Measurement of the semileptonic decays $\bar{B} \to D\tau^-\bar{\nu}_\tau$ and $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$

B. Aubert,[1] M. Bona,[1] Y. Karyotakis,[1] J. P. Lees,[1] V. Poireau,[1] E. Prencipe,[1] X. Prudent,[1] V. Tisserand,[1] J. Garra Tico,[2] E. Grauges,[2] L. Lopez,[3a,3b] A. Palano,[3a,3b] M. Pappagallo,[3a,3b] G. Eigen,[4] B. Stugu,[4] L. Sun,[4] G. S. Abrams,[5] M. Battaglia,[5] D. N. Brown,[5] R. G. Jacobsen,[5] L. T. Kerth,[5] Yu. G. Kolomensky,[5] G. Lynch,[5] I. L. Osipenkov,[5] M. T. Ronan,[5,*] K. Tackmann,[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] B. G. Fulsom,[8] C. Hearty,[8] T. S. Mattison,[8] J. A. McKenna,[8] M. Barrett,[9] A. Khan,[9] V. E. Blinov,[10] A. D. Bukin,[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] M. Bondioli,[11] S. Curry,[11] I. Eschrich,[11] D. Kirkby,[11] A. J. Lankford,[11] P. Lund,[11] M. Mandelkern,[11] E. C. Martin,[11] D. P. Stoker,[11] S. Abachi,[12] C. Buchanan,[12] H. Atmacan,[13] J. W. Gary,[13] F. Liu,[13] O. Long,[13] G. M. Vitug,[13] Z. Yasin,[13] L. Zhang,[13] V. Sharma,[14] C. Campagnari,[15] T. M. Hong,[15] D. Kovalskyi,[15] M. A. Mazur,[15] J. D. Richman,[15] T. W. Beck,[16] A. M. Eisner,[16] C. J. Flacco,[16] C. A. Heusch,[16] J. Kroseberg,[16] W. S. Lockman,[16] A. J. Martinez,[16] T. Schalk,[16] B. A. Schumm,[16] A. Seiden,[16] M. G. Wilson,[16] L. O. Winstrom,[16] C. H. Cheng,[17] D. A. Doll,[17] B. Echenard,[17] F. Fang,[17] D. G. Hitlin,[17] I. Narsky,[17] T. Piatenko,[17] F. C. Porter,[17] R. Andreassen,[18] G. Mancinelli,[18] B. T. Meadows,[18] K. Mishra,[18] M. D. Sokoloff,[18] P. C. Bloom,[19] W. T. Ford,[19] A. Gaz,[19] J. F. Hirschauer,[19] M. Nagel,[19] U. Nauenberg,[19] J. G. Smith,[19] S. R. Wagner,[19] R. Ayad,[20,†] A. Soffer,[20,‡] W. H. Toki,[20] R. J. Wilson,[20] E. Feltresi,[21] A. Hauke,[21] H. Jasper,[21] M. Karbach,[21] J. Merkel,[21] A. Petzold,[21] B. Spaan,[21] K. Wacker,[21] M. J. Kobel,[22] R. Nogowski,[22] K. R. Schubert,[22] R. Schwierz,[22] A. Volk,[22] D. Bernard,[23] G. R. Bonneaud,[23] E. Latour,[23] M. Verderi,[23] P. J. Clark,[24] S. Playfer,[24] J. E. Watson,[24] M. Andreotti,[25a,25b] D. Bettoni,[25a] C. Bozzi,[25a] R. Calabrese,[25a,25b] A. Cecchi,[25a,25b] G. Cibinetto,[25a,25b] P. Franchini,[25a,25b] E. Luppi,[25a,25b] M. Negrini,[25a,25b] A. Petrella,[25a,25b] L. Piemontese,[25a] V. Santoro,[25a,25b] R. Baldini-Ferroli,[26] A. Calcaterra,[26] R. de Sangro,[26] G. Finocchiaro,[26] S. Pacetti,[26] P. Patteri,[26] I. M. Peruzzi,[26,§] M. Piccolo,[26] M. Rama,[26] A. Zallo,[26] A. Buzzo,[27a] R. Contri,[27a,27b] M. Lo Vetere,[27a,27b] M. M. Macri,[27a] M. R. Monge,[27a,27b] S. Passaggio,[27a] C. Patrignani,[27a,27b] E. Robutti,[27a] A. Santroni,[27a,27b] S. Tosi,[27a,27b] K. S. Chaisanguanthum,[28] M. Morii,[28] A. Adametz,[29] J. Marks,[29] S. Schenk,[29] U. Uwer,[29] V. Klose,[30] H. M. Lacker,[30] D. J. Bard,[31] P. D. Dauncey,[31] M. Tibbetts,[31] P. K. Behera,[32] X. Chai,[32] M. J. Charles,[32] U. Mallik,[32] 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] C. K. Lae,[34] N. Arnaud,[35] J. Béquilleux,[35] A. D'Orazio,[35] M. Davier,[35] J. Firmino da Costa,[35] G. Grosdidier,[35] F. Le Diberder,[35] V. Lepeltier,[35] A. M. Lutz,[35] S. Pruvot,[35] P. Roudeau,[35] M. H. Schune,[35] J. Serrano,[35] V. Sordini,[35,‖] A. Stocchi,[35] G. Wormser,[35] D. J. Lange,[36] D. M. Wright,[36] I. Bingham,[37] J. P. Burke,[37] C. A. Chavez,[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] C. K. Clarke,[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] K. E. Alwyn,[42] D. Bailey,[42] R. J. Barlow,[42] G. Jackson,[42] G. D. Lafferty,[42] T. J. West,[42] J. I. Yi,[42] J. Anderson,[43] C. Chen,[43] A. Jawahery,[43] D. A. Roberts,[43] G. Simi,[43] J. M. Tuggle,[43] C. Dallapiccola,[44] X. Li,[44] E. Salvati,[44] S. Saremi,[44] R. Cowan,[45] D. Dujmic,[45] P. H. Fisher,[45] S. W. Henderson,[45] G. Sciolla,[45] M. Spitznagel,[45] F. Taylor,[45] R. K. Yamamoto,[45] M. Zhao,[45] P. M. Patel,[46] S. H. Robertson,[46] A. Lazzaro,[47a,47b] V. Lombardo,[47a] F. Palombo,[47a,47b] J. M. Bauer,[48] L. Cremaldi,[48] R. Godang,[48,¶] R. Kroeger,[48] D. J. Summers,[48] H. W. Zhao,[48] M. Simard,[49] P. Taras,[49] H. Nicholson,[50] G. De Nardo,[51a,51b] L. Lista,[51a] D. Monorchio,[51a,51b] G. Onorato,[51a,51b] C. Sciacca,[51a,51b] G. Raven,[52] H. L. Snoek,[52] C. P. Jessop,[53] K. J. Knoepfel,[53] J. M. LoSecco,[53] W. F. Wang,[53] L. A. Corwin,[54] K. Honscheid,[54] H. Kagan,[54] R. Kass,[54] J. P. Morris,[54] A. M. Rahimi,[54] J. J. Regensburger,[54] S. J. Sekula,[54] Q. K. Wong,[54] N. L. Blount,[55] J. Brau,[55] R. Frey,[55] O. Igonkina,[55] J. A. Kolb,[55] M. Lu,[55] R. Rahmat,[55] N. B. Sinev,[55] D. Strom,[55] J. Strube,[55] E. Torrence,[55] G. Castelli,[56a,56b] N. Gagliardi,[56a,56b] M. Margoni,[56a,56b] M. Morandin,[56a] M. Posocco,[56a] M. Rotondo,[56a] F. Simonetto,[56a,56b] R. Stroili,[56a,56b] C. Voci,[56a,56b] P. del Amo Sanchez,[57] E. Ben-Haim,[57] H. Briand,[57] G. Calderini,[57] J. Chauveau,[57] O. Hamon,[57] Ph. Leruste,[57] J. Ocariz,[57] A. Perez,[57] J. Prendki,[57] S. Sitt,[57] L. Gladney,[58] M. Biasini,[59a,59b] E. Manoni,[59a,59b] C. Angelini,[60a,60b] G. Batignani,[60a,60b] S. Bettarini,[60a,60b] M. Carpinelli,[60a,60b,**] A. Cervelli,[60a,60b] F. Forti,[60a,60b] M. A. Giorgi,[60a,60b] A. Lusiani,[60a,60c] G. Marchiori,[60a,60b] M. Morganti,[60a,60b] N. Neri,[60a,60b] E. Paoloni,[60a,60b] G. Rizzo,[60a,60b] J. J. Walsh,[60a] D. Lopes Pegna,[61] C. Lu,[61] J. Olsen,[61] A. J. S. Smith,[61] A. V. Telnov,[61] F. Anulli,[62a] E. Baracchini,[62a,62b] G. Cavoto,[62a] R. Faccini,[62a,62b] F. Ferrarotto,[62a] F. Ferroni,[62a,62b] M. Gaspero,[62a,62b] P. D. Jackson,[62a] L. Li Gioi,[62a] M. A. Mazzoni,[62a] S. Morganti,[62a] G. Piredda,[62a] F. Renga,[62a,62b] C. Voena,[62a] M. Ebert,[63] T. Hartmann,[63] H. Schröder,[63] R. Waldi,[63] T. Adye,[64] B. Franek,[64] E. O. Olaiya,[64] F. F. Wilson,[64] S. Emery,[65] M. Escalier,[65] L. Esteve,[65] G. Hamel de Monchenault,[65] W. Kozanecki,[65] G. Vasseur,[65] Ch. Yèche,[65] M. Zito,[65] X. R. Chen,[66] H. Liu,[66] W. Park,[66] M. V. Purohit,[66] R. M. White,[66] J. R. Wilson,[66]

M. T. Allen,[67] D. Aston,[67] R. Bartoldus,[67] J. F. Benitez,[67] R. Cenci,[67] J. P. Coleman,[67] M. R. Convery,[67] J. C. Dingfelder,[67] J. Dorfan,[67] G. P. Dubois-Felsmann,[67] W. Dunwoodie,[67] R. C. Field,[67] A. M. Gabareen,[67] M. T. Graham,[67] P. Grenier,[67] C. Hast,[67] W. R. Innes,[67] J. Kaminski,[67] M. H. Kelsey,[67] H. Kim,[67] P. Kim,[67] M. L. Kocian,[67] D. W. G. S. Leith,[67] S. Li,[67] B. Lindquist,[67] S. Luitz,[67] V. Luth,[67] H. L. Lynch,[67] D. B. MacFarlane,[67] H. Marsiske,[67] R. Messner,[67] D. R. Muller,[67] H. Neal,[67] S. Nelson,[67] C. P. O'Grady,[67] I. Ofte,[67] M. Perl,[67] B. N. Ratcliff,[67] A. Roodman,[67] A. A. Salnikov,[67] R. H. Schindler,[67] J. Schwiening,[67] A. Snyder,[67] D. Su,[67] M. K. Sullivan,[67] K. Suzuki,[67] S. K. Swain,[67] J. M. Thompson,[67] J. Va'vra,[67] A. P. Wagner,[67] M. Weaver,[67] C. A. West,[67] W. J. Wisniewski,[67] M. Wittgen,[67] D. H. Wright,[67] H. W. Wulsin,[67] A. K. Yarritu,[67] K. Yi,[67] C. C. Young,[67] V. Ziegler,[67] P. R. Burchat,[68] A. J. Edwards,[68] T. S. Miyashita,[68] S. Ahmed,[69] M. S. Alam,[69] J. A. Ernst,[69] B. Pan,[69] M. A. Saeed,[69] S. B. Zain,[69] S. M. Spanier,[70] B. J. Wogsland,[70] R. Eckmann,[71] J. L. Ritchie,[71] A. M. Ruland,[71] C. J. Schilling,[71] R. F. Schwitters,[71] B. W. Drummond,[72] J. M. Izen,[72] X. C. Lou,[72] F. Bianchi,[73a,73b] D. Gamba,[73a,73b] M. Pelliccioni,[73a,73b] M. Bomben,[74a,74b] L. Bosisio,[74a,74b] C. Cartaro,[74a,74b] G. Della Ricca,[74a,74b] L. Lanceri,[74a,74b] L. Vitale,[74a,74b] V. Azzolini,[75] N. Lopez-March,[75] F. Martinez-Vidal,[75] D. A. Milanes,[75] A. Oyanguren,[75] J. Albert,[76] Sw. Banerjee,[76] B. Bhuyan,[76] H. H. F. Choi,[76] K. Hamano,[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] 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]

(*BABAR* Collaboration)

[1]*Laboratoire de Physique des Particules, IN2P3/CNRS et Université de Savoie, F-74941 Annecy-Le-Vieux, France*
[2]*Universitat de Barcelona, Facultat de Fisica, Departament ECM, E-08028 Barcelona, Spain*
[3a]*INFN Sezione di Bari, I-70126 Bari, Italy*
[3b]*Dipartmento di Fisica, Università di Bari, 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, V6T 1Z1 Canada*
[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 Los Angeles, Los Angeles, California 90024, USA*
[13]*University of California at Riverside, Riverside, California 92521, USA*
[14]*University of California at San Diego, La Jolla, California 92093, USA*
[15]*University of California at Santa Barbara, Santa Barbara, California 93106, USA*
[16]*University of California at Santa Cruz, Institute for Particle Physics, Santa Cruz, California 95064, USA*
[17]*California Institute of Technology, Pasadena, California 91125, USA*
[18]*University of Cincinnati, Cincinnati, Ohio 45221, USA*
[19]*University of Colorado, Boulder, Colorado 80309, USA*
[20]*Colorado State University, Fort Collins, Colorado 80523, USA*
[21]*Technische Universität Dortmund, Fakultät Physik, D-44221 Dortmund, Germany*
[22]*Technische Universität Dresden, Institut für Kern- und Teilchenphysik, D-01062 Dresden, Germany*
[23]*Laboratoire Leprince-Ringuet, CNRS/IN2P3, Ecole Polytechnique, F-91128 Palaiseau, France*
[24]*University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom*
[25a]*INFN Sezione di Ferrara, I-44100 Ferrara, Italy*
[25b]*Dipartimento di Fisica, Università di Ferrara, I-44100 Ferrara, Italy*
[26]*INFN Laboratori Nazionali di Frascati, I-00044 Frascati, Italy*
[27a]*INFN Sezione di Genova, I-16146 Genova, Italy*
[27b]*Dipartimento di Fisica, Università di Genova, I-16146 Genova, Italy*
[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, H3A 2T8 Canada
[47a]INFN Sezione di Milano, I-20133 Milano, Italy
[47b]Dipartimento di Fisica, Università di Milano, I-20133 Milano, Italy
[48]University of Mississippi, University, Mississippi 38677, USA
[49]Université de Montréal, Physique des Particules, Montréal, Québec, H3C 3J7 Canada
[50]Mount Holyoke College, South Hadley, Massachusetts 01075, USA
[51a]INFN Sezione di Napoli, I-80126 Napoli, Italy
[51b]Dipartimento di Scienze Fisiche, Università di Napoli Federico II, I-80126 Napoli, Italy
[52]NIKHEF, National Institute for Nuclear Physics and High Energy Physics, NL-1009 DB Amsterdam, The Netherlands
[53]University of Notre Dame, Notre Dame, Indiana 46556, USA
[54]Ohio State University, Columbus, Ohio 43210, USA
[55]University of Oregon, Eugene, Oregon 97403, USA
[56a]INFN Sezione di Padova, I-35131 Padova, Italy
[56b]Dipartimento di Fisica, Università di Padova, I-35131 Padova, Italy
[57]Laboratoire de Physique Nucléaire et de Hautes Energies, IN2P3/CNRS, Université Pierre et Marie Curie-Paris6, Université Denis Diderot-Paris7, F-75252 Paris, France
[58]University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA
[59a]INFN Sezione di Perugia, I-06100 Perugia, Italy
[59b]Dipartimento di Fisica, Università di Perugia, I-06100 Perugia, Italy
[60a]INFN Sezione di Pisa, I-56127 Pisa, Italy
[60b]Dipartimento di Fisica, Università di Pisa, I-56127 Pisa, Italy
[60c]Scuola Normale Superiore di Pisa, I-56127 Pisa, Italy
[61]Princeton University, Princeton, New Jersey 08544, USA
[62a]INFN Sezione di Roma, I-00185 Roma, Italy
[62b]Dipartimento di Fisica, Università di Roma La Sapienza, I-00185 Roma, Italy
[63]Universität Rostock, D-18051 Rostock, Germany
[64]Rutherford Appleton Laboratory, Chilton, Didcot, Oxon, OX11 0QX, United Kingdom
[65]CEA, Irfu, SPP, Centre de Saclay, F-91191 Gif-sur-Yvette, France
[66]University of South Carolina, Columbia, South Carolina 29208, USA
[67]Stanford Linear Accelerator Center, Stanford, California 94309, USA
[68]Stanford University, Stanford, California 94305-4060, USA
[69]State University of New York, Albany, New York 12222, USA
[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
[73a]INFN Sezione di Torino, I-10125 Torino, Italy
[73b]Dipartimento di Fisica Sperimentale, Università di Torino, I-10125 Torino, Italy
[74a]INFN Sezione di Trieste, I-34127 Trieste, Italy
[74b]Dipartimento di Fisica, Università di Trieste, I-34127 Trieste, Italy
[75]IFIC, Universitat de Valencia-CSIC, E-46071 Valencia, Spain
[76]University of Victoria, Victoria, British Columbia, V8W 3P6 Canada
[77]Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom

[*]Deceased.
[†]Now at Temple University, Philadelphia, Pennsylvania 19122, USA.
[‡]Now at Tel Aviv University, Tel Aviv, 69978, Israel.
[§]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 Università di Sassari, Sassari, Italy.

