25	[41, 45]

TABLE II: Form factors used in the simulation of $B \to D\ell\nu$ and $B \to D^*\ell\nu$ decays, based on the parameterization of Caprini, Lellouch and Neubert [40]. The B^0 lifetime, the B^0 to B^+ lifetime ratio, and relative branching fraction at the $\Upsilon(4S)$ resonance, used in the simulation.

Parameter	Value	Ref.
$B \to D\ell\nu\,FF : \rho_D^2$	$1.18 \pm 0.04 \pm 0.04$	[41]
$B \to D^*\ell\nu\,FF : \rho_{D^*}^2$	$1.191 \pm 0.048 \pm 0.028$	[44]
$B \to D^*\ell\nu\,FF : R_1$	$1.429 \pm 0.061 \pm 0.044$	[44]
$B \to D^*\ell\nu\,FF : R_2$	$0.827 \pm 0.038 \pm 0.022$	[44]
B^0 life time $\tau_0\,(\mathrm{ps})$	1.530 ± 0.009	[7]
B life time ratio τ_+/τ_0	1.071 ± 0.009	[7]
$\Upsilon(4S)$ ratio f_{+-}/f_{00}	1.065 ± 0.026	[41]

describe the dynamics of the b quark inside the B meson we use HQE parameters extracted from global fits to moments of inclusive lepton-energy and hadron-mass distributions in $B \to X_c\ell\nu$ decays and moments of inclusive photon-energy distributions in $B \to X_s\gamma$ decays [38]. The specific values of the HQE parameters in the shape-function scheme are $m_b = 4.631 \pm 0.034$ GeV and $\mu_\pi^2 = 0.184 \pm 0.36\,\mathrm{GeV}^2$; they have a correlation of $\rho = -0.27$. Samples of exclusive semileptonic decays involving low-mass charmless mesons $(\pi, \rho, \omega, \eta, \eta')$ are simulated separately and then combined with samples of decays to non-resonant and higher-mass resonant states, so that the cumulative distributions of the hadron mass, the momentum transfer squared, and the lepton momentum reproduce the HQE predictions. The generated distributions are reweighted to accommodate variations due to specific choices of the parameters for the inclusive and exclusive decays. The overall normalization is adjusted to reproduce the measured inclusive $B \to X_u\ell\nu$ branching fraction.

For the generation of decays involving charmless pseudo-scalar mesons we choose two approaches. For the signal decay $B \to \pi\ell\nu$ we use the ansatz by Becirevic and Kaidalov [26] for the q^2 dependence, with the single parameter α_{BK} set to the value determined in a previous

BABAR analysis [9] of $B \to \pi \ell \nu$ decays, $\alpha_{BK} = 0.52 \pm 0.06$. For decays to η and η' we use the form factor parameterization of Ball and Zwicky with specific values reported in [18].

Decays involving charmless vector mesons (ρ, ω) are generated based on form factors determined from LCSR by Ball and Zwicky [17]. We use the parameterization proposed by the authors to describe the q^2 dependence of the form factors in terms of a modified pole ansatz using up to three independent parameters r_1, r_2 and m_{fit}. Table III shows the suggested values for these parameters. m_{fit} refers to an effective pole mass that accounts for contributions from higher-mass B mesons with $J^P = 1^-$, and r_1, and r_2 give the relative scale of the two pole terms.

TABLE III: Parameterization of the LCSR form-factor calculations [15, 17] for decays to pseudo-scalar mesons η and η' (f_+) and vector mesons ρ and ω (A_1, A_2, V).

Form factor	f_+	A_1^ρ	A_2^ρ	V^ρ	A_1^ω	A_2^ω	V^ω
$F(0)$	0.273	0.242	0.221	0.323	0.219	0.198	0.293
r_1	0.122	–	0.009	1.045	–	0.006	1.006
r_2	0.155	0.240	0.212	-0.721	0.217	0.192	-0.713
$m_{\text{fit}}^2(\text{GeV}^2)$	31.46	37.51	40.82	38.34	37.01	41.24	37.45

For the simulation of the dominant $B \to X_c \ell \nu$ decays, we have chosen a variety of models. For $B \to D\ell\nu$ and $B \to D^*\ell\nu$ decays we use parameterizations [39, 40] of the form factors based on heavy quark effective theory (HQET). In the limit of negligible lepton masses, decays to pseudoscalar mesons are described by a single form factor for which the q^2 dependence is given by a slope parameter. We use the world average [41], updated for recent precise measurements by the BABAR Collaboration [42, 43]. Decays to vector mesons are described by three form factors, of which the axial vector form factor dominates. In the limit of heavy quark symmetry, their q^2 dependence can be described by three parameters: $\rho_{D^*}^2$, R_1, and R_2. We use the most precise BABAR measurement [44] of these parameters.

For the generation of the semileptonic decays to D^{**} resonances (four $L = 1$ states), we use the ISGW2 [14] model. At present, the sum of the branching fractions for these four decays modes is measured to be 1.7%, but so far only the decays $D^{**} \to D\pi$ and $D^{**} \to D^*\pi$ have been reconstructed, while the total individual branching fractions for these four states remain unknown. Since the measured inclusive branching fraction for $B \to X_c \ell \nu$ exceeds the sum of the measured branching fractions of all exclusive semileptonic decays by about 1.0%, and since non-resonant $B \to D^{(*)}\pi\ell\nu$ decays have not been observed [45], we assume that the missing decays are due to $B \to D^{**}\ell\nu$, involving hadronic decays of the D^{**} mesons that have not yet been measured. To account for the observed deficit, we increase the $B \to D^{**}\ell\nu$ branch-

ing fractions by 60% and inflate the errors by a factor of three.

IV. EVENT RECONSTRUCTION AND CANDIDATE SELECTION

In the following, we describe the selection and kinematic reconstruction of signal candidates, the definition of the various background classes, and the application of neural networks to suppress these backgrounds.

A. Signal-Candidate Selection

Signal candidates are selected from events having at least four charged tracks. The reconstruction of the four signal decay modes, $B^0 \to \pi^-\ell^+\nu$, $B^+ \to \pi^0\ell^+\nu$, $B^0 \to \rho^-\ell^+\nu$ and $B^+ \to \rho^0\ell^+\nu$, requires the identification of a charged lepton, the reconstruction of the hadronic state consisting of one or more charged or neutral pions, and the reconstruction of the neutrino from the missing energy and missing momentum of the whole event.

1. Lepton and Hadron Selection

Candidates for leptons, both $e^\pm$ and $\mu^\pm$, are required to have high c.m. momenta, $p_\ell^* \geq 1.0 \, \text{GeV}$ for the $B \to \pi\ell\nu$, and $p_\ell^* \geq 1.8 \, \text{GeV}$ for the $B \to \rho\ell\nu$ sample. This requirement significantly reduces the background from hadrons that are misidentified as leptons, and also removes a large fraction of true leptons from secondary decays or photon conversions, and from $B \to X_c\ell\nu$ decays.

To suppress Bhabha scattering and two-photon processes in which an electron or a photon from initial-state or final state radiation interacts in the material of the detector and generates additional charged tracks and photons at small angles to the beam axis, we require $\xi_z < 0.65$ for events with a candidate electron. Here $\xi_z = \sum_i p_i^z / \sum_i E_i$, where the sum runs over all charged particles in the event and p_i^z and E_i are their longitudinal momentum components and energies measured in the laboratory frame.

For the reconstruction of the signal hadron, we consider all charged tracks that are not consistent with a signal lepton and not identified as a kaon. Neutral pions are reconstructed from pairs of photon candidates and the π^0 c.m. momentum is required to exceed 0.2 GeV. Candidate $\rho^\pm \to \pi^\pm\pi^0$ or $\rho^0 \to \pi^+\pi^-$ decays are required to have a two-pion mass within one full width of the nominal ρ mass, $0.650 < M_{\pi\pi} < 0.850 \, \text{GeV}$. To reduce the combinatorial background, we also require that the c.m. momentum of one of the pions exceed 0.4 GeV, and that the c.m. momentum of the other pion be larger than 0.2 GeV.

Each charged lepton candidate is combined with a hadron candidate to form a so-called Y candidate of charge zero or one. At this stage in the analysis we allow for more than one candidate per event. Two or three charged tracks associated with the Y candidate are fitted to a common vertex. This vertex fit must yield a χ^2 probability of at least 0.1%. To remove background from $J/\psi \to \ell^+\ell^-$ decays, we reject a Y candidate if the invariant mass of the lepton and any oppositely charged track in the event is consistent with this decay.

To further reduce backgrounds without significant signal losses, we impose additional restrictions on the c.m. momenta of the lepton and hadron candidates by requiring at least one of the following conditions to be satisfied, for $B \to \pi\ell\nu$

$$p^*_{\mathrm{hadron}} \geq 1.3\,\mathrm{GeV} \quad \text{or}$$
$$p^*_\ell \geq 2.2\,\mathrm{GeV} \quad \text{or}$$
$$p^*_{\mathrm{hadron}} + p^*_\ell \geq 2.8\,\mathrm{GeV},$$

and for $B \to \rho\ell\nu$

$$p^*_{\mathrm{hadron}} \geq 1.3\,\mathrm{GeV} \quad \text{or}$$
$$p^*_\ell \geq 2.0\,\mathrm{GeV} \quad \text{or}$$
$$p^*_{\mathrm{hadron}} + p^*_\ell \geq 2.65\,\mathrm{GeV}.$$

These additional requirements on the lepton and hadron c.m. momenta primarily reject background candidates that are inconsistent with the phase space of the signal decay modes.

If a Y candidate originates from a signal decay mode, the cosine of the angle between the momentum vectors of the B meson and the Y candidate, $\cos\theta_{BY}$, can be calculated as follows,

$$\cos\theta_{\mathrm{BY}} = \frac{2E^*_B E^*_Y - M^2_B - M^2_Y}{2p^*_B\, p^*_Y}, \qquad (14)$$

and the condition $|\cos\theta_{\mathrm{BY}}| \leq 1.0$ should be fulfilled. The energy E^*_B and momentum p^*_B of the B meson are not measured event-by-event. Specifically, $E^*_B = \sqrt{s}/2$ is given by the average c.m. energy of the colliding beams, and the B momentum is derived as $p^*_B = \sqrt{E^{*2}_B - M^2_B}$. To allow for the finite resolution in this variable, we impose the requirement $-1.2 < \cos\theta_{\mathrm{BY}} < 1.1$.

2. Neutrino Reconstruction

The neutrino four-momentum, $P_\nu \simeq (E_{\mathrm{miss}}, \vec{p}_{\mathrm{miss}})$, is inferred from the difference between the net four-momentum of the colliding-beam particles, $P_{e^+e^-} = (E_{e^+e^-}, \vec{p}_{e^+e^-})$, and the sum of the measured four-vectors of all detected particles in the event,

$$(E_{\mathrm{miss}}, \vec{p}_{\mathrm{miss}}) = (E_{e^+e^-}, \vec{p}_{e^+e^-}) - (\sum_i E_i, \sum_i \vec{p}_i), \quad (15)$$

where E_i and $\vec{p}_i$ are the energy and three-momentum of the ith track or EMC shower, measured in the laboratory frame. The energy calculation depends on the correct mass assignments for charged tracks. For this reason we choose to calculate the missing momentum and energy in the laboratory frame rather than in the rest frame of the $\Upsilon(4S)$. By doing so, we keep this uncertainty confined to the missing energy.

If all particles in the event, except the neutrino, are well measured, $P_\nu \simeq (E_{\mathrm{miss}}, \vec{p}_{\mathrm{miss}})$ is a good approximation. However, particles that are undetected because of inefficiency or acceptance losses, in particular K_L mesons and additional neutrinos, or spurious tracks or photons that do not originate from the $B\overline{B}$ event, impact the accuracy of this approximation. To reduce the effect of losses due to the limited detector acceptance, we require that the polar angle of the missing momentum in the laboratory frame be in the range $0.3 < \theta_{miss} < 2.2\,\mathrm{rad}$. We also require the missing momentum in the laboratory frame to exceed 0.5 GeV.

For the rejection of background events and signal decays that are poorly reconstructed as well as events with more than one missing particle, we make use of the missing mass squared of the whole event,

$$P^2_\nu \simeq m^2_{\mathrm{miss}} = E^2_{\mathrm{miss}} - |\vec{p}_{\mathrm{miss}}|^2. \qquad (16)$$

For a correctly reconstructed event with a single semileptonic B decay, m^2_{miss} should be consistent with zero within measurement errors. Failure to detect one or more particles in the event creates a substantial tail at large positive values. Since the resolution in m^2_{miss} increases linearly with E_{miss}, we use the variable $m^2_{\mathrm{miss}}/2E_{\mathrm{miss}} \simeq E_{\mathrm{miss}} - p_{\mathrm{miss}}$ as a discriminator and require $m^2_{\mathrm{miss}}/2E_{\mathrm{miss}} < 2.5\,\mathrm{GeV}$.

3. Variables Used for Signal Extraction

The kinematic consistency of the candidate decay with a signal B decay is ascertained using two variables, the beam-energy substituted B mass m_{ES}, and the difference between the reconstructed and expected energy of the B candidate ΔE. In the laboratory frame, they are defined as

$$\Delta E = \frac{P_B \cdot P_{e^+e^-} - s/2}{\sqrt{s}} \qquad (17)$$

and

$$m_{\mathrm{ES}} = \sqrt{\frac{(s/2 + \vec{p}_B \cdot \vec{p}_{e^+e^-})^2}{E^2_{e^+e^-}} - p^2_B}\,, \qquad (18)$$

where $P_B = (E_B, \vec{p}_B)$ and $P_{e^+e^-}$ denote the four-momenta of the B meson and the colliding beam particles, respectively. The B-meson momentum vector $\vec{p}_B$ is determined from the measured three-momenta of the decay products, and $P_{e^+e^-}$ is derived from the calibration and monitoring of the energies and angles of the stored beams. We extract the signal yields by a fit to the

two-dimensional $\Delta E - m_{\mathrm{ES}}$ distributions in bins of the momentum transfer squared q^2. We define a region in the $\Delta E - m_{\mathrm{ES}}$ plane that contains almost all of the signal events and leaves sufficient phase space to constrain the different background contributions. This *fit region* is defined as

$$|\Delta E| < 0.95\,\mathrm{GeV}\ ,\ \ 5.095 < m_{\mathrm{ES}} < 5.295\,\mathrm{GeV}. \quad (19)$$

Only candidates that fall inside the fit region are considered in the analysis. We also define a smaller region where the signal contribution is much enhanced relative to the background. This *signal region* is defined as

$$-0.15 < \Delta E < 0.25\,\mathrm{GeV}\ ,\ \ 5.255 < m_{\mathrm{ES}} < 5.295\,\mathrm{GeV}. \quad (20)$$

The signal region is chosen to be slightly asymmetric in ΔE to avoid sizable $B \to X_c \ell \nu$ background, which peaks near $-0.2\,\mathrm{GeV}$. In the following, we refer to the phase space outside the signal region, but inside the fit region, as the *side bands*.

As a measure of the momentum transfer squared q^2 we adopt the mass squared of the virtual W, *i.e.*, the invariant mass squared of the four-vector sum of the reconstructed lepton and neutrino,

$$q^2_{\mathrm{raw}} = [(E_\ell, \vec{p}_\ell) + (E_{\mathrm{miss}}, \vec{p}_{\mathrm{miss}})]^2. \quad (21)$$

The resolution in q^2_{raw} is dominated by the measurement of the missing energy which tends to have a poorer resolution than the measured missing momentum, because the missing momentum is a vector sum and contributions from particle losses (or additional tracks and EMC showers) do not add linearly as is the case for E_{miss}. Thus for the definition of q^2_{raw} it is advantageous to replace E_{miss} by p_{miss}, the absolute value of the measured missing momentum,

$$q^2_{\mathrm{raw}} = [(E_\ell, \vec{p}_\ell) + (p_{\mathrm{miss}}, \vec{p}_{\mathrm{miss}})]^2. \quad (22)$$

The resolution of q^2_{raw} can be further improved by scaling p_{miss} by a factor of α, such that ΔE of the B candidate is forced to zero,

$$\vec{p}_\nu = \alpha \vec{p}_{\mathrm{miss}} \quad \text{with } \alpha = 1 - \frac{\Delta E}{E_{\mathrm{miss}}}\ , \quad (23)$$

and substituting $\vec{p}_\nu$ for $\vec{p}_{\mathrm{miss}}$ to obtain q^2_{corr}. Any candidates for which this q^2 correction yields unphysical values, $q^2_{\mathrm{corr}} < 0\,\mathrm{GeV}^2$, are rejected. This is the case for about 1% of the background not associated with semileptonic decays. The quantity q^2_{corr} is used as the measured q^2 throughout this analysis.

The q^2 resolution is critical for the measurements of the form factors. Figure 4 shows the correlation between the true q^2 and the reconstructed q^2_{corr} for simulated samples of $B \to \pi \ell \nu$ and $B \to \rho \ell \nu$ candidates passing the entire event selection, which is described below. Correctly reconstructed signal events and combinatorial signal events, for which the hadron has been incorrectly

selected, are shown. For correctly reconstructed signal decays the resolution improves with higher q^2 and can be well described by the sum of two Gaussian resolution functions, see Table IV. In the signal region, the widths of the core resolution are in the range $0.18 - 0.34\,\mathrm{GeV}^2$, and the tails can be approximated by a second Gaussian function with widths in the range $0.6 - 0.8\,\mathrm{GeV}^2$. As expected, the resolution is significantly worse in the larger fit region. Combinatorial signal events contribute primarily at high q^2. We rely on the Monte Carlo simulation to reproduce the resolution in the reconstructed q^2_{corr} variable.

B. Background Suppression

1. Signal and Background Sources

A variety of processes contribute to the four samples of selected candidates for the charmless semileptonic decay modes $B^0 \to \pi^- \ell^+ \nu$, $B^+ \to \pi^0 \ell^+ \nu$, $B^0 \to \rho^- \ell^+ \nu$, and $B^+ \to \rho^0 \ell^+ \nu$. We divide the signal and background for each of the four candidate samples into a set of sources based on the origin of the charged lepton candidate.

- Signal: We differentiate four classes of signal events; for all of them the lepton originates from a signal decay under study:

  - *True signal:* the hadron originates from the signal decay under study;

  - *Combinatorial Signal:* the hadron is incorrectly selected, in many cases from decay products of the second B meson in the event;

  - *Isospin-conjugate signal:* the lepton originates from the isospin conjugate of the signal decay;

  - *Cross-feed signal:* the lepton originates from another signal decay mode, for instance $B \to \rho \ell \nu$ in a $B \to \pi \ell \nu$ sample.

- Continuum background: We differentiate two classes of continuum backgrounds:

  - *True leptons:* the lepton candidate originates from a leptonic or semileptonic decay of a hadron produced in $e^+ e^- \to q \bar{q}$ (mostly $c \bar{c}$) or $e^+ e^- \to \ell^+ \ell^- (\gamma)$ processes, where $\ell^+ \ell^-$ stands for $e^+ e^-, \mu^+ \mu^-$ or $\tau^+ \tau^-$, or $e^+ e^- \to \gamma \gamma$;

  - *Fake leptons:* the lepton candidate is a misidentified hadron; this is a sizable contribution to the muon sample.

- $B \to X_u \ell \nu$ background: We differentiate two different sources of $B \to X_u \ell \nu$ background:

  - *Exclusive $B \to X_u \ell \nu$ decays involving a single hadron with mass below 1 GeV:* decays
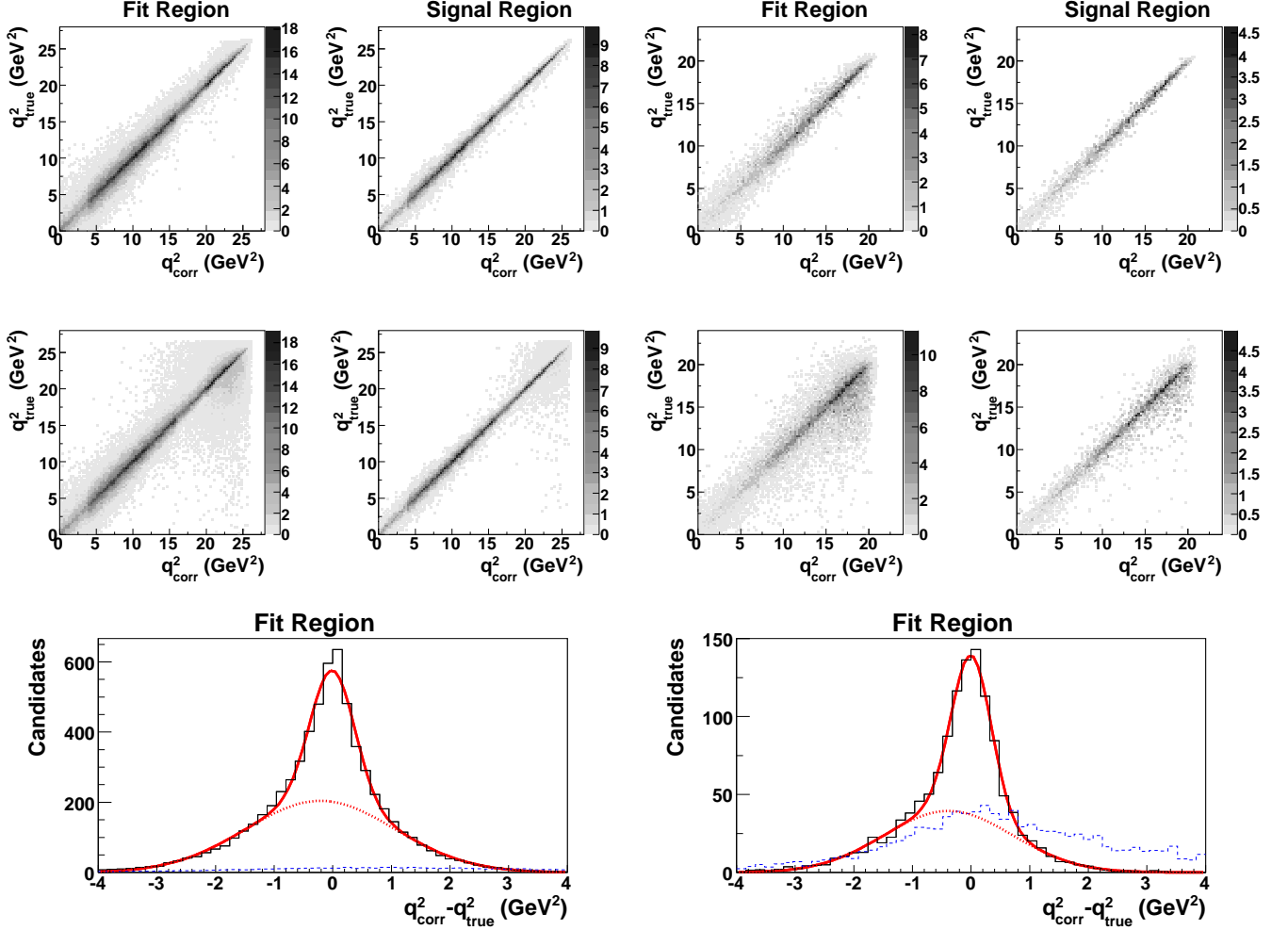
FIG. 4: q^2 resolution for $B^0 \rightarrow \pi^- \ell^+ \nu$ (left) and $B^0 \rightarrow \rho^- \ell^+ \nu$ (right) samples after the full event selection: two-dimensional distribution of q^2_{true} versus q^2_{corr} in the fit region and in the signal region. Top row: true signal decays, middle row: all signal decays (true and combinatorial), bottom row: distribution of $q^2_{\text{corr}} - q^2_{\text{true}}$ for true signal (black, solid histogram) and combinatorial signal (blue, dashed histogram) in the fit region. The fit of the sum of two Gaussian functions to the true signal distribution is shown as a solid red line, the contribution of the broader of the two functions is shown as a dotted red line.

that are not analyzed as signal ($B^+ \rightarrow \omega \ell^+ \nu$, $B^+ \rightarrow \eta \ell^+ \nu$, and $B^+ \rightarrow \eta' \ell^+ \nu$) ;

– *Inclusive $B \rightarrow X_u \ell \nu$ decays:* decays involving more than one hadron or a single hadron with mass above 1 GeV.

- $B\overline{B}$ background: We differentiate three classes of $B\overline{B}$ background, excluding $B \rightarrow X_u \ell \nu$ decays:

  – *Primary leptons, i.e., $B \rightarrow X_c \ell \nu$ decays:* the lepton originates from a charm semileptonic B decay, either $B \rightarrow D\ell\nu$, $B \rightarrow D^*\ell\nu$, or $B \rightarrow D^{(*)}(n\pi)\ell\nu$ with $n \geq 1$ additional pions; this class is dominated by $B \rightarrow D^*\ell\nu$ decays; the largest contributions involve hadrons that do not originate from the semileptonic decay;

  – *Secondary leptons:* the lepton originates from the decay of a particle other than a B

meson, for instance charm mesons, τ leptons, J/ψ, or from photon conversions;

– *Fake leptons:* the lepton candidate is not a lepton, but a misidentified charged hadron; this background is dominated by fake muons.

Given that the secondary-lepton and fake-lepton $B\overline{B}$ background contributions are relatively small in this analysis, we combine them into one class (other $B\overline{B}$).