[78]*University of Wisconsin, Madison, Wisconsin 53706, USA*

(Received 16 February 2009; published 13 May 2009)

We present measurements of the semileptonic decays $B^- \to D^0\tau^-\bar{\nu}_\tau$, $B^- \to D^{*0}\tau^-\bar{\nu}_\tau$, $\bar{B}^0 \to D^+\tau^-\bar{\nu}_\tau$, and $\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$, which are sensitive to non-standard model amplitudes in certain scenarios. The data sample consists of 232×10^6 $\Upsilon(4S) \to B\bar{B}$ decays collected with the *BABAR* detector at the PEP-II e^+e^- collider. We select events with a D or D^* meson and a light lepton ($\ell = e$ or μ) recoiling against a fully reconstructed B meson. We perform a fit to the joint distribution of lepton momentum and missing mass squared to distinguish signal $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau(\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau)$ events from the backgrounds, predominantly $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$. We measure the branching-fraction ratios $R(D) \equiv \mathcal{B}(\bar{B} \to D\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B} \to D\ell^-\bar{\nu}_\ell)$ and $R(D^*) \equiv \mathcal{B}(\bar{B} \to D^*\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B} \to D^*\ell^-\bar{\nu}_\ell)$ and, from a combined fit to B^- and $\bar{B}^0$ channels, obtain the results $R(D) = (41.6 \pm 11.7 \pm 5.2)\%$ and $R(D^*) = (29.7 \pm 5.6 \pm 1.8)\%$, where the uncertainties are statistical and systematic. Normalizing to measured $B^- \to D^{(*)0}\ell^-\bar{\nu}_\ell$ branching fractions, we obtain $\mathcal{B}(\bar{B} \to D\tau^-\bar{\nu}_\tau) = (0.86 \pm 0.24 \pm 0.11 \pm 0.06)\%$ and $\mathcal{B}(\bar{B} \to D^*\tau^-\bar{\nu}_\tau) = (1.62 \pm 0.31 \pm 0.10 \pm 0.05)\%$, where the additional third uncertainty is from the normalization mode. We also present, for the first time, distributions of the lepton momentum, $|\mathbf{p}^*_\ell|$, and the squared momentum transfer, q^2.

DOI: 10.1103/PhysRevD.79.092002 PACS numbers: 12.15.Hh, 13.20.−v, 13.20.He, 14.40.Nd

I. INTRODUCTION

Semileptonic decays of B mesons to the τ lepton—the heaviest of the three charged leptons—provide a new source of information on standard model (SM) processes [1–3], as well as a new window on physics beyond the SM [4–9]. In the SM, semileptonic decays occur at tree level and are mediated by the W boson, but the large mass of the τ lepton provides sensitivity to additional amplitudes, such as those mediated by a charged Higgs boson. Experimentally, $b \to c\tau^-\bar{\nu}_\tau$ decays[1] are challenging to study because the final state contains not just one, but two or three neutrinos as a result of the τ decay.

Theoretical predictions for semileptonic decays to exclusive final states require knowledge of the form factors, which parametrize the hadronic current as functions of $q^2 = [p_B - p_{D^{(*)}}]^2$. For light leptons $\ell \equiv e$, μ,[2] there is effectively one form factor for $\bar{B} \to D\ell^-\bar{\nu}_\ell$, while there are three for $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$. If a τ lepton is produced instead, one additional form factor enters in each mode. The form factors for $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays[3] involving the light leptons have been measured [10–12], providing direct information on four of the six form factors. Heavy quark symmetry (HQS) relations [13] allow one to express the two additional form factors for $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ in terms of the form factors measurable from decays with the light leptons. With sufficient data, one could probe the additional form factors and test the HQS relations.

Branching fractions for semileptonic B decays to τ leptons are predicted to be smaller than those to light leptons. Calculations based on the SM predict $\mathcal{B}(\bar{B}^0 \to D^+\tau^-\bar{\nu}_\tau) = (0.69 \pm 0.04)\%$ and $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau) = (1.41 \pm 0.07)\%$ [8], which account for most of the predicted inclusive rate $\mathcal{B}(\bar{B} \to X_c\tau^-\bar{\nu}_\tau) = (2.30 \pm 0.25)\%$ [2] (here, X_c represents all hadronic final states from the $b \to c$ transition). In multi-Higgs doublet models [4–8], substantial departures, either positive or negative, from the SM decay rate could occur for $\mathcal{B}(\bar{B} \to D\tau^-\bar{\nu}_\tau)$, while smaller departures are expected for $\mathcal{B}(\bar{B} \to D^*\tau^-\bar{\nu}_\tau)$. Thus, measurements of $\mathcal{B}(\bar{B} \to D\tau^-\bar{\nu}_\tau)$ are more sensitive to non-SM contributions than either $\mathcal{B}(\bar{B} \to D^*\tau^-\bar{\nu}_\tau)$ or the inclusive rate. In addition to the branching fractions, several other observables are sensitive to possible non-SM contributions, including q^2 distributions and D^* and τ polarization [4–6,8,14].

The first measurements of semileptonic b-hadron decays to τ leptons were performed by the LEP experiments [15] operating at the Z^0 resonance, yielding an average [16] inclusive branching fraction $\mathcal{B}(b_{\rm had} \to X\tau^-\bar{\nu}_\tau) = (2.48 \pm 0.26)\%$, where $b_{\rm had}$ represents the mixture of b-hadrons produced in $Z^0 \to b\bar{b}$ decays. The Belle experiment has reported $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau) = (2.02^{+0.40}_{-0.37} \pm 0.37)\%$ [17].

The *BABAR* Collaboration has presented a measurement of the branching fractions for $\bar{B} \to D\tau^-\bar{\nu}_\tau$ and $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ for both charged and neutral B mesons [18]. In this article, we describe the analysis in greater detail, with particular emphasis on several novel features of the event selection and fit technique. We also present distributions of two important kinematic variables, the lepton momentum $|\mathbf{p}^*_\ell|$ and the squared momentum transfer q^2.

A. Analysis overview and strategy

We determine the branching fractions of four exclusive decay modes: $B^- \to D^0\tau^-\bar{\nu}_\tau$, $B^- \to D^{*0}\tau^-\bar{\nu}_\tau$, $\bar{B}^0 \to D^+\tau^-\bar{\nu}_\tau$, and $\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$, each of which is measured as a branching-fraction ratio R relative to the corresponding e and μ modes. To reconstruct the τ, we use the decays $\tau^- \to e^-\bar{\nu}_e\nu_\tau$ and $\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$, which are experimentally the most accessible. The main challenge of the measurement is to distinguish $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays, which

[1]Charge-conjugate modes are implied throughout.

[2]Throughout this article, we use the symbol ℓ to refer only to the light charged leptons e and μ.

[3]The symbol $D^{(*)}$ refers either to a D or a D^* meson.

have three neutrinos, from $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays, which have the same observable final-state particles but only one neutrino.

The analysis strategy is to reconstruct the decays of both B mesons in the $\Upsilon(4S) \to B\bar{B}$ event, providing powerful constraints on unobserved particles. One B meson, denoted $B_{\rm tag}$, is fully reconstructed in a purely hadronic decay chain. The remaining charged particles and photons are required to be consistent with the products of a $b \to c$ semileptonic B decay: the daughter charm meson (either a D or D^*) and a lepton (e or μ). The lepton may be either primary or from $\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau$. To distinguish signal events from the normalization modes $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$, we calculate the missing four-momentum,

$$p_{\rm miss} = p_{e^+e^-} - p_{\rm tag} - p_{D^{(*)}} - p_\ell \qquad (1)$$

of any particles recoiling against the observed $B_{\rm tag} + D^{(*)}\ell$ system. A large peak at zero in $m^2_{\rm miss} = p^2_{\rm miss}$ corresponds to semileptonic decays with one neutrino, whereas signal events produce a broad tail out to $m^2_{\rm miss} \sim 8\ ({\rm GeV}/c^2)^2$.

To separate signal and background events, we perform a fit (described in Sec. VII) to the joint distribution of $m^2_{\rm miss}$ and the lepton momentum ($|\mathbf{p}^*_\ell|$) in the rest frame of the B meson. In signal events, the observed lepton is the daughter of the τ and typically has a soft spectrum; for most background events, this lepton typically has higher momentum. The fit is performed simultaneously in eight channels, with a set of constraints relating the event yields between the channels. The fit is designed to maximize the sensitivity to the $\bar{B} \to D\tau^-\bar{\nu}_\tau$ signals by using events in the $D^*\ell^-$ channels to constrain the dominant backgrounds, $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ feed-down, in which the final-state D^* meson is not completely reconstructed. Similarly, we use a set of D^{**} control samples to constrain the feed-down background to both the $D\tau^-\bar{\nu}_\tau$ and $D^*\tau^-\bar{\nu}_\tau$ signals.[4]

We perform a relative measurement, extracting both signal $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ and normalization $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ yields from the fit to obtain the four branching-fraction ratios $R(D^0)$, $R(D^+)$, $R(D^{*0})$, and $R(D^{*+})$, where, for example, $R(D^{*0}) \equiv \mathcal{B}(B^- \to D^{*0}\tau^-\bar{\nu}_\tau)/\mathcal{B}(B^- \to D^{*0}\ell^-\bar{\nu}_\ell)$. In the ratio, many systematic uncertainties cancel, either partially or completely. These ratios are normalized such that ℓ represents only one of e or μ; however, both light lepton species are included in the measurement. We multiply these branching-fraction ratios by previous measurements of $\mathcal{B}(\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell)$ to derive absolute branching fractions.

[4]Throughout this paper, we use the symbol D^{**} to represent all charm resonances heavier than the $D^*(2010)$, as well as nonresonant $D^{(*)}n\pi$ systems with $n \geq 1$.

II. THE *BABAR* DETECTOR AND DATA SETS

We analyze data collected with the *BABAR* detector at the PEP-II e^+e^- storage rings at the Stanford Linear Accelerator Center. PEP-II is an asymmetric-energy B factory, colliding 9.0 GeV e^- with 3.1 GeV e^+ at a center-of-mass energy of 10.58 GeV, corresponding to the $\Upsilon(4S)$ resonance. The data sample used consists of 208.9 fb^{-1} of integrated luminosity recorded on the $\Upsilon(4S)$ resonance between 1999 and 2004, yielding 232×10^6 $\Upsilon(4S) \to B\bar{B}$ decays. This data sample can be divided into two major periods: Runs 1–3, comprising 109.0 fb^{-1} taken from 1999 to June 2003, and Run 4, comprising 99.9 fb^{-1} taken from September 2003 to July 2004. The accelerator background conditions were significantly different between Runs 1–3 and Run 4, which could affect missing-energy analyses such as this one; for this reason, the two running periods have been independently validated, and the fraction of signal-like events found in the Run 4 sample is used as a cross-check of the results, as described in Sec. X.

The *BABAR* detector is a large, general-purpose magnetic spectrometer and is described in detail elsewhere [19]. Charged particle trajectories are measured in a tracking system consisting of a five-layer double-sided silicon strip detector and a 40-layer drift chamber, both of which operate in the 1.5 T magnetic field of a superconducting solenoid. A detector of internally reflected Cherenkov light (DIRC) is used to measure charged particle velocity for particle identification (PID). An electromagnetic calorimeter (EMC), consisting of 6580 CsI(Tl) crystals, is used to reconstruct photons and in electron identification. The steel flux return of the solenoid is segmented and instrumented with resistive plate chambers (IFR) for muon and neutral hadron identification.

All detector systems contribute to charged particle identification. Ionization energy losses in the tracking systems and the Cherenkov light signature in the DIRC are used for all charged particle types. Electrons are also identified on the basis of shower shape in the EMC and the ratio of energy deposited in the EMC to the track momentum. Muon identification is based on a minimum-ionization energy deposit in the EMC and on the measured interaction length in the IFR.

This analysis relies on measurement of the missing momentum carried off by multiple neutrinos, and the large solid angle coverage (hermeticity) of the detector is therefore crucial. The tracking system, calorimeter, and IFR cover the full azimuthal range and the polar angle range from approximately $0.3 < \theta < 2.7$ rad in the laboratory frame, corresponding to a $\Upsilon(4S)$ center-of-mass coverage of approximately 90% [the direction $\theta = 0$ corresponds to the direction of the high-energy beam, and therefore to the $\Upsilon(4S)$ boost]. The DIRC fiducial volume is slightly smaller, corresponding to a center-of-mass frame coverage of about 84%.

Within the active detector volume, the efficiency for reconstructing charged tracks and photons is very high, typically greater than 95% over most of the momentum range. At low momenta, however, the reconstruction efficiency drops off, leading to an increased contribution from feed-down processes to which special attention is paid throughout this analysis. Feed-down occurs when the photon from $D^* \to D\gamma$ or the π^0 from $D^* \to D\pi^0$ is not reconstructed (in the case of the π^0, either one or both of the photons from $\pi^0 \to \gamma\gamma$ may be missed). Care must therefore be taken to avoid confusing D^* feed-down events for D signals.

We use a Monte Carlo simulation (MC) of the production and decay of signal and background events based on EVTGEN [20]. A sample of simulated inclusive $B\bar{B}$ events equivalent to about five times the integrated luminosity is used to study backgrounds and to optimize event selection criteria. Large samples of many individual semileptonic B decays (discussed in Sec. III) are used to parameterize the distributions of variables used in the fit. Final-state radiation is simulated using PHOTOS [21]. Simulation of the detector response is performed with GEANT [22] and the resulting efficiencies and resolutions are validated in multiple data control samples.

III. SEMILEPTONIC DECAY MODELS

In the SM, the matrix element for a semileptonic B meson decay can be written as

$$\mathcal{M}(\bar{B} \to D^{(*)}(\ell^-/\tau^-)\bar{\nu}) = -i\frac{g^2}{8m_W^2} V_{cb} L^\mu H_\mu, \quad (2)$$

where g is the weak coupling constant, m_W is the W mass, V_{cb} is the quark mixing matrix element, and L^μ and H_μ are the leptonic and hadronic currents, respectively. Here, we have used a simplified form for the W propagator appropriate for energies much less than m_W. The leptonic current is exactly known,

$$L^\mu = \bar{u}_\ell \gamma^\mu (1 - \gamma_5) v_\nu, \quad (3)$$

and the hadronic current is given by

$$H_\mu = \langle D^{(*)} | \bar{c}\gamma_\mu (1 - \gamma_5) b | B \rangle. \quad (4)$$

In the case of a $\bar{B} \to D$ transition, the axial-vector part of the current does not contribute to the decay, and we may write the hadronic current in terms of two form factors $f_+(q^2)$ and $f_-(q^2)$:

$$\langle D | V^\mu | B \rangle = (p + p')^\mu f_+(q^2) + (p - p')^\mu f_-(q^2), \quad (5)$$

with $V^\mu \equiv \bar{c}\gamma^\mu b$ and where p and p' are the four-momenta of the B and D mesons, respectively. For the $\bar{B} \to D^*$ transition, the axial-vector term contributes to the decay as well, and we write the hadronic current in terms of form factors $V(q^2)$, $A_1(q^2)$, $A_2(q^2)$, $A_3(q^2)$, and $A_0(q^2)$:

$$\begin{aligned}\langle D^* | V^\mu - A^\mu | B \rangle = {} & \frac{2i\epsilon^{\mu\nu\alpha\beta}}{m_B + m_{D^*}} \varepsilon^*_\nu p'_\alpha p_\beta V(q^2) \\ & - (m_B + m_{D^*}) \varepsilon^{*\mu} A_1(q^2) + \frac{\varepsilon^* \cdot q}{m_B + m_{D^*}} \\ & \times (p + p')^\mu A_2(q^2) + 2m_{D^*} \frac{\varepsilon^* \cdot q}{q^2} \\ & \times q^\mu A_3(q^2) - 2m_{D^*} \frac{\varepsilon^* \cdot q}{q^2} q^\mu A_0(q^2),\end{aligned} \quad (6)$$

where $A^\mu \equiv \bar{c}\gamma^\mu\gamma_5 b$ and ε is the D^* polarization vector. The form factor $A_3(q^2)$ is related to two other form factors as

$$A_3(q^2) = \frac{m_B + m_{D^*}}{2m_{D^*}} A_1(q^2) - \frac{m_B - m_{D^*}}{2m_{D^*}} A_2(q^2) \quad (7)$$

so that there are only four independent form factors.

In the limit of massless leptons, any terms proportional to $q^\mu \equiv (p - p')^\mu$ vanish when the hadronic current is contracted with the leptonic current. For this reason, the contributions from the form factors $f_-(q^2)$ and $A_0(q^2)$ are essentially negligible for electrons and muons, as mentioned above.

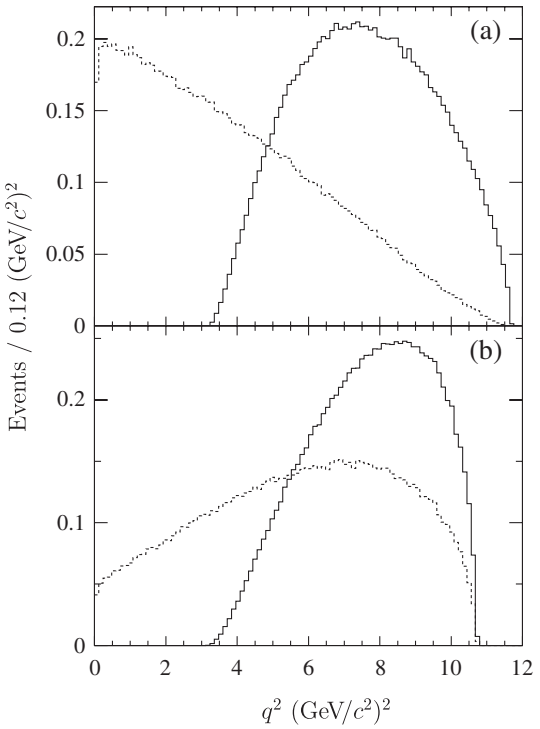

FIG. 1. Generated q^2 distributions for (a) $\bar{B} \to D\ell^-\bar{\nu}_\ell$ and $\bar{B} \to D\tau^-\bar{\nu}_\tau$; (b) $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ and $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$. The two curves in each plot show q^2 for the light lepton (dashed line) and for the τ (solid line). All distributions use the CLN form factor model with experimentally measured shape parameters. The distributions are normalized to equal areas.

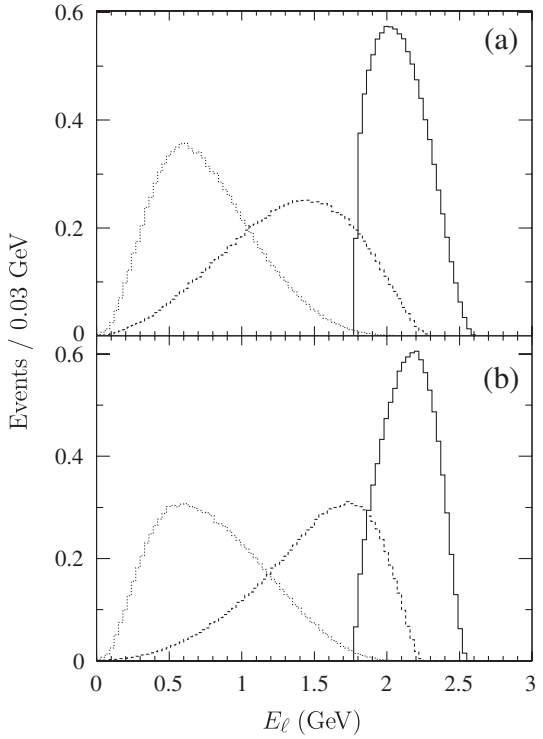

FIG. 2. Generated lepton energy distributions for (a) $\bar{B} \to D\ell^-\bar{\nu}_\ell$ and $\bar{B} \to D\tau^-\bar{\nu}_\tau$; (b) $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ and $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$. The three curves in each plot show the ℓ^- energy in $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ (dashed line), the τ^- energy in $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ (solid line), and the secondary lepton energy in $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ (dotted line), all defined in the B meson rest frame. All distributions use the CLN form factor model with experimentally measured shape parameters. The distributions are normalized to equal areas.

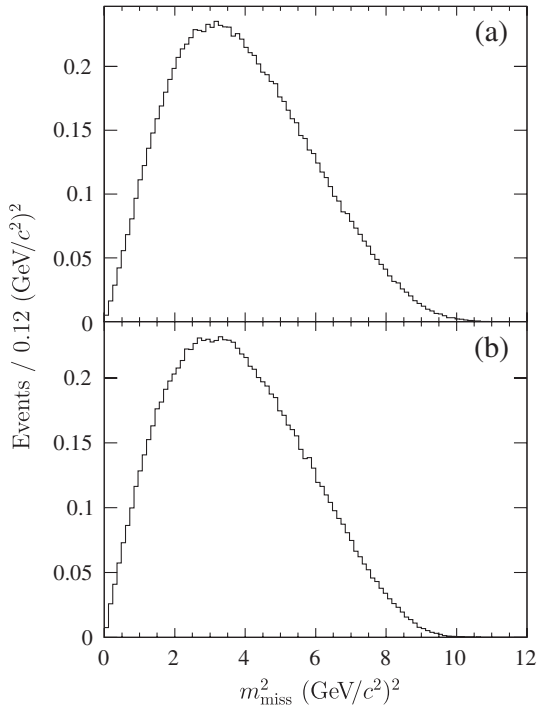

FIG. 3. Generated m^2_{miss} distributions for (a) $\bar{B} \to D\tau^-\bar{\nu}_\tau$ and (b) $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$. Both distributions use the CLN form factor model with experimentally measured shape parameters.

Semileptonic decays are simulated using the ISGW2 model [23], except for $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ decays, which use an HQET model with a linear form factor expansion [24], and nonresonant $\bar{B} \to D^{(*)}\pi\ell^-\bar{\nu}_\ell$ decays, which use the model of Goity and Roberts [25]. We reweight both signal $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ and normalization $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ events [26] so that the decay distributions follow the Caprini-Lellouch-Neubert (CLN) form factor model [27] with parameters measured in data. We use $\rho^2_+ = 1.17 \pm 0.18$ [28] for $\bar{B} \to D\ell^-\bar{\nu}_\ell$ and $\bar{B} \to D\tau^-\bar{\nu}_\tau$ decays, and we use $R_1 = 1.417 \pm 0.061 \pm 0.044$, $R_2 = 0.836 \pm 0.037 \pm 0.022$, and $\rho^2_{A_1} = 1.179 \pm 0.048 \pm 0.028$ [11] for $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ and $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decays.[5] Variation of these form factors is taken into account as a systematic uncertainty, including the correlations between the three $\bar{B} \to D^*$ form factor parameters.

Figures 1–3 show distributions of three kinematic variables important to this analysis, all generated using the CLN form factor parameterization with parameters given above. Figure 1 compares q^2 distributions between the signal and normalization modes. Signal events must satisfy $q^2 > m^2_\tau$, leading to qualitatively different q^2 spectra for signal and normalization events; this feature is exploited in the event selection and in validation studies. Figure 2 shows distributions of lepton energy in the B meson rest frame. While the τ^- lepton in signal events typically has high energy (due to its mass), the secondary lepton ℓ^- typically has much lower energy than either the τ^- or the primary lepton in $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ events. This low lepton energy leads to a lower reconstruction efficiency for signal leptons than those in the normalization modes. Figure 3 shows distributions of m^2_{miss} for the two signal modes, which, due to the three neutrinos in these events, forms a broad structure up to very large m^2_{miss}.

IV. EVENT RECONSTRUCTION AND SELECTION

All event selection requirements (as well as the fit procedure described in Sec. VII) are defined using simulated events or using control samples in data that exclude the signal region in order to avoid any potential sources of bias. About 60% of the $B\bar{B}$ MC sample is used in optimizing the event selection, while the remaining 40% is used as

[5]The parameters R_1 and R_2 are not included in the model of Caprini, Lellouch, and Neubert [27]; to model the $\bar{B} \to D^*$ form factors, we adopt the formalism used in [12], Eqs. (13) and (14), where the leading terms in these form factor ratio expansions are taken as free parameters. We use independent slope parameters ρ^2_+ and $\rho^2_{A_1}$ for the $\bar{B} \to D$ and $\bar{B} \to D^*$ form factors, respectively, treating the two sets of form factors as uncorrelated.

an independent validation of the selection and fitting procedures.

Most of the selection criteria described here are optimized to maximize the quantity $S/\sqrt{S+B}$, where S and B are the expected signal and background yields in the large $m^2_{\rm miss}$ region of our data sample, assuming standard model branching fractions for signal decays. The requirement on ΔE of the $B_{\rm tag}$ candidate (defined below) was initially optimized in the same way, but was tightened because fits to MC samples indicated that events at large $|\Delta{\rm E}|$ contributed to biases in the signal extraction. The final selection corresponds to a compromise between the statistical $S/\sqrt{S+B}$ optimization and the systematic effects due to this bias.

A. $B_{\rm tag}$ Reconstruction

We reconstruct $B_{\rm tag}$ candidates in 1114 final states $B_{\rm tag} \to D^{(*)}Y^{\pm}$ with an algorithm that has been used previously at *BABAR* for a number of analyses, especially those dependent on measuring missing momentum [29]. These final states arise from the large number of ways to reconstruct the D and D^* mesons within the $B_{\rm tag}$ candidate and the possible pion and kaon combinations within the $Y^{\pm}$ system. Tag-side D candidates are reconstructed as $D^0_{\rm tag} \to K^-\pi^+$, $K^-\pi^+\pi^0$, $K^-\pi^+\pi^+\pi^-$, and $K^0_S\pi^+\pi^-$, and as $D^+_{\rm tag} \to K^-\pi^+\pi^+$, $K^-\pi^+\pi^+\pi^0$, $K^0_S\pi^+$, $K^0_S\pi^+\pi^-\pi^+$, and $K^0_S\pi^+\pi^0$. Tag-side D^* candidates are reconstructed as $D^{*0}_{\rm tag} \to D^0_{\rm tag}\pi^0$ and $D^0_{\rm tag}\gamma$ and as $D^{*+}_{\rm tag} \to D^0_{\rm tag}\pi^+$. The $Y^{\pm}$ system may consist of up to six light hadrons ($\pi^{\pm}$, π^0, $K^{\pm}$, or K^0_S). In both the $D^{(*)}_{\rm tag}$ and $Y^{\pm}$ systems, we reconstruct $\pi^0 \to \gamma\gamma$ and $K^0_S \to \pi^+\pi^-$ and require charged kaon candidates to satisfy PID criteria (loose criteria for $D^0 \to K^-\pi^+$, tight for all other modes.[6]) $D^{(*)}_{\rm tag}$ candidates are selected within about 2σ (standard deviations) of the nominal mass, with σ depending on the reconstruction mode and typically 5–10 MeV/c^2 for the $D_{\rm tag}$ mass and 1–2 MeV/c^2 for the $D^*_{\rm tag} - D_{\rm tag}$ mass difference.

We use two kinematic variables to identify $B_{\rm tag}$ candidates,

$$m_{\rm ES} = \sqrt{s/4 - |\mathbf{p}_{\rm tag}|^2} \tag{8}$$

and

$$\Delta E = E_{\rm tag} - \sqrt{s}/2, \tag{9}$$

where $\sqrt{s}$ is the total e^+e^- energy, $|\mathbf{p}_{\rm tag}|$ is the magnitude of the $B_{\rm tag}$ momentum, and $E_{\rm tag}$ is the $B_{\rm tag}$ energy, all defined in the e^+e^- center-of-mass frame. For correctly reconstructed $B_{\rm tag}$ candidates, $m_{\rm ES}$ is equal to the B meson mass, with a resolution of about 2.5 MeV/c^2, and ΔE is equal to zero, with a resolution of about 18 MeV.

For each $D^{(*)}_{\rm tag}$ "seed" candidate, we use a recursive algorithm to identify candidate $Y^{\pm}$ systems. Light hadrons from the remaining tracks and photons in the event are added to the $Y^{\pm}$ system, one at a time. If the resulting values of $m_{\rm ES}$ and ΔE for the $D^{(*)}_{\rm tag}Y^{\pm}$ candidate are close to the nominal values, the $B_{\rm tag}$ candidate is accepted. If the value of ΔE is too large, the light hadron just added is removed from the $Y^{\pm}$ system, since continuing to add particles to this $Y^{\pm}$ candidate will increase ΔE further. The algorithm then continues recursively with the remaining particles in the event, adding and removing light hadrons to the $Y^{\pm}$ system according to $m_{\rm ES}$, ΔE, and the $Y^{\pm}$ system topology. This algorithm is semiexclusive, meaning that particles in the $Y^{\pm}$ system are not constrained to intermediate resonance states. Because of this, the yield is significantly higher than exclusive B reconstruction, while the purity is somewhat lower. In this analysis, however, since we exclusively reconstruct the second B meson in the event, the purity of our final sample is substantially improved with respect to the raw $B_{\rm tag}$ sample.

We require $m_{\rm ES} > 5.27$ GeV/c^2 and $|\Delta E| < 72$ MeV, corresponding to $\pm 4\sigma$ in ΔE and -4σ in $m_{\rm ES}$ (the kinematic limit $m_{\rm ES} < \sqrt{s}/2$ provides an effective $+4\sigma$ requirement). We reconstruct $B_{\rm tag}$ candidates with an efficiency of 0.2% to 0.3%. Figure 4 shows distributions of $m_{\rm ES}$ for selected $B_{\rm tag}$ candidates both before and after the signal-side reconstruction. We make no attempt at this stage to select a single $B_{\rm tag}$ among multiple reconstructed candidates: this decision is made after reconstructing the signal side as well.

B. Reconstruction of the Signal B

For the B meson decaying semileptonically, we reconstruct $D^{(*)}$ candidates in the modes $D^0 \to K^-\pi^+$, $K^-\pi^+\pi^0$, $K^-\pi^+\pi^+\pi^-$, $K^0_S\pi^+\pi^-$; $D^+ \to K^-\pi^+\pi^+$, $K^-\pi^+\pi^+\pi^0$, $K^0_S\pi^+$, $K^-K^+\pi^+$; $D^{*0} \to D^0\pi^0$, $D^0\gamma$; and $D^{*+} \to D^0\pi^+$, $D^+\pi^0$. We reconstruct K^0_S mesons as $K^0_S \to \pi^+\pi^-$ with $491 < m_{\pi^+\pi^-} < 506$ MeV/c^2, corresponding to $\pm 3\sigma$. We reconstruct π^0 mesons as $\pi^0 \to \gamma\gamma$, requiring $90 < m_{\gamma\gamma} < 165$ MeV/c^2 for the soft π^0 used to reconstruct $D^* \to D\pi^0_{\rm soft}$, and requiring $E_{\gamma\gamma} > 200$ MeV and $125 < m_{\gamma\gamma} < 145$ MeV/c^2 for a π^0 used to reconstruct a D meson; the mass intervals correspond to $\pm 3\sigma$ in both cases and are different because the resolution is poorer at low energies. Charged kaon candidates are required to satisfy tight PID criteria with a typical efficiency of 85% while rejecting 98% of pions. Charged pion candidates are required to satisfy loose PID criteria with a typical effi-

[6]The terms "loose" and "tight" refer to the relative signal-to-background discrimination of various PID criteria. Loose criteria are chosen to have high efficiency, and have relatively high background rates as well; tight criteria have lower background but also lower signal efficiency. The optimal choice of criteria depends on the particle type and on the *a priori* purity of the sample, and is therefore different for each reconstruction channel.