For intermediate values of q^2 (in the range $4 < q^2 < 20\,\text{GeV}^2$), $B \rightarrow X_c \ell \nu$ decays are by far the dominant background, whereas continuum background contributes mostly at low and high q^2. The $B \rightarrow X_u \ell \nu$ decays have much smaller branching fractions, but their properties are very similar to the signal decays and thus they are difficult to discriminate against. They contribute mostly at high q^2, where they are the dominant background.

TABLE IV: Description of the q^2 resolution in terms of a sum of two Gaussian resolution functions for true signal decays in the $\Delta E - m_{\mathrm{ES}}$ fit region and in the signal region, integrated over q^2; σ_1, μ_1 and σ_2, μ_2 denote the means and the widths of the two Gaussian functions, and the last column lists the fraction of the events characterized by the narrower resolution function.

| | | Gaussian Fct. 1 | | Gaussian Fct. 2 | | |
	Signal mode	$\mu_1(\,\mathrm{GeV}^2)$	$\sigma_1(\,\mathrm{GeV}^2)$	$\mu_2(\,\mathrm{GeV}^2)$	$\sigma_2(\,\mathrm{GeV}^2)$	Fraction
Fit region	$B^0 \to \pi^- \ell^+ \nu$	-0.005	0.380	-0.021	1.270	0.35
	$B^+ \to \pi^0 \ell^+ \nu$	0.076	0.468	-0.039	1.343	0.43
	$B^0 \to \rho^- \ell^+ \nu$	0.005	0.343	-0.386	1.094	0.45
	$B^+ \to \rho^0 \ell^+ \nu$	-0.032	0.311	-0.498	1.086	0.46
Signal region	$B^0 \to \pi^- \ell^+ \nu$	0.006	0.242	-0.020	0.720	0.45
	$B^+ \to \pi^0 \ell^+ \nu$	0.058	0.338	0.172	0.807	0.58
	$B^0 \to \rho^- \ell^+ \nu$	0.042	0.246	0.036	0.647	0.50
	$B^+ \to \rho^0 \ell^+ \nu$	0.010	0.177	-0.078	0.586	0.46

2. Neural Networks

To separate signal events from the background sources, continuum events, non-signal $B \to X_u \ell \nu$ decays and the remaining $B\overline{B}$ events, we employ a neural-network technique based on a multi-layer perceptron (MLP) [46]. We have set up a network structure with seven input neurons and one hidden layer with three neurons and have adopted the method introduced by Broyden, Fletcher, Goldfarb, and Shanno [47] to train the network. Some of the input variables are used as part of the event preselection that is designed to reduce the $B\overline{B}$ and continuum backgrounds by cutting out regions where the signal contribution is small or where there are spikes in distributions, which the neural network may not deal with effectively. The following variables are input to the neural networks:

- $R2$, the second normalized Fox-Wolfram moment [48] determined from all charged and neutral particles in the event; we require $R2 < 0.5$;

- $L2 = \sum_i p_i^* \cos^2 \theta_i^*$, where the sum runs over all tracks in the event excluding the Y candidate, and p_i^* and θ_i^* refer to the c.m. momenta and the angles measured with respect to the thrust axis of the Y candidate; we set a loose restriction, $L2 < 3.0$ GeV.

- $\cos \Delta \theta_{\mathrm{thrust}}$, where $\Delta \theta_{\mathrm{thrust}}$ is the angle between the thrust axis of the Y candidate and thrust axis of all other detected particles in the event; there is no preselection requirement for this variable;

- $m_{\mathrm{miss}}^2/(2E_{\mathrm{miss}}) = (E_{\mathrm{miss}}^2 - p_{\mathrm{miss}}^2)/(2E_{\mathrm{miss}})$; we require $|m_{\mathrm{miss}}^2/2E_{\mathrm{miss}}| < 2.5$ GeV;

- $\cos \theta_{\mathrm{BY}}$; we require $-1.2 < \cos \theta_{\mathrm{BY}} < 1.1$;

- $\cos \theta_{W\ell}$, the helicity angle of the lepton; we require $|\cos \theta_{W\ell}| < 0.8$ for the $B^0 \to \pi^- \ell^+ \nu$ and $B^+ \to \pi^0 \ell^+ \nu$ modes;

- θ_{miss}, the polar angle of the missing momentum in the laboratory frame; we require $0.3 < \theta_{\mathrm{miss}} < 2.2$ rad.

The first three input variables are sensitive to the topological difference between the jet-like continuum events and the more spherical $B\overline{B}$ events. Restrictions on these variables do not bias the q^2 distribution significantly.

The restrictions placed on $\cos \theta_{\mathrm{BY}}$, $m_{\mathrm{miss}}^2/(2E_{\mathrm{miss}})$, and θ_{miss} do not significantly bias the q^2 distribution either. However, the variable $\cos \theta_{W\ell}$ is correlated with the lepton momentum and thereby q^2. To ensure that the selection does not adversely affect the measurement of the q^2 spectrum, we have chosen rather moderate restrictions on $\cos \theta_{W\ell}$.

Figure 5 shows the ΔE and m_{ES} distributions for samples of $B^0 \to \pi^- \ell^+ \nu$ and $B^0 \to \rho^- \ell^+ \nu$ candidates (integrated over q^2) that have been preselected by the criteria described above. The stacked histograms show the signal and background contributions compared to the data, prior to the fit. The three dominant backgrounds are $B \to X_c \ell \nu$ decays (including $B \to D\ell\nu$, $B \to D^* \ell \nu$ and $B \to D^{(*)}(n\pi)\ell\nu)$, $q\bar{q}$ continuum and $B \to X_u \ell \nu$ decays. The signal contributions are very small by comparison and difficult to observe.

The neural networks are trained separately for the three background categories and for different q^2 intervals. We introduce six bins in q^2 for $B \to \pi \ell \nu$ and three bins for $B \to \rho \ell \nu$. The bin sizes are $4 \,\mathrm{GeV}^2$ for $B \to \pi \ell \nu$ and $8 \,\mathrm{GeV}^2$ for $B \to \rho \ell \nu$, except for the last bin, which extends to the kinematic limit of $26.4 \,\mathrm{GeV}^2$ and $20.3 \,\mathrm{GeV}^2$, respectively. Thus in total we train $3 \times (2 \times 6 + 2 \times 3) = 54$ neural networks. Since we aim for a good signal-to-background ratio in the region where most of the signal is located, we do not train the neural network with events in the whole fit region, but in an extended signal region, $-0.25 < \Delta E < 0.35 \,\mathrm{GeV}$, $5.240 < m_{\mathrm{ES}} < 5.295 \,\mathrm{GeV}$.

For the training of the neural networks we use MC simulated events containing correctly reconstructed signal decays and the following simulated background samples:

1. a sample of continuum events, $e^+ e^- \to q\bar{q}$ with $q = u, d, s, c$ ($q\bar{q}$ neural network);

2. a combined sample of $B \to X_c \ell \nu$ decays ($B \to X_c \ell \nu$ neural network); and

3. a sample of inclusive $B \to X_u \ell \nu$ decays ($B \to X_u \ell \nu$

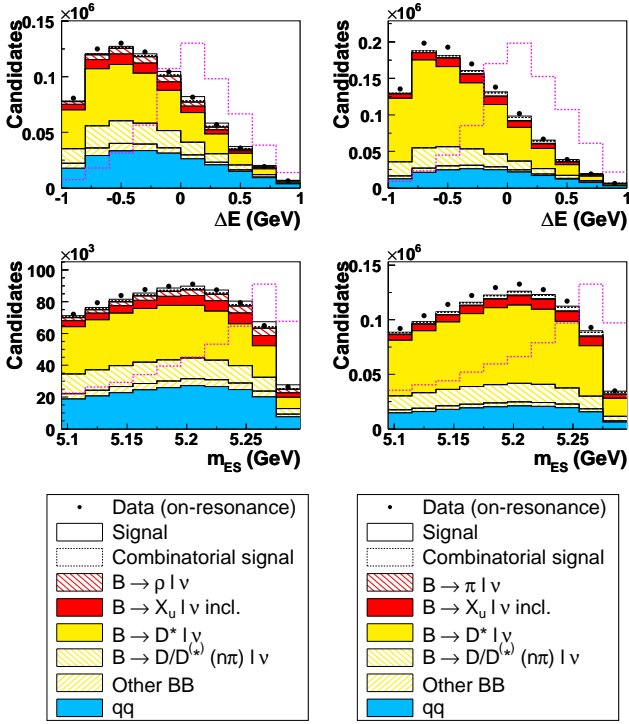FIG. 5: (color online) Distributions of ΔE and m_{ES} for $B^0 \rightarrow \pi^- \ell^+ \nu$ (left) and $B^0 \rightarrow \rho^- \ell^+ \nu$ (right) candidates after the preselection, i.e., prior to the neural-network application. The stacked histograms show the predicted signal and background contributions prior to the fit. The expected signal distribution (with arbitrary normalization) is indicated as a magenta dashed histogram.

neural network).

The training of the neural networks and the subsequent background reduction is performed sequentially for the three background samples. We use subsamples of typically less than half the total MC samples for training and validation of the neural networks. Of these subsamples, one half of the events is used as training sample, and the other half for validation.

Studies of the neural-network performance for the $B \rightarrow X_u \ell \nu$ background indicate that the separation of this background from the signal is very difficult because of the similarity in the shape of the distributions, especially for the $B^+ \rightarrow \pi^0 \ell^+ \nu$ and the $B \rightarrow \rho \ell \nu$ samples. Given these difficulties, we use the $B \rightarrow X_u \ell \nu$ neural network only for the $B^0 \rightarrow \pi^- \ell^+ \nu$ sample, and only for $q^2 > 12 \, \text{GeV}^2$, where the $B \rightarrow X_u \ell \nu$ background becomes significant.

Figure 6 shows, for the sample of $B^0 \rightarrow \pi^- \ell^+ \nu$ candidates, the distributions of the seven input variables to the neural networks. The distributions are shown sequentially after application of the preselection, the $q\bar{q}$ neural network and the $B \rightarrow X_c \ell \nu$ neural network to illustrate the change in the sample composition. Figures 7 to 9 show the distributions of the three neural-network discriminators for the $B^0 \rightarrow \pi^- \ell^+ \nu$ sample in four of the six

q^2 bins. Figures 10 and 11 show the distributions of the two neural-network discriminators for the $B^0 \rightarrow \rho^- \ell^+ \nu$ sample in all three q^2 bins. The discriminator cuts are chosen to minimize the total error on the signal yield for each channel, using the sum in quadrature of the error obtained from the maximum-likelihood fit described in Section VI and the estimated total systematic error of the partial signal branching fraction in each q^2 bin (see Section VII). The data-MC agreement is reasonably good for the input distributions and the neural-network discriminators. One should keep in mind that at this stage the distributions are taken directly from the simulation, without any adjustments or fit.

TABLE V: Overview of the selection efficiencies for the four signal decays (true and combinatorial signal combined) and their primary background sources, $B \rightarrow X_u \ell \nu$, $B \rightarrow X_c \ell \nu$, and non-$B\overline{B}$ background.

	Selection	Signal	$X_u \ell \nu$	$X_c \ell \nu$	$q\bar{q}$
	Units	10^{-2}	10^{-3}	10^{-4}	10^{-5}
$B^0 \rightarrow \pi^- \ell^+ \nu$					
	Preselection	18.1	25.6	26.4	19.4
	NN $q\bar{q}$	11.6	13.4	18.3	2.6
	NN $B \rightarrow X_c \ell \nu$	7.8	8.1	3.4	1.3
	NN $B \rightarrow X_u \ell \nu$	6.8	5.3	2.5	1.0
	Signal region	1.8	0.5	0.1	<0.1
$B^+ \rightarrow \pi^0 \ell^+ \nu$					
	Preselection	12.8	20.0	17.4	15.1
	NN $q\bar{q}$	8.4	11.3	13.3	2.7
	NN $B \rightarrow X_c \ell \nu$	5.9	6.6	1.9	1.6
	Signal region	1.6	0.5	<0.1	<0.1
$B^0 \rightarrow \rho^- \ell^+ \nu$					
	Preselection	8.9	23.9	35.8	13.2
	NN $q\bar{q}$	4.8	11.9	18.5	1.0
	NN $B \rightarrow X_c \ell \nu$	1.1	1.9	0.3	0.2
	Signal region	0.3	0.3	<0.1	<0.1
$B^+ \rightarrow \rho^0 \ell^+ \nu$					
	Preselection	11.1	22.1	30.0	12.6
	NN $q\bar{q}$	6.8	12.4	17.7	1.5
	NN $B \rightarrow X_c \ell \nu$	2.5	3.2	0.7	0.5
	Signal region	0.8	0.6	<0.1	<0.1

Table V shows the selection efficiencies for the four signal samples compared to the efficiencies for the dominant background sources for these samples. The total signal efficiencies are typically $6 - 7\%$ for $B \rightarrow \pi \ell \nu$ decays and roughly $1 - 2.5\%$ for $B \rightarrow \rho \ell \nu$ decays in the fit region. The dominant $B\overline{B}$ and $q\bar{q}$ backgrounds are suppressed by factors of order 10^4 and 10^5, respectively.

3. Candidate Multiplicity

After the neural-network selection there are on average 1.14 candidates per event in the $B^0 \rightarrow \pi^- \ell^+ \nu$ sample, 1.46 in the $B^+ \rightarrow \pi^0 \ell^+ \nu$ sample, 1.30 in the $B^0 \rightarrow \rho^- \ell^+ \nu$ sample, and 1.17 in the $B^+ \rightarrow \rho^0 \ell^+ \nu$ sample. We observe fewer candidates for decay modes without neutral pions
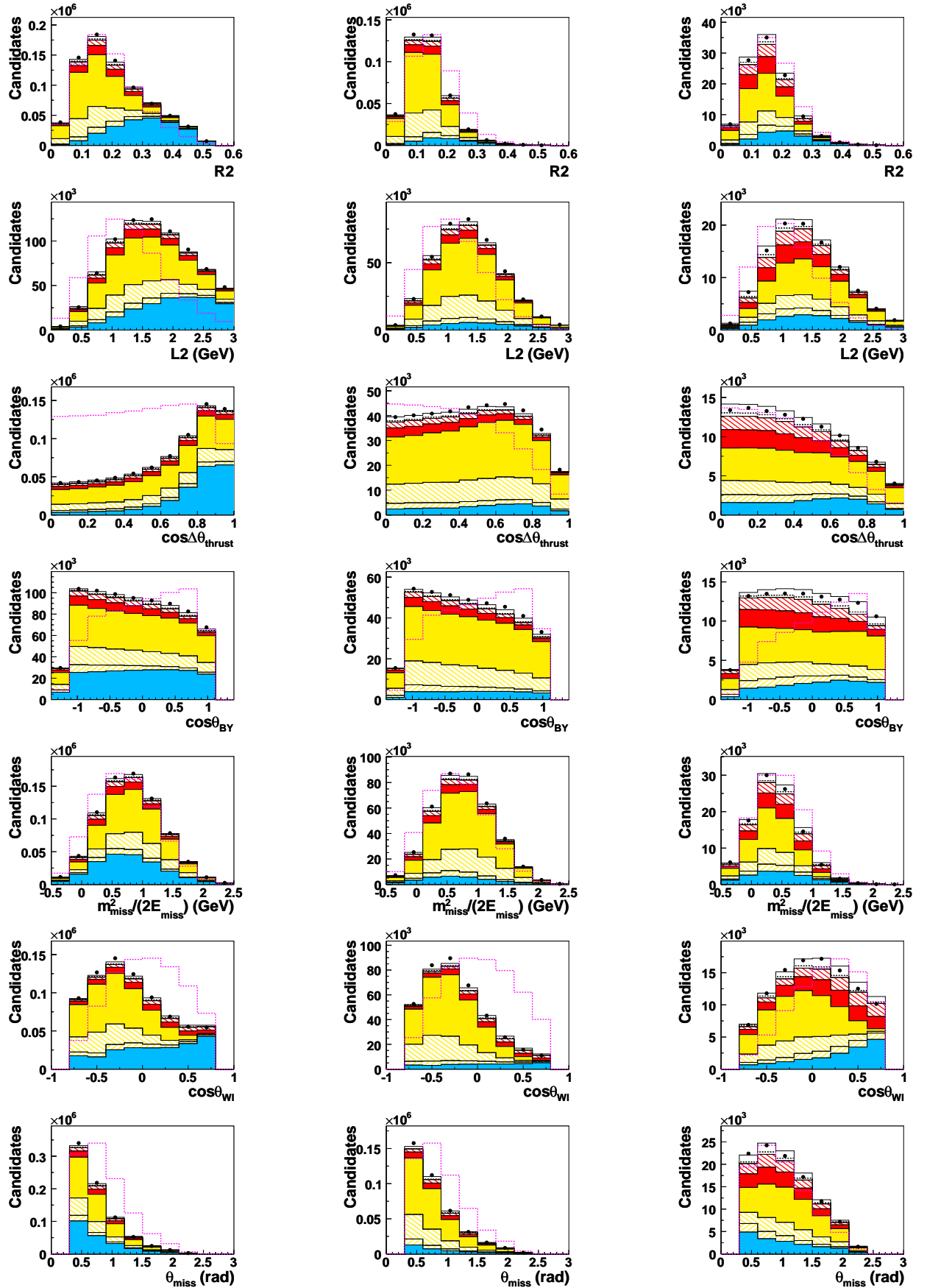
FIG. 6: (color online) Background suppression for $B^0 \to \pi^- \ell^+ \nu$ candidates. Distributions of the seven input variables to the neural network: after the preselection (left column), after the $q\bar{q}$ neural network (center column), and after the $B \to X_c \ell \nu$ neural network (right column). The data are compared to the sum of the MC-simulated background contributions; for a legend see Figure 5. The expected signal distribution is overlaid as a magenta, dashed histogram with arbitrary normalization.

FIG. 7: (color online) The $q\bar{q}$ neural-network discriminators for $B^0 \to \pi^- \ell^+ \nu$ candidates in the signal region, $-0.15 < |\Delta E| < 0.25 \,\text{GeV}$; $5.255 < m_{\text{ES}} < 5.295 \,\text{GeV}$. The distributions are shown for four different q^2 bins, columns from left to right: $0 < q^2 < 4 \,\text{GeV}^2$, $4 < q^2 < 8 \,\text{GeV}^2$, $12 < q^2 < 16 \,\text{GeV}^2$, $q^2 > 20 \,\text{GeV}^2$. Top row: Discriminator distributions for signal (magneta, dashed) and $q\bar{q}$ background (blue, solid), normalized to the same area. The arrows indicate the chosen cuts. Bottom row: Discriminator distributions for data compared with MC-simulated signal and background contributions. For a legend see Figure 5.

FIG. 8: (color online) The $B \to X_c \ell \nu$ neural-network discriminators for $B^0 \to \pi^- \ell^+ \nu$ candidates in the signal region, $-0.15 < |\Delta E| < 0.25 \,\text{GeV}$; $5.255 < m_{\text{ES}} < 5.295 \,\text{GeV}$. The distributions are shown for four different q^2 bins, columns from left to right: $0 < q^2 < 4 \,\text{GeV}^2$, $4 < q^2 < 8 \,\text{GeV}^2$, $12 < q^2 < 16 \,\text{GeV}^2$, $q^2 > 20 \,\text{GeV}^2$. Top row: Discriminator distributions for signal (magneta, dashed) and $B \to X_c \ell \nu$ background (blue, solid), normalized to the same area. The arrows indicate the chosen cuts. Bottom row: Discriminator distributions for data compared with MC-simulated signal and background contributions. For a legend see Figure 5.

FIG. 9: (color online) The $B \to X_u \ell \nu$ neural-network discriminators for $B^0 \to \pi^- \ell^+ \nu$ candidates in the signal region, $-0.15 < |\Delta E| < 0.25 \, \text{GeV}$; $5.255 < m_{\text{ES}} < 5.295 \, \text{GeV}$. The distributions are shown for the three highest q^2 bins, columns from left to right: $12 < q^2 < 16 \, \text{GeV}^2$, $16 < q^2 < 20 \, \text{GeV}^2$, $q^2 > 20 \, \text{GeV}^2$. Top row: Discriminator distributions for signal (magneta, dashed) and $B \to X_u \ell \nu$ background (blue, solid), normalized to the same area. The arrows indicate the chosen cuts. Bottom row: Discriminator distributions for data compared with MC-simulated signal and background contributions. For a legend see Figure 5.

FIG. 10: (color online) The $q\bar{q}$ neural-network discriminators for $B^0 \to \rho^- \ell^+ \nu$ candidates in the signal region, $-0.15 < |\Delta E| < 0.25 \, \text{GeV}$; $5.255 < m_{\text{ES}} < 5.295 \, \text{GeV}$. The distributions are shown for three different q^2 bins, columns from left to 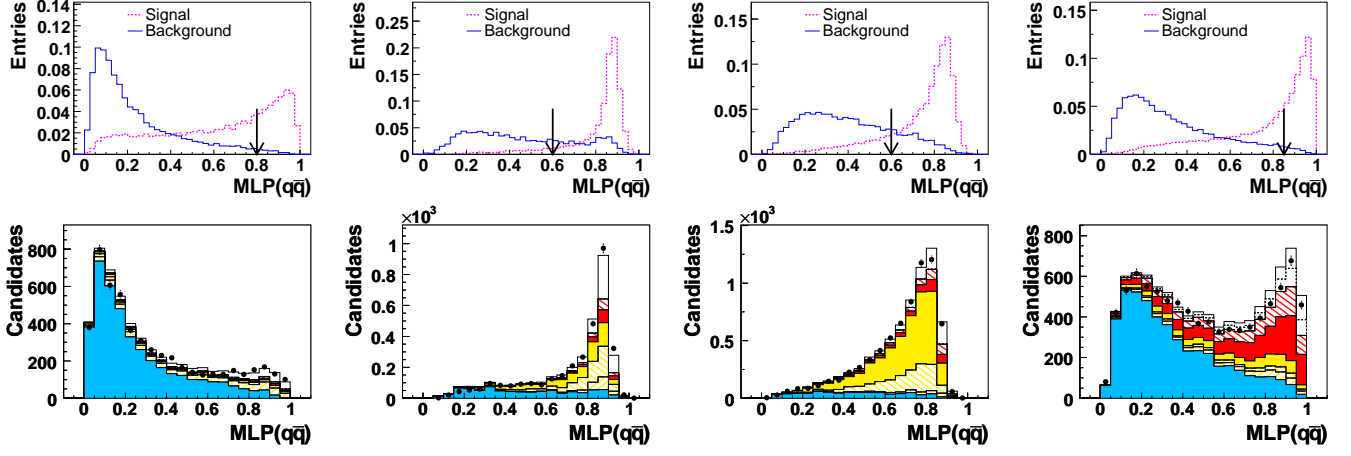right: $0 < q^2 < 8 \, \text{GeV}^2$, $8 < q^2 < 16 \, \text{GeV}^2$, $q^2 > 16 \, \text{GeV}^2$. Top row: Discriminator distributions for signal (magneta, dashed) and $q\bar{q}$ background (blue, solid), normalized to the same area. The arrows indicate the chosen cuts. Bottom row: Discriminator distributions for data compared with MC-simulated signal and background contributions. For a legend see Figure 5.

in the final state. For all four samples, the observed candidate multiplicity is well reproduced by MC simulation.

In case of multiple candidates for a given decay mode, we select the one with the highest probability of the vertex fit for the Y candidate. Since this is not an option for $B^+ \to \pi^0 \ell^+ \nu$ decays, we select the photon pair with an invariant mass closest to the π^0 mass. Simulations of signal events indicate that this procedure selects the correct signal decay in $60 - 65\%$ of the cases. By this selection we do not allow a single event to contribute more than one candidate to a given decay-mode sample, though we do allow an event to contribute candidates to more than one decay-mode sample.

V. DATA-MONTE CARLO COMPARISONS

The determination of the number of signal events relies heavily on the MC simulation to correctly describe the distributions for signal and background sources. Therefore a significant effort has been devoted to detailed comparisons of data and MC distributions for samples that have been selected to enhance a given source of background.