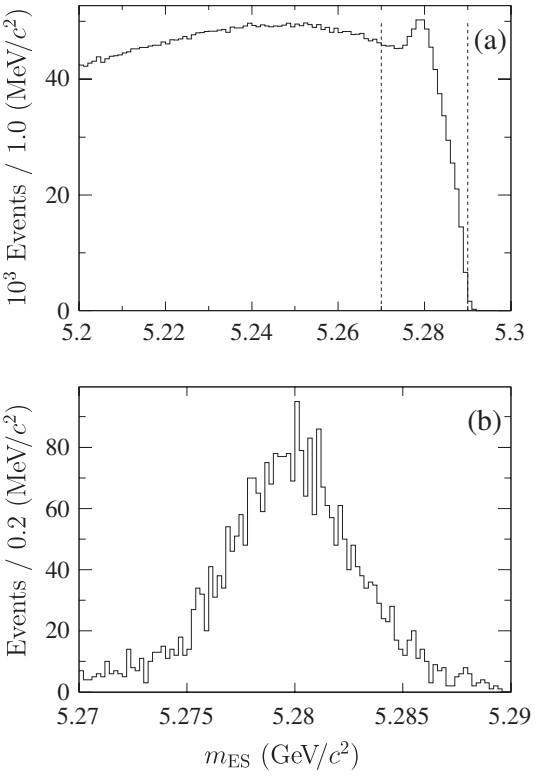

FIG. 4. Distributions of $m_{\rm ES}$ in selected events. (a) shows all $B_{\rm tag}$ candidates reconstructed in 20 fb^{-1} of data, and the purity of this plot has been increased by requiring $|\Delta E| < 50$ MeV. (b) shows the distribution for the complete data sample after the signal B and total-event selection requirements. Note the substantial improvement in purity due to the complete reconstruction.

ciency of 97% while rejecting 88% of kaons. D (D^*) candidates are selected within 4σ of the D mass ($D^* - D$ mass difference); as on the tag side, σ is typically 5–10 MeV/c^2 (1–2 MeV/c^2).

Electron candidates are required to satisfy tight PID criteria and to have lab-frame momentum $|\mathbf{p}_e| > 300$ MeV/c, with an efficiency that rises from 85% at the lowest momenta to 95% for $|\mathbf{p}_e| > 1.0$ GeV/c. Muon candidates are required to satisfy tight PID criteria; since muon PID relies on the hit pattern in the IFR, this effectively requires $|\mathbf{p}_\mu| \gtrsim 600$ MeV/c, and results in an efficiency of 40%–60% over the allowed momentum range. The energy of electron candidates is corrected for bremsstrahlung energy loss if photons are found close to the electron direction. Lepton candidates of either flavor are required to have at least 12 hits in the drift chamber and to have a laboratory-frame polar angle $0.4 < \theta < 2.6$ rad (excluding the very forward and very backward regions of the tracking system) in order to ensure a well-measured momentum, since mismeasured lepton momenta distort the $m^2_{\rm miss}$ distribution and tend to move background events into the signal-like region. Approximately 5% of selected lepton candidates are misidentified, almost all of which are pions misreconstructed as muons.

C. Total-event and single-candidate selection

We form whole-event candidates by combining $B_{\rm tag}$ candidates with $D^{(*)}\ell^-$ candidate systems. We combine charged $B_{\rm tag}$ candidates with $D^{(*)0}\ell^-$ systems and neutral $B_{\rm tag}$ candidates with both $D^{(*)+}\ell^-$ and $D^{(*)-}\ell^+$ systems, where the inclusion of both charge combinations allows for neutral B mixing.

In correctly reconstructed signal and normalization events, all of the stable final-state particles, with the exception of the neutrinos, are associated with either the $B_{\rm tag}$, $D^{(*)}$, or ℓ^- candidate. Events with additional particles in the final state must therefore have been misreconstructed, and we suppress these backgrounds with two selection requirements on the ''extra'' particle content in the event. We require that all observed charged tracks be associated with either the $B_{\rm tag}$, $D^{(*)}$, or ℓ candidate. We compute $E_{\rm extra}$, the sum of the energies of all photon candidates not associated with the $B_{\rm tag} + D^{(*)}\ell$ candidate system, and we require $E_{\rm extra} < 150$–300 MeV, depending on the $D^{(*)}$ channel. When considering these extra tracks and extra photons, care is taken to reject track and photon candidates which are likely to be due to accelerator background, electronics noise, or reconstruction software failures; fake photons in the EMC are, to some degree, unavoidable, which is why we can not simply require $E_{\rm extra} = 0$. The different D modes have very different levels of combinatorial background, which the $E_{\rm extra}$ cut is particularly effective at rejecting. Figure 5 shows distributions of $E_{\rm extra}$ for simulated signal and normalization events. Excellent agreement is seen in the two distributions, indicating that the efficiency of a cut on $E_{\rm extra}$ will largely cancel when we measure the branching-fraction ratio; we observe the same level of agreement in the four $D^{(*)}\ell^-$ channels separately, as well as in the e and μ final states separately.

We suppress hadronic events and combinatorial backgrounds by requiring $|\mathbf{p}_{\rm miss}| > 200$ MeV/c, where $|\mathbf{p}_{\rm miss}|$ is the magnitude of the missing momentum. This requirement mainly rejects hadronic events such as $\bar{B} \to D^{(*)}\pi^-$, where the π^- is misidentified as a μ^-. Our selection rejects more than 99% of $\bar{B} \to D^{(*)}\pi^-$ background, while rejecting less than 1% of signal and other semileptonic events.

We further suppress background by requiring $q^2 > 4$ (GeV/c^2)^2, where q^2 is calculated as

$$q^2 = [p_B - p_{D^{(*)}}]^2 = [p_{e^+e^-} - p_{\rm tag} - p_{D^{(*)}}]^2. \quad (10)$$

This requirement preferentially rejects combinatorial backgrounds from two-body B decays such as $B \to D^{(*)}D$, where one D meson decays semileptonically (or, in the case of a D_s^+, leptonically as $D_s^+ \to \tau^+\nu_\tau$). Our

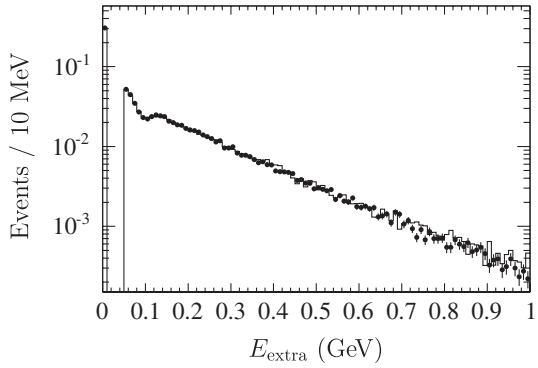

FIG. 5. Distributions of $E_{\rm extra}$ in simulated events, shown after all event selection except the cut on $E_{\rm extra}$ itself. Signal $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ events are shown as points, while $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ normalization events are shown as the histogram. The gap between the $E_{\rm extra} = 0$ bin and the remainder of the distribution corresponds to the minimum allowed photon energy, 50 MeV, while the bump at 100 MeV is due to events with two minimum-energy photons. The normalization is arbitrary. The agreement between the two distributions indicates that the efficiency of a cut on $E_{\rm extra}$ will cancel when we measure the branching-fraction ratio.

selection rejects about 25% of these backgrounds, while the signal efficiency is about 98% because signal events automatically satisfy $q^2 > m_\tau^2 \approx 3.16\ (\mathrm{GeV}/c^2)^2$. For $\bar{B} \to D\ell^-\bar{\nu}_\ell$ decays, the q^2 distribution peaks near zero (see Fig. 1), so this selection has an efficiency of about 60% for this normalization mode; for $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ decays, the q^2 distribution peaks at higher values, so our efficiency is about 70%. The q^2 requirement is the main reason why the reconstruction efficiency is different for signal and normalization modes, as seen below.

If multiple candidate systems pass our selection in a given event, we select the one with the lowest value of $E_{\rm extra}$. This scheme preferentially selects the candidate that is least likely to have lost additional particles. The main effect of this algorithm is that a candidate in one of the $D^*\ell^-$ channels will be selected before a candidate in one of the $D\ell^-$ channels when both candidates are present in an event. Because $D^* \to D$ feed-down is a dominant background while $D \to D^*$ feed-up is comparatively rare, keeping as many true D^* events in the $D^*\ell^-$ reconstruction channels helps to increase the sensitivity to the $D\tau^-\bar{\nu}_\tau$ signals.

To improve the resolution on the missing momentum, we perform a kinematic fit [30] to all $\Upsilon(4S) \to B_{\rm tag}D^{(*)}\ell^-$ candidates. We constrain charged track daughters of K_S^0, D, and B mesons to originate from common vertices, and we constrain the $\Upsilon(4S) \to B\bar{B}$ vertex to be consistent with the measured *BABAR* beamspot location. We constrain the mass of the signal D meson (and D^* meson, if there is one) to the measured value [16], and the combined momentum of the two B mesons to be consistent with the measured beam energy.

D. D^{**} Control sample selection

We select four control samples to constrain the poorly known $\bar{B} \to D^{**}(\ell^-/\tau^-)\bar{\nu}$ background. The selection is identical to that of the signal channels, but we require the presence of a π^0 meson in addition to the $B_{\rm tag} + D^{(*)}\ell$ system. The π^0 candidate must have momentum greater than 400 MeV/c, and the event must satisfy $E_{\rm extra} <$ 500 MeV, where the two photons from $\pi^0 \to \gamma\gamma$ are excluded from the calculation of $E_{\rm extra}$.

Most of the D^{**} background in the four signal channels occurs when the π^0 from $D^{**} \to D^{(*)}\pi^0$ is not reconstructed, so these control samples provide a direct normalization of the background source. Similar D^{**} decays in which a $\pi^\pm$ is lost contribute very little to the background since they do not have the correct charge correlation between the $B_{\rm tag}$ and $D^{(*)}$ candidate, and decays with two missing charged pions, which may have the correct charge correlation, have very low reconstruction efficiencies. The feed-down probabilities for the $D^{**}(\ell^-/\tau^-)\bar{\nu}$ background are determined from simulation, with uncertainties in the D^{**} content treated as a systematic error as described in Sec. IX A 3.

V. SELECTED EVENT SAMPLES

After applying all of the criteria above, we select a total of 3196 data events, 2886 in the four signal channels and 310 in the D^{**} control samples, as listed in Table I. Since most of the events at large $m_{\rm miss}^2$ are either $D\tau^-\bar{\nu}_\tau$ or $D^*\tau^-\bar{\nu}_\tau$ signal events, the third column in the table gives a first indication of where our sensitivity comes from. There are more events in the two B^- channels, $D^0\ell^-$ and $D^{*0}\ell^-$, due to a larger efficiency to reconstruct charged $B_{\rm tag}$ candidates than neutral ones and, to a lesser extent, a larger efficiency to reconstruct D^0 mesons on the signal side than D^+ mesons. There are more events in the D channels than the D^* channels, particularly at large $m_{\rm miss}^2$, because these channels contain both D mesons and D^* feed-down. The greatest signal sensitivity therefore comes from the $D^0\ell^-$ channel.

Figure 6 shows distributions of $|\mathbf{p}_\ell^*|$ versus $m_{\rm miss}^2$ for the selected data samples. One-dimensional distributions of $m_{\rm miss}^2$ and $|\mathbf{p}_\ell^*|$ for these samples are shown when we discuss the signal fit in Sec. VII.

TABLE I. Number of selected data events in the four signal channels, $N_{\rm ev}$, and in the four D^{**} control samples, $N_{D^{**}\rm CS}$. Here, the large $m_{\rm miss}^2$ region is taken to be $m_{\rm miss}^2 > 2\ (\mathrm{GeV}/c^2)^2$ and corresponds to the region with greatest signal sensitivity.

Channel	$N_{\rm ev}$	$N_{\rm ev}$ (large $m_{\rm miss}^2$)	$N_{D^{**}\rm CS}$
$D^0\ell^-$	1403	121	137
$D^{*0}\ell^-$	790	43	77
$D^+\ell^-$	295	36	66
$D^{*+}\ell^-$	398	14	30

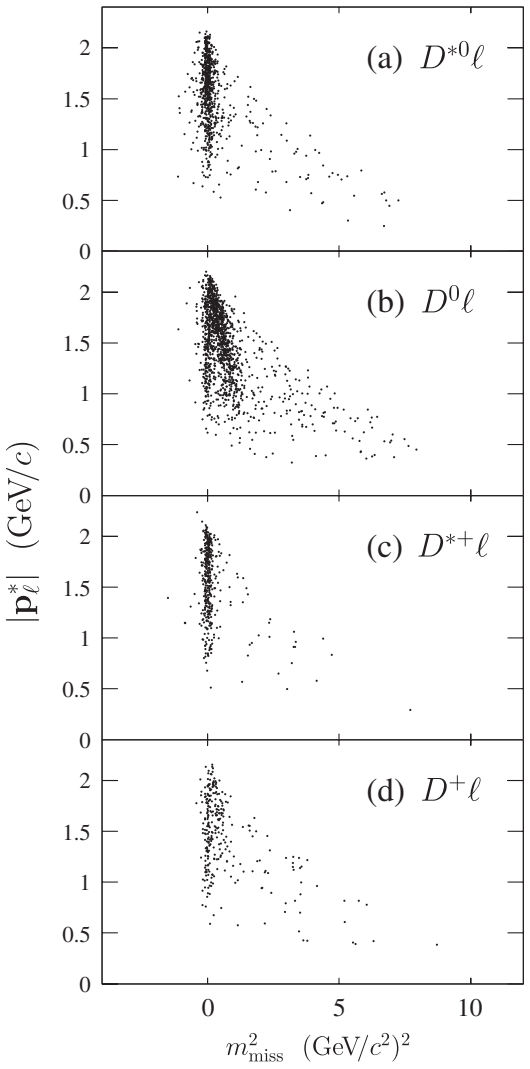

FIG. 6. Distributions of $|\mathbf{p}^*_\ell|$ versus $m^2_{\rm miss}$ for selected data events in the four signal channels. The structures seen here can be compared to the distributions from simulation seen in Fig. 7 and described in the text.

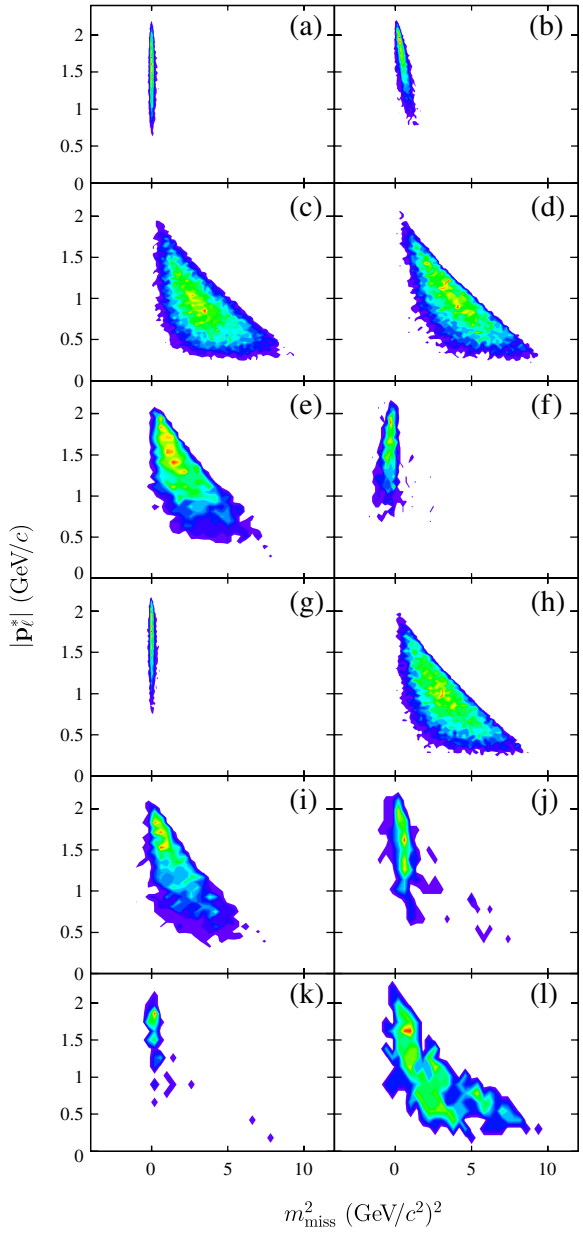

FIG. 7 (color online). Distributions of $|\mathbf{p}^*_\ell|$ versus $m^2_{\rm miss}$ for several MC samples after all event selection. Red (light grey) regions indicate relatively high density of reconstructed events, while blue (dark grey) indicate relatively low density. Shown are (a) $D^0\ell^-\bar{\nu}_\ell \Rightarrow D^0\ell^-$, (b) $D^{*0}\ell^-\bar{\nu}_\ell \Rightarrow D^0\ell^-$, (c) $D^0\tau^-\bar{\nu}_\tau \Rightarrow D^0\ell^-$, (d) $D^{*0}\tau^-\bar{\nu}_\tau \Rightarrow D^0\ell^-$, (e) $D^{**}\ell^-\bar{\nu}_\ell \Rightarrow D^0\ell^-$, (f) $D^0\ell^-\bar{\nu}_\ell \Rightarrow D^{*0}\ell^-$, (g) $D^{*0}\ell^-\bar{\nu}_\ell \Rightarrow D^{*0}\ell^-$, (h) $D^{*0}\tau^-\bar{\nu}_\tau \Rightarrow D^{*0}\ell^-$, (i) $D^{**}\ell^-\bar{\nu}_\ell \Rightarrow D^{*0}\ell^-$, (j) charge-cross feed reconstructed in the $D^0\ell^-$ and $D^+\ell^-$ channels, (k) charge-cross feed reconstructed in the $D^{*0}\ell^-$ and $D^{*+}\ell^-$ channels, and (l) combinatorial background in the four $D^{(*)}\ell^-$ channels. The reconstruction channel notation $\Rightarrow$ and the features of these distributions are discussed in the text.

Figure 7 shows distributions of $|\mathbf{p}^*_\ell|$ versus $m^2_{\rm miss}$ for several MC samples after applying all event selection criteria. While the composition of the event sample will be discussed in greater detail in the following section, these distributions exhibit the qualitative features of the data sample which are most relevant to our signal extraction. Figure 7(a) shows $D^0\ell^-\bar{\nu}_\ell \Rightarrow D^0\ell^-$, where we introduce the $\Rightarrow$ notation to mean that these are true $B^- \rightarrow D^0\ell^-\bar{\nu}_\ell$ events reconstructed in the $D^0\ell^-$ channel. The $m^2_{\rm miss}$ distribution is very narrowly peaked around zero, as expected for one-neutrino events. Figure 7(b) shows $D^{*0}\ell^-\bar{\nu}_\ell \Rightarrow D^0\ell^-$, feed-down events where a D^{*0} is misreconstructed as a D^0. In this case, the center of the $m^2_{\rm miss}$ distribution is offset from zero, and this offset decreases with increasing $|\mathbf{p}^*_\ell|$; this kinematic feature is common to all feed-down processes, and is due to the fact that higher $|\mathbf{p}^*_\ell|$ correspond to lower D^* momenta and therefore to lower momenta for the lost π^0 or γ. The width of the $m^2_{\rm miss}$ distribution is also observed to decrease with increasing $|\mathbf{p}^*_\ell|$, a feature which

is also common to most distributions; this narrowing is partly due to the same kinematic effect as before, the reduced D^* phase space at large $|\mathbf{p}^*_\ell|$, and partly due to the fact that the lepton momentum resolution improves at higher momenta.

Figures 7(c) and 7(d) show similar distributions for signal events, (c) showing correctly reconstructed $B^- \to D^0\tau^-\bar{\nu}_\tau$ and (d) showing D^{*0} feed-down; in both plots, the large values of $m^2_{\rm miss}$ due to the three neutrinos are clearly observed. Again, the $m^2_{\rm miss}$ distributions move towards zero and become narrower at high $|\mathbf{p}^*_\ell|$, in this case due to the reduced phase space for the multiple neutrinos, although, in Fig. 7(d), the effect of the lost π^0 or γ can also be seen as a deficit along the lower-left edge of the distribution. Figure 7(e) shows feed-down from $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ into the $D^0\ell^-$ channel, where, in addition to the neutrino, one or more π^0 mesons or photons from the D^{**} decay have been lost. Since π^0 mesons from D^{**} decay typically have higher momentum than those from D^* decay, the $m^2_{\rm miss}$ distribution is much broader than that in Fig. 7(b). Figure 7(f) shows the feed-up process $D^0\ell^-\bar{\nu}_\ell \Rightarrow D^{*0}\ell^-$, where a true D^0 meson is paired with a combinatorial π^0 or γ to fake a D^{*0} candidate. In this case, the $m^2_{\rm miss}$ distribution is shifted in the opposite direction from Fig. 7(b). Figures 7(g)–7(i), show three additional distributions for events reconstructed in the $D^{*0}\ell^-$ channel, $D^{*0}\ell^-\bar{\nu}_\ell$, $D^{*0}\tau^-\bar{\nu}_\tau$ signal, and $D^{**}\ell^-\bar{\nu}_\ell$ background, respectively; each of these distributions is similar to the corresponding one in the $D^0\ell^-$ channel, Figs. 7(a), 7(c), and 7(e), respectively.

Figures 7(j) and 7(k) show charge-cross feed backgrounds: true $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ events reconstructed with the wrong charge for both the $B_{\rm tag}$ and $D^{(*)}$ meson. Typically this occurs when a low momentum $\pi^\pm$ is swapped between the two mesons. Note that, even though the event is misreconstructed, this particle misassignment does not substantially alter the total missing momentum, so that the $m^2_{\rm miss}$ distribution still peaks at or near zero. While the events in Fig. 7(k), which are reconstructed in the $D^*\ell^-$ channels, are very strongly peaked at $m^2_{\rm miss} = 0$, Fig. 7(j) includes a large feed-down component, and therefore exhibits the same sloping behavior seen in Fig. 7(b).

Figure 7(l) shows the distribution for combinatorial background for all four signal channels. This background is dominated by hadronic B decays such as $B \to D^{(*)}D_s^{(*)}$ that produce a secondary lepton, including events with τ leptons from D_s decay.

In our $B\bar{B}$ MC sample, our criteria select D^{**} control samples which are 60%–80% pure $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ events, of which more than 90% involve true $D^{**} \to D^{(*)}\pi^0$ transitions. The remaining events are split between feed-up from $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ and combinatorial background. In these control samples, the $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ component peaks at or near zero in $m^2_{\rm miss}$, just as $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ does in the four signal channels. The qualitative features of the other contributions are similar to what is seen in the signal channels.

VI. KINEMATIC CONTROL SAMPLES

The event selection criteria described in Sec. IV are more complicated than those used in a typical *BABAR* analysis, due to the full-event reconstruction of a high-multiplicity final state and the need to veto events with extra tracks and neutral clusters. We use two data control samples to validate our simulation with respect to the observed behavior in data. The control samples are kinematically selected, with no requirement on $m^2_{\rm miss}$, to be high purity samples of $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ events, with little or no contamination from signal decays.

The first control sample is defined by requiring the reconstructed lepton to satisfy $|\mathbf{p}^*_\ell| > 1.5\,{\rm GeV}/c$, and is therefore a subset of our full analysis sample. In simulation, 95% of the selected sample is $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ (the two $D^*\ell^-$ channels are approximately 95% $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$, while the two $D\ell^-$ channels include both $\bar{B} \to D\ell^-\bar{\nu}_\ell$ and large feed-down from $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$). The remaining 5% of the sample is composed of about 1%–3% $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$, less than 1% $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$, and 1%–2% of combinatorial backgrounds.

For the second control sample, we remove the standard $q^2 > 4\ ({\rm GeV}/c^2)^2$ selection and require that events instead satisfy $q^2 < 5\ ({\rm GeV}/c^2)^2$, with q^2 calculated according to Eq. (10). This control sample has very little overlap with our final event sample, where we require $q^2 > 4\ ({\rm GeV}/c^2)^2$. Although the two control samples do have some overlap, this q^2 control sample has the advantage over the first of allowing us to examine events with low $|\mathbf{p}^*_\ell|$, as expected for signal events. In simulation, approximately 90% of this control sample is $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ (as in the first sample, the two $D^*\ell^-$ channels are approximately 90% $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$, while the two $D\ell^-$ channels include D^* feed-down). The remainder of the sample is composed of about 3% $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$, 3% $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$, and 4%–5% combinatorial backgrounds.

Figure 8 shows several data-simulation comparisons in the two control samples. The four $D^{(*)}\ell^-$ channels have been combined in these plots as have the two control samples, and this union of the two control samples is responsible for the large steps visible in (a) and (b). We see good agreement between data and simulation in these plots, as well as in similar studies where the two control samples are examined separately, the four $D^{(*)}\ell^-$ channels are examined separately, the two lepton types are examined separately, and where the data are split according to *BABAR* running period. We have examined variables related to $B_{\rm tag}$ reconstruction, signal-side reconstruction, hermeticity and whole-event reconstruction, and missing momentum. In all cases, we observe that the simulation does a reasonable job describing the data. Because of the relative normalization scheme, small differences between simulation and data have no detrimental effect on the analysis.

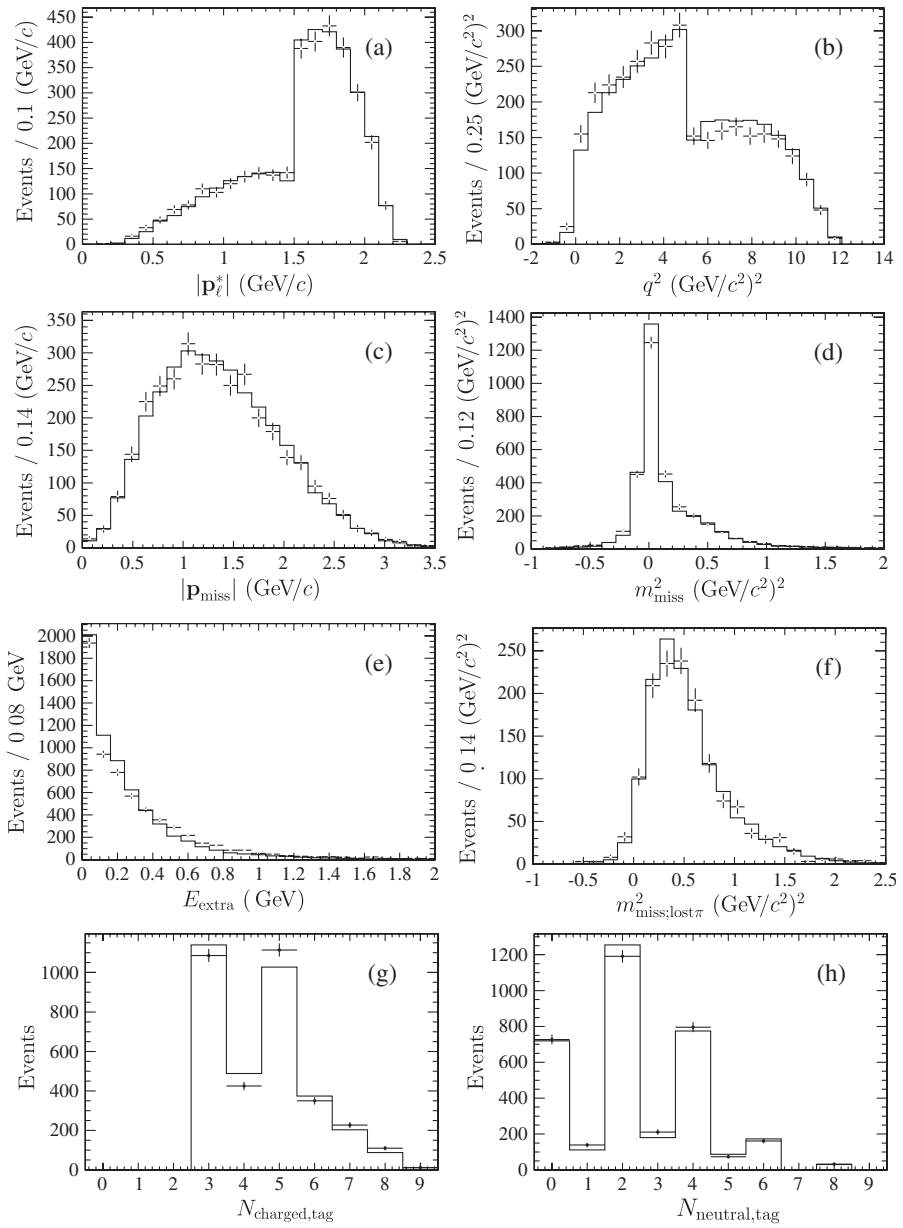

FIG. 8. Kinematic control sample plots: (a) $|\mathbf{p}^*_\ell|$; (b) q^2; (c) $|\mathbf{p}_{\rm miss}|$; (d) $m^2_{\rm miss}$; (e) $E_{\rm extra}$; (f) $m^2_{\rm miss}$ in the D^* channels, assuming that the soft π/γ had been lost; multiplicity of (g) charged tracks and (h) neutral clusters used to reconstruct the $B_{\rm tag}$. In all plots, the points with error bars are the data and the solid histogram is the simulation, scaled to the data luminosity. Good agreement is seen between data and simulation in a variety of variables corresponding to reconstruction, kinematics, and hermeticity requirements. Small differences between data and simulation cancel in the relative measurement and have no detrimental effect on the analysis. The large steps in (a) and (b) are due to the combination of two control samples, as described in the text. The structure in (g) is caused by the larger efficiency to reconstruct charged $B_{\rm tag}$ candidates—with an odd number of charged tracks—than neutral candidates, while the prominent even-odd structure in (h) is due to the fact that most neutral clusters correspond to the process $\pi^0 \to \gamma\gamma$ and so appear in pairs.

VII. FIT OF SIGNAL AND BACKGROUND YIELDS

A. Fit Overview

Signal and background yields are extracted using an extended unbinned maximum likelihood fit to the joint $(m^2_{\rm miss}, |\mathbf{p}^*_\ell|)$ distribution. The fit is performed simultaneously in the four signal channels and the four D^{**} control samples. Two two-dimensional probability density functions (PDFs) are presented in Sec. VII B; each component in the fit (listed below) is described by one of these two PDFs, with parameters determined from fits to simulated event samples. A set of constraints, described in Sec. VII C, relate fit components in different reconstruction channels. These constraints are also determined from MC samples,

except for parameters describing the amount of D^* feed-down into the $D\ell^-$ signal channels, which are determined directly by the fit to data.

Tables II and III summarize the parameterization of the fit in the four signal channels and the four D^{**} control samples, respectively. In each of the four signal channels, we describe the data as the sum of seven components: $D\tau^-\bar{\nu}_\tau$, $D^*\tau^-\bar{\nu}_\tau$, $D\ell^-\bar{\nu}_\ell$, $D^*\ell^-\bar{\nu}_\ell$, $D^{**}(\ell^-/\tau^-)\bar{\nu}$, charge-cross feed, and combinatorial background. The four D^{**} control samples are described as the sum of five components: $D^{**}(\ell^-/\tau^-)\bar{\nu}$, $D(\ell^-/\tau^-)\bar{\nu}$, $D^*(\ell^-/\tau^-)\bar{\nu}$, charge-cross feed, and combinatorial background. Each of these components is described by one of the two PDFs given in Sec. VII B, with the numerical parameters of the 32 PDFs determined from independent MC samples. The charge-cross feed components in the two $D\ell^-$ signal channels are described by a single PDF, with common parameters for $D^0\ell^-$ and $D^+\ell^-$, as are the two $D^*\ell^-$ charge-cross feed components and the four $D^{(*)}\ell^-\pi^0$ components; the four combinatorial background components in the signal channels are described by a single PDF with common parameters, as are the four in the D^{**} control samples.

TABLE II. Components of the signal extraction fit in the signal channels, and their approximate abundances in our $B\bar{B}$ MC sample. The structure of the fit is identical between the B^- and $\bar{B}^0$ channels. There are seven components in each of the four signal channels.

Channel	Source	Abundance in $B\bar{B}$ MC (%)
$D^{*0}\ell^-$	$D^{*0}\tau^-\bar{\nu}_\tau$ signal	5
	$D^0\tau^-\bar{\nu}_\tau$ signal feed-up	0.5
	$D^{*0}\ell^-\bar{\nu}_\ell$ normalization	90
	$D^0\ell^-\bar{\nu}_\ell$ feed-up	2
	$D^{**}(\ell^-/\tau^-)\bar{\nu}$ feed-down	3
	Charge-cross feed	0.5
	Combinatorial background	1
$D^0\ell^-$	$D^0\tau^-\bar{\nu}_\tau$ signal	3
	$D^{*0}\tau^-\bar{\nu}_\tau$ signal feed-down	3
	$D^0\ell^-\bar{\nu}_\ell$ normalization	25
	$D^{*0}\ell^-\bar{\nu}_\ell$ feed-down	60
	$D^{**}\ell^-\bar{\nu}_\ell$ feed-down	2
	Charge-cross feed	2
	Combinatorial background	2
$D^{*+}\ell^-$	$D^{*+}\tau^-\bar{\nu}_\tau$ signal	5
	$D^+\tau^-\bar{\nu}_\tau$ signal feed-up	0.1
	$D^{*+}\ell^-\bar{\nu}_\ell$ normalization	90
	$D^+\ell^-\bar{\nu}_\ell$ feed-up	0.5
	$D^{**}\ell^-\bar{\nu}_\ell$ feed-down	3
	Charge-cross feed	0.1
	Combinatorial background	2
$D^+\ell^-$	$D^+\tau^-\bar{\nu}_\tau$ signal	5
	$D^{*+}\tau^-\bar{\nu}_\tau$ signal feed-down	2
	$D^+\ell^-\bar{\nu}_\ell$ normalization	45
	$D^{*+}\ell^-\bar{\nu}_\ell$ feed-down	40
	$D^{**}\ell^-\bar{\nu}_\ell$ feed-down	6
	Charge-cross feed	1
	Combinatorial background	3

TABLE III. Components of the signal extraction fit in the D^{**} control sample channels, and their approximate abundances in our $B\bar{B}$ MC sample. The structure of the fit is identical between the B^- and $\bar{B}^0$ channels. There are five components in each of the four D^{**} control sample channels.