FIG. 11: (color online) The $B \to X_c\ell\nu$ neural-network discriminators for $B^0 \to \rho^-\ell^+\nu$ candidates in the signal region, $-0.15 < |\Delta E| < 0.25\,\text{GeV}$; $5.255 < m_{\text{ES}} < 5.295\,\text{GeV}$. The distributions are shown for three different q^2 bins, columns from left to right: $0 < q^2 < 8\,\text{GeV}^2$, $8 < q^2 < 16\,\text{GeV}^2$, $q^2 > 16\,\text{GeV}^2$. Top row: Discriminator distributions for signal (magneta, dashed) and $B \to X_c\ell\nu$ background (blue, solid), normalized to the same area. The arrows indicate the chosen cuts. Bottom row: Discriminator distributions for data compared with MC-simulated signal and background contributions. For a legend see Figure 5.
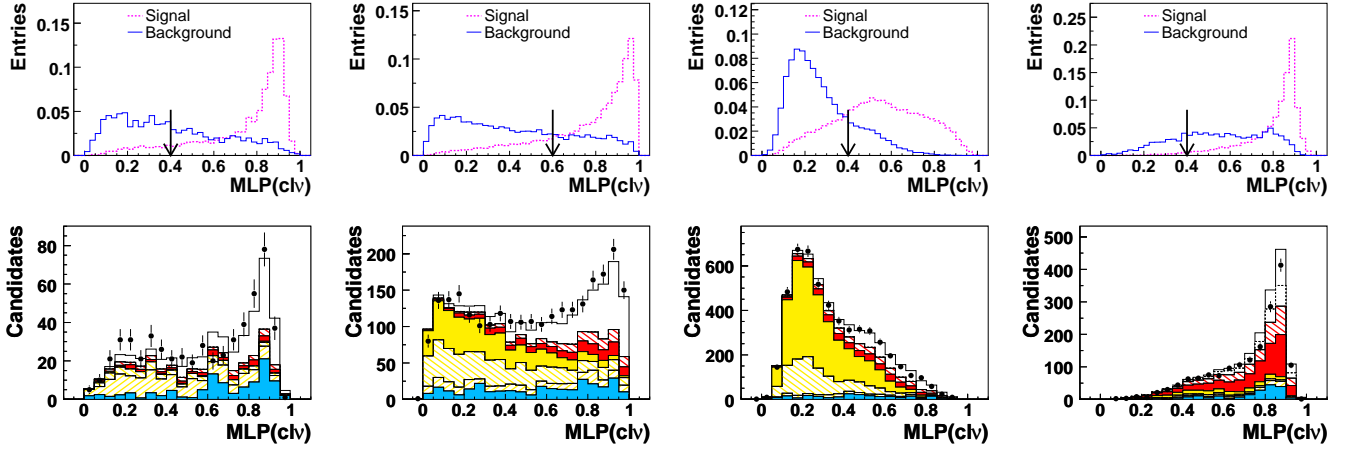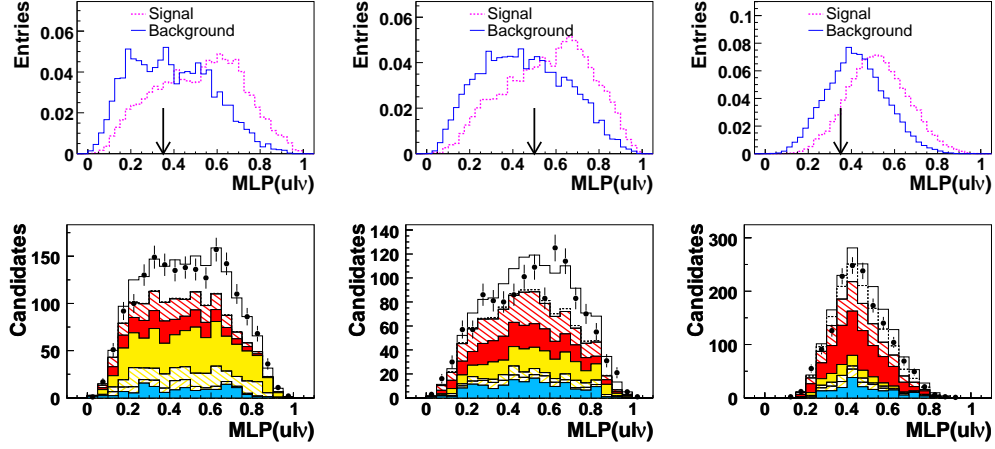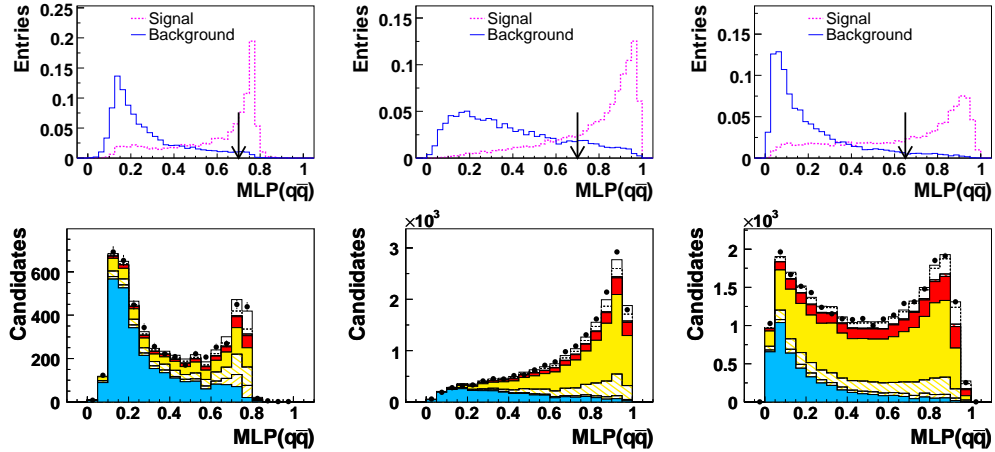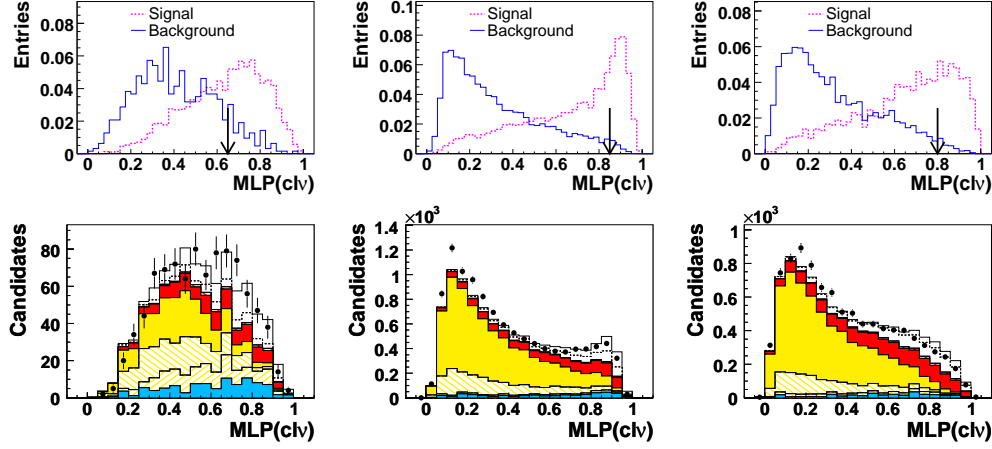
A. Comparison of Off-Resonance Data with $q\overline{q}$ MC

Though we record data below $B\overline{B}$ threshold (off-resonance data) the total luminosity of this sample is only about 10% of the $\Upsilon(4S)$ data sample (on-resonance data), and thus we need to rely on MC simulation to predict the shapes of these background distributions.

To study the simulation of $q\overline{q}$ events, we scale the MC sample to match the integrated luminosity of the off-resonance data. The study is performed separately for samples with electrons and muons. This background contains events with true leptons from leptonic or semileptonic decays of hadrons, as well as hadrons misidentified as leptons. The muon sample is dominated by misidentified hadrons, whereas the electron sample contains small contributions from Dalitz pairs and photon conversions, as well as some residual background from non-$q\overline{q}$ processes. We observe a clear difference in the normalization, not only in the relatively small event sample passing the neural-network selection, but also for the much larger sample available before the neural-network suppression. To correct for this difference, we apply additional scale factors to the simulated $q\overline{q}$ samples; they are different for electrons and muons.

In addition to correcting the normalization, we also examine the shapes of the m_{ES}, ΔE, and q^2 distributions that are used to extract the signal yield. Since the size of the off-resonance data set is small, we study samples with a looser selection, namely we bypass the $q\overline{q}$ neural-network discrimination. The comparison of these $q\overline{q}$-enriched samples reveals small differences between data and simulation. We derive linear corrections from the bin-by-bin ratios and apply these corrections to the m_{ES}, ΔE, and q^2 distributions. Figures 12 and 13 show a comparison of the rescaled and corrected $q\overline{q}$ MC samples with the off-resonance data for the ΔE, m_{ES}, and q^2 distributions. Within the relatively large statistical errors of the off-resonance data the simulation agrees well with the data. The uncertainties in the shape of the simulated distributions will be assessed as a systematic uncertainty.

B. $B \to X_c\ell\nu$ Enhanced Sample

The overall dominant background source in this analysis is $B \to X_c\ell\nu$ decays. Therefore it is important to verify that these decays are correctly simulated. This has been done in two ways, a) by relaxing the $B \to X_c\ell\nu$ suppression to obtain a charm-enhanced sample, and b) by reconstructing a specific decay mode, such as $B^0 \to D^{*-}\ell^+\nu$, in the same way we reconstruct the signal decays, and comparing the kinematic distributions with MC simulations (see Section V C).

We select a charm-enhanced sample by inverting the cut on the $B \to X_c\ell\nu$ neural-network discriminator. Figures 14 and 15 show the ΔE and m_{ES} distributions in the signal region and the side bands, as well as the q^2 distribution in the signal region. All distributions show good agreement in shape; the absolute yields differ at a level that is expected, considering that the MC distributions have not been adjusted.

C. $B^0 \to D^{*-}\ell^+\nu$ Control Sample

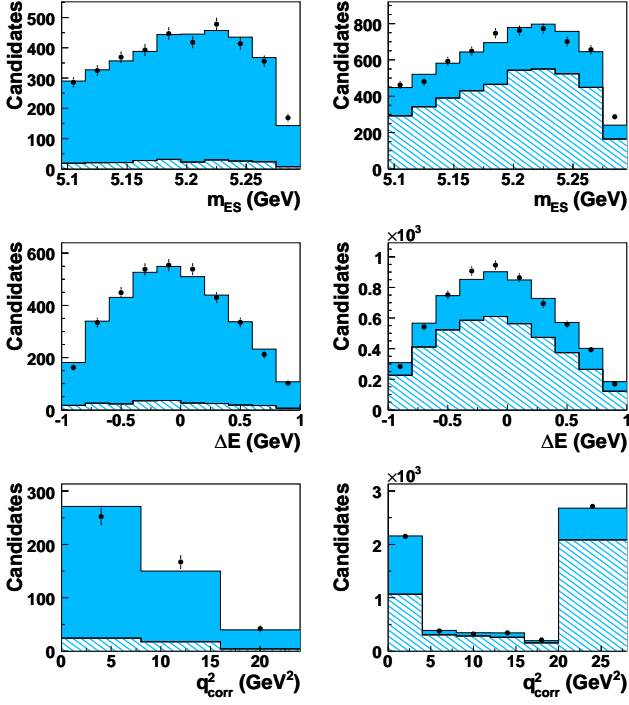To study the Monte Carlo simulation of the neutrino reconstruction employed in this analysis, we use a con-

FIG. 12: (color online) Comparison of off-resonance background to $B^0 \to \pi^- \ell^+ \nu$ samples for data and MC distributions. Top row: m_{ES}, center row: ΔE, and bottom row: q^2, separately for the electron (left column) and muon (right column) samples. The shaded histograms indicate the true leptons, the hatched histograms indicate the fake leptons. The distributions are obtained from the full event selection, except for the $q\bar{q}$ neural-network discrimination. Linear corrections have been applied to the simulation.

FIG. 13: (color online) Comparison of off-resonance background to $B^0 \to \rho^- \ell^+ \nu$ samples for data and MC distributions. Top row: m_{ES}, center row: ΔE, and bottom row: q^2, separately for the electron (left column) and muon (right column) samples. The shaded histograms indicate the true leptons, the hatched histograms indicate the fake leptons. The distributions are obtained from the full event selection, except for the $q\bar{q}$ neural-network discrimination. Linear corrections have been applied to the simulation.

trol sample of exclusively reconstructed $B^0 \to D^{*-}\ell^+\nu$ decays with $D^{*-} \to \overline{D}{}^0\pi_s^-$ and $\overline{D}{}^0 \to K^+\pi^-$. Since the $B^0 \to D^{*-}\ell^+\nu$ decay rate exceeds the rate for $B^+ \to \rho^0\ell^+\nu$ by a factor of about 30 (including the $\overline{D}{}^0$ branching fraction), this control sample represents a high-statistics and high-purity sample of exclusive semileptonic decays. Except for the low-momentum pion (π_s^-), this final state has the same number of tracks, and very similar kinematics, as the $B^+ \to \rho^0\ell^+\nu$ signal decay. Furthermore, since about 50% of the $B \to X_c\ell\nu$ background in the $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ samples comes from $B^0 \to D^{*-}\ell^+\nu$ decays, this $B^0 \to D^{*-}\ell^+\nu$ sample can provide important tests of the shapes of the distributions that are used to discriminate the $B \to X_c\ell\nu$ background from signal.

Moreover, the distributions of the primary background suppression variables, in particular $R2$, $L2$, $\cos\theta_{\mathrm{BY}}$, θ_{miss}, and $m^2_{\mathrm{miss}}/(2E_{\mathrm{miss}})$, are relatively insensitive to the specific semileptonic decay mode. Likewise, the resolutions for the fit variables ΔE, m_{ES}, and q^2 are dominated by the resolution of the reconstructed neutrino, and thus depend little on the decay mode under study.

The reconstruction of the D^{*-} from its decay prod-

ucts is straightforward. Except for the selection of the D^{*-}, we apply the same preselection as for the signal charmless decays. We require the $K^+\pi^-$ invariant mass to be within 17 MeV of the nominal $\overline{D}{}^0$ mass, and restrict the mass difference, $\Delta m_{D^*} = m_{D\pi} - m_D$ to $0.1432 < \Delta m_{D^*} < 0.1478\,\mathrm{GeV}$. The number of events in this data control sample exceeds the MC prediction by $3.8 \pm 1.7\%$, a result consistent with the uncertainties in the efficiency for the very-low-momentum charged pion from the $D^{*-} \to \overline{D}{}^0\pi^-$ decay. We correct the MC yield and sequentially place requirements on the same seven variables we use in the neural networks to both the data and MC samples. We compare the step-by-step reduction in the number of events; the largest difference is $0.9 \pm 0.7\%$, for the cut on $\cos\theta_{W\ell}$. For all other critical requirements the agreement is better than 0.5% and one standard deviation. The remaining background is at the level of 10%.
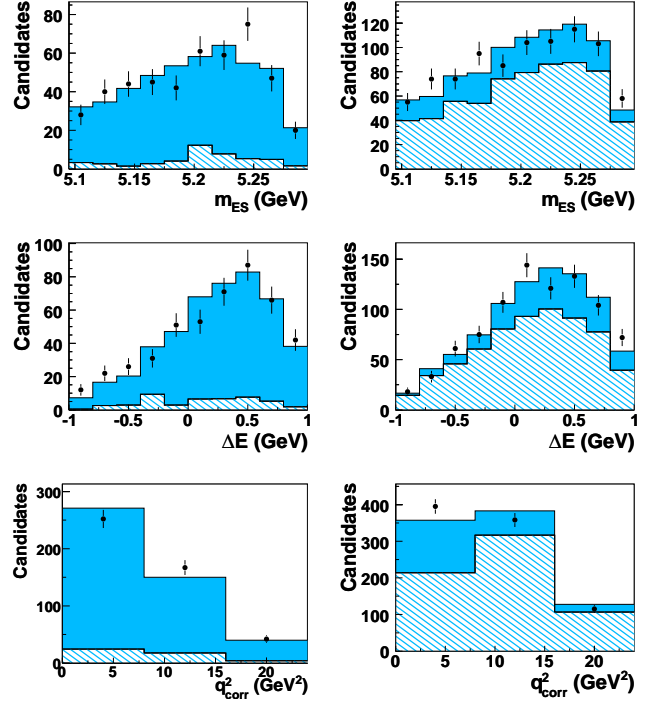
We have compared the MC-generated distributions for the control sample with the selected $B^0 \to D^{*-}\ell^+\nu$ data sample and find very good agreement for the basic event variables, i.e., the multiplicity of charged particles and photons, and the total charge per event, indicating that
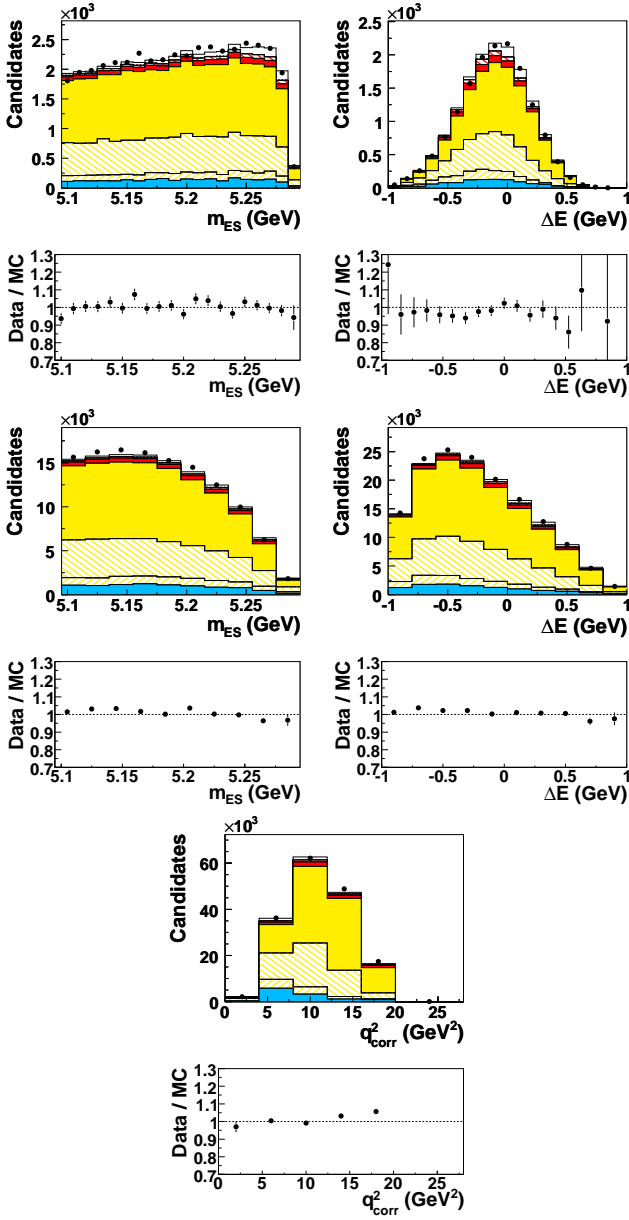
FIG. 14: (color online) Comparison of data with MC distributions for the charm-enhanced selection for the $B^0 \to \pi^- \ell^+ \nu$ sample. Top row: m_{ES} and ΔE for the signal bands, center row: m_{ES} and ΔE for the side bands, and bottom row: q^2_{corr} for the whole fit region. The bin-by-bin ratio of data over the sum of all MC contributions is given in the plots below each histogram.

FIG. 15: (color online) Comparison of data with MC distributions for the charm-enhanced selection for the $B^0 \to \rho^- \ell^+ \nu$ sample. Top row: m_{ES} and ΔE for the signal bands, center row: m_{ES} and ΔE for the side bands, and bottom row: q^2_{corr} for the whole fit region. The bin-by-bin ratios of data over the sum of all MC contributions is given in the plots below each histogram.

the efficiency losses are well reproduced by the simulation. The distributions of the topological event variables $R2$ and $L2$ match well. Figure 16 shows the distributions of the variables critical for the neutrino reconstruction, $p_{miss}, m^2_{miss}/(2E_{miss})$, $\cos\theta_{BY}$, and θ_{miss}; they are also well reproduced.

Figure 17 shows distributions of ΔE and m_{ES} for events in the signal region and in the side bands. Again, the agreement between data and the MC simulation is

reasonable.

We have also compared the q^2 distributions of the simulation and the data control sample and find good agreement for both the raw and the corrected spectra, as illustrated in Figure 18. After corrections, no events appear above the kinematic limit of $10.7\,\mathrm{GeV}^2$. The q^2_{corr} resolution function can be described by the sum of two Gaussian resolution functions, with widths of $0.27\,\mathrm{GeV}^2$ and $0.67\,\mathrm{GeV}^2$, close to the values obtained for events in the
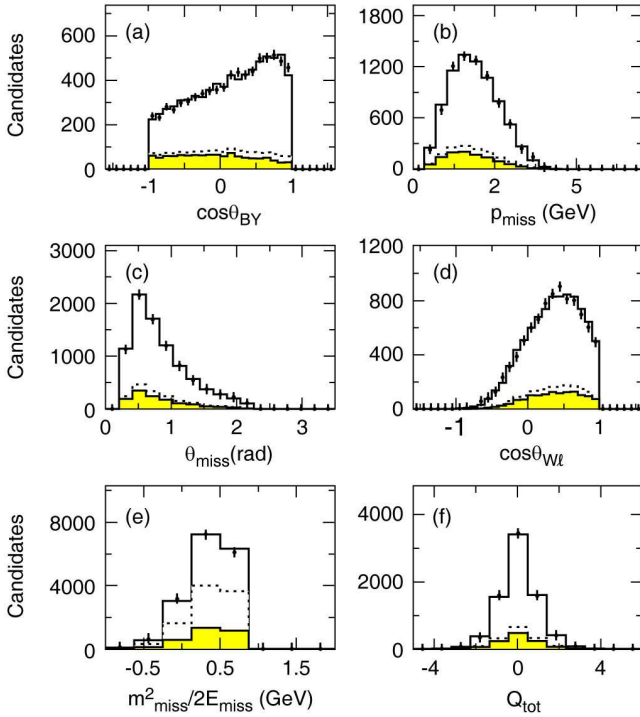
FIG. 16: (color online) Comparison of data and MC-simulated distributions for the $B^0 \to D^{*-}\ell^+\nu$, $\bar{D}^0 \to K^+\pi^-$ sample, after selection criteria have been applied on all variables except the one presented. a) $\cos\theta_{\mathrm{BY}}$, b) p_{miss}, c) θ_{miss}, d) $\cos\theta_{W\ell}$, e) $m_{\mathrm{miss}}^2/2E_{\mathrm{miss}}$, and f) the total charge per event Q_{tot}. The background to the sample is indicated as a shaded (yellow) histogram. The combinatorial signal contribution is indicated as dashed histogram.

fit region for the signal $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ decays, respectively.

VI. MAXIMUM-LIKELIHOOD FIT

A. Overview

We determine the yields for the signal decay modes, $B^0 \to \pi^-\ell^+\nu$, $B^+ \to \pi^0\ell^+\nu$, $B^0 \to \rho^-\ell^+\nu$, and $B^+ \to \rho^0\ell^+\nu$, by performing a maximum-likelihood fit to the three-dimensional $\Delta E - m_{\mathrm{ES}} - q^2$ distributions for the four selected data samples corresponding to the four exclusive decay modes. The fit technique employed in this analysis is an extended binned maximum-likelihood fit that accounts for the statistical fluctuations not only of the data samples but also of the MC samples by allowing the MC-simulated distributions to fluctuate in each bin according to the statistical uncertainty given by the number of events in the bin. This method was introduced by Barlow and Beeston [49].

The parameters of the fit are scale factors for the signal and background yields of the four selected event samples. We use the following nomenclature for the fit parame-
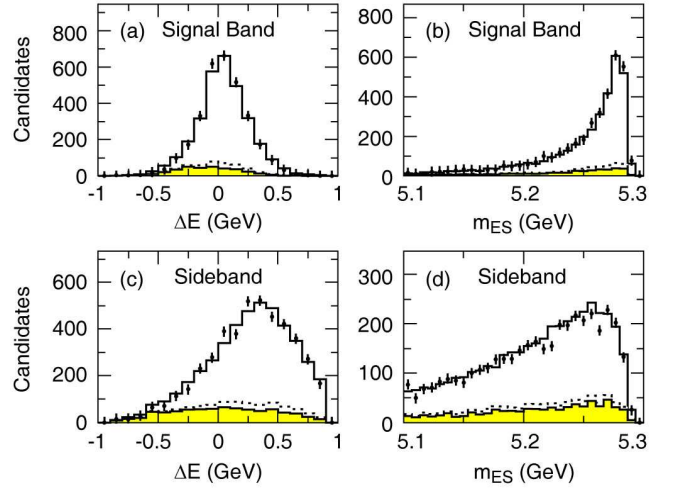
FIG. 17: (color online) Comparison of data and MC-simulated distributions for the $B^0 \to D^{*-}\ell^+\nu$, $\bar{D}^0 \to K^+\pi^-$ sample, after all selection cuts have been applied, a) ΔE for events in the m_{ES} signal band, b) m_{ES} for events in the ΔE signal band, c) ΔE for events in the m_{ES} side band, and d) m_{ES} for events in the ΔE side bands. The background to the sample is indicated as a shaded (yellow) histogram. The combinatorial signal contribution is indicated as dashed histogram.
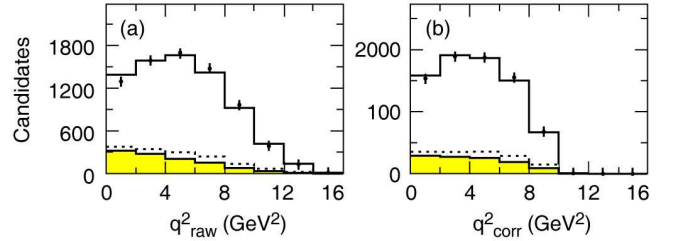
FIG. 18: (color online) Comparison of the data and MC simulation of q^2 distributions for the $B^0 \to D^{*-}\ell^+\nu$, $\bar{D}^0 \to K^+\pi^-$ sample after all selection criteria have been applied, a) the raw q^2, and b) the corrected q^2. The background to the sample is indicated as a shaded (yellow) histogram. The combinatorial signal contribution is indicated as dashed histogram.

ters: p_j^{source}, where the superscript *source* denotes the fit source (signal or background type) and the subscript j labels the q^2_{corr} bin (if no subscript is given, the same fit parameter is used across all q^2 bins). Predictions for the shape of the $\Delta E - m_{\mathrm{ES}}$ distributions are taken from simulation of both signal and the various background sources, separately for each bin in q^2. The branching fractions for the four signal decays are obtained by multiplying the fitted values of the scale factors with the branching fractions that are implemented in the MC simulation.

The choice of a two-dimensional distribution in ΔE and m_{ES} is mandated because the two variables are correlated for both signal, in particular the combinatorial signal events, and for some of the background sources. Since it would be difficult to determine reliable analytic expressions for these two-dimensional distributions, a binned

maximum-likelihood method is used, with the bin sizes chosen to obtain a good signal and background separation while retaining adequate statistics in all bins. The bin sizes are small in the region where most of the signal is located and larger in the side bands. There are 47 $\Delta E - m_{\mathrm{ES}}$ bins for each bin in q^2_{corr}. Figure 19 shows the $\Delta E - m_{\mathrm{ES}}$ distribution for signal events and the binning used in the fit. As mentioned in Section IV B 2, for the two $B \to \pi \ell \nu$ samples the q^2 range $0 < q^2_{\mathrm{corr}} < 26.4 \, \mathrm{GeV}^2$ is divided into six bins, and for the two $B \to \rho \ell \nu$ samples the range $0 < q^2_{\mathrm{corr}} < 20.3 \, \mathrm{GeV}^2$ is divided into three bins.
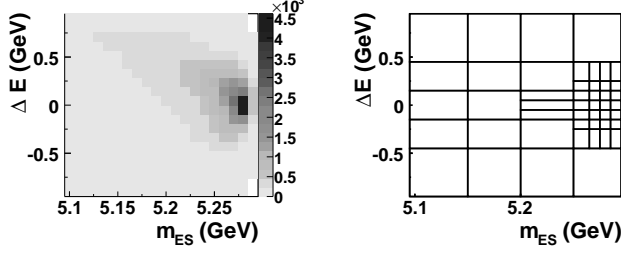
FIG. 19: Distribution of ΔE vs m_{ES} for true $B^0 \to \pi^- \ell^+ \nu$ signal events (left) and definition of bins in the $\Delta E - m_{\mathrm{ES}}$ plane used in the fit for all samples (right).

B. Fit Method

Since the MC samples available to create the probability density functions (PDFs) for the individual sources that are input to the fits are rather limited in size, it is necessary to take into account the statistical uncertainties, given by the number of events generated for each bin. For this reason we have adopted a generalized binned maximum-likelihood fit method. The MC samples that are used to define the PDFs are to a good approximation statistically independent of those used to train the neural networks for background suppression, since for the latter relatively small subsamples of the full MC samples have been used.

As mentioned above, the data are divided into n bins in a three-dimensional array in $\Delta E - m_{\mathrm{ES}} - q^2_{\mathrm{corr}}$.

If d_i is the number of selected events in bin i for a given single data sample corresponding to candidates for a specific decay mode, and a_{ji} is the number of MC events from source j in this bin, then

$$N_D = \sum_{i=1}^{n} d_i, \qquad N_j = \sum_{i=1}^{n} a_{ji}, \qquad (24)$$

where N_D is the total number of events in the data sample, and N_j is the total number in the MC sample for source j. We assume that there are m different MC-generated source distributions that add up to describe the data. The predicted number of events in each bin

$f_i(P_j)$ can be written in terms of the strength of the individual contributions P_j $(j = 1, .., m)$ as

$$f_i = N_D \sum_{j=1}^{m} P_j w_{ji} a_{ji} / N_j = \sum_{j=1}^{m} p_j w_{ji} a_{ji}, \qquad (25)$$

with $p_j = N_D P_j / N_j$. In each bin, the weights w_{ji} account for the relative normalization of the samples and various other corrections.

Since the MC samples are limited in size, the generated numbers of events a_{ji} have statistical fluctuations relative to the value A_{ji} expected for infinite statistics, and thus the more correct prediction for each bin is

$$f_i = \sum_{j=1}^{m} p_j w_{ji} A_{ji}. \qquad (26)$$

If we assume Poisson statistics for both the data and MC samples, the total likelihood function $\mathcal{L}$ is the combined probability for the observed d_i and a_{ji} [49],

$$\ln \mathcal{L} = \sum_{i=1}^{n} (d_i \ln f_i - f_i) + \sum_{i=1}^{n} \sum_{j=1}^{m} (a_{ji} \ln A_{ji} - A_{ji}). \quad (27)$$

The first sum has the usual form associated with the uncertainty of the data and the second term refers to the MC statistics and is not dependent on data. There are $(n + 1) \times m$ unknown parameters that need to be determined: the m relative normalization factors p_j, which are of interest to the signal extraction, and $n \times m$ values A_{ji}.

The problem can be significantly simplified. The $n \times m$ quantities A_{ji} can be determined by solving n simultaneous equations for A_{ji} of the form

$$f_i = \sum_{j=1}^{m} p_j w_{ji} A_{ji} = \sum_{j=1}^{m} \frac{p_j w_{ji} a_{ji}}{1 + p_j w_{ji} t_i}, \qquad (28)$$

with $A_{ji} = a_{ji}/(1 + p_j w_{ji} t_i)$ and $t_i = 1 - d_i / f_i$ (for $d_i = 0$ we define $t_i = 1$). At every step in the minimization of $-2\ln\mathcal{L}$ these n independent equations need to be solved. This procedure results not only in the determination of the parameters p_j, but also in improved estimates for the various contributions A_{ji} in each bin.

For fits to the individual data samples corresponding to the four signal decay modes, there is a specific likelihood function (Eq. 27). To perform a simultaneous fit to all four data samples the log-likelihood function is the sum of the individual ones. Some of the parameters p_j may be shared among the four likelihood functions,

$$\ln \mathcal{L} = \sum_{h=1}^{4} \ln \mathcal{L}_h = \sum_{h=1}^{4} \sum_{i=1}^{n} (d_{hi} \ln f_{hi} - f_{hi}) \qquad (29)$$

$$+ \sum_{h=1}^{4} \sum_{i=1}^{n} \sum_{j=1}^{m} (a_{hji} \ln A_{hji} - A_{hji}).$$

C. Fit Parameters and Inputs

The fits can be performed separately for each of the four data samples or combined for all four data samples, and where possible, with common fit parameters. The nominal fit in this analysis is a simultaneous fit of all four data samples: $B^0 \to \pi^- \ell^+ \nu$, $B^+ \to \pi^0 \ell^+ \nu$, $B^0 \to \rho^- \ell^+ \nu$, and $B^+ \to \rho^0 \ell^+ \nu$. A signal decay in one data sample may contribute to the background in another sample, and therefore these sources share a common fit parameter. For example, the scale factor for the $B^0 \to \rho^- \ell^+ \nu$ signal in the $B^0 \to \rho^- \ell^+ \nu$ sample is also applied in the $B^0 \to \pi^- \ell^+ \nu$ sample, where it represents cross-feed background.

We impose isospin invariance for the signal decay modes,

$$\Gamma(B^0 \to \pi^- \ell^+ \nu) = 2\Gamma(B^+ \to \pi^0 \ell^+ \nu),$$
$$\Gamma(B^0 \to \rho^- \ell^+ \nu) = 2\Gamma(B^+ \to \rho^0 \ell^+ \nu) . \quad (30)$$

The yields of the true and combinatorial signal decays as well as isospin-conjugate decays are related to the same branching fraction and therefore share the same fit parameter. The $B \to X_u \ell \nu$ background, which contains exclusive and non-resonant decays, is scaled by two parameters, one for low and intermediate q^2 ($q^2 < 20\,\mathrm{GeV}^2$) and one for high q^2 ($q^2 > 20\,\mathrm{GeV}^2$), for the fits to the $B \to \pi \ell \nu$ samples. Because of the large correlation between $B \to X_u \ell \nu$ background and $B \to \rho \ell \nu$ signal ($> 90\%$ for both $B \to \rho \ell \nu$ modes), we rely on MC simulation for the $B \to X_u \ell \nu$ background and keep it fixed in the fits to the $B \to \rho \ell \nu$ samples. The $B\overline{B}$ background is split into two sources. Among the $B \to X_c \ell \nu$ decays, we treat the dominant decay mode, $B \to D^* \ell \nu$, as a separate source and combine the other semileptonic decays ($B \to D\ell\nu$, $B \to D^{(*)}(n\pi)\ell\nu$) and the remaining (or "other") $B\overline{B}$ background (secondary leptons and fake leptons) into a single source. The continuum $q\overline{q}$ background sources containing true and fake leptons are combined into one fit source and scaled by a single fit parameter.

The complete list of fit sources and corresponding fit parameters is given in Table VI. The $\pi \leftrightarrow \rho$ cross feed is a free fit parameter in the four-mode fit; for one-mode fits, it is fixed to the value obtained from the four-mode fit. In the four-mode fit, all background sources that are not fixed are fit separately for each signal mode, since the different hadrons of the signal decays lead to different combinatorial backgrounds.

D. Fit Results

The fits are performed both separately and simultaneously for the four signal decay modes, $B^0 \to \pi^- \ell^+ \nu$, $B^+ \to \pi^0 \ell^+ \nu$, $B^0 \to \rho^- \ell^+ \nu$, and $B^+ \to \rho^0 \ell^+ \nu$. Figures 20–23 show projections of the fitted $\Delta E - m_{\mathrm{ES}}$ distributions in the signal bands for these decays, separately

TABLE VI: List of fit parameters representing scale factors for the different signal samples and background sources. Parameters with index j are free parameters in the fit, one for each q^2 bin j. The $\pi \leftrightarrow \rho$ crossfeed parameter is free only in the four-mode fit; for one-mode fits, it is fixed to the values obtained from the four-mode fit. There are independent scale factors for $q\overline{q}$ background, $B \to D^* \ell \nu$ decays and for all other background sources from $B\overline{B}$ events for all four signal modes (subscripts $\pi^\pm, \pi^0, \rho^\pm, \rho^0$). For the $B \to \pi \ell \nu$ decays, the $B \to X_u \ell \nu$ background is fit in two q^2 intervals (index $k = 1, 2$); for the $B \to \rho \ell \nu$ decays it is fixed.

Source / sample	$B \to \pi\ell\nu$	$B \to \rho\ell\nu$
Signal	$p_j^{\pi\ell\nu}$	$p_j^{\rho\ell\nu}$
Combinatorial signal	$p_j^{\pi\ell\nu}$	$p_j^{\rho\ell\nu}$
Isospin-conjugate signal	$p_j^{\pi\ell\nu}$	$p_j^{\rho\ell\nu}$
Cross feed $\pi \leftrightarrow \rho$	$p_j^{\rho\ell\nu}$	$p_j^{\pi\ell\nu}$
$B \to X_u\ell\nu$ background	$p_{\pi^\pm,k}^{u\ell\nu}, p_{\pi^0,k}^{u\ell\nu}$	fixed
$B \to D^*\ell\nu$ background	$p_{\pi^\pm}^{D^*\ell\nu}, p_{\pi^0}^{D^*\ell\nu}$	$p_{\rho^\pm}^{D^*\ell\nu}, p_{\rho^0}^{D^*\ell\nu}$
Other $B\overline{B}$ background	$p_{\pi^\pm}^{otherB\overline{B}}, p_{\pi^0}^{otherB\overline{B}}$	$p_{\rho^\pm}^{otherB\overline{B}}, p_{\rho^0}^{otherB\overline{B}}$
$q\overline{q}$ background	$p_{\pi^\pm}^{q\overline{q}}, p_{\pi^0}^{q\overline{q}}$	$p_{\rho^\pm}^{q\overline{q}}, p_{\rho^0}^{q\overline{q}}$

for each bin in q_{corr}^2. As a measure of the goodness-of-fit we use χ^2 per degree of freedom; all fits have values in the range $1.05 - 1.11$ (for details see Table VII).

The scale factors for the signal contributions, which are determined by the fits, can be translated to numbers of background-subtracted signal events for the four signal decays. These signal yields are listed in Table VII with errors that are a combination of the statistical uncertainties of the data and MC samples and the uncertainties of the fitted yields of the various backgrounds. For each signal decay mode, the table specifies the number of true and combinatorial signal decays. Their relative fraction is taken from simulation. This fraction is larger for decays with a π^0 in the final state. For all signal modes, the fraction of combinatorial signal events is small at low q^2, increases with q^2, and at the highest q^2 it is similar to or exceeds the one of true signal decays. This leads to larger errors in the measurement of q^2, m_{ES} and ΔE.

In Table XVII in Appendix XI B the correlation matrix of the four-mode fit is presented. We observe correlations of about $40 - 60\%$ between the $q\overline{q}$ and the other $B\overline{B}$ backgrounds and between the $B \to D^* \ell \nu$ and the other $B\overline{B}$ backgrounds for all signal modes. For $B \to \pi\ell\nu$, the correlation between the $B \to X_u\ell\nu$ background and the signal at high q^2 is also sizable ($\simeq 60\%$). For $B \to \rho\ell\nu$, this correlation is larger than 90%, which is why we choose to fix the $B \to X_u\ell\nu$ background normalization for these two samples. As a test, we let the $B \to X_u\ell\nu$ background normalization in the $B \to \rho\ell\nu$ modes vary as free parameter in the four-mode fit. This results in a $B \to X_u\ell\nu$ contribution that is lower by a factor of 0.85 ± 0.15 for $B^0 \to \rho^- \ell^+ \nu$ and 0.90 ± 0.14 for $B^+ \to \rho^0 \ell^+ \nu$ and an increase of the $B \to \rho\ell\nu$ signal yields by 10% in the first two q^2 bins and by 15% in the last q^2 bin. These changes are covered by the systematic uncertainties due

to the $B \to X_u \ell \nu$ background stated in Section VII.

To cross-check the results of the nominal four-mode fit, we also perform fits for each signal mode separately. The contributions from the other signal decay modes are fixed to the result obtained from the four-mode fit. Since the shape of the $\pi \leftrightarrow \rho$ cross-feed contribution is very similar to the other $B \to X_u \ell \nu$ background, we fix its normalization to the one obtained from the four-mode fit. A comparison of the results of the one-mode fits with the combined four-mode fit shows agreement within the fit errors of the $B^0 \to \pi^- \ell^+ \nu$ and $B^+ \to \pi^0 \ell^+ \nu$ modes and the $B^0 \to \rho^- \ell^+ \nu$ and $B^+ \to \rho^0 \ell^+ \nu$ modes in all q^2 bins.

The partial branching fractions for the different q^2_{corr} bins are derived as the products of the fitted signal scale factors and the signal decay branching fractions used in the simulation. The total branching fraction integrated over the entire q^2 range and its error are calculated as the sum of all partial branching fractions, taking into account the correlations of the fitted yields in different q^2 bins. The branching fraction for B^0 decays, $\mathcal{B}^0_{\mathrm{signal}}$, is related to the fitted signal yields, N^0_{signal}, in the following way,

$$ \mathcal{B}^0_{\mathrm{signal}} = \frac{N^0_{\mathrm{signal}}}{4 \times \epsilon^0_{\mathrm{signal}} f_{00} N_{B\overline{B}}} \ , \tag{31} $$

where $f_{00} = 0.484 \pm 0.006$ [41] denotes the fraction of $B^0 \overline{B}^0$ events produced in $\Upsilon(4S)$ decays and $\epsilon^0_{\mathrm{signal}}$ is the total signal efficiency (averaged over the electron and muon samples) as predicted by the MC simulation. The factor of four accounts for the fact that each event contains two B mesons, and that the branching fraction is quoted for a single charged lepton, not for the sum of the decays to electrons or muons. The branching fraction results are presented in Section VIII.

E. Fit Validation and Consistency

The fit procedure is validated several ways. First of all, the implementation of the Barlow-Beeston fit technique allowing statistical fluctuations of the MC distributions to be incorporated is checked by verifying the consistency of the fit variations with the statistical error of the input distributions. Secondly, a large number of simulated experiments are generated based on random samples drawn from the three-dimensional histograms used in the standard fit. Specifically, we create 500 sets of distributions by fluctuating each simulated source distribution bin-by-bin using Poisson statistics. For each of the sets, we add the source distributions to make up to total distribution that corresponds to the data distribution ("toy data"), which are then fitted by the standard procedure. In addition, we create independent fluctuations for the distributions that make up the source PDFs for the fit, in the same way as for the toy data described above. For a compilation of these 500 "toy experiments", we study

the distributions of the deviation of the fit result from the input value divided by the fit error. These distributions show no significant bias for any of the free parameters and confirm that the errors are correctly estimated.

Additional fits are performed to check the consistency of the data. For instance, the data samples are divided into subsamples, $i.e.$, the electron sample separated from the muon sample or the data separated into different run periods. These subsamples are fitted separately; the results agree within the statistical uncertainties.

VII. SYSTEMATIC UNCERTAINTIES

Many sources of systematic uncertainties have been assessed for the measurement of the exclusive branching fractions as a function of q^2. Since this analysis does not depend only on the reconstruction of the charged lepton and hadron from the signal decay mode, but also on the measurement of all remaining tracks and photons in the event, the uncertainties in the detection efficiencies of all particles as well as the uncertainties in the background yields and shapes enter into the systematic errors.

Tables VIII and IX summarize the systematic uncertainties for $B \to \pi \ell \nu$ and $B \to \rho \ell \nu$ for the four-mode fit. In Appendix XI A the systematic error tables for the one-mode fits are presented. The individual sources are, to a good approximation, uncorrelated and can therefore be added in quadrature to obtain the total systematic errors for each decay mode. In the following, we discuss the assessment of the systematic uncertainties in detail.

For the estimation of the systematic errors of the fitted branching fractions, we compare the differential branching fractions obtained from the nominal fit with results obtained after changes to the MC simulation that reflect the uncertainty in the parameters that impact the detector efficiency and resolution or the simulation of signal and background processes. For instance, we vary the tracking efficiency, reprocess the MC samples, reapply the fit to the data, and take the difference compared to the results obtained with the nominal MC simulation as an estimate of the systematic error. The sources of systematic errors are not identical for all four signal decay modes, and the size of their impact on the event yields depends on the sample composition and q^2.

A. Detector Effects

Uncertainties in the reconstruction efficiencies for charged and neutral particles and in the rate of tracks and photons from beam background, fake tracks, failures in the matching of EMC clusters to charged tracks, and showers split off from hadronic interactions, undetected K_L, and additional neutrinos, all contribute to the quality of the neutrino reconstruction and impact the variables that are used in the preselection and the neural
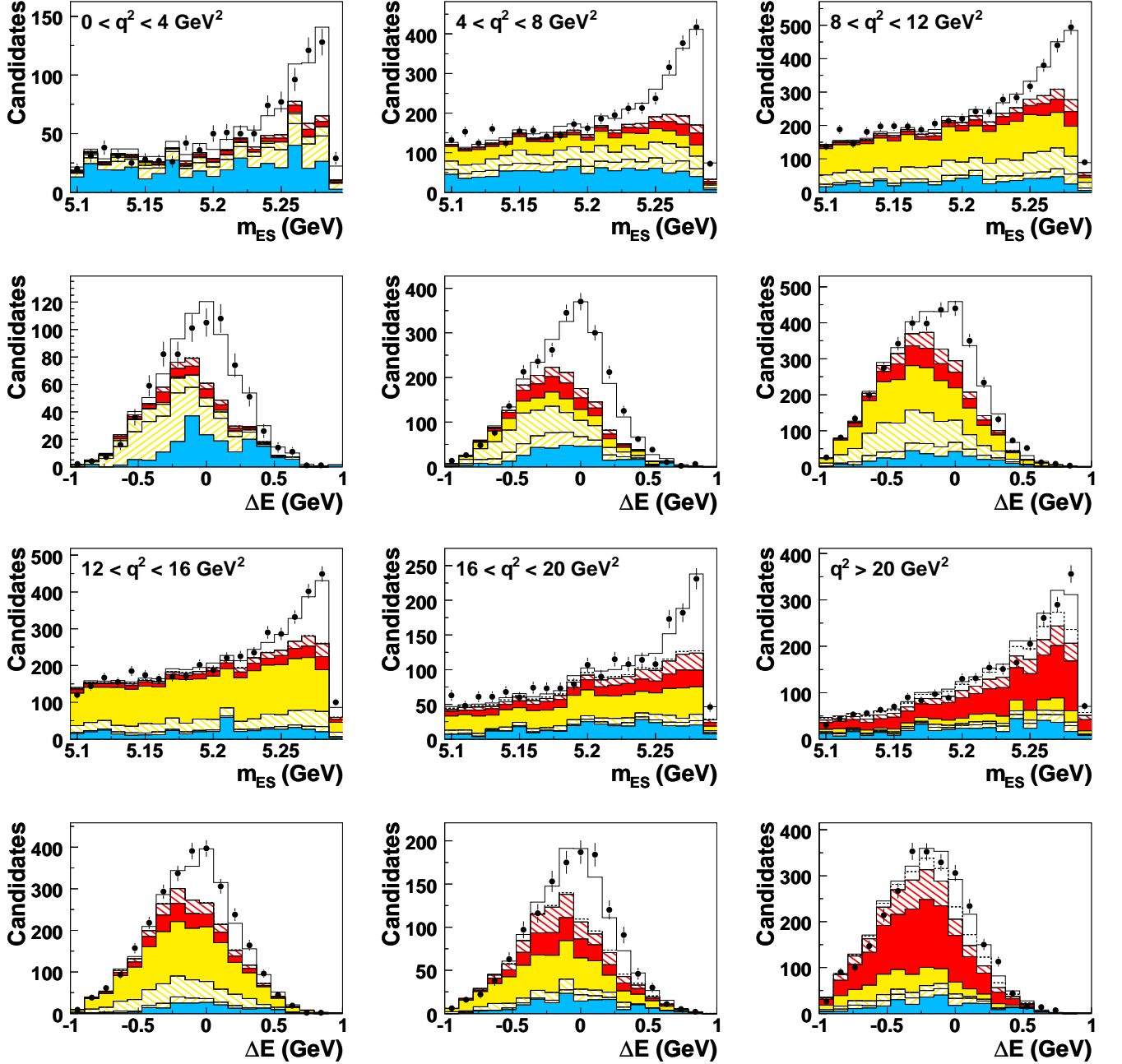
FIG. 20: (color online) m_{ES} and ΔE distributions in each q^2 bin for $B^0 \to \pi^- \ell^+ \nu$ after the fit. The distributions are shown in the ΔE and m_{ES} signal bands, respectively. Legend: see Figure 5.

networks. For all these effects the uncertainties in the efficiencies and resolution have been derived independently from comparisons of data and MC simulation for selected control samples.

1. Track, Photon, and Neutral-Pion Reconstruction

We evaluate the impact of uncertainties in the tracking efficiency by randomly eliminating tracks with a probability that is given by the uncertainty ranging from 0.25% to 0.5% per track, as measured with data control samples.

Similarly, we evaluate the uncertainty due to photon efficiency by eliminating photons at random with an energy-dependent probability, ranging from 0.7% per photon above 1 GeV to 1.8% at lower energies. This estimate includes the uncertainty in the π^0 efficiency for signal decays with a π^0, since photons originating from the signal hadron are also eliminated.
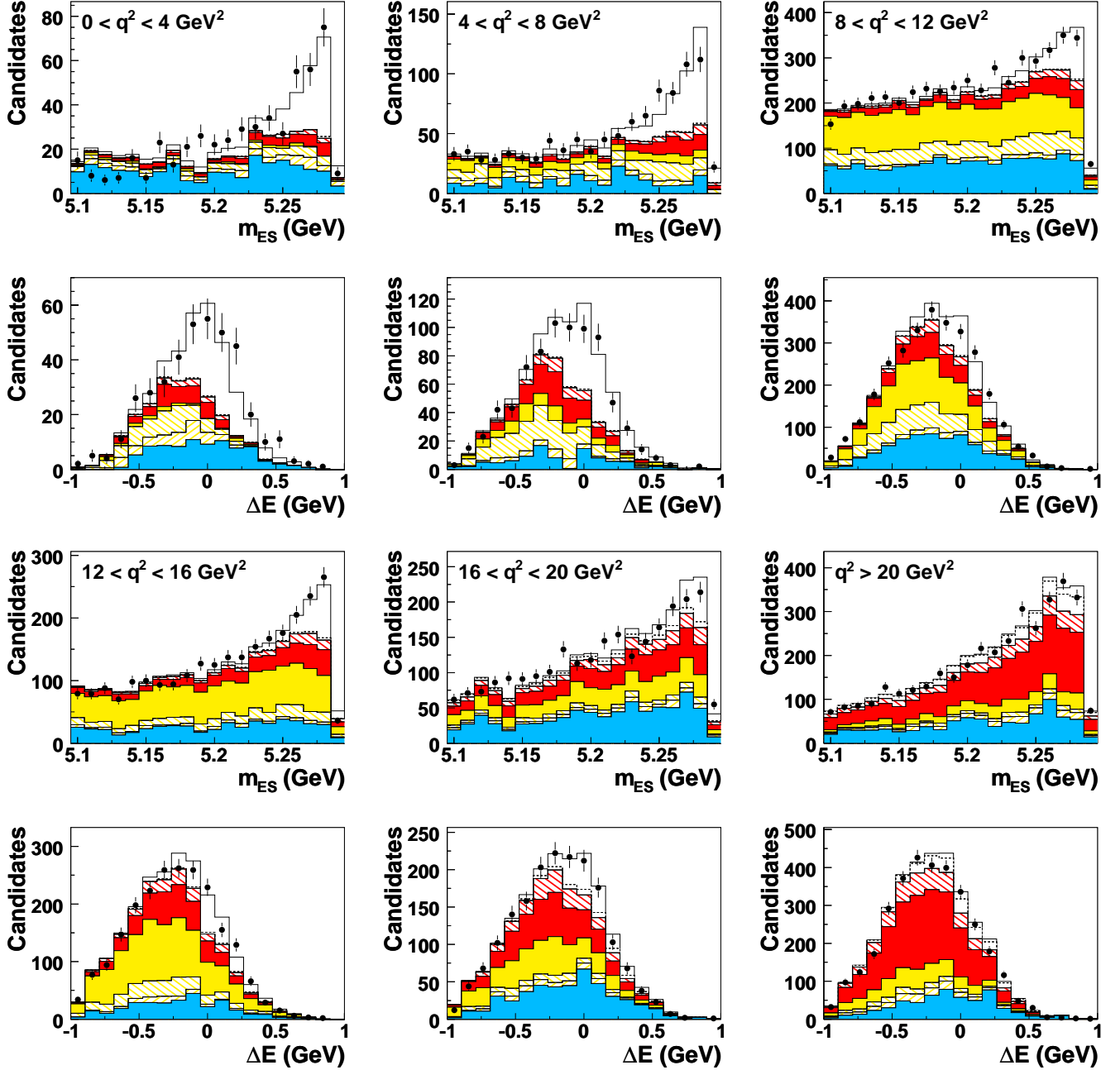
FIG. 21: (color online) m_{ES} and ΔE distributions in each q^2 bin for $B^+ \to \pi^0 \ell^+ \nu$ after the fit. The distributions are shown in the ΔE and m_{ES} signal bands, respectively. Legend: see Figure 5.