Channel	Source	Abundance in $B\bar{B}$ MC (%)
$D^{*0}\pi^0\ell^-$	$D^{**}\ell^-\bar{\nu}_\ell$	60
	$D^{*0}\ell^-\bar{\nu}_\ell$ feed-up	18
	$D^0\ell^-\bar{\nu}_\ell$ feed-up	2
	Charge-cross feed	1
	Combinatorial background	20
$D^0\pi^0\ell^-$	$D^{**}\ell^-\bar{\nu}_\ell$	70
	$D^{*0}\ell^-\bar{\nu}_\ell$ feed-up	10
	$D^0\ell^-\bar{\nu}_\ell$ feed-up	3
	Charge-cross feed	2
	Combinatorial background	15
$D^{*+}\pi^0\ell^-$	$D^{**}\ell^-\bar{\nu}_\ell$	65
	$D^{*+}\ell^-\bar{\nu}_\ell$ feed-up	20
	$D^+\ell^-\bar{\nu}_\ell$ feed-up	0.1
	Charge-cross feed	0.1
	Combinatorial background	15
$D^+\pi^0\ell^-$	$D^{**}\ell^-\bar{\nu}_\ell$	60
	$D^{*+}\ell^-\bar{\nu}_\ell$ feed-up	6
	$D^+\ell^-\bar{\nu}_\ell$ feed-up	3
	Charge-cross feed	1
	Combinatorial background	30

$\bar{B}\to D^{(*)}\tau^-\bar{\nu}_\tau$ events feeding up into the D^{**} control samples are expected to contribute 1.8 ± 0.6 events in the four channels together, so these events are combined with the light lepton contribution. In both the control samples and in the signal channels, $\bar{B}\to D^{**}\tau^-\bar{\nu}_\tau$ events are expected to contribute 3.5%–4.5% of the total D^{**} yield; these events are combined with the light lepton contribution, and the amount of $D^{**}\tau^-\bar{\nu}_\tau$ is varied as a systematic uncertainty.

The fit has 18 free parameters: four signal branching-fraction ratios R, one for each $D^{(*)}$ meson; four $\bar{B}\to D^{(*)}\ell^-\bar{\nu}_\ell$ normalization yields; four $\bar{B}\to D^{**}\ell^-\bar{\nu}_\ell$ background yields; four combinatorial background yields, one in each of the four D^{**} control samples; two parameters describing $D^*\Rightarrow D$ feed-down, one for charged B modes and one for neutral B modes. The combinatorial background yields in the four signal channels are fixed in the fit to the expected value from simulation, as are the charge-cross feed backgrounds in both the signal channels and D^{**} control samples; variation of these backgrounds is treated as a systematic uncertainty below.

We also perform a second, $B^- - \bar{B}^0$ constrained, fit, by requiring $R(D^+) = R(D^0)$ and $R(D^{*+}) = R(D^{*0})$,[7] reducing the number of free parameters to 16.

B. Probability density functions

We construct an empirical model of the two-dimensional $(m^2_{\rm miss}, |\mathbf{p}^*_\ell|)$ PDF as the product of two terms: a one-dimensional function to describe the $|\mathbf{p}^*_\ell|$ distribution, discussed in Sec. VII B 1; and a $|\mathbf{p}^*_\ell|$-dependent "resolution" function to describe the $m^2_{\rm miss}$ distribution, to be discussed in Sec. VII B 2. For processes in which the only missing particle is a single neutrino, the true $m^2_{\rm miss}$ spectrum is a delta function located at zero and the observed distribution is a pure resolution function. For components with multiple missing particles, the observed $m^2_{\rm miss}$ distribution is the convolution of the physical $m^2_{\rm miss}$ spectrum with our detector resolution. The PDFs presented below are used to describe both of these physical cases, with different numerical parameters describing the different behaviors; these two PDFs are flexible enough to describe the variety of physical and resolution processes needed in this analysis.

1. One-dimensional $|\mathbf{p}^*_\ell|$ parameterization

We use a generalized form of a Gaussian to model the $|\mathbf{p}^*_\ell|$ distribution. The Gaussian distribution,

$$\mathcal{G}(|\mathbf{p}^*_\ell|) \propto \exp\left(-\frac{1}{2}\cdot\left|\frac{|\mathbf{p}^*_\ell| - p_0}{\sigma}\right|^2\right), \tag{11}$$

has the same general properties as our distributions: it rises smoothly from zero to a peak value and then falls smoothly back to zero again. Here, p_0 represents the value of $|\mathbf{p}^*_\ell|$ for which $\mathcal{G}$ peaks and σ represents the width of the Gaussian distribution.

This gross agreement is not enough, however, so we define a modified Gaussian function,

$$\mathcal{H}(|\mathbf{p}^*_\ell|) \propto \exp\left(-\left|\frac{|\mathbf{p}^*_\ell| - p_0}{\sigma(|\mathbf{p}^*_\ell|)}\right|^{\nu(|\mathbf{p}^*_\ell|)}\right), \tag{12}$$

where, for convenience, we have absorbed the constant factor of 2 into the definition of $\sigma(|\mathbf{p}^*_\ell|)$. By allowing the width and exponent of the Gaussian to be functions of $|\mathbf{p}^*_\ell|$, we are able to describe a greater variety of shapes. Specifically, we take $\nu(|\mathbf{p}^*_\ell|)$ to be a linear function,

$$\nu(|\mathbf{p}^*_\ell|) = \nu_L + \frac{\nu_H - \nu_L}{2.4\ \text{GeV}/c}\cdot|\mathbf{p}^*_\ell|, \tag{13}$$

where ν_L and ν_H are the values of the exponential term at the low and high endpoints of $|\mathbf{p}^*_\ell|$, fixed at zero and 2.4 GeV/c, respectively. Similarly, we parameterize $\sigma(|\mathbf{p}^*_\ell|)$ as a bilinear function,

$$\sigma(|\mathbf{p}^*_\ell|) = \begin{cases} \sigma_L + \frac{\sigma_0-\sigma_L}{p_0}\cdot|\mathbf{p}^*_\ell| & |\mathbf{p}^*_\ell| < p_0 \\ \sigma_0 + \frac{\sigma_H-\sigma_0}{2.4\ \text{GeV}/c - p_0}\cdot(|\mathbf{p}^*_\ell| - p_0) & |\mathbf{p}^*_\ell| > p_0, \end{cases} \tag{14}$$

where σ_L, σ_0, and σ_H represent the widths of the Gaussian at $|\mathbf{p}^*_\ell| = 0$, $|\mathbf{p}^*_\ell| = p_0$, and $|\mathbf{p}^*_\ell| = 2.4$ GeV/c, respectively. Even though this parameterization is discontinuous at the point $|\mathbf{p}^*_\ell| = p_0$, the resulting function $\mathcal{H}(|\mathbf{p}^*_\ell|)$ remains smooth since the numerator in the exponent, $(|\mathbf{p}^*_\ell| - p_0)$, goes to zero at the same point.

The $|\mathbf{p}^*_\ell|$ parameterization therefore has six free parameters: p_0, the peak; ν_L and ν_H, describing the exponential term; and σ_L, σ_0, and σ_H, describing the width. When performing fits using this PDF, we integrate $\mathcal{H}$ numerically to compute the normalization.

2. Two-dimensional PDF parameterization

We construct two types of two-dimensional PDF, $\mathcal{P}_1(|\mathbf{p}^*_\ell|, m^2_{\rm miss})$ and $\mathcal{P}_2(|\mathbf{p}^*_\ell|, m^2_{\rm miss})$ by multiplying the model of the lepton spectrum above by a resolution function in $m^2_{\rm miss}$, where the resolution is a function of $|\mathbf{p}^*_\ell|$. Allowing the parameters of the resolution function to be functions of $|\mathbf{p}^*_\ell|$ produces a correlation between the two fit variables, and it is these parameters which allow the PDFs to describe such a wide variety of shapes.

Using the model of the lepton spectrum $\mathcal{H}(|\mathbf{p}^*_\ell|)$ introduced above, we construct the PDFs as

$$\begin{aligned}\mathcal{P}_1(|\mathbf{p}^*_\ell|, m^2_{\rm miss}) \equiv\ & \mathcal{H}(|\mathbf{p}^*_\ell|)\times[f_1(|\mathbf{p}^*_\ell|)\mathcal{G}_1(|\mathbf{p}^*_\ell|; m^2_{\rm miss}) \\ & + (1 - f_1(|\mathbf{p}^*_\ell|))\mathcal{G}_2(|\mathbf{p}^*_\ell|; m^2_{\rm miss})]\end{aligned} \tag{15}$$

and

$$\begin{aligned}\mathcal{P}_2(|\mathbf{p}^*_\ell|, m^2_{\rm miss}) \equiv\ & \mathcal{H}(|\mathbf{p}^*_\ell|)\times[f_1(|\mathbf{p}^*_\ell|)\mathcal{G}_1(|\mathbf{p}^*_\ell|; m^2_{\rm miss}) \\ & + f_2(|\mathbf{p}^*_\ell|)\mathcal{G}_b(|\mathbf{p}^*_\ell|; m^2_{\rm miss}) + (1 - f_1(|\mathbf{p}^*_\ell|) \\ & - f_2(|\mathbf{p}^*_\ell|))\mathcal{G}_2(|\mathbf{p}^*_\ell|; m^2_{\rm miss})].\end{aligned} \tag{16}$$

Here, the functions $\mathcal{G}_1$ and $\mathcal{G}_2$ are Gaussians and $\mathcal{G}_b$ is a bifurcated Gaussian (Gaussian with different σ parameters on either side of the mean), respectively; all are functions of $m^2_{\rm miss}$, with parameters dependent on $|\mathbf{p}^*_\ell|$.

The $|\mathbf{p}^*_\ell|$ dependence of the various parameters of $\mathcal{G}_{1,2}$ and $\mathcal{G}_b$ is listed in Table IV. The total number of free parameters for $\mathcal{P}_1$ is 18: six for $\mathcal{H}(|\mathbf{p}^*_\ell|)$, five each for $\mathcal{G}_1$ and $\mathcal{G}_2$, and two for f_1. The total number of free parameters for $\mathcal{P}_2$ is 24: six for $\mathcal{H}(|\mathbf{p}^*_\ell|)$, five each for $\mathcal{G}_1$ and $\mathcal{G}_2$, four for $\mathcal{G}_b$, and two each for f_1 and f_2.

We use the simpler PDF, $\mathcal{P}_1$, to model most of the semileptonic fit components (22 out of 32), as well as the charge-cross feed and combinatorial backgrounds. For the remaining ten components, however, the more complicated parameterization $\mathcal{P}_2$ is required to adequately describe the $m^2_{\rm miss}$ tail. Eight of these components are the ones in which

[7] This constraint follows from isospin symmetry in both the signal and normalization modes but is more general.

TABLE IV. $|\mathbf{p}^*_\ell|$ dependence of the $m^2_{\rm miss}$ PDF parameterization. The form of f_2 is chosen to allow the $\mathcal{G}_b$ term to contribute at low $|\mathbf{p}^*_\ell|$, but to drive this term rapidly to zero as $|\mathbf{p}^*_\ell|$ increases. The form of σ_H is chosen to allow for a long tail towards high $m^2_{\rm miss}$ at low $|\mathbf{p}^*_\ell|$, but to drive this term rapidly to zero as $|\mathbf{p}^*_\ell|$ increases (note that there is no problem having σ approach zero since the amplitude of this term goes to zero as well; the result is finite and well behaved). $N_{\rm par}$ gives the number of free parameters for each term separately.

Function	Parameter	Dependence on $\lvert\mathbf{p}^*_\ell\rvert$	$N_{\rm par}$
$\mathcal{G}_{1,2}$	Mean	Quadratic	3
$\mathcal{G}_{1,2}$	σ	Linear	2
$\mathcal{G}_b$	Mean	Constant	1
$\mathcal{G}_b$	σ_L	Constant	1
$\mathcal{G}_b$	σ_H	$\sigma_{H0}\cdot[1-(\frac{\lvert\mathbf{p}^*_\ell\rvert}{2.4\ \mathrm{GeV}/c})^\phi]$	2
$\mathcal{P}_{1,2}$	f_1	Linear	2
$\mathcal{P}_2$	f_2	$F\cdot(\frac{2.4\ \mathrm{GeV}/c-\lvert\mathbf{p}^*_\ell\rvert}{2.2\ \mathrm{GeV}/c})^\theta$	2

the only missing particle is a single neutrino,

$$D^{*0}\ell^-\bar{\nu}_\ell \Rightarrow D^{*0}\ell^- \qquad D^0\ell^-\bar{\nu}_\ell \Rightarrow D^0\ell^-$$
$$D^{*+}\ell^-\bar{\nu}_\ell \Rightarrow D^{*+}\ell^- \qquad D^+\ell^-\bar{\nu}_\ell \Rightarrow D^+\ell^-$$
$$D^{**}\ell^-\bar{\nu}_\ell \Rightarrow D^{*0}\pi^0\ell^- \qquad D^{**}\ell^-\bar{\nu}_\ell \Rightarrow D^0\pi^0\ell^-$$
$$D^{**}\ell^-\bar{\nu}_\ell \Rightarrow D^{*+}\pi^0\ell^- \qquad D^{**}\ell^-\bar{\nu}_\ell \Rightarrow D^+\pi^0\ell^-,$$

and the remaining two are components in which a single neutrino and a soft π^0 or γ are missing,

$$D^{*0}\ell^-\bar{\nu}_\ell \Rightarrow D^0\ell^- \qquad D^{*+}\ell^-\bar{\nu}_\ell \Rightarrow D^+\ell^-.$$

Figure 9 shows projections of the PDF fit to several simulated event samples, showing the range of shapes described by our PDFs.

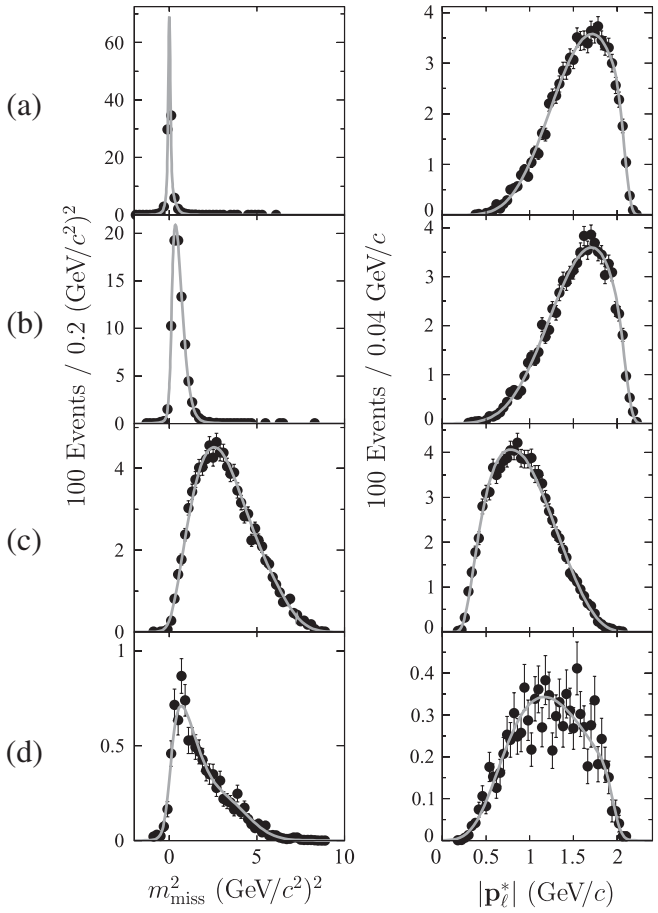

FIG. 9. Projections of the PDF from fits to MC samples. The left plots show projections onto $m^2_{\rm miss}$, while the right plots show projections onto $|\mathbf{p}^*_\ell|$. Shown are projections for four of the PDFs used in the fit: (a) $D^{*0}\ell^-\bar{\nu}_\ell \Rightarrow D^{*0}\ell^-$, (b) $D^{*0}\ell^-\bar{\nu}_\ell \Rightarrow D^0\ell^-$, (c) $D^{*0}\tau^-\bar{\nu}_\tau \Rightarrow D^{*0}\ell^-$, and (d) $D^{**}\ell^-\bar{\nu}_\ell \Rightarrow D^{*0}\ell^-$. The MC sample is shown as points, and the projection of the fit is shown as a curve. Note the sharp peak at $m^2_{\rm miss}=0$ in (a), while the peak in (b) is somewhat spread out and shifted to larger values of $m^2_{\rm miss}$ because of the lost π^0 or γ from D^{*0} decay.