2. Lepton Identification

The average uncertainties in the identification of electrons and muons have been assessed to be 1.4% and 3%, respectively. The uncertainty in the misidentification of hadrons as electrons or muons is about 15%.

3. K_L^0 Production and Interactions

Events containing a K_L^0 have a significant impact on the neutrino reconstruction, because only a small fraction of the K_L^0 energy is deposited in the electromagnetic calorimeter. Based on detailed studies of data control samples of $D^0 \to K^0 \pi^+ \pi^-$ decays and inclusive K_S^0 samples in data and MC, corrections to the efficiency, shower deposition and the production rates have been derived and applied to the simulation as a function of the K_L^0 mo-
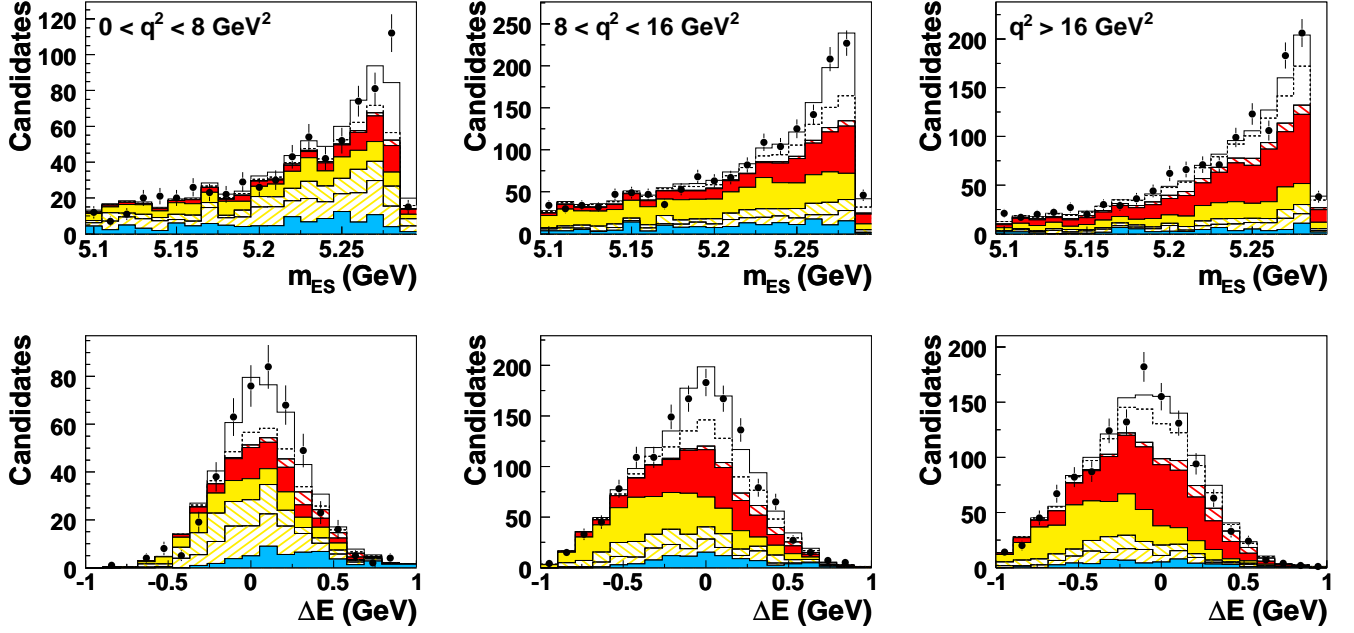
FIG. 22: (color online) $m_{\rm ES}$ and ΔE distributions in each q^2 bin for $B^0 \to \rho^- \ell^+ \nu$ after the fit. The distributions are shown in the ΔE and $m_{\rm ES}$ signal bands, respectively. Legend: see Figure 5.
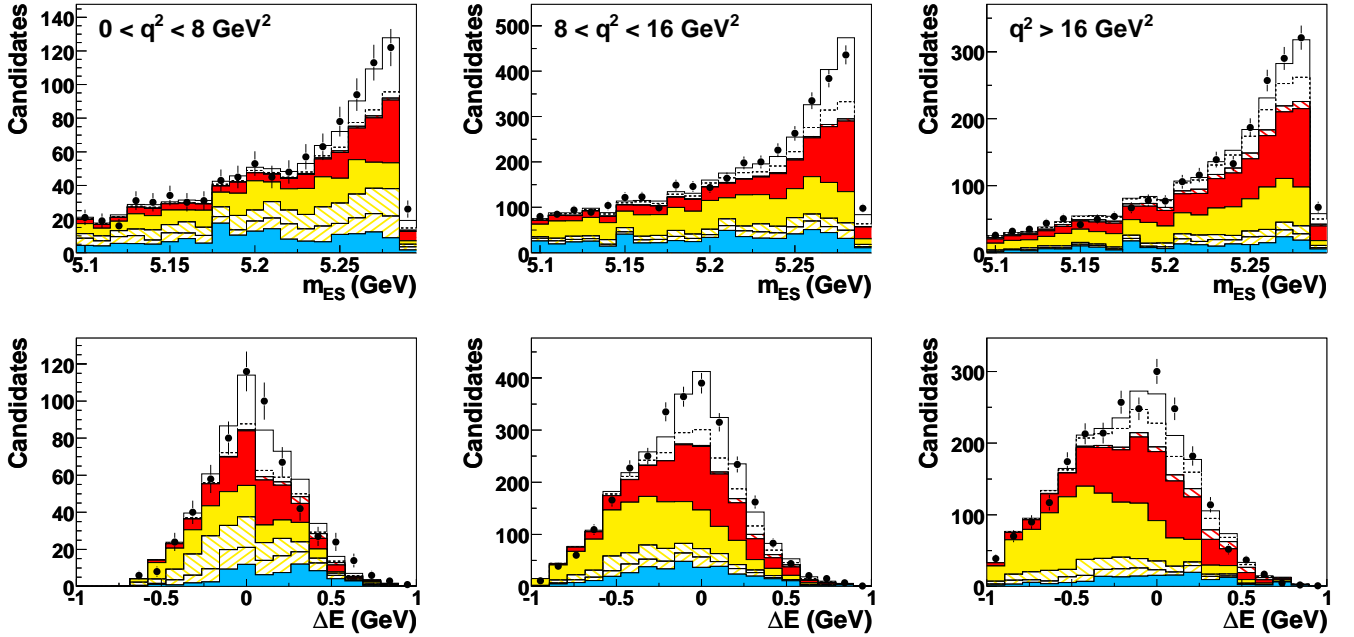
FIG. 23: (color online) $m_{\rm ES}$ and ΔE distributions in each q^2 bin for $B^+ \to \rho^0 \ell^+ \nu$ after the fit. The distributions are shown in the ΔE and $m_{\rm ES}$ signal bands, respectively. Legend: see Figure 5.

mentum and angles (see Section III). To determine the systematic uncertainties in the MC simulations we vary the scale factors within their statistical and systematic uncertainties. The average uncertainty of the energy deposition in the EMC due to K_L^0 interactions is estimated to be 7.5%. Above 0.7 GeV, the K_L^0 detection efficiency is well reproduced by the simulation, with an estimated average uncertainty of 2%. At lower momenta, the simulation is corrected to match the data, and the uncertainty increases to 25% below 0.4 GeV.

The production rates for K_S^0 in data and MC agree within errors, except for momenta below 0.4 GeV where

TABLE VII: Results of fits performed separately for each of the four signal decay modes and simultaneously for all four decay modes in bins of q^2: χ^2 per degree of freedom, signal yields for true signal decays, N^{sig}, and combinatorial signal, N^{comb}. The stated errors are the fit errors.

q^2 range (GeV^2)	χ^2/ndf	0-4	4-8	8-12	12-16	16-20	>20	$0 < q^2 < 26.4$
$N^{sig}_{\pi\pm}$	259/268	701 ± 64	1950 ± 104	1552 ± 113	1184 ± 104	732 ± 80	541 ± 68	$6660 \pm 2\,78$
$N^{comb}_{\pi\pm}$		1 ± 0.1	4 ± 0.2	9 ± 1	30 ± 3	77 ± 8	401 ± 51	521 ± 22
$N_{\pi\pm} = N^{sig}_{\pi\pm} + N^{comb}_{\pi\pm}$		702 ± 64	1954 ± 104	1561 ± 113	1214 ± 104	809 ± 80	942 ± 85	7181 ± 279
$N^{sig}_{\pi0}$	237/268	315 ± 42	576 ± 54	904 ± 107	471 ± 68	414 ± 83	159 ± 55	2840 ± 203
$N^{comb}_{\pi0}$		5 ± 0.7	13 ± 1	36 ± 4	36 ± 5	119 ± 24	397 ± 137	606 ± 43
$N_{\pi0} = N^{sig}_{\pi0} + N^{comb}_{\pi0}$		320 ± 42	589 ± 54	940 ± 107	507 ± 68	533 ± 86	556 ± 148	3446 ± 208
$N_{\pi\pm} + N_{\pi0}$ 4-mode	799/819	1012 ± 76	2535 ± 128	2485 ± 157	1729 ± 130	1291 ± 125	1552 ± 180	10604 ± 376

q^2 range (GeV^2)		0-8	8-16	>16	$0 < q^2 < 20.3$
$N^{sig}_{\rho\pm}$	147/131	237 ± 56	459 ± 44	170 ± 17	866 ± 101
$N^{comb}_{\rho\pm}$		56 ± 13	287 ± 27	368 ± 38	711 ± 82
$N_{\rho\pm} = N^{sig}_{\rho\pm} + N^{comb}_{\rho\pm}$		293 ± 57	746 ± 52	538 ± 42	1577 ± 130
$N^{sig}_{\rho0}$	162/131	253 ± 63	856 ± 74	294 ± 29	1403 ± 143
$N^{comb}_{\rho0}$		31 ± 8	267 ± 23	270 ± 27	567 ± 58
$N_{\rho0} = N^{sig}_{\rho0} + N^{comb}_{\rho0}$		284 ± 64	1123 ± 77	564 ± 40	1970 ± 154
$N_{\rho\pm} + N_{\rho0}$ 4-mode	799/819	471 ± 101	1754 ± 120	1105 ± 86	3332 ± 286

the data spectrum is low by $22 \pm 7\%$ compared to the MC simulation and a correction is applied. To assess the impact of the uncertainty of the correction procedure, the size of the correction is varied by its estimated uncertainty.

B. Simulation of Signal and Background

1. Signal Form Factors

To assess the impact of the form-factor (FF) uncertainty on the shape of the simulated signal distributions, we vary the $B \to \pi$ form factor within the uncertainty of the previous *BABAR* measurement [9] and the $B \to \rho$ form factors within the uncertainties of the LCSR calculation assessed by Ball and Zwicky [17]. For the latter we assume uncertainties on the form factors A_1, A_2 and V of 10% at $q^2 = 0$. They rise linearly to 13% at $q^2 = 14 \, \mathrm{GeV}^2$ and are extrapolated up to the kinematic endpoint. We add the uncertainties due to the three form factors in quadrature. For $B \to \pi\ell\nu$, the form-factor uncertainty is small, since we extract the signal in six bins of q^2. In contrast, for $B \to \rho\ell\nu$ the form-factor uncertainty is one of the dominant sources of systematic error. This is partly due to the stricter requirement on the lepton momentum, $p^*_\ell > 1.8 \, \mathrm{GeV}$, which is imposed to suppress the large $B \to X_c\ell\nu$ background. We refrain from using

the difference between LCSR and ISGW2 as systematic uncertainty, but this difference is comparable to the estimate we obtain from the uncertainties in the LCSR calculation.

2. $B \to X_u\ell\nu$ Background

The $B \to X_u\ell\nu$ background contribution is composed of the sum of exclusive decays, $B^+ \to \omega\ell^+\nu$, $B^+ \to \eta\ell^+\nu$, and $B^+ \to \eta'\ell^+\nu$ decays, and the remaining resonant and non-resonant $B \to X_u\ell\nu$ decays that make up the total $B \to X_u\ell\nu$ branching fraction. We estimate the total error of the $B \to X_u\ell\nu$ background composition by repeating the fit with branching fractions for various exclusive and non-resonant decays varied independently within their current measurement errors. The uncertainty of the branching fraction for non-resonant decays is dominant; it is equal to the error on the total $B \to X_u\ell\nu$ branching fraction, $\mathcal{B}(B \to X_u\ell\nu) = (2.33 \pm 0.22) \times 10^{-3}$ [41].

In addition, the analysis is sensitive to the mass and composition of the charmless hadronic states. We assess the uncertainty of the predictions by varying the QCD parameters that define the mass, the lepton spectrum, and the q^2 distributions predicted by calculations [36] based on HQE. We vary the shape-function (SF) parameters m_b and μ^2_π within the uncertainties (error ellipse) given in Ref. [38].

TABLE VIII: Systematic errors in % for $\mathcal{B}(B^0 \to \pi^- \ell^+ \nu)$ from the four-mode fit for bins in q^2 and the total q^2 range. The total errors are derived from the individual contributions taking into account the complete covariance matrix.

	$B \to \pi \ell \nu$						
q^2 range (GeV^2)	0-4	4-8	8-12	12-16	16-20	>20	0-26.4
Track efficiency	3.4	1.5	2.3	0.1	1.5	2.8	1.9
Photon efficiency	0.1	1.4	1.0	4.6	2.8	0.3	1.8
Lepton identification	3.8	1.6	1.9	1.8	1.9	3.0	1.8
K_L efficiency	1.0	0.1	0.5	4.5	0.4	2.0	1.4
K_L shower energy	0.1	0.1	0.1	0.8	0.9	3.8	0.7
K_L spectrum	1.6	1.9	2.2	3.1	4.4	2.3	2.5
$B \to \pi \ell \nu$ FF f_+	0.5	0.5	0.5	0.6	1.0	1.0	0.6
$B \to \rho \ell \nu$ FF A_1	1.7	1.2	3.4	2.0	0.1	1.6	1.7
$B \to \rho \ell \nu$ FF A_2	1.3	0.8	2.6	1.0	0.1	0.4	1.1
$B \to \rho \ell \nu$ FF V	0.2	0.3	0.9	0.7	0.1	0.5	0.5
$\mathcal{B}(B^+ \to \omega \ell^+ \nu)$	0.1	0.1	0.1	0.2	0.3	1.5	0.2
$\mathcal{B}(B^+ \to \eta \ell^+ \nu)$	0.1	0.1	0.2	0.2	0.2	0.5	0.2
$\mathcal{B}(B^+ \to \eta' \ell^+ \nu)$	0.1	0.1	0.1	0.1	0.1	0.3	0.1
$\mathcal{B}(B \to X_u \ell \nu)$	0.2	0.1	0.1	0.1	1.1	1.6	0.4
$B \to X_u \ell \nu$ SF param.	0.4	0.1	0.2	0.2	0.5	4.2	0.7
$B \to D \ell \nu$ FF ρ_D^2	0.2	0.1	0.5	0.3	0.2	0.7	0.3
$B \to D^* \ell \nu$ FF R_1	0.1	0.4	0.8	0.6	0.3	0.6	0.5
$B \to D^* \ell \nu$ FF R_2	0.5	0.2	0.1	0.2	0.1	0.4	0.2
$B \to D^* \ell \nu$ FF $\rho_{D^*}^2$	0.7	0.2	0.6	0.8	0.4	1.1	0.6
$\mathcal{B}(B \to D \ell \nu)$	0.2	0.2	0.3	0.4	0.5	0.5	0.3
$\mathcal{B}(B \to D^* \ell \nu)$	0.4	0.1	0.3	0.3	0.3	0.7	0.3
$\mathcal{B}(B \to D^{**} \ell \nu)_{\mathrm{narrow}}$	0.4	0.1	0.1	0.3	0.1	0.5	0.2
$\mathcal{B}(B \to D^{**} \ell \nu)_{\mathrm{broad}}$	0.1	0.1	0.1	0.5	0.1	0.2	0.2
Secondary leptons	0.5	0.2	0.3	0.2	0.2	0.7	0.3
Continuum	5.3	1.0	2.6	1.8	3.1	6.1	2.0
Bremsstrahlung	0.3	0.1	0.1	0.1	0.1	0.4	0.2
Radiative corrections	0.5	0.1	0.1	0.2	0.2	0.6	0.3
$N_{B\overline{B}}$	1.2	1.0	1.2	1.2	1.1	1.6	1.2
B lifetimes	0.3	0.3	0.3	0.3	0.3	0.7	0.3
$f_\pm / f_{00}$	1.0	0.4	0.8	0.8	0.5	1.3	0.8
Total syst. error	8.2	3.9	6.7	8.3	6.9	10.6	5.0

TABLE IX: Systematic errors in % for $\mathcal{B}(B^0 \to \rho^- \ell^+ \nu)$ from the four-mode fit for three bins in q^2 and the total q^2 range. The total errors are derived from the individual contributions taking into account the complete covariance matrix.

	$B \to \rho \ell \nu$			
q^2 range (GeV^2)	0-8	8-16	>16	0-20.3
Track efficiency	3.2	2.9	0.3	2.5
Photon efficiency	2.6	2.0	2.6	2.4
Lepton Identification	5.7	3.0	4.0	3.4
K_L efficiency	10.3	1.2	4.9	4.8
K_L shower energy	1.6	0.8	1.0	1.1
K_L spectrum	4.2	6.1	7.0	5.7
$B \to \pi \ell \nu$ FF f_+	0.1	0.1	0.7	0.2
$B \to \rho \ell \nu$ FF A_1	10.7	6.6	4.5	7.5
$B \to \rho \ell \nu$ FF A_2	8.5	3.8	0.8	4.7
$B \to \rho \ell \nu$ FF V	3.4	3.0	3.6	3.2
$\mathcal{B}(B^+ \to \omega \ell^+ \nu)$	0.7	0.7	3.4	1.2
$\mathcal{B}(B^+ \to \eta \ell^+ \nu)$	0.8	0.1	0.6	0.4
$\mathcal{B}(B^+ \to \eta' \ell^+ \nu)$	0.8	0.5	1.2	0.7
$\mathcal{B}(B \to X_u \ell \nu)$	7.4	7.3	10.6	8.0
$B \to X_u \ell \nu$ SF param.	11.9	7.6	12.8	10.0
$B \to D \ell \nu$ FF ρ_D^2	0.9	0.2	0.1	0.4
$B \to D^* \ell \nu$ FF R_1	0.7	0.1	0.3	0.3
$B \to D^* \ell \nu$ FF R_2	1.7	0.1	0.2	0.6
$B \to D^* \ell \nu$ FF $\rho_{D^*}^2$	2.0	0.2	0.1	0.7
$\mathcal{B}(B \to D \ell \nu)$	1.6	0.3	0.1	0.7
$\mathcal{B}(B \to D^* \ell \nu)$	0.5	0.1	0.3	0.3
$\mathcal{B}(B \to D^{**} \ell \nu)_{\mathrm{narrow}}$	1.3	0.1	0.1	0.5
$\mathcal{B}(B \to D^{**} \ell \nu)_{\mathrm{broad}}$	0.7	0.1	0.1	0.3
Secondary leptons	1.5	0.1	0.1	0.5
Continuum	8.9	3.8	5.0	4.0
Bremsstrahlung	0.9	0.1	0.2	0.4
Radiative corrections	1.3	0.1	0.7	0.6
$N_{B\overline{B}}$	2.7	2.0	2.5	2.3
B lifetimes	1.5	0.4	0.4	0.7
$f_\pm / f_{00}$	1.2	0.1	0.1	0.4
Total syst. error	26.1	16.1	21.3	15.7

For the two $B \to \rho \ell \nu$ samples, the $B \to X_u \ell \nu$ background is large compared to the signal and very difficult to separate. Consequently, the fit shows very high correlations between the fitted yields for signal and this background. We therefore choose to fix the background yields and shapes to those provided by the simulation, and account for the uncertainty by assessing the sensitivity of the fitted signal yield to variations of the $B \to X_u \ell \nu$ branching fraction and the shapes of the background distributions, corresponding to the estimated error of the shape-function parameters. The resulting estimated errors are the two dominant contributions to the systematic errors of the $B \to \rho \ell \nu$ partial and total branching fractions.

3. $B \to X_c \ell \nu$ Background

The systematic error related to the shapes of the $B \to X_c \ell \nu$ background distributions is dominated by the uncertainties in the branching fractions and form factors for the various semileptonic decays. We vary the composition of the $B \to X_c \ell \nu$ background based on a compilation of the individual branching fractions of $B \to D \ell \nu$, $B \to D^* \ell \nu$ and $B \to D^{**} \ell \nu$ (narrow and broad D^{**} states) decays within the ranges given by their errors, see Table I. Since we scaled up the four $B \to D^{**} \ell \nu$ branching fractions to take into account the unknown D^{**} partial branching fractions, the errors were increased by a factor of three relative to the published values.

To evaluate the effect of uncertainties in the form-factor parameters for the dominant $B \to D^* \ell \nu$ component, we repeat the fit with $\pm 1\sigma$ variations in each of the three form-factor parameters, $\rho_{D^*}^2$, R_1 and R_2. The impact of the form factor for the $B \to D \ell \nu$ background is evaluated by varying the parameter ρ_D^2 within its uncertainty.

4. Continuum Background

In Section V A, we have described the correction of the simulated shapes of the m_{ES}, ΔE, and q^2 distributions for the continuum using linear functions derived from comparison with off-resonance data. The uncertainties of the fitted slopes of these correction functions are used to evaluate the errors due to modeling of the shape of the continuum background distributions. They represent a sizable contribution to the systematic error, which is mainly due to the low statistics of the off-resonance data sample.

C. Other Systematic Uncertainties

1. Final-State Radiation and Bremsstrahlung

The kinematics of the signal decays are corrected for radiative effects such as final-state radiation and bremsstrahlung in detector material.

In the MC simulation, final-state radiation (FSR) is modeled using PHOTOS [35], which is based on $\mathcal{O}(\alpha)$ calculations but includes multiple-photon emission from the electron. We have studied the effects of FSR on the q^2 dependence of the measured signal and background yields by comparing events generated with and without PHOTOS. The observed change is largest, up to 5%, for electron momenta of about $0.6\,\mathrm{GeV}$ (i.e. well below our cut-off at $1\,\mathrm{GeV}$ for $B \to \pi\ell\nu$ and $1.8\,\mathrm{GeV}$ for $B \to \rho\ell\nu$). Comparisons of the PHOTOS simulation with semi-analytical calculations [50] show excellent agreement. Allowing for the fact that non-leading terms from possible electromagnetic corrections to the strong interactions of the quarks in the initial and final state have not been calculated to any precision [51], we adopt an uncertainty in the PHOTOS calculations of 20%.

The uncertainty of the bremsstrahlung correction is determined by the uncertainty of the amount of detector material in the inner detector. We have adopted as the systematic uncertainty due to bremsstrahlung the impact of a change in the thickness of the detector material by 0.14% radiation lengths, the estimated uncertainty in the thickness of inner detector and the beam vacuum pipe. As for final-state radiation, the uncertainty in the effective radiator thickness impacts primarily the electron spectrum.

The uncertainties due to final-state radiation and bremsstrahlung combined amount to far less than 1% for most of the q^2 range.

2. Number of $B\overline{B}$ Events

The determination of the on-resonance luminosity and the number of $B\overline{B}$ events is described in detail elsewhere [52]. The uncertainty of the total number of $B\overline{B}$ pairs is estimated to be 1.1%.

At the $\Upsilon(4S)$ resonance, the fraction of $B^0\bar{B}^0$ events is measured to be $f_{00} = 0.484 \pm 0.006$, with the ratio $f_{+-}/f_{00} = 1.065 \pm 0.026$ [41]. This error impacts the branching-ratio measurements by 0.8%.

3. B^0 and B^+ Lifetimes

Since we combine fits to decays of charged and neutral B mesons and make use of isospin relations, the B-meson lifetimes enter into the four-mode fit. We use the PDG [7] value for the B lifetime, $\tau_0 = 1.530 \pm 0.009$ ps , and the lifetime ratio, $\tau_+/\tau_0 = 1.071 \pm 0.009$. These uncertainties lead to a systematic error of 0.3% for $B \to \pi\ell\nu$ and 0.7% for $B \to \rho\ell\nu$ decays.

VIII. RESULTS

Based on the signal yields obtained in the four-mode fit, integrated over the full q^2 range (see Table VII), we derive the following total branching fractions, constrained by the isospin relations stated in Eqs. 30,

$$\mathcal{B}(B^0 \to \pi^-\ell^+\nu) = (1.41 \pm 0.05 \pm 0.07) \times 10^{-4} \ ,$$
$$\mathcal{B}(B^0 \to \rho^-\ell^+\nu) = (1.75 \pm 0.15 \pm 0.27) \times 10^{-4} \ .$$

Here and in the following, the first error reflects the statistical (fit) error and the second the estimated systematic error. The total branching fractions obtained from the single-mode fits for the charged and neutral $B \to \pi\ell\nu$ samples are

$$\mathcal{B}(B^0 \to \pi^-\ell^+\nu) = (1.44 \pm 0.06 \pm 0.07) \times 10^{-4} \ ,$$
$$\mathcal{B}(B^+ \to \pi^0\ell^+\nu) \times 2\frac{\tau_0}{\tau_+} = (1.40 \pm 0.10 \pm 0.11) \times 10^{-4} \ .$$

For the charged and neutral $B \to \rho\ell\nu$ samples, we obtain

$$\mathcal{B}(B^0 \to \rho^-\ell^+\nu) = (1.98 \pm 0.21 \pm 0.38) \times 10^{-4} \ ,$$
$$\mathcal{B}(B^+ \to \rho^0\ell^+\nu) \times 2\frac{\tau_0}{\tau_+} = (1.87 \pm 0.19 \pm 0.32) \times 10^{-4} \ .$$

The single-mode fits result in higher values for $\mathcal{B}(B^0 \to \rho^-\ell^+\nu)$ and $\mathcal{B}(B^+ \to \rho^0\ell^+\nu)$ than the average branching fraction obtained from the four-mode fit. This may be explained by different treatments of the isospin-conjugate signal and the $\pi \leftrightarrow \rho$ cross feed in the single- and four-mode fits. In contrast to the four-mode fit, the isospin-conjugate signal contribution in the single-mode fits is not constrained by the isospin-conjugate mode. In addition, the four-mode fit uses the same fit parameter for the signal and the cross feed from the signal mode into other modes, which leads to a slight decrease in the $B \to \rho\ell\nu$ branching fraction compared to the single-mode fits. Since the $\rho \to \pi$ cross feed is significantly larger than the $\pi \to \rho$ cross feed, the effect on the $B \to \rho\ell\nu$ results is larger than for $B \to \pi\ell\nu$.

Both the $B \to \pi\ell\nu$ and the $B \to \rho\ell\nu$ results are consistent within errors with the isospin relations,

$$\frac{\mathcal{B}(B^0 \to \pi^-\ell^+\nu)}{\mathcal{B}(B^+ \to \pi^0\ell^+\nu) \times 2\frac{\tau_0}{\tau_+}} = 1.03 \pm 0.09 \pm 0.06 \ ,$$

$$\frac{\mathcal{B}(B^0 \to \rho^-\ell^+\nu)}{\mathcal{B}(B^+ \to \rho^0\ell^+\nu) \times 2\frac{\tau_0}{\tau_+}} = 1.06 \pm 0.16 \pm 0.08 \ .$$

By extracting the signal in several q^2 bins we also measure the q^2 spectra of $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ decays. These spectra need to be corrected for effects such as detector resolution, bremsstrahlung, and final-state radiation.

A. Partial Branching Fractions

We correct the measured q^2 spectra for resolution, radiative effects and bremsstrahlung by applying an unfolding technique that is based on singular-value decomposition of the detector response matrix [53]. The detector response matrix in the form of a two-dimensional histogram of the reconstructed versus the true q^2 values (see Figure 4) is used as input to the unfolding algorithm. This algorithm contains a regularization term to suppress spurious oscillations originating from statistical fluctuations. To find the best choice of the regularization parameter κ we have studied the systematic bias on the partial branching fractions compared to the statistical uncertainty as a function of κ using a set of simulated distributions. The data samples in this analysis are large enough that no severe distortions due to statistical fluctuations are expected. We choose the largest possible value of κ, *i.e.*, we set κ equal to the number of q^2 bins, to minimize a potential bias.

The $\Delta\mathcal{B}/\Delta q^2$ distributions resulting from the unfolding procedure are presented in Figure 24 for $B \to \pi\ell\nu$ and in Figure 25 for $B \to \rho\ell\nu$. Tables X and XI list the partial branching fractions $\Delta\mathcal{B}$ for $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$, respectively.

B. Form-factor Shape

For $B \to \pi\ell\nu$ decays, we extract the shape of the form factor $f_+(q^2)$ directly from data. For $B \to \rho\ell\nu$ decays, we restrict ourselves to the measurement of the q^2 dependence, since the current experimental precision is not adequate to extract the three different form factors involved.

Several parameterizations of $f_+(q^2)$ are used to interpolate between results of various form-factor calculations or to extrapolate these calculations from a partial to the whole q^2 range. The four most common parameterizations, the BK [26], BZ [15], BGL [24, 25] and BCL [29] parameterizations, have been introduced in Section II. For the BGL and BCL parameterizations, we consider a linear ($k_{max} = 2$) and a quadratic ($k_{max} = 3$) ansatz.

We perform χ^2 fits to the measured q^2 spectrum to determine the free parameters for each of these parameterizations. The fit employs the following χ^2 definition, with integration of the fit function over the q^2 bins,

$$\chi^2 = \sum_{i,j=1}^{N_{\text{bins}}} \Delta_i \ V_{ij}^{-1} \Delta_j \ , \tag{32}$$

where $V_{i,j}^{-1}$ is the inverse covariance matrix of the partial-branching-fraction measurements. Δ_k for q^2 bin k is defined as

$$\Delta_k = \left(\frac{\Delta B}{\Delta q^2}\right)_k^{\text{data}} - \frac{C}{\Delta q_k^2} \int\limits_{\Delta q_k^2} p_\pi^3 |f_+(q^2; \alpha)|^2 dq^2, \tag{33}$$

where α denotes the set of parameters for a chosen parameterization of $f_+(q^2)$, and $C = |V_{ub}|^2 \tau_0 G_F^2/(24\pi^3)$ is an overall normalization factor whose value is irrelevant for these fits since the data can only constrain the shape of the form factor, but not its normalization.

In Table XII and Figure 24 we present the results of these fits to the $B \to \pi\ell\nu$ samples. All parameterizations describe the data well, with χ^2 probabilities ranging from 10% to 18%. Thus, within the current experimental precision, all parameterizations are valid choices, and the central values for $|V_{ub}|f_+(0)$ agree with each other. We choose the quadratic BGL parameterization as the default, though even a linear parameterization results in a very good fit to the data. The error band represents the uncertainties of the fit to data, based on the quadratic BGL parameterization (solid line in Figure 24). It has been computed using standard error propagation, taking the correlation between the fit parameters into account.

We compare the measured q^2 spectra with the shapes predicted by form-factor calculations based on lattice QCD [23], light-cone sum rules [15, 19], and the ISGW2 [14] relativistic quark model. Among the available calculations for $B \to \pi\ell\nu$ decays, the HPQCD lattice calculation agrees best with the data. It should be noted that the LQCD predictions are only valid for $q^2 > 16\,\text{GeV}^2$, the earlier LCSR calculation (LCSR 1) for $q^2 < 16\,\text{GeV}^2$, and the more recent LCSR calculation (LCSR 2) for $q^2 < 12\,\text{GeV}^2$; their extrapolation is impacted by sizable uncertainties.

In Table XI and Figure 25 we present the results of the fits to the $B \to \rho\ell\nu$ samples. The LCSR calculation and the ISGW2 model are in good agreement with the data. However, the errors of the measured $B \to \rho\ell\nu$ partial branching fractions are relatively large, at the level of 15-30%, depending on the q^2 interval.

It should be noted that the theoretical calculations differ most for low and high q^2. In these regions of phase space, the measurements are impacted significantly by higher levels of backgrounds, specifically continuum events at low q^2 and other $B \to X_u\ell\nu$ decays that are difficult to separate from the signal modes at higher q^2. These two background sources have been examined

TABLE X: Partial and total branching fractions (corrected for radiative effects) for $B^0 \to \pi^- \ell^+ \nu$ and $B^+ \to \pi^0 \ell^+ \nu$ decays obtained from the single-mode fits and $B \to \pi \ell \nu$ decays from the four-mode fit with statistical (fit), systematic and total errors. The branching fraction for $B^+ \to \pi^0 \ell^+ \nu$ has been scaled by twice the lifetime ratio of neutral and charged B mesons. All branching fractions and associated errors are given in units of 10^{-4}.

q^2 range (GeV^2)	0-4	4-8	8-12	12-16	16-20	20-26.4	Total	<16	>16
$\Delta\mathcal{B}(B^0 \to \pi^- \ell^+ \nu)$	0.313	0.329	0.241	0.222	0.206	0.124	1.435	1.105	0.330
Fit error	0.030	0.018	0.018	0.020	0.020	0.018	0.061	0.049	0.027
Syst. error	0.025	0.016	0.015	0.015	0.013	0.010	0.068	0.059	0.019
Total error	0.039	0.024	0.023	0.025	0.024	0.021	0.092	0.077	0.033
$\Delta\mathcal{B}(B^+ \to \pi^0 \ell^+ \nu) \times 2\tau_0/\tau_+$	0.357	0.294	0.234	0.210	0.206	0.099	1.401	1.096	0.305
Fit error	0.049	0.031	0.031	0.033	0.039	0.043	0.102	0.075	0.062
Syst. error	0.050	0.015	0.028	0.019	0.024	0.028	0.106	0.089	0.037
Total error	0.070	0.035	0.041	0.038	0.046	0.051	0.147	0.117	0.072
$\Delta\mathcal{B}(B^0 \to \pi^- \ell^+ \nu)$ 4-mode	0.320	0.321	0.235	0.220	0.201	0.118	1.414	1.095	0.319
Fit error	0.025	0.017	0.015	0.017	0.018	0.016	0.050	0.041	0.024
Syst. error	0.027	0.012	0.016	0.018	0.014	0.014	0.074	0.061	0.024
Total error	0.037	0.021	0.022	0.025	0.023	0.022	0.089	0.074	0.034
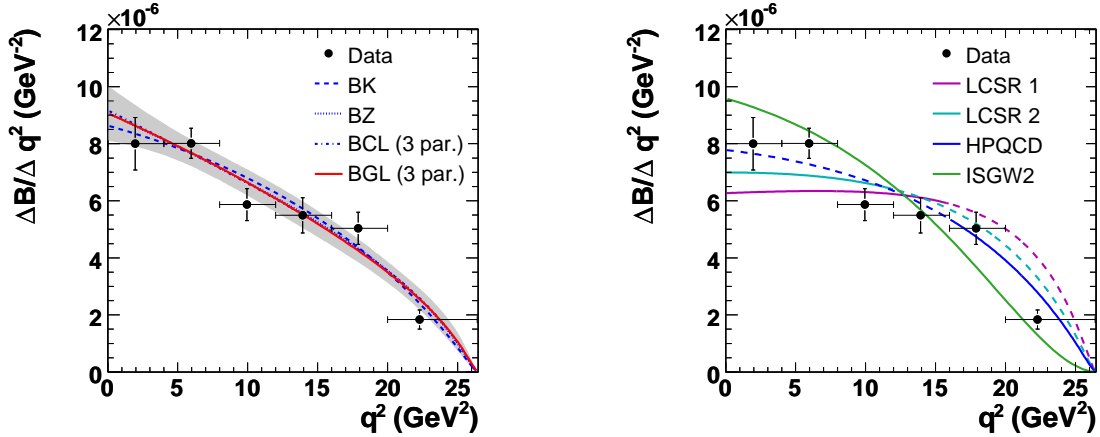
FIG. 24: Measured $\Delta\mathcal{B}/\Delta q^2$ distribution for $B \to \pi\ell\nu$. The vertical error bars correspond to the combined statistical and systematic uncertainties. The positions of the data points have been adjusted to correspond to the mean q^2 value in each bin, based on the quadratic BGL ansatz. Left: fits of four different form-factor parameterizations to the $\Delta\mathcal{B}/\Delta q^2$ data spectrum. The fit result for the BZ and BCL parameterizations are barely visible, since they overlap almost completely with the BGL result. The shaded band illustrates the uncertainty of the quadratic BGL fit to data. Right: shape comparisons of the data to various $B \to \pi\ell\nu$ form-factor predictions (LCSR 1 [15], LCSR 2 [19], HPQCD [23], ISGW2 [14]), which have been normalized to the measured total branching fraction. The extrapolations of the QCD predictions to the full q^2 range are marked as dashed lines.

in detail, and the uncertainties in their normalization and shape are included in the systematic uncertainties. For the inclusive $B \to X_u\ell\nu$ background, the q^2 and the hadronic mass spectra are derived from theoretical predictions that depend on non-perturbative parameters that are not well measured [38]. For $B \to \rho\ell\nu$ the correlation between the signal and the $B \to X_u\ell\nu$ background is so large that they cannot both be fitted simultaneously.

Thus the $B \to X_u\ell\nu$ background scale factor and shape are fixed to the MC predictions, which have large uncertainties. MC studies indicate that this may introduce a bias affecting the signal yield. The stated errors account for this potential bias.

TABLE XI: Partial and total branching fractions (corrected for radiative effects) for $B^0 \to \rho^- \ell^+ \nu$ and $B^+ \to \rho^0 \ell^+ \nu$ decays obtained from the single-mode fits and for $B \to \rho \ell \nu$ decays from the four-mode fit with statistical (fit), systematic and total errors. The branching fractions for $B^+ \to \rho^0 \ell^+ \nu$ have been scaled by twice the ratio of the lifetimes of neutral and charged B mesons. All branching fractions and associated errors are given in units of 10^{-4}.

q^2 range (GeV^2)	0-8	8-16	16-20.3	Total
$\Delta\mathcal{B}(B^0 \to \rho^- \ell^+ \nu)$	0.747	0.980	0.256	1.984
Fit error	0.151	0.087	0.030	0.214
Syst. error	0.178	0.165	0.066	0.379
Total error	0.234	0.187	0.072	0.435
$\Delta\mathcal{B}(B^+ \to \rho^0 \ell^+ \nu) \times 2\tau_0/\tau_+$	0.627	0.977	0.265	1.871
Fit error	0.136	0.079	0.028	0.190
Syst. error	0.152	0.161	0.061	0.320
Total error	0.204	0.179	0.068	0.373
$\Delta\mathcal{B}(B^0 \to \rho^- \ell^+ \nu)$ 4-mode	0.564	0.912	0.268	1.745
Fit error	0.107	0.059	0.022	0.149
Syst. error	0.126	0.135	0.058	0.272
Total error	0.166	0.147	0.062	0.310
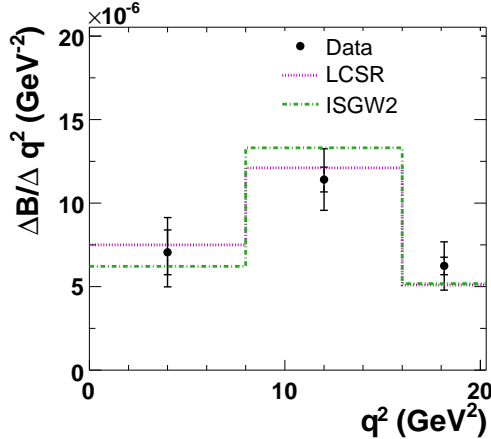
FIG. 25: (color online) Measured $\Delta\mathcal{B}/\Delta q^2$ distribution for $B \to \rho \ell \nu$. The inner and outer error bars correspond to the statistical uncertainty and the combined statistical and systematic uncertainty, respectively. The data are compared with the $B \to \rho \ell \nu$ form-factor predictions from LCSR [17] and from the ISGW2 quark model [14].

C. Determination of $|V_{ub}|$

We choose two different approaches to determine the magnitude of the CKM matrix element V_{ub}.

First, we use the traditional method to derive $|V_{ub}|$. As in previous publications [6, 8–11], we combine the measured partial branching fractions with integrals of the form-factor calculations over a certain q^2 range using the relation

$$|V_{ub}| = \sqrt{\frac{\Delta\mathcal{B}(q^2_{min}, q^2_{max})}{\tau_0 \, \Delta\zeta(q^2_{min}, q^2_{max})}} \, , \qquad (34)$$

where $\tau_0 = (1.530 \pm 0.009)$ ps is the B^0 lifetime and $\Delta\zeta$ is defined as

$$\Delta\zeta(q^2_{min}, q^2_{max}) = \frac{G_F^2}{24\pi^3} \int_{q^2_{min}}^{q^2_{max}} p_\pi^3 |f_+(q^2)|^2 dq^2 \, . \qquad (35)$$

The values of $\Delta\zeta$ are derived from theoretical form-factor calculations for different q^2 ranges. Table XIII summarizes the $\Delta\zeta$ values, the partial branching fractions and the $|V_{ub}|$ results.

For $B \to \rho \ell \nu$, values of $\Delta\zeta$ are taken from the LCSR calculation in the range $q^2 < 16$ GeV and the quark model predictions of ISGW2 over the full q^2 range. The results are also presented in Table XIII. Estimates of the uncertainties for $\Delta\zeta$ are not given in Refs. [17] and [14] .

Second, we perform a simultaneous fit to the most recent lattice results and BABAR data to make best use of the available information on the form factor from data (shape) and theory (shape and normalization). A fit of this kind was first presented by the FNAL/MILC Collaboration [22] using the earlier BABAR results on $B^0 \to \pi^- \ell^+ \nu$ decays [9].

To perform this fit, we translate the $f_+(q^2)$ predictions from LQCD to $1/(\tau_0 |V_{ub}|^2)\Delta\mathcal{B}/\Delta q^2$. We simultaneously fit this distribution and the $\Delta\mathcal{B}/\Delta q^2$ distribution from data as a function of q^2. We use the BGL form-factor parameterization as the fit function, with the additional normalization parameter $a_{\mathrm{norm}} = \tau_0 |V_{ub}|^2$, which allows us to determine $|V_{ub}|$ from the relative normalization of data and LQCD predictions.

The χ^2 for this fit is given by

$$\begin{aligned} \chi^2 &= \chi^2(\mathrm{data}) + \chi^2(\mathrm{lattice}) \\ &= \sum_{i,j=1}^{N_{\mathrm{bins}}} \Delta_i^{\mathrm{data}} (V_{ij}^{\mathrm{data}})^{-1} \Delta_j^{\mathrm{data}} \\ &+ \sum_{l,m=1}^{N_{\mathrm{points}}} \Delta_l^{\mathrm{lat}} (V_{lm}^{\mathrm{lat}})^{-1} \Delta_m^{\mathrm{lat}} \end{aligned} \qquad (36)$$

where

$$\Delta_i^{\mathrm{data}} = \left(\frac{\Delta B}{\Delta q^2}\right)_i^{\mathrm{data}} - \frac{1}{\Delta q_i^2} \int_{\Delta q_i^2} g(q^2; \alpha) dq^2 \, , \quad (37)$$

$$\Delta_l^{\mathrm{lat}} = \frac{G_F^2}{24\pi^3} p_\pi^3(q_l^2) |f_+^{\mathrm{lat}}(q_l^2)|^2 - g(q_l^2; \alpha) \qquad (38)$$

TABLE XII: Results of fits to the measured $\Delta\mathcal{B}/\Delta q^2$ for $B \to \pi\ell\nu$ decays, based on different form-factor parameterizations.

Parametrization	χ^2/ndf	Prob(χ^2/ndf)	Fit parameters	$f_+(0)|V_{ub}|$ [10^{-3}]
BK	6.8/4	0.148	$\alpha_{BK} = +0.310 \pm 0.085$	1.052 ± 0.042
BZ	6.0/3	0.112	$r_{BZ} = +0.170 \pm 0.124$	1.079 ± 0.046
			$\alpha_{BZ} = +0.761 \pm 0.337$	
BCL (2 par.)	6.3/4	0.179	$b_1/b_0 = -0.67 \pm 0.18$	1.065 ± 0.042
BCL (3 par.)	6.0/3	0.112	$b_1/b_0 = -0.90 \pm 0.46$	1.086 ± 0.055
			$b_2/b_0 = +0.47 \pm 1.49$	
BGL (2 par.)	6.6/4	0.156	$a_1/a_0 = -0.94 \pm 0.20$	1.103 ± 0.042
BGL (3 par.)	6.3/3	0.100	$a_1/a_0 = -0.82 \pm 0.29$	1.080 ± 0.056
			$a_2/a_0 = -1.14 \pm 1.81$	

TABLE XIII: $|V_{ub}|$ derived from $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ decays for various q^2 regions and form-factor calculations. Quoted errors are experimental uncertainties and theoretical uncertainties of the form-factor integral $\Delta\zeta$. No uncertainties on $\Delta\zeta$ for $B \to \rho\ell\nu$ are given in Refs. [17] and [14].

	q^2 Range (GeV^2)	$\Delta\mathcal{B}$ (10^{-4})	$\Delta\zeta$ (ps^{-1})	$|V_{ub}|$ (10^{-3})
$B \to \pi\ell\nu$				
LCSR 1 [15]	$0 - 16$	1.10 ± 0.07	5.44 ± 1.43	$3.63 \pm 0.12^{+0.59}_{-0.40}$
LCSR 2 [19]	$0 - 12$	0.88 ± 0.06	$4.00^{+1.01}_{-0.95}$	$3.78 \pm 0.13^{+0.55}_{-0.40}$
HPQCD [23]	$16 - 26.4$	0.32 ± 0.03	2.02 ± 0.55	$3.21 \pm 0.17^{+0.55}_{-0.36}$
$B \to \rho\ell\nu$				
LCSR [17]	$0 - 16.0$	1.48 ± 0.28	13.79	2.75 ± 0.24
ISGW2 [14]	$0 - 20.3$	1.75 ± 0.31	14.20	2.83 ± 0.24

and

$$g(q^2; \alpha) = \frac{G_F^2}{24\pi^3} p_\pi^3(q^2)|f_+(q^2)|^2$$
$$\times \begin{cases} a_{\text{norm}} & \text{for data} \\ 1 & \text{for LQCD} \end{cases}, \quad (39)$$

$$f_+(q^2) = \frac{1}{\mathcal{P}(q^2)\phi(q^2, q_0^2)} \sum_{k=0}^{k_{max}} a_k(q_0^2)[z(q^2, q_0^2)]^k \quad (40)$$

Here $(\Delta B/\Delta q^2)^{\text{data}}$ is the measured spectrum, $f_+^{\text{lat}}(q_l^2)$ are the form-factor predictions from LQCD, and $(V_{ij}^{\text{data}})^{-1}$ and $(V_{ij}^{\text{lat}})^{-1}$ are the corresponding inverse covariance matrices for $(\Delta B/\Delta q^2)^{\text{data}}$ and $G_F^2/(24\pi^3)p_\pi^3(q_l^2)|f_+^{\text{lat}}(q_l^2)|^2$, respectively. The set of free parameters α of the fit function $g(q^2; \alpha)$ contains the coefficients a_k of the BGL parameterization and the normalization parameter a_{norm}.

From the FNAL/MILC [22] lattice calculations, we use only subsets with six, four or three of the twelve predictions at different values of q^2, since neighboring points are very strongly correlated. All chosen subsets of LQCD points contain the point at lowest q^2. It has been checked that alternative choices of subsets give compatible re-

sults. From the HPQCD [23] lattice calculations, we use only the point at lowest q^2 since the correlation matrix for the four predicted points is not available. For comparison, we also perform the corresponding fit using only the point at lowest q^2 from FNAL/MILC. The data, the lattice predictions, and the fitted functions are shown in Figure 26. Table XIV shows the numerical results of the fit.

For the nominal fit we use the subset with four FNAL/MILC points and assume a quadratic BGL parameterization. We refer to this fit as 3+1-parameter BGL fit (three coefficients a_k and the normalization parameter a_{norm}). As can be seen in Table XII for the fit to data alone, the data are well described by a linear function with the normalization a_0 and a slope a_1/a_0. This indicates that most of the variation of the form factor is due to well-understood QCD effects that are parameterized by the functions $\mathcal{P}(q^2)$ and $\phi(q^2, q_0^2)$ in the BGL parameterization. If we include a curvature term in the fit, the slope $a_1/a_0 = -0.82 \pm 0.29$ is fully consistent with the linear fit; the curvature a_2/a_0 is negative and consistent with zero. Since the z distribution is almost linear, we also perform a linear fit (2+1-parameter BGL fit) for comparison. The results of the linear fits are also shown in Table XIV.