C. Cross feed constraints

We apply a number of constraints in the fit, relating the event yields between different reconstruction channels in order to make use of all available information. These constraints help to maximize our sensitivity, particularly to the $\bar{B}\to D\tau^-\bar{\nu}_\tau$ signals where the dominant backgrounds are due to feed-down. There are 20 such constraints in the fit, corresponding to 20 different ways in which a true $\bar{B}\to D/D^*/D^{**}\ell^-\bar{\nu}_\ell$ event can be reconstructed with the wrong final-state meson, either as feed-down ($D^*\Rightarrow D$ and $D^{**}\Rightarrow D/D^*$) or as feed-up ($D\Rightarrow D^*/D^{(*)}\pi^0$ and $D^*\Rightarrow D^{(*)}\pi^0$).

These constraints are implemented in the fit by requiring that the number of events of type j correctly reconstructed in the ith channel (N_{ij}) is related to the number of events of type j reconstructed in a cross feed channel i' ($N_{i'j}$) by

$$N_{i'j} \equiv N_{ij}\cdot f_{i\to i',j}, \tag{17}$$

where $f_{i\to i',j}$ is a cross feed constraint relating the two yields. The cross feed constraints $f_{i\to i',j}$ are linearly related to the misreconstruction probability. For feed-down processes, in which the probability to lose a low-momentum π^0 or γ is high, $f_{i\to i',j}$ typically takes values between 0.2 and 1.0; for feed-up processes, in which the probability to reconstruct a fake π^0 or γ in a narrow mass window is low, $f_{i\to i',j}$ typically takes values between 0.01 and 0.1.

The values for most of the $f_{i\to i',j}$ terms are taken from simulation, but, in order to reduce systematic effects, the values of the dominant feed-down components, $\bar{B}\to D^*\ell^-\bar{\nu}_\ell$ reconstructed in the $D\ell^-$ signal channels, are left free in the fit to data. We also use the floating values of these D^* feed-down constraints to apply a small first-order correction to the corresponding signal feed-down constraints describing $\bar{B}\to D^*\tau^-\bar{\nu}_\tau$ reconstructed in the $D\ell^-$ channels; in this way, we use the high-statistics

$D^*\ell^-\bar{\nu}_\ell$ samples to improve our knowledge of the signal feed-down probability.

D. Projections of the fit to data

Figures 10–14 show projections of the $B^- - \bar{B}^0$-constrained fit. Figure 10 shows projections in $m^2_{\rm miss}$ for the four signal channels, showing both the low $m^2_{\rm miss}$ region, which is dominated by the normalization modes $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$, and the high $m^2_{\rm miss}$ region, which is dominated by the signal modes $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$. Figures 11 and 12 show projections in $|\mathbf{p}^*_\ell|$ for the normalization and signal regions, respectively, and Figs. 13 and 14 show projections of both $m^2_{\rm miss}$ and $|\mathbf{p}^*_\ell|$ for the four D^{**} control samples. In all cases, we see that the fit does a reasonable job of describing the observed event sample, both in background-dominated and signal-dominated regions.

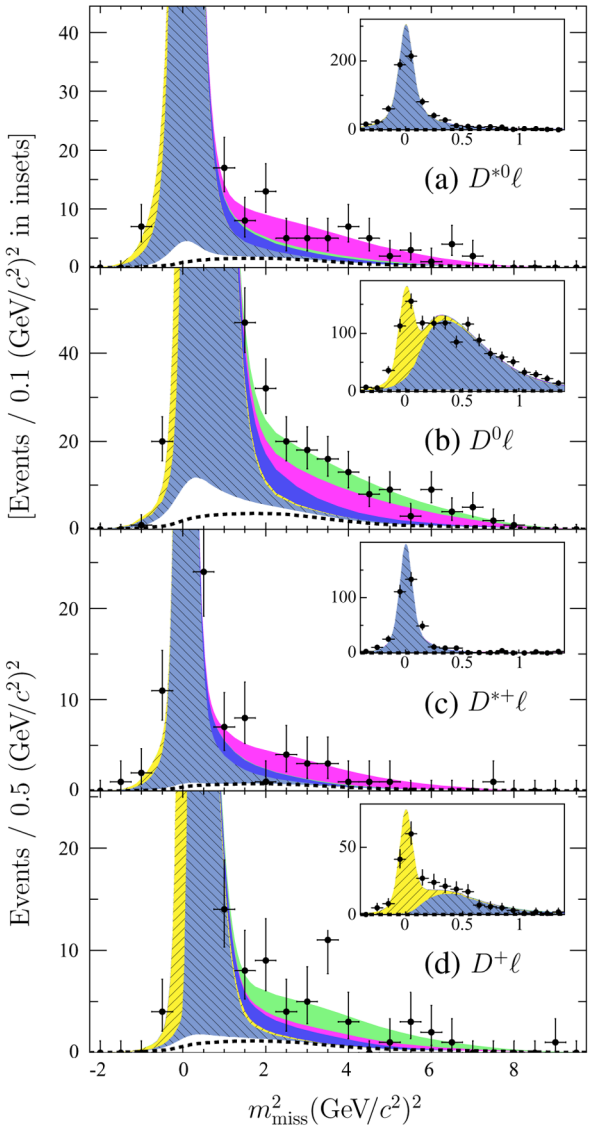

FIG. 10 (color online). Distributions of events and fit projections in $m^2_{\rm miss}$ for the four final states: $D^{*0}\ell^-$, $D^0\ell^-$, $D^{*+}\ell^-$, and $D^+\ell^-$. The normalization region $m^2_{\rm miss} \approx 0$ is shown with finer binning in the insets. The fit components are combinatorial background (white, below dashed line), charge-cross feed background (white, above dashed line), the $\bar{B} \to D\ell^-\bar{\nu}_\ell$ normalization mode (forward slanted hatching, yellow), the $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ normalization mode (backward slanted hatching, light blue), $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ background (dark, or blue), the $\bar{B} \to D\tau^-\bar{\nu}_\tau$ signal (light grey, green), and the $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ signal (medium grey, magenta). The fit shown incorporates the $B^- - \bar{B}^0$ constraints.

VIII. SIGNAL EXTRACTION AND NORMALIZATION

The fit described in Sec. VII directly measures, for each signal mode, the ratio of the number of signal events in the data sample, $N_{\rm sig}$, to the number of corresponding normal-

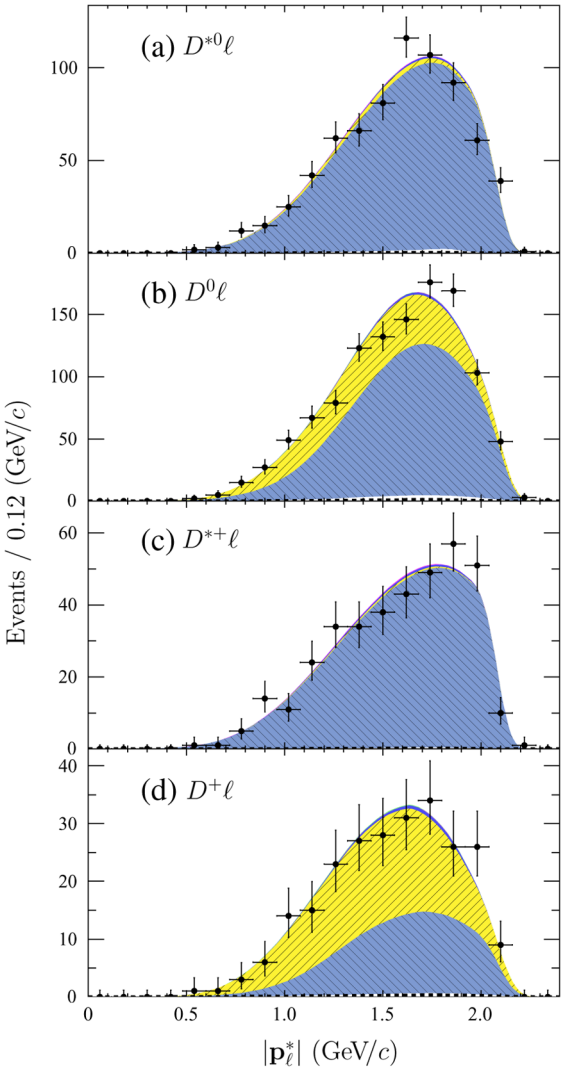

FIG. 11 (color online). Distributions of events and fit projections in $|\mathbf{p}^*_\ell|$ for the four final states $D^{*0}\ell^-$, $D^0\ell^-$, $D^{*+}\ell^-$, and $D^+\ell^-$, shown in the normalization region, $m^2_{\rm miss} < 1$ (GeV/c^2)^2. The fit components are shaded as in Fig. 10.

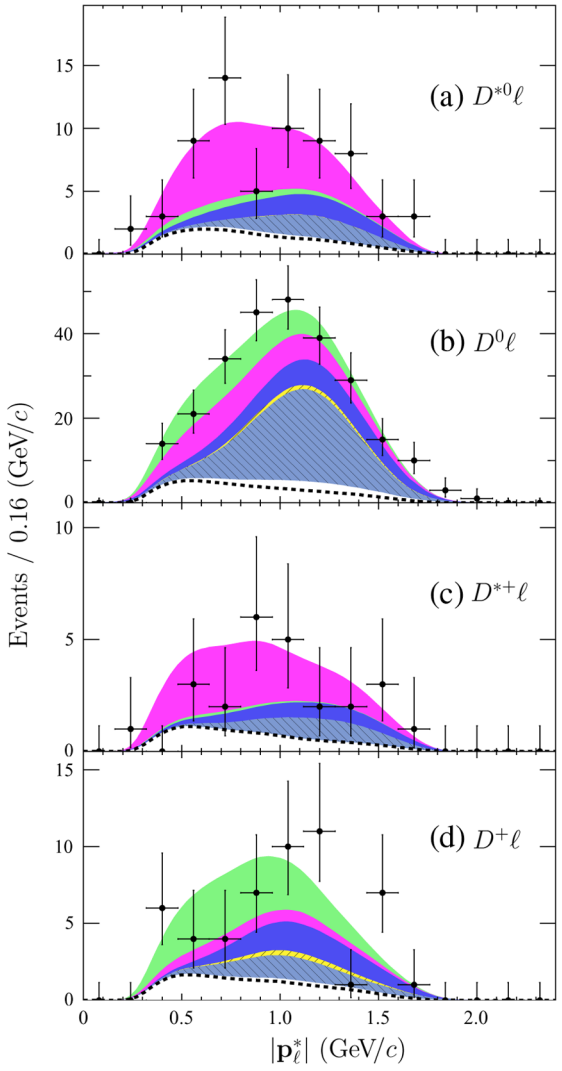

FIG. 12 (color online). Distributions of events and fit projections in $|\mathbf{p}^*_\ell|$ for the four final states $D^{*0}\ell^-$, $D^0\ell^-$, $D^{*+}\ell^-$, and $D^+\ell^-$, shown in the signal region, $m^2_{\rm miss} > 1$ (GeV/c^2)^2. The fit components are shaded as in Fig. 10.

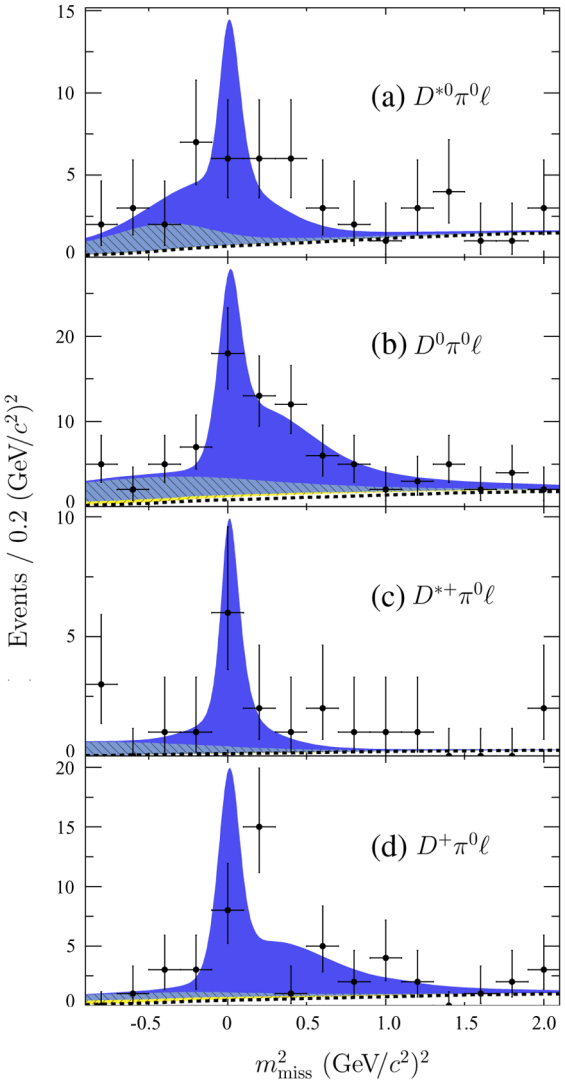

FIG. 13 (color online). Distributions of events and fit projections in $m^2_{\rm miss}$ for the four D^{**} control samples $D^{*0}\pi^0\ell^-$, $D^0\pi^0\ell^-$, $D^{*+}\pi^0\ell^-$, and $D^+\pi^0\ell^-$. The fit components are shaded as in Fig. 10.

ization events, $N_{\rm norm}$. We measure the signal branching-fraction ratios R as

$$R \equiv \frac{N_{\rm sig}}{N_{\rm norm}} \cdot \frac{1}{\varepsilon_{\rm sig}/\varepsilon_{\rm norm}} \cdot \frac{1}{\mathcal{B}(\tau^- \to \ell^- \bar{\nu}_\ell \nu_\tau)}, \qquad (18)$$

where the relative efficiency $\varepsilon_{\rm sig}/\varepsilon_{\rm norm}$ is calculated from signal MC samples as

$$\varepsilon_{\rm sig}/\varepsilon_{\rm norm} \equiv \frac{N^{\rm reco}_{\rm sig}/N^{\rm gen}_{\rm sig}}{N^{\rm reco}_{\rm norm}/N^{\rm gen}_{\rm norm}}. \qquad (19)$$

Here, the $N^{\rm gen}$ are the numbers of simulated events, and the $N^{\rm reco}$ are the numbers of reconstructed events, including both correctly reconstructed events and contributions from feed-up or feed-down. Cross feed is not a large effect, however, because both the numerator and denominator in this relative efficiency receive cross feed contributions, and the net result tends to cancel (this cancellation is not exact, since the D^* momentum spectra are not identical between signal and normalization modes, but these differences are already accounted for in our normalization procedure).

Signal efficiencies are given in Table V. The relative efficiencies for the two $\bar{B} \to D\tau^-\bar{\nu}_\tau$ modes are much larger than unity because of the q^2 cut, which is $\approx 98\%$ efficient for signal events but rejects about 50% of the $\bar{B} \to D\ell^-\bar{\nu}_\ell$ normalization events, as seen in Fig. 1(a). The q^2 cut has a similar, but less pronounced, effect on the D^* modes, but, due to the lower efficiency for identifying secondary leptons in the signal modes, the resulting relative efficiency is close to unity.

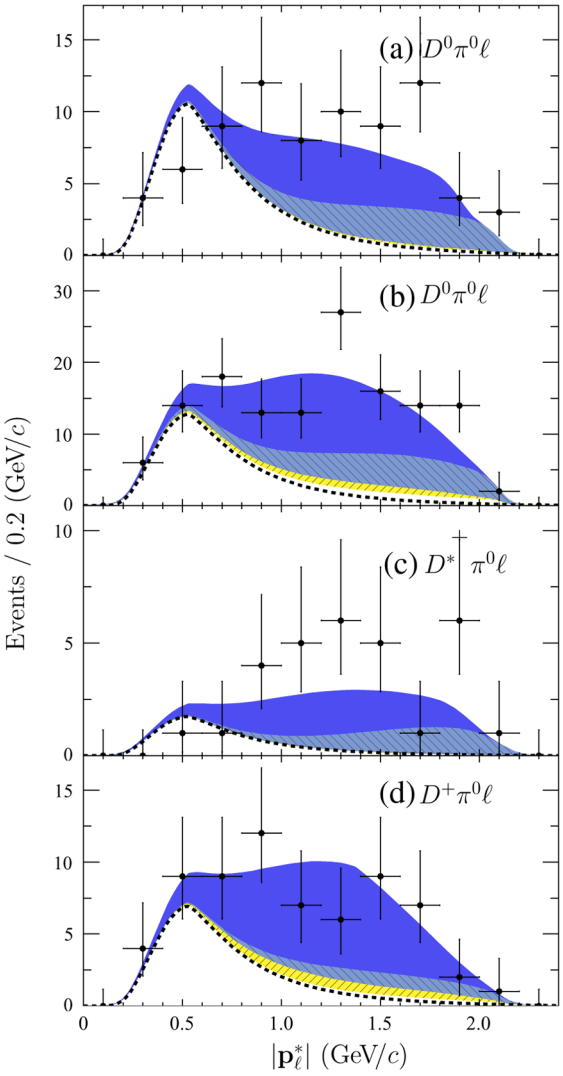

FIG. 14 (color online). Distributions of events and fit projections in $|\mathbf{p}^*_\ell|$ for the four D^{**} control samples $D^{*0}\pi^0\ell^-$, $D^0\pi^0\ell^-$, $D^{*+}\pi^0\ell^-$, and $D^+\pi^0\ell^-$. The fit components are shaded as in Fig. 10.