The simultaneous fits provide very similar results, both for the BGL expansion coefficients, which determine the shape of the spectrum, and for $|V_{ub}|$. The fitted values for the form-factor parameters are very similar to those obtained from the fits to data alone. This is not surprising, since the data dominate the fit results. Unfortunately the decay rate is lowest and the experimental errors are largest at large q^2, where the lattice calculation can make predictions. We obtain from these simultaneous fits

$$|V_{ub}| = (2.87 \pm 0.28) \times 10^{-3} \quad \text{FNAL/MILC (6 points)},$$
$$|V_{ub}| = (2.95 \pm 0.31) \times 10^{-3} \quad \text{FNAL/MILC (4 points)},$$
$$|V_{ub}| = (2.93 \pm 0.31) \times 10^{-3} \quad \text{FNAL/MILC (3 points)},$$
$$|V_{ub}| = (2.92 \pm 0.37) \times 10^{-3} \quad \text{FNAL/MILC (1 point)},$$
$$|V_{ub}| = (2.99 \pm 0.35) \times 10^{-3} \quad \text{HPQCD (1 point)},$$

where the stated error is the combined experimental and theoretical error obtained from the fit. The coefficients
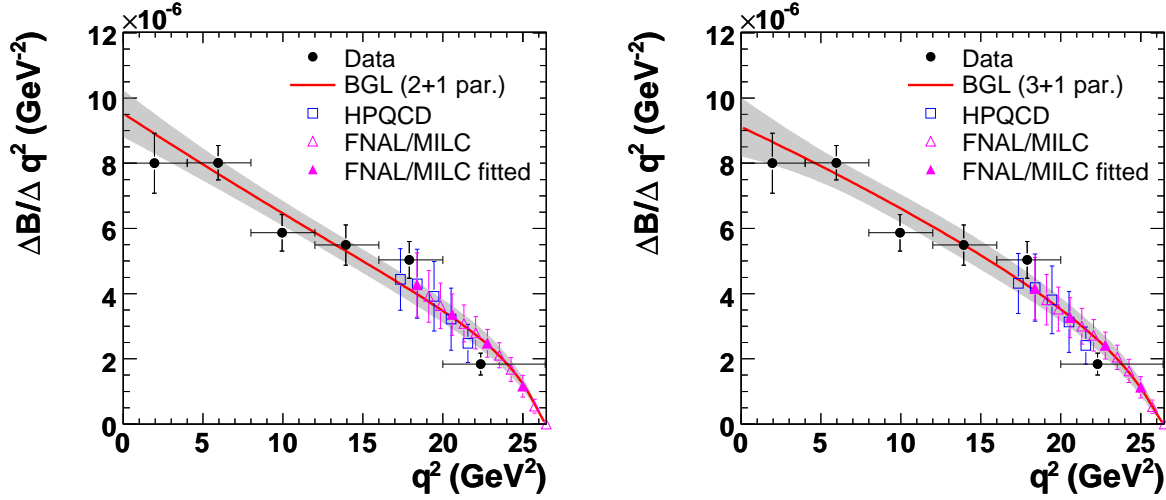
FIG. 26: (color online) Simultaneous fits of the BGL parameterization to data (solid points with vertical error bars representing the total experimental uncertainties) and to four of the twelve points of the FNAL/MILC lattice prediction (magenta, closed triangles). Left: linear (2+1-parameter) BGL fit, right: quadratic (3+1-parameter) BGL fit. The LQCD results are rescaled to the data according to the $|V_{ub}|$ value obtained in the fit. The shaded band illustrates the uncertainty of the fitted function. For comparison, the HPQCD (blue, open squares) lattice results are also shown. They are used in an alternate fit.

a_k are significantly smaller than 1, as predicted. The sum of the squares of the first two coefficients, $\sum_{k=0}^{1} a_k^2 = (0.85 \pm 0.20) \times 10^{-3}$, is consistent with the tighter bounds set by Becher and Hill [25].

Since the total error of 10% on $|V_{ub}|$ results from the simultaneous fit to data and LQCD predictions, it is non-trivial to separate the error into contributions from experiment and theory. We have estimated that the error contains contributions of 3% from the branching-fraction measurement, 5% from the shape of the q^2 spectrum determined from data, and 8.5% from the form-factor normalization obtained from theory.

We study the effect of variations of the isospin relations imposed in the combined four-mode fit as stated in Eqs. 30. These relations are not expected to be exact, though the comparison of the single-mode fit results provides no indication for isospin breaking. The isospin-breaking effects are primarily due to $\pi^0 - \eta$ and $\rho^0 - \omega$ mixing in $B^+ \to \pi^0 \ell^+ \nu$ and $B^+ \to \rho^0 \ell^+ \nu$ decays, respectively. They are expected to increase the branching fractions of the B^+ relative to the B^0 meson. Given the masses and widths of the mesons involved, the impact of $\pi^0 - \eta$ mixing is expected to be smaller than that of $\rho^0 - \omega$ mixing.

Detailed calculations have been performed to correct form-factor measurements and to extract V_{us} from semileptonic decays of charged and neutral kaons [54]. These calculations account for isospin breaking due to $\pi^0 - \eta$ mixing and should also be applicable to $B^+ \to \pi^0 \ell^+ \nu$ decays. For $B^+ \to \pi^0 \ell^+ \nu$ decays the effect is expected to be smaller by a factor of three, i.e., the predicted increase is $(1.5 \pm 0.2)\%$ [55]. For $B^+ \to \rho^0 \ell^+ \nu$ decays, calculations have not been carried out to the same precision. Based on the change in the $\pi^+ \pi^-$ rate at the peak of the ρ mass distribution, the branching fraction is predicted to increase by as much as 34% [56]. However, an integration over the resonances weighted by the proper Breit-Wigner function and taking into account the masses and finite ρ and ω widths results in a much smaller effect, an increase in the $\pi^+ \pi^-$ branching fraction of 6% [57].

We have assessed the impact of changes in the ratios of the branching fractions for charged and neutral B mesons on the extraction of the differential decay rates due to adjustments of the MC default branching fractions of the B^+ decays in the combined four-mode fit. For a 1.5% increase in the $B^+ \to \pi^0 \ell^+ \nu$ branching fraction, the fitted $B \to \pi \ell \nu$ partial branching fraction decreases by 0.5%, while the $B \to \rho \ell \nu$ rate increases by less than 0.1%. A 6% increase in the $B^+ \to \rho^0 \ell^+ \nu$ branching fraction results in a decrease of the $B \to \rho \ell \nu$ rate by 3.1% and a 0.14% increase for the fitted $B \to \pi \ell \nu$ rate. We observe a partial compensation to the change in the simulated $B^+ \to \pi^0 \ell^+ \nu$ rate due to changes in the $B^+ \to \rho^0 \ell^+ \nu$ background contribution, and vice versa. The observed changes in the fitted yields depend linearly on the imposed branching-fraction changes and are independent of q^2.

For a 1.5% variation of the $B^+ \to \pi^0 \ell^+ \nu$ branching fraction, the value for $|V_{ub}|$ extracted from the measured $B \to \pi \ell \nu$ spectrum decreases by 0.2%. A +6% variation of the $B^+ \to \rho^0 \ell^+ \nu$ branching fraction increases the value of $|V_{ub}|$ extracted from the same measured spectrum by 0.3%.

TABLE XIV: Results of simultaneous fits to data and LQCD calculations, based on the linear or quadratic BGL parameterizations. $|V_{ub}|$ is determined from the relative normalization of data and theory prediction.

Parameterization	Input	χ^2/ndf	Prob(χ^2/ndf)	Fit parameters
BGL (2+1 par.)	Data	18.2/9	0.033	$a_0 = (2.07 \pm 0.21) \times 10^{-2}$
	+ FNAL/MILC			$a_1/a_0 = -0.78 \pm 0.22$
	(6 points)			$|V_{ub}| = (3.04 \pm 0.38) \times 10^{-3}$
BGL (2+1 par.)	Data	7.1/7	0.415	$a_0 = (2.16 \pm 0.19) \times 10^{-2}$
	+ FNAL/MILC			$a_1/a_0 = -0.93 \pm 0.20$
	(4 points)			$|V_{ub}| = (2.99 \pm 0.32) \times 10^{-3}$
BGL (2+1 par.)	Data	6.8/6	0.341	$a_0 = (2.17 \pm 0.19) \times 10^{-2}$
	+ FNAL/MILC			$a_1/a_0 = -0.93 \pm 0.20$
	(3 points)			$|V_{ub}| = (2.97 \pm 0.32) \times 10^{-3}$
BGL (2+1 par.)	Data	6.6/4	0.156	$a_0 = (2.23 \pm 0.26) \times 10^{-2}$
	+ FNAL/MILC			$a_1/a_0 = -0.94 \pm 0.20$
	(1 point)			$|V_{ub}| = (2.90 \pm 0.36) \times 10^{-3}$
BGL (2+1 par.)	Data	6.6/4	0.156	$a_0 = (2.19 \pm 0.23) \times 10^{-2}$
	+ HPQCD			$a_1/a_0 = -0.94 \pm 0.20$
	(1 point)			$|V_{ub}| = (2.94 \pm 0.34) \times 10^{-3}$
BGL (3+1 par.)	Data	9.8/8	0.276	$a_0 = (2.31 \pm 0.20) \times 10^{-2}$
	+ FNAL/MILC			$a_1/a_0 = -0.71 \pm 0.19$
	(6 points)			$a_2/a_0 = -2.33 \pm 0.84$
				$|V_{ub}| = (2.87 \pm 0.28) \times 10^{-3}$
BGL (3+1 par.)	Data	6.6/6	0.355	$a_0 = (2.22 \pm 0.21) \times 10^{-2}$
	+ FNAL/MILC			$a_1/a_0 = -0.86 \pm 0.23$
	(4 points)			$a_2/a_0 = -0.97 \pm 1.36$
				$|V_{ub}| = (2.95 \pm 0.31) \times 10^{-3}$
BGL (3+1 par.)	Data	6.3/5	0.279	$a_0 = (2.24 \pm 0.22) \times 10^{-2}$
	+ FNAL/MILC			$a_1/a_0 = -0.84 \pm 0.23$
	(3 points)			$a_2/a_0 = -1.01 \pm 1.40$
				$|V_{ub}| = (2.93 \pm 0.31) \times 10^{-3}$
BGL (3+1 par.)	Data	6.3/3	0.100	$a_0 = (2.24 \pm 0.26) \times 10^{-2}$
	+ FNAL/MILC			$a_1/a_0 = -0.82 \pm 0.29$
	(1 point)			$a_2/a_0 = -1.14 \pm 1.81$
				$|V_{ub}| = (2.92 \pm 0.37) \times 10^{-3}$
BGL (3+1 par.)	Data	6.3/3	0.100	$a_0 = (2.19 \pm 0.23) \times 10^{-2}$
	+ HPQCD			$a_1/a_0 = -0.82 \pm 0.29$
	(1 point)			$a_2/a_0 = -1.14 \pm 1.81$
				$|V_{ub}| = (2.99 \pm 0.35) \times 10^{-3}$

IX. CONCLUSIONS

In summary, we have measured the exclusive branching fractions $\mathcal{B}(B^0 \to \pi^- \ell^+ \nu)$ and $\mathcal{B}(B^0 \to \rho^- \ell^+ \nu)$ as a function of q^2 and have determined $|V_{ub}|$ using recent form-factor calculations. We measure the total branching fractions, based on samples of charged and neutral B mesons and isospin constraints, to be

$$\mathcal{B}(B^0 \to \pi^- \ell^+ \nu) = (1.41 \pm 0.05 \pm 0.07) \times 10^{-4} ,$$
$$\mathcal{B}(B^0 \to \rho^- \ell^+ \nu) = (1.75 \pm 0.15 \pm 0.27) \times 10^{-4} ,$$

where the first error is the statistical uncertainty of the fit employed to determine the signal and background yields

and the second is the systematic uncertainty. The separate measurements of the branching fractions for charged and neutral B mesons are consistent within errors with the assumed isospin relations,

$$\frac{\mathcal{B}(B^0 \to \pi^- \ell^+ \nu)}{\mathcal{B}(B^+ \to \pi^0 \ell^+ \nu) \times 2\frac{\tau_0}{\tau_+}} = 1.03 \pm 0.09 \pm 0.06 ,$$

$$\frac{\mathcal{B}(B^0 \to \rho^- \ell^+ \nu)}{\mathcal{B}(B^+ \to \rho^0 \ell^+ \nu) \times 2\frac{\tau_0}{\tau_+}} = 1.06 \pm 0.16 \pm 0.08 .$$

We have assessed the sensitivity of the combined branching-fraction measurements to isospin violations due to $\pi^0 - \eta$ and $\rho^0 - \omega$ mixing in B^+ decays. Based

on the best estimates currently available, the impact on the branching fractions is small compared to the total systematic errors. We refrain from applying corrections, given the uncertainties in the size of the effects.

The measured branching fraction for $B \to \pi\ell\nu$ is more precise than any previous measurement and agrees well with the current world average $\mathcal{B}(B^0 \to \pi^-\ell^+\nu) = (1.36 \pm 0.05 \pm 0.05) \times 10^{-4}$ [41]. The branching fraction for $B \to \rho\ell\nu$ is also the most precise single measurement to date based on a large signal event sample, although the Belle Collaboration [11] has reported a smaller systematic error (by a factor of two), based on a small signal sample of hadronically-tagged events [11]. The $B \to \rho\ell\nu$ branching fraction presented here is significantly lower (by about 2.5 σ) compared to the current world average $\mathcal{B}(B^0 \to \rho^-\ell^+\nu) = (2.77 \pm 0.18 \pm 0.16) \times 10^{-4}$ [41]. The dominant uncertainty of this $B \to \rho\ell\nu$ measurement is due to the limited knowledge of the normalization and shape of the irreducible background from other $B \to X_u\ell\nu$ decays.

Within the sizable errors, the measured q^2 spectrum for $B \to \rho\ell\nu$ agrees well with the predictions from light-cone sum rules [17] and the ISGW2 [14] quark model. Neither of these calculations includes an estimate of their uncertainties. In the future, it will require much cleaner data samples and considerably better understanding of other $B \to X_u\ell\nu$ decays to achieve significant improvements in the measurements of the form factors in B decays to vector mesons.

For $B \to \pi\ell\nu$ decays, the measured q^2 spectrum agrees best with the one predicted by the HPQCD lattice calculations [23]. The measurement of the differential decay rates is consistent with earlier *BABAR* measurements [6, 9] within the stated errors, though the yield at low q^2 is somewhat higher than previously measured. This results in a smaller value of α_{BK}, the parameter introduced by Becirevic and Kaidalov [26], namely $\alpha_{BK} = 0.31 \pm 0.09$. Using the BGL ansatz, we determine a value $f_+(0)|V_{ub}| = (1.08 \pm 0.06) \times 10^{-3}$, which is larger than the value, $f_+(0)|V_{ub}| = (0.91 \pm 0.06 \pm 0.3) \times 10^{-3}$ [16], based on the earlier *BABAR* decay rate measurement [9] and an average branching fraction of $(1.37 \pm 0.06 \pm 0.07) \times 10^{-4}$ [41].

We determine the CKM matrix element $|V_{ub}|$ using two different approaches. First, we use the traditional method to derive $|V_{ub}|$ by combining the measured partial branching fractions with the form-factor predictions based on different QCD calculations. The results, presented in Table XIII, agree within the sizable uncertainties of the form-factor predictions. For this approach we quote as a result the value of

$$|V_{ub}| = (3.78 \pm 0.13^{+0.55}_{-0.40}) \times 10^{-3},$$

based on the most recent LCSR calculation for $q^2 < 12\,\mathrm{GeV}^2$. Second, we extract $|V_{ub}|$ from simultaneous fits to data and lattice predictions using the quadratic BGL parameterization for the whole q^2 range. These fits to data and the two most recent lattice calculations by the FNAL/MILC [22] and HPQCD [23] Collaborations agree very well. We quote as a result the fitted value of

$$|V_{ub}| = (2.95 \pm 0.31) \times 10^{-3},$$

based on the normalization predicted by the FNAL/MILC Collaboration. The total error of 10% is dominated by the theory error of 8.5%. This value of $|V_{ub}|$ is smaller by one standard deviation compared to the results of a combined fit to earlier *BABAR* measurements and the same recent FNAL/MILC lattice calculations [22].

The values of $|V_{ub}|$ presented here appear to be sensitive to the q^2 range for which theory predictions and the measured spectrum can be compared. LCSR calculations are restricted to low values of q^2 and result in values of $|V_{ub}|$ in the range of $(3.63 - 3.78) \times 10^{-3}$ with theoretical uncertainties of $^{+16}_{-11}\%$ and experimental errors of $3 - 4\%$. LQCD predictions are available for $q^2 > 16\,\mathrm{GeV}^2$ and result in $|V_{ub}|$ in the range of $(2.95 - 3.21) \times 10^{-3}$ and experimental errors of $5 - 6\%$ for both the traditional method and the simultaneous fit to LQCD predictions and the measured spectrum. This fit combines the measured shape of the spectrum over the full q^2 range with the lattice QCD form-factor predictions at high q^2 and results in a reduced theoretical uncertainty of 8.5%, as compared to $^{+17}_{-11}\%$ for the traditional method.

Both $|V_{ub}|$ values quoted as results are also lower than most determinations of $|V_{ub}|$ based on inclusive $B \to X_u\ell\nu$ decays, which are typically in the range $(4.0 - 4.5) \times 10^{-3}$. These inclusive measurements are very sensitive to the mass of the b quark, which is extracted from fits to moments of inclusive $B \to X_c\ell\nu$ and $B \to X_s\gamma$ decay distributions [38] and depends on higher-order QCD corrections. Estimated theoretical uncertainties are typically 6%.

Global fits constraining the parameters of the CKM unitarity triangle performed by the CKMfitter [58] and UTfit [59] Collaborations currently predict values for $|V_{ub}|$ that fall between the two results presented here, $|V_{ub}| = 3.51^{+0.14}_{-0.16} \times 10^{-3}$ and $|V_{ub}| = 3.41 \pm 0.18 \times 10^{-3}$, respectively.

To permit more stringent tests of the CKM framework and its consistency with the standard model of electroweak interactions, further reductions in the experimental and theoretical uncertainties will be necessary. For $B \to \pi\ell\nu$ decays this will require a reduction in the statistical errors and improved detector hermeticity to more effectively reconstruct the neutrino, which will reduce backgrounds from all sources. Further improvements in the precision of lattice and other QCD calculations will also be beneficial.

[1] BABAR Collaboration, B. Aubert et al., Phys. Rev. Lett. **99**, 171803 (2007).

[2] Belle Collaboration, K. F. Chen et al., Phys. Rev. Lett. **98**, 031802 (2007).

[3] N. Cabibbo, Phys. Lett. **10**, 531 (1963); M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[4] R. Kowalewski and T. Mannel in Review of Particle Physics, Phys. Lett. **B667**, 951 (2008).

[5] Charge conjugate decay modes are implicitly included.

[6] BABAR Collaboration, B. Aubert et al., Phys. Rev. **D72**, 051102 (2005).

[7] Particle Data Group, C. Amsler et al., Phys. Lett. **B667**, 1 (2008).

[8] CLEO Collaboration, S. B. Athar et al., Phys. Rev. **D68**, 072003 (2003); N. E. Adam et al., Phys. Rev. Lett. **99**, 041802 (2007).

[9] BABAR Collaboration, B. Aubert et al., Phys. Rev. Lett. **98**, 091801 (2007).

[10] BABAR Collaboration, B. Aubert et al., Phys. Rev. Lett. **97**, 211801 (2006); Phys. Rev. Lett. **101**, 081801 (2008).

[11] Belle Collaboration, T. Hokuue et al., Phys. Lett. **B648**, 139 (2007); Belle Collaboration, I. Adachi et al., Measurement of exclusive $B \to X_u \ell \nu$ decays using full-reconstruction tagging at Belle, e-Print: arXiv:0812:1414 (2008).

[12] M. Neubert, Physics Reports **245**, 259 (1994).

[13] J. D. Richman and P. R. Burchat, Rev. Mod. Phys. **67**, 893 (1995).

[14] N. Isgur, D. Scora, B. Grinstein, and M. B. Wise, Phys. Rev. **D39**, 799 (1989); D. Scora, N. Isgur, Phys. Rev. **D52**, 2783 (1995).

[15] P. Ball and R. Zwicky, JHEP **0110**, 019 (2001); Phys. Rev. **D71**, 014015 (2005).

[16] P. Ball, Phys. Lett. **B644**, 38 (2007).

[17] P. Ball and V. M. Braun, Phys. Rev. **D58**, 094016 (1998); P. Ball and R. Zwicky, Phys. Rev. **D71**, 014029 (2005).

[18] P. Ball and G. Jones, JHEP **08**, 025 (2007).

[19] G. Duplancic, A. Khodjamirian, T. Mannel, B. Melic and N. Offen, JHEP **804**, 14 (2008); private communication, A. Khodjamirian (2010). The extrapolation of this calculation to higher q^2 uses a BK parameterization with shape parameter $\alpha = 0.53$, based on a fit to earlier BABAR data.

[20] Fermilab Lattice Collaboration, Aida X. El-Khadra et al., Phys. Rev. **D64**, 014502 (2001).

[21] JLQCD Collaboration, S. Aoki et al., Phys. Rev. **D64**, 114505 (2001).

[22] Fermilab Lattice and MILC Collaboration, J. Bailey et al., Phys. Rev. **D79**, 054507 (2009).

[23] HPQCD Collaboration, E. Gulez, et al., Phys. Rev. **D73**, 074502 (2006) and Erratum ibid. **D75**, 119906 (2007).

[24] C. G. Boyd, B. Grinstein, and R.F. Lebed, Phys. Rev. Lett. **74**, 4603 (1995); C.G. Boyd and M.J. Savage, Phys. Rev. **D56**, 303 (1997).

[25] T. Becher and R. J. Hill, Phys. Lett. **B633**, 61 (2006).

[26] D. Becirevic and A. B. Kaidalov, Phys. Lett. **B478**, 417 (2000).

[27] For the BGL parameterization, we choose the value $q_0^2 = -0.65m_-^2$ as proposed in Ref. [22]. For the BCL parameterization, we choose $q_0^2 = (M_B + m_\pi)(\sqrt{M_B} - \sqrt{m_\pi})^2$, as proposed in Ref. [29]. For the B-meson and pion masses, we use $M_B = 5.279\,\mathrm{GeV}$ and $m_\pi = 0.1396\,\mathrm{GeV}$.

[28] The function $\phi(q^2, q_0^2)$ in the BGL parameterization is given by the expression

$$
\begin{aligned}
\phi(q^2, q_0^2) \; = \; & \sqrt{\frac{1}{32\pi\chi_J^{(0)}}} (\sqrt{m_+^2 - q^2} + \sqrt{m_+^2 - q_0^2}) \\
& \times \; (\sqrt{m_+^2 - q^2} + \sqrt{m_+^2 - m_-^2})^{3/2} \\
& \times \; (\sqrt{m_+^2 - q^2} + \sqrt{m_+^2})^{-5} \\
& \times \; \frac{(m_+^2 - q^2)}{(m_+^2 - q_0^2)^{1/4}}
\end{aligned}
$$

with numerical factor $\chi_J^{(0)} = 6.889 \times 10^{-4}$.

[29] C. Bourrely, I. Caprini, and L. Lellouch, Phys. Rev. **D79**, 013008 (2009).

[30] BABAR Collaboration, B. Aubert et al., Nucl. Instr. and Methods **A479**, 1 (2002).

[31] BABAR Collaboration, B. Aubert et al., Phys. Rev. **D66**, 032003 (2002).

[32] The current inclusive limit on the branching fraction for non-$B\overline{B}$ decays is 4% at 95 C.L., based on $2fb^1$: CLEO Collaboration, B. Barish et al., Phys. Rev. Lett. **76**, 1570 (1996). Measured exclusive $\Upsilon(4S)$ branching fractions to $\Upsilon(1S)\pi^+\pi^-$, $\Upsilon(2S)\pi^+\pi^-$, and $\Upsilon(1S)\eta$ sum to 4×10^{-4}: BABAR Collaboration, B. Aubert et al., Phys. Rev. Lett. **96**, 342001 (2006); Phys. Rev. **D78**, 112002 (2008).

[33] D. J. Lange, Nucl. Instr. and Methods **A462**, 152 (2001).

[34] GEANT4 Collaboration, S. Agostinelli *et al.*, Nucl. Instr. and Methods, Phys. Res., **A506**, 250 (2003).

[35] E. Barberio and Z. Was, Comput. Phys. Commun. **79**, 291 (1994).

[36] F. DeFazio and M. Neubert, JHEP **06**, 017 (1999).

[37] T. Sjöstrand, Comput. Phys. Commun. **82**, 74 (1994).

[38] O. Buchmüller and H. Flächer, Phys. Rev. **D73**, 073008 (2006) and Update for LP07, referenced by HFAG in e-Print: arXiv:0808.1297 [hep-ex] (2008).

[39] N. Isgur and M. B. Wise, Phys. Lett. **232**, 113 (1989); *ibid.* **237**, 527 (1990).

[40] I. Caprini, L. Lellouch and M. Neubert, Nucl. Phys. **B530** 153 (1998).

[41] Heavy Flavor Averaging Group, E. Barbieri *et al.*, e-Print: arXiv:0808.1297 [hep-ex], and recent updates.

[42] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. **D79**, 012002 (2009).

[43] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **104**, 011802 (2010).

[44] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. **D77**, 032002 (2008).

[45] These branching fractions have been derived from averages for B^0 and B^+ branching fractions obtained from recent updates for HFAG 2008, enforcing isospin relations and scaling the measured $B \to D^{**}\ell\nu$ branching fractions to make up for the missing semileptonic decays.

[46] F. Rosenblatt, Psychological Review **65**, 386 (1958).

[47] R. Fletcher, Practical Methods of Optimization (Wiley, Hoboken, 2000).

[48] G. C. Fox and S. Wolfram, Nucl. Phys. **B149**, 413 (1979), Erratum-ibid. **B157**, 543 (1979).

[49] R. J. Barlow and C. Beeston, Comput. Phys. Commun. **77**, 219–228 (1993).

[50] E. Richter-Was, Phys. Lett. **B303**, 163 (1993).

[51] E. S. Ginsberg, Phys. Rev. **142**, 1035 (1966); D. Atwood and J. W. Marciano, Phys. Rev. **D41**, 1736 (1990).

[52] G. D. McGregor, *B Counting at BABAR*, MSc Thesis, University of British Columbia, CANADA, SLAC-R-912 (2008), e-Print: arXiv:0812.1954 [hep-ex].

[53] A. Höcker and V. Kartvelishvili, Nucl. Instrum. Meth. **A37**, 469 (1996).

[54] H. Leutwyler and M. Roos, Z. Phys. **C25**, 91 (1984); V. Cirigliano, M. Knecht, H. Neufeld, H. Rubertsberger, and P. Talavera, Eur. Phys. J. **C23**, 121 (2002); A. Kastner and H. Neufeld, Eur. Phys. J. **C57**, 541 (2008).

[55] This difference accounts for the fact that the K^+ and the B^+ are members of different SU_3 multiplets; private communication, V. Cirigliano (2010).

[56] J. L. Diaz-Cruz, G. Lopez Castro, and J. H. Munoz, Phys. Rev. **D54**, 2388 (1996); G. Lopez Castro, J. H. Munoz, and G. Toledo Sanchez, Phys. Rev. **D56**, 6025 (1997);

[57] D. J. Lange, *Rare Exclusive Semileptonic Deacys at CLEO*, PhD Thesis, U.C. Santa Barbara; CORNELL-LEPP-THESIS-99-3.

[58] CKMfitter Collaboration, J. Charles *et al.*, Eur. Phys. J. **C41** 1 (2005); Updated results at the 12th International Conference on B-Physics at Hadron Machines, Heidelberg (2009); http://ckmfitter.in2p3.fr.

[59] UTfit Collaboration, M. Bona *et al.*, Updated results at the 12th International Conference on B-Physics at Hadron Machines, Heidelberg 2009; http://www.utfit.org/.

XI. APPENDIX

A. Systematic Uncertainties for One-Mode Fits

TABLE XV: Systematic errors in % for $\mathcal{B}(B^0 \to \pi^- \ell^+ \nu)$ from the $B^0 \to \pi^- \ell^+ \nu$ and $B^+ \to \pi^0 \ell^+ \nu$ one-mode fits.

	$B^0 \to \pi^- \ell^+ \nu$							$B^+ \to \pi^0 \ell^+ \nu$						
q^2 range (GeV^2)	0-4	4-8	8-12	12-16	16-20	>20	0-26.4	0-4	4-8	8-12	12-16	16-20	>20	0-26.4
Track efficiency	2.0	1.7	2.9	1.3	0.1	2.3	1.7	7.3	1.8	3.3	1.6	1.8	6.8	3.7
Photon efficiency	0.7	0.9	1.9	3.5	0.7	1.0	1.6	2.9	1.8	5.3	2.4	8.9	10.2	4.8
Lepton identification	4.1	1.7	1.8	1.8	2.1	2.4	1.8	3.3	1.3	2.2	1.4	2.7	2.8	1.4
K_L efficiency	1.0	0.1	0.3	3.5	1.2	1.7	1.2	1.2	0.5	2.4	3.6	1.8	1.1	1.8
K_L shower energy	0.1	0.1	0.1	0.7	1.4	2.8	0.7	2.6	0.5	0.4	0.6	1.4	4.1	1.4
K_L spectrum	1.1	2.0	2.5	2.6	5.1	1.1	2.4	2.7	0.7	2.0	4.5	4.4	4.2	2.9
$B \to \pi\ell\nu FF\ f_+$	0.5	0.3	0.4	0.5	0.6	0.8	0.5	0.5	0.5	0.6	0.7	1.8	2.6	1.0
$B \to \rho\ell\nu FF A_1$	1.6	1.8	2.0	0.9	2.0	1.5	1.6	2.7	2.2	3.6	2.0	1.6	5.1	2.8
$B \to \rho\ell\nu FF A_2$	1.5	1.7	1.7	0.5	2.1	0.6	1.4	2.5	2.0	3.0	1.4	1.9	2.7	2.2
$B \to \rho\ell\nu FF V$	0.3	0.4	0.5	0.3	0.4	0.8	0.4	1.1	0.7	1.2	1.0	0.9	3.0	1.2
$\mathcal{B}(B^0 \to \rho^- \ell^+ \nu)$	0.3	0.2	0.5	0.8	0.7	0.4	0.5	0.4	0.2	0.3	0.4	0.9	1.6	0.5
$\mathcal{B}(B^+ \to \omega\ell^+ \nu)$	0.1	0.1	0.1	0.1	0.3	1.4	0.2	0.6	0.3	0.6	0.6	0.9	4.8	1.1
$\mathcal{B}(B^+ \to \eta\ell^+ \nu)$	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.5	0.6	0.2
$\mathcal{B}(B^+ \to \eta'\ell^+ \nu)$	0.1	0.1	0.1	0.1	0.1	0.3	0.1	0.1	0.1	0.1	0.1	0.4	0.7	0.2
$\mathcal{B}(B \to X_u\ell\nu)$	0.1	0.1	0.1	0.1	0.3	0.9	0.2	0.7	0.2	0.3	0.4	1.2	1.4	0.6
$B \to X_u\ell\nu$ SF param.	0.4	0.2	0.1	0.4	0.6	2.3	0.5	0.6	0.3	0.4	1.1	1.3	6.2	1.3
$B \to D\ell\nu$ FF ρ_D^2	0.1	0.1	0.4	0.3	0.1	0.3	0.2	0.4	0.1	0.6	0.3	0.2	0.8	0.4
$B \to D^*\ell\nu$ FF R_1	0.2	0.4	0.8	0.9	0.3	0.9	0.6	0.1	0.5	1.3	0.5	0.7	1.8	0.7
$B \to D^*\ell\nu$ FF R_2	0.5	0.1	0.2	0.1	0.1	0.2	0.2	0.5	0.3	0.6	0.5	0.9	1.4	0.6
$B \to D^*\ell\nu$ FF $\rho_{D^*}^2$	0.9	0.4	0.5	1.0	0.2	0.8	0.6	0.7	0.4	0.6	0.6	0.6	2.4	0.8
$\mathcal{B}(B \to D\ell\nu)$	0.2	0.1	0.2	0.3	0.4	0.1	0.2	1.1	0.3	0.8	0.3	0.3	0.9	0.6
$\mathcal{B}(B \to D^*\ell\nu)$	0.5	0.1	0.3	0.4	0.3	0.3	0.3	0.9	0.4	0.6	0.7	1.7	1.8	0.9
$\mathcal{B}(B \to D^{**}\ell\nu)_{\text{narrow}}$	0.3	0.1	0.1	0.4	0.1	0.1	0.2	0.2	0.2	0.3	0.3	0.2	0.2	0.2
$\mathcal{B}(B \to D^{**}\ell\nu)_{\text{broad}}$	0.2	0.1	0.1	0.5	0.1	0.2	0.2	0.3	0.3	0.3	0.4	0.1	0.6	0.3
Secondary leptons	0.8	0.1	0.2	0.1	0.1	0.2	0.2	0.5	0.2	0.3	0.4	0.8	0.9	0.5
Continuum	5.0	1.6	1.6	1.3	1.8	4.1	1.7	8.8	1.8	6.1	4.0	4.8	13.1	4.2
Bremsstrahlung	0.2	0.1	0.1	0.1	0.1	0.2	0.1	0.2	0	0	0.1	0.2	0.2	0.1
Radiative corrections	0.4	0.1	0.1	0.3	0.5	0.9	0.3	0.5	0.1	0.1	0.3	0.5	0.9	0.4
$N_{B\overline{B}}$	1.1	1.2	1.1	1.2	1.2	1.0	1.1	1.2	1.1	1.2	1.2	1.1	1.2	1.2
$f_\pm/f_{00}$	1.5	1.3	1.3	1.3	1.0	0.8	1.3	1.3	1.1	1.0	1.0	0.9	0.2	1.0
Total	7.8	4.8	6.0	6.8	7.1	7.6	4.4	13.7	5.0	11.1	8.6	12.7	22.4	7.5

TABLE XVI: Systematic errors in % for $\mathcal{B}(B^0 \to \rho^- \ell^+ \nu)$ from the $B^0 \to \rho^- \ell^+ \nu$ and $B^+ \to \rho^0 \ell^+ \nu$ one-mode fits.

	$B^0 \to \rho^- \ell^+ \nu$				$B^+ \to \rho^0 \ell^+ \nu$			
q^2 range (GeV^2)	0-8	8-16	> 16	0-20.3	0-8	8-16	> 16	0-20.3
Track efficiency	2.4	0.8	3.0	1.8	2.4	4.0	1.0	2.9
Photon efficiency	1.4	1.5	3.9	2.0	5.5	3.5	3.5	4.2
Lepton Identification	2.6	2.9	4.7	3.0	4.3	3.2	4.6	3.7
K_L efficiency	8.2	1.0	6.5	4.3	10.1	1.8	5.8	5.2
K_L shower energy	4.4	0.2	0.4	1.6	1.0	0.7	0.8	0.8
K_L spectrum	9.3	8.2	9.7	8.8	1.0	5.2	7.3	4.3
$B \to \pi \ell \nu$ FF f_+	1.2	0.3	2.8	1.1	0.8	0.2	2.4	0.8
$B \to \rho \ell \nu$ FF A_1	14.8	8.3	4.8	9.6	14.3	8.3	4.3	9.4
$B \to \rho \ell \nu$ FF A_2	11.3	5.1	0.8	6.2	11.0	5.0	0.6	6.0
$B \to \rho \ell \nu$ FF V	3.7	3.5	3.9	3.6	3.8	3.6	3.5	3.6
$\mathcal{B}(B^0 \to \pi^- \ell^+ \nu)$	0.8	0.6	1.8	0.9	0.6	0.1	1.5	0.5
$\mathcal{B}(B^+ \to \omega \ell^+ \nu)$	0.1	0.4	3.3	0.8	0.9	0.9	3.3	1.4
$\mathcal{B}(B^+ \to \eta \ell^+ \nu)$	0.1	0.1	0.8	0.2	0.1	0.1	0.6	0.1
$\mathcal{B}(B^+ \to \eta' \ell^+ \nu)$	0.1	0.1	1.2	0.3	1.2	0.8	1.1	1.0
$\mathcal{B}(B \to X_u \ell \nu)$	4.0	5.3	10.3	5.9	7.9	8.0	11.0	8.6
$B \to X_u \ell \nu$ SF param.	6.9	6.5	13.1	7.9	10.9	7.8	12.7	9.7
$B \to D \ell \nu$ FF ρ_D^2	0.1	0.2	0.1	0.1	0.6	0.1	0.1	0.2
$B \to D^* \ell \nu$ FF R_1	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.1
$B \to D^* \ell \nu$ FF R_2	0.2	0.1	0.3	0.2	0.6	0.1	0.2	0.3
$B \to D^* \ell \nu$ FF $\rho_{D^*}^2$	0.4	0.2	0.1	0.2	1.4	0.1	0.2	0.5
$\mathcal{B}(B \to D \ell \nu)$	0.6	0.3	0.2	0.4	0.4	0.2	0.2	0.3
$\mathcal{B}(B \to D^* \ell \nu)$	0.3	0.1	0.2	0.2	0.2	0.1	0.2	0.2
$\mathcal{B}(B \to D^{**} \ell \nu)_{\mathrm{narrow}}$	0.7	0.1	0.1	0.2	0.6	0.1	0.1	0.2
$\mathcal{B}(B \to D^{**} \ell \nu)_{\mathrm{broad}}$	0.9	0.1	0.1	0.3	0.9	0.1	0.1	0.4
Secondary leptons	0.5	0.1	0.2	0.2	0.7	0.1	0.1	0.2
Continuum	5.9	3.4	6.3	3.2	8.2	3.8	6.4	3.4
Bremsstrahlung	0.2	0.1	0.2	0.2	0.2	0.1	0.2	0.2
Radiative corrections	0.6	0.1	0.8	0.4	0.6	0.1	0.8	0.4
$N_{B\bar{B}}$	1.7	1.9	2.9	2.0	2.2	2.1	2.9	2.3
$f_\pm/f_{00}$	1.9	1.4	1.0	1.5	1.6	1.4	0.7	1.3
Total	25.6	16.5	23.9	19.4	27.6	18.0	22.5	17.1

B. Correlation and Covariance Matrices

Table XVII shows the full correlation matrix for all signal and background fit parameters in the four-mode maximum-likelihood fit used to determine the signal yields, described in Section VI. This appendix also contains all statistical, systematic and total correlation and covariance matrices for the $B \rightarrow \pi \ell \nu$ and $B \rightarrow \rho \ell \nu$ $\Delta B / \Delta q^2$ measurements. The total correlation matrix is shown before and after unfolding of the q^2 spectrum. All covariance matrices are shown after q^2 unfolding. The total covariance matrix for $B \rightarrow \pi \ell \nu$ in Table XXVIII is used in the form-factor fits described in Eq. 32 or 36.

TABLE XVII: Full correlation matrix of the four-mode maximum-likelihood fit, for all fit parameters (signal and background). The first column indicates the fit parameter that corresponds to a certain row/column of the matrix.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
$p_{\pi\pm}^{q\bar{q}}$	1	1.000																								
$p_{\pi\pm}^{D^*\ell\nu}$	2	0.147	1.000																							
$p_{\pi\pm}^{otherB\bar{B}}$	3	-0.654	-0.449	1.000																						
$p_{\pi\pm,1}^{u\ell\nu}$	4	-0.096	-0.449	-0.225	1.000																					
$p_{\pi\pm,2}^{u\ell\nu}$	5	-0.175	-0.119	0.033	0.214	1.000																				
$p_1^{\rho\ell\nu}$	6	0.021	0.197	-0.148	-0.158	-0.080	1.000																			
$p_2^{\rho\ell\nu}$	7	0.055	0.060	-0.066	-0.158	-0.111	0.264	1.000																		
$p_3^{\rho\ell\nu}$	8	0.033	0.002	-0.015	-0.079	-0.564	0.137	0.189	1.000																	
$p_1^{\pi\ell\nu}$	9	-0.263	0.134	0.030	-0.013	0.044	-0.144	-0.033	-0.019	1.000																
$p_2^{\pi\ell\nu}$	10	-0.105	0.126	-0.007	-0.120	0.032	-0.429	-0.050	-0.033	0.191	1.000															
$p_3^{\pi\ell\nu}$	11	0.086	0.150	-0.110	-0.271	-0.044	0.085	-0.156	0.024	0.050	0.089	1.000														
$p_4^{\pi\ell\nu}$	12	0.004	-0.073	0.110	-0.232	-0.025	-0.001	-0.267	0.026	-0.005	0.058	0.197	1.000													
$p_5^{\pi\ell\nu}$	13	-0.162	0.116	0.244	-0.432	0.216	0.061	0.059	-0.469	0.068	0.085	0.127	0.135	1.000												
$p_6^{\pi\ell\nu}$	14	-0.124	0.036	0.079	-0.033	-0.602	0.007	-0.005	0.035	0.057	0.011	0.005	-0.008	0.032	1.000											
$p_{\pi0}^{q\bar{q}}$	15	0.055	-0.015	-0.049	0.029	0.103	0.031	0.037	-0.020	-0.119	0.059	-0.050	0.056	-0.079	-0.163	1.000										
$p_{\pi0}^{D^*\ell\nu}$	16	-0.075	0.104	-0.009	-0.082	-0.043	0.128	-0.019	0.024	0.156	0.109	0.086	-0.021	0.095	0.073	-0.193	1.000									
$p_{\pi0}^{otherB\bar{B}}$	17	0.003	-0.086	0.092	0.007	-0.082	-0.143	-0.039	0.059	-0.064	-0.127	-0.078	0.033	0.031	0.067	-0.438	-0.532	1.000								
$p_{\pi0,1}^{u\ell\nu}$	18	0.024	-0.052	-0.045	0.200	0.098	-0.118	-0.119	-0.198	0.006	-0.094	-0.078	-0.177	-0.142	0.028	-0.445	-0.235	0.018	1.000							
$p_{\pi0,2}^{u\ell\nu}$	19	0.022	-0.002	-0.033	0.066	0.557	-0.077	-0.104	-0.617	0.034	0.003	0.020	-0.052	0.259	-0.377	-0.237	-0.020	-0.089	0.431	1.000						
$p_{\rho\pm}^{q\bar{q}}$	20	-0.006	-0.022	-0.002	0.084	0.015	-0.066	-0.230	-0.019	0.011	-0.013	0.032	-0.007	-0.091	0.001	-0.017	-0.001	0.009	0.060	0.017	1.000					
$p_{\rho\pm}^{D^*\ell\nu}$	21	-0.004	0.113	-0.071	-0.088	-0.002	0.528	0.069	-0.000	-0.063	-0.201	0.076	0.023	0.077	0.005	0.011	0.080	-0.085	-0.048	0.005	-0.145	1.000				
$p_{\rho\pm}^{otherB\bar{B}}$	22	0.002	-0.109	0.072	0.080	0.064	-0.537	-0.069	-0.111	0.061	0.216	-0.085	-0.006	-0.006	-0.010	-0.001	-0.082	0.076	0.055	0.061	-0.514	-0.671	1.000			
$p_{\rho0}^{q\bar{q}}$	23	-0.004	-0.005	-0.025	0.106	0.040	0.013	-0.353	-0.058	-0.012	-0.049	0.043	0.015	-0.103	-0.001	-0.006	0.015	-0.011	0.068	0.038	0.086	0.027	-0.028	1.000		
$p_{\rho0}^{D^*\ell\nu}$	24	0.012	0.116	-0.093	-0.079	0.020	0.567	0.156	-0.036	-0.082	-0.232	0.041	-0.007	0.075	-0.002	0.026	0.071	-0.097	-0.045	0.027	-0.044	0.304	-0.294	-0.064	1.000	
$p_{\rho0}^{otherB\bar{B}}$	25	-0.013	-0.119	0.103	0.057	0.022	-0.598	-0.067	-0.043	0.088	0.257	-0.058	0.008	-0.005	-0.002	-0.021	-0.080	0.096	0.044	0.021	0.021	-0.319	0.319	-0.472	-0.781	1.000

TABLE XVIII: Statistical (fit) correlation matrix of the $B \to \pi\ell\nu$ $\Delta B/\Delta q^2$ measurement for the four-mode fit.

q^2 range (GeV^2)	0-4	4-8	8-12	12-16	16-20	>20
0-4	1.000	0.191	0.050	-0.005	0.068	0.057
4-8		1.000	0.089	0.058	0.085	0.011
8-12			1.000	0.197	0.127	0.005
12-16				1.000	0.135	-0.008
16-20					1.000	0.032
>20						1.000

TABLE XIX: Systematic correlation matrix of the $B \to \pi\ell\nu$ $\Delta B/\Delta q^2$ measurement for the four-mode fit.

q^2 range (GeV^2)	0-4	4-8	8-12	12-16	16-20	>20
0-4	1.000	0.521	0.705	0.394	-0.052	0.075
4-8		1.000	0.853	0.687	0.605	0.478
8-12			1.000	0.652	0.366	0.439
12-16				1.000	0.637	0.367
16-20					1.000	0.509
>20						1.000

TABLE XX: Total correlation matrix of the $B \to \pi\ell\nu$ $\Delta B/\Delta q^2$ measurement for the four-mode fit.

q^2 range (GeV^2)	0-4	4-8	8-12	12-16	16-20	>20
0-4	1.000	0.337	0.401	0.212	0.015	0.066
4-8		1.000	0.430	0.343	0.272	0.205
8-12			1.000	0.443	0.227	0.219
12-16				1.000	0.350	0.180
16-20					1.000	0.221
>20						1.000

TABLE XXI: Total correlation matrix of the $B \to \pi\ell\nu$ $\Delta B/\Delta q^2$ measurement for the four-mode fit after unfolding of the q^2 spectrum.

q^2 range (GeV^2)	0-4	4-8	8-12	12-16	16-20	>20
0-4	1.000	0.272	0.331	0.216	-0.037	0.045
4-8		1.000	0.390	0.273	0.252	0.172
8-12			1.000	0.475	0.194	0.170
12-16				1.000	0.462	0.042
16-20					1.000	0.195
>20						1.000

TABLE XXII: Statistical (fit) correlation matrix of the $B \to \rho\ell\nu$ $\Delta B/\Delta q^2$ measurement for the four-mode fit.

q^2 range (GeV^2)	0-8	8-16	>16
0-8	1.000	0.264	0.137
8-16		1.000	0.189
>16			1.000

TABLE XXIII: Systematic correlation matrix of the $B \to \rho\ell\nu$ $\Delta B/\Delta q^2$ measurement for the four-mode fit.

q^2 range (GeV^2)	0-8	8-16	>16
0-8	1.000	0.339	0.692
8-16		1.000	0.296
>16			1.000

TABLE XXIV: Total correlation matrix of the $B \to \rho\ell\nu$ $\Delta B/\Delta q^2$ measurement for the four-mode fit.

q^2 range (GeV^2)	0-8	8-16	>16
0-8	1.000	0.307	0.532
8-16		1.000	0.281
>16			1.000

TABLE XXV: Total correlation matrix of the $B \to \rho\ell\nu$ $\Delta B/\Delta q^2$ measurement for the four-mode fit after unfolding of the q^2 spectrum.

q^2 range (GeV^2)	0-8	8-16	>16
0-8	1.000	0.574	0.380
8-16		1.000	0.389
>16			1.000

TABLE XXVI: Statistical (fit) covariance matrix of the $B \to \pi\ell\nu$ $\Delta B/\Delta q^2$ measurement for the four-mode fit after unfolding of the q^2 spectrum in units of 10^{-13}.

q^2 range (GeV^2)	0-4	4-8	8-12	12-16	16-20	>20
0-4	4.039	0.436	-0.134	-0.020	-0.015	0.116
4-8		1.861	0.135	-0.027	0.104	0.018
8-12			1.462	0.404	0.110	-0.018
12-16				1.720	0.534	-0.157
16-20					1.995	0.014
>20						0.650

TABLE XXVII: Systematic covariance matrix of the $B \to \pi\ell\nu$ $\Delta B/\Delta q^2$ measurement for the four-mode fit after unfolding of the q^2 spectrum in units of 10^{-13}.

q^2 range (GeV^2)	0-4	4-8	8-12	12-16	16-20	>20
0-4	4.425	0.888	1.836	1.254	-0.176	0.022
4-8		0.931	1.018	0.920	0.653	0.289
8-12			1.666	1.244	0.504	0.338
12-16				2.123	1.091	0.245
16-20					1.228	0.359
>20						0.488

TABLE XXVIII: Total covariance matrix of the $B \to \pi\ell\nu$ $\Delta B/\Delta q^2$ measurement for the four-mode fit after unfolding of the q^2 spectrum in units of 10^{-13}.

q^2 range (GeV^2)	0-4	4-8	8-12	12-16	16-20	>20
0-4	8.463	1.324	1.702	1.234	-0.191	0.139
4-8		2.792	1.152	0.894	0.757	0.307
8-12			3.129	1.648	0.615	0.320
12-16				3.843	1.625	0.089
16-20					3.223	0.373
>20						1.138

TABLE XXIX: Statistical (fit) covariance matrix of the $B \to \rho\ell\nu$ $\Delta B/\Delta q^2$ measurement for the four-mode fit after unfolding of the q^2 spectrum in units of 10^{-12}.

q^2 range (GeV^2)	0-8	8-16	>16
0-8	1.798	0.536	-0.093
8-16		0.543	0.066
>16			0.269

TABLE XXX: Systematic covariance matrix of the $B \to \rho\ell\nu$ $\Delta B/\Delta q^2$ measurement for the four-mode fit after unfolding of the q^2 spectrum in units of 10^{-12}.

q^2 range (GeV^2)	0-8	8-16	>16
0-8	2.500	1.653	1.230
8-16		2.837	0.968
>16			1.816

TABLE XXXI: Total covariance matrix of the $B \to \rho\ell\nu$ $\Delta B/\Delta q^2$ measurement for the four-mode fit after unfolding of the q^2 spectrum in units of 10^{-12}.

q^2 range (GeV^2)	0-8	8-16	>16
0-8	4.298	2.188	1.137
8-16		3.381	1.034
>16			2.086