IX. SYSTEMATIC UNCERTAINTIES

Table VI summarizes all of the systematic uncertainties considered in this analysis. Because our signal is extracted and normalized relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$, many sources of systematic uncertainty—especially those related to reconstruction efficiency—are expected to cancel, either partially or completely, when we take the ratio.

TABLE V. Relative signal efficiencies $\varepsilon_{\rm sig}/\varepsilon_{\rm norm}$ for the four signal modes.

Signal mode	$\varepsilon_{\rm sig}/\varepsilon_{\rm norm}$
$B^- \to D^0\tau^-\bar{\nu}_\tau$	1.85 ± 0.02
$B^- \to D^{*0}\tau^-\bar{\nu}_\tau$	0.99 ± 0.01
$\bar{B}^0 \to D^+\tau^-\bar{\nu}_\tau$	1.83 ± 0.03
$\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$	0.91 ± 0.01

TABLE VI. Contributions to the total systematic uncertainty. The additive systematic uncertainties represent uncertainties on the fit yield, and therefore reduce the statistical significance of the results. The multiplicative systematic uncertainties represent uncertainties on the normalization, so they affect the numerical results but not the statistical significance. The first four columns summarize errors on the individual branching-fraction ratios; the last two columns summarize errors on the $B^- - \bar{B}^0$ constrained measurement. The totals here refer to errors on the branching-fraction ratios R; the errors on $\mathcal{B}(\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell)$ (discussed in Sec. X) only apply to the absolute branching fractions, and are not included in the quoted total error.

Source	Fractional uncertainty (%)					
	$D^0\tau\nu$	$D^{*0}\tau\nu$	$D^+\tau\nu$	$D^{*+}\tau\nu$	$D\tau\nu$	$D^*\tau\nu$
	Additive systematic uncertainties					
MC statistics (PDF shape)	11.5	8.4	4.5	1.8	6.9	4.7
MC statistics (constraints)	4.2	1.9	6.1	1.3	3.6	1.4
Combinatorial BG modeling	7.5	4.1	11.5	2.6	9.1	2.9
D^{**} modeling	5.7	0.5	1.6	0.2	3.0	0.4
$\bar{B} \to D^*$ form factors	1.9	0.7	0.8	0.2	1.4	0.4
$\bar{B} \to D$ form factors	0.2	0.7	0.6	0.2	0.3	0.4
$m^2_{\rm miss}$ tail modeling	1.5	0.5	1.2	0.4	1.6	0.1
π^0 cross feed constraints	0.5	1.1	0.5	0.9	0.5	1.0
D^{**} feed-down	0.4	0.1	0.1	0.3	0.2	0.2
$D^{**}\tau^-\bar{\nu}_\tau$ abundance	0.4	1.3	0.3	0.2	0.3	0.8
Total additive	15.6	9.7	14.0	3.6	12.5	5.8
	Multiplicative systematic uncertainties					
MC statistics (efficiency)	1.2	1.1	1.5	1.1	1.0	0.8
Bremsstrahlung/FSR	0.6	0.5	0.3	0.4	0.4	0.5
Tracking ε	0.0	0.0	0.0	0.0	0.0	0.0
e PID ε	0.5	0.5	0.6	0.6	0.6	0.6
μ PID ε	0.5	0.6	0.7	0.6	0.6	0.6
K PID ε	0.2	0.1	0.2	0.0	0.2	0.0
π PID ε	0.1	0.1	0.2	0.0	0.1	0.1
K^0_S ε	0.1	0.0	0.1	0.1	0.1	0.0
Neutral (π^0 and γ) ε	0.0	0.0	0.0	0.1	0.0	0.0
Daughter $\mathcal{B}$'s	0.1	0.3	0.0	0.1	0.1	0.3
$\mathcal{B}(\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau)$	0.2	0.2	0.2	0.2	0.2	0.2
Total multiplicative	1.6	1.5	1.8	1.4	1.4	1.3
Total	15.6	9.9	14.0	3.9	12.5	6.0
$\mathcal{B}(\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell)$	10.2	7.7	9.4	3.7	6.8	3.4

We describe the individual contributions to the systematic uncertainty below. We divide the systematics into two broad categories: additive and multiplicative. Additive systematic uncertainties are those which affect the fit yields and therefore reduce the significance of the measured signals. Multiplicative uncertainties affect the normalization of the signals and the numerical results but not the significance.

A. Additive systematic uncertainties

In order to estimate additive systematic uncertainties, we perform an ensemble of fits to MC event samples. For each source of uncertainty, we perform a number of tests where we modify, as appropriate, the fit shapes, cross feed con-

straints, and the combinatorial background yields (all of which are fixed to MC-derived values in the nominal fit) and perform a signal fit. By doing a large number of such tests and studying the distribution of fit results in these ensembles, we are able to estimate the systematic uncertainties. In all of these ensembles, we take the RMS of the observed distribution, relative to the corresponding mean fit value, as the systematic uncertainty.

1. Monte Carlo statistics

In order to study the systematic uncertainties due to limited Monte Carlo statistics, we perform two ensembles of fits. In the first ensemble, we perform a variation of the PDF shapes. Each of the 37 PDFs are independently varied by generating new values for each of the 18 or 24 shape parameters according to the uncertainties in the PDF fit, taking into account correlations between the fitted parameters. In the second ensemble, we vary each of the feed-up and feed-down constraints according to their statistical uncertainties.

Figure 15 shows distributions of fit results for the ensemble of PDF shape fits.

2. Combinatorial background modeling

Table VII summarizes the physical sources of combinatorial background considered in this analysis, including their approximate abundances in our $B\bar{B}$ MC sample after all event selection. In order to study systematic effects, we perform an ensemble of fits, reweighting events from the various combinatorial sources.

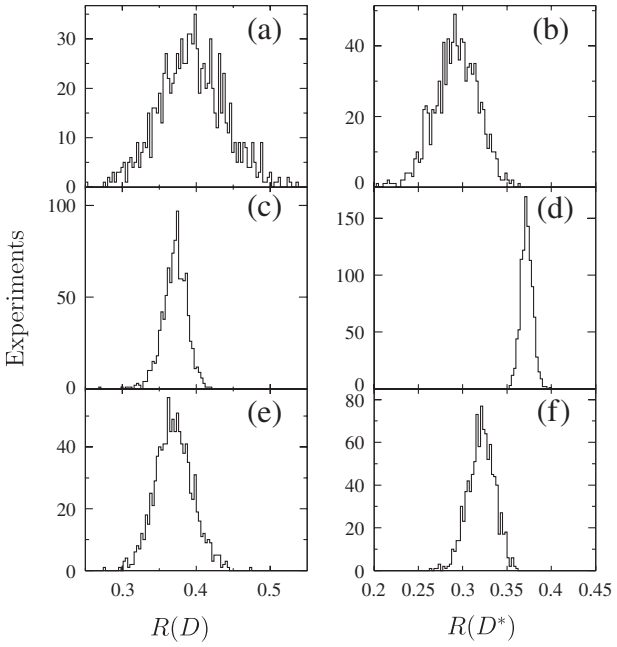

FIG. 15. Distributions of fit results for systematic uncertainties due to Monte Carlo statistics, shown for (a) $B^- \to D^0\tau^-\bar{\nu}_\tau$, (b) $B^- \to D^{*0}\tau^-\bar{\nu}_\tau$, (c) $\bar{B}^0 \to D^+\tau^-\bar{\nu}_\tau$, (d) $\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$, (e) $\bar{B} \to D\tau^-\bar{\nu}_\tau$, and (f) $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$. In all figures, the branching-fraction ratio R is shown.

TABLE VII. Sources and approximate abundances of combinatorial background in our $B\bar{B}$ MC sample. All four signal channels are combined here. The third and fourth columns show what fraction of the B decays in each group have previously been observed. The fourth column is the product of the second and third, and indicates how much of the estimated combinatorial background is known from other measurements.

Source	% of total BG	% $\mathcal{B}$ measured (Relative)	(Absolute)
$B \to D_s^{(*)+} D^{(*)}$ (+ light hadrons)			
...with $D_s^+ \to \tau\nu$	30	90	27
...with $D_s^+ \to \ell\nu(\phi/\eta/\eta')$	10	90	9
$B \to D^{(*)} D^{(*)}$ (+ light hadrons)	35	65	25
Both $\bar{B} \to D^{(*)}\ell\nu$	15	100	15
Fake lepton	5	0	0
Other misreconstructed	5	0	0
Total			75

In total, the two-body B decays $B \to D_s^{(*)+} D^{(*)(*)}$ and $B \to D^{(*)} D^{(*)}$ constitute approximately 45% of the total combinatorial background yield, while the three-body decays $B \to D^{(*)} D^{(*)} K$ contribute another 15%. Branching fractions of most of the relevant two-body B decays (and some of the three-body decays as well) have previously been measured. These branching fractions are listed in Table VIII, along with relevant branching fractions of the D_s^+ meson.

To study systematic uncertainties related to combinatorial background modeling, we perform an ensemble of fits. In each fit, we reweight events in the simulation. For modes listed in Table VIII, we reweight the branching fraction, generating random weights from a Gaussian distribution based on the measured value (for decays involving a D_s^+ meson, the weight is the product of weights for both the B and D_s^+ decays). For charge-cross feed events (true $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ events where the $B_{\rm tag}$ and signal $D^{(*)}$ swap a charged particle), the dominant systematic uncertainty is not the branching fraction, but rather the efficiency to reconstruct the $B_{\rm tag}$ with the wrong charge. We estimate a 10% uncertainty on the modeling of this process, i.e., we generate weights for these events using a Gaussian with a mean of 1 and a width of 0.1. For double-semileptonic events, with both B mesons decaying to $D^{(*)}\ell^-\bar{\nu}_\ell$, again, the dominant uncertainty comes from the probability to misreconstruct a $B_{\rm tag}$ candidate in this event, and we assume a 10% uncertainty on this number as well. For events in which the signal lepton is misidentified, we assign a 10% uncertainty; the typical fake rate measured in data is 2%–3%, with data-simulation discrepancies generally 10% or less in the momentum ranges of interest. For all remaining sources of combinatorial background, including high-multiplicity $B \to Dhh(h\ldots)$ and $B \to DDhh(h\ldots)$ (where D here represents any charm meson

and h any light meson) and other misreconstructed events, we assume a 50% uncertainty in the relevant rates.

In each test, we fit the reweighted MC sample to generate new PDF shapes and recalculate the expected yield of combinatorial events in each channel. Figure 16 shows the effect of this reweighting on the combinatorial BG PDF in the signal channels. We note that the reweighting affects the normalization of the charge-cross feed backgrounds but not the shape.

3. D^{**} Modeling

We generate an ensemble of D^{**} models, sampling from the distribution in Table IX. This model is based on the current world averages [16,28] but includes information from selected recent measurements [33] and imposes isospin symmetry between charged and neutral B modes.

TABLE VIII. Branching fractions of D_s^+ and two- and three-body B decays contributing to combinatorial background. Measurements are taken from [31], except (†) which are taken from [32]. The last column gives the branching fraction used to generate the *BABAR* MC sample, where each number is shown in the same scale as the corresponding number in the second column.

Mode		$\mathcal{B}$	MC
$D_s^+ \to \tau\nu$		$(6.4 \pm 1.5) \times 10^{-2}$	7.0
$D_s^+ \to \eta\ell\nu$		$(2.5 \pm 0.7) \times 10^{-2}$	2.6
$D_s^+ \to \eta'\ell\nu$		$(8.9 \pm 3.3) \times 10^{-3}$	8.9
$D_s^+ \to \phi\ell\nu$		$(2.0 \pm 0.5) \times 10^{-2}$	2.0
$D_s^+ \to \mu\nu$		$(5.0 \pm 1.9) \times 10^{-3}$	4.6
$B^+ \to \bar{D}^0 D_s^+$		$(1.3 \pm 0.4) \times 10^{-2}$	1.06
$B^+ \to \bar{D}^0 D_s^{*+}$		$(9 \pm 4) \times 10^{-3}$	9.1
$B^+ \to \bar{D}^{*0} D_s^+$		$(1.2 \pm 0.5) \times 10^{-2}$	1.02
$B^+ \to \bar{D}^{*0} D_s^{*+}$		$(2.7 \pm 1.0) \times 10^{-2}$	2.28
$B^+ \to D^{**0} D_s^{(*)+}$		$(2.7 \pm 1.2) \times 10^{-2}$	3.0
$B^+ \to \bar{D}^0 D^{*+} K^0$		$(5.2 \pm 1.2) \times 10^{-3}$	5.2
$B^+ \to \bar{D}^{*0} D^{*+} K^0$		$(7.8 \pm 2.6) \times 10^{-3}$	7.8
$B^+ \to D^0 \bar{D}^0 K^+$		$(1.37 \pm 0.32) \times 10^{-3}$	1.9
$B^+ \to D^{*0} \bar{D}^{*0} K^+$		$(5.3 \pm 1.6) \times 10^{-3}$	5.3
$B^+ \to D^{*0} \bar{D}^0 K^+$		$(4.7 \pm 1.0) \times 10^{-3}$	4.8
$B^+ \to D^{*-} D^+ K^+$		$(1.5 \pm 0.4) \times 10^{-3}$	0.5
$B^0 \to D^- D_s^+$		$(8.0 \pm 3.0) \times 10^{-3}$	7.4
$B^0 \to D^{*-} D_s^+$		$(1.07 \pm 0.29) \times 10^{-2}$	1.03
$B^0 \to D^- D_s^{*+}$		$(1.0 \pm 0.5) \times 10^{-2}$	0.74
$B^0 \to D^{*-} D_s^{*+}$		$(1.9 \pm 0.5) \times 10^{-2}$	1.97
$B^0 \to D^- D^0 K^+$		$(1.7 \pm 0.4) \times 10^{-3}$	1.7
$B^0 \to D^- D^{*0} K^+$		$(3.1 \pm 0.6) \times 10^{-3}$	3.1
$B^0 \to D^{*-} D^{*0} K^+$		$(1.18 \pm 0.20) \times 10^{-2}$	1.18
$B^0 \to D^{*-} D^+ K^0$		$(6.5 \pm 1.6) \times 10^{-3}$	8.1
$B^0 \to D^{*-} D^{*+} K^0$		$(8.8 \pm 1.9) \times 10^{-3}$	8.8
$B^0 \to D^{*+} D^{*-}$	(†)	$(8.1 \pm 1.2) \times 10^{-4}$	8.3
$B^0 \to D^+ D^{*-}$	(†)	$(10.4 \pm 2.0) \times 10^{-4}$	6.7
$B^0 \to D^+ D^-$	(†)	$(2.8 \pm 0.7) \times 10^{-4}$	2.7

For each test, we generate random numbers for the six exclusive modes (D, D^*, and the resonant D^{**} states), independently for B^+ and B^0 decays. We then saturate the remaining inclusive $b \to c\ell^-\bar{\nu}_\ell$ rate with the four nonresonant states, maintaining the Monte Carlo ratio of 0.1:0.3:0.2:0.6. Even though we are only interested in the $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ states, we need to generate distributions of the $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ branching fractions to allow for sufficient variations in the nonresonant states which are used to saturate the total rate.

For each test, we reweight both the $D^{**}\ell^-\bar{\nu}_\ell$ PDFs and cross feed constraints to estimate the systematic uncertainty.

4. $\bar{B} \to D^{(*)}$ Form factors

We reweight the form factors of both signal $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ and normalization $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays. In both cases, we use the form factor parameterization of Caprini, Lellouch, and Neubert [27], with numerical parameters given in Sec. III. We reweight signal and normal-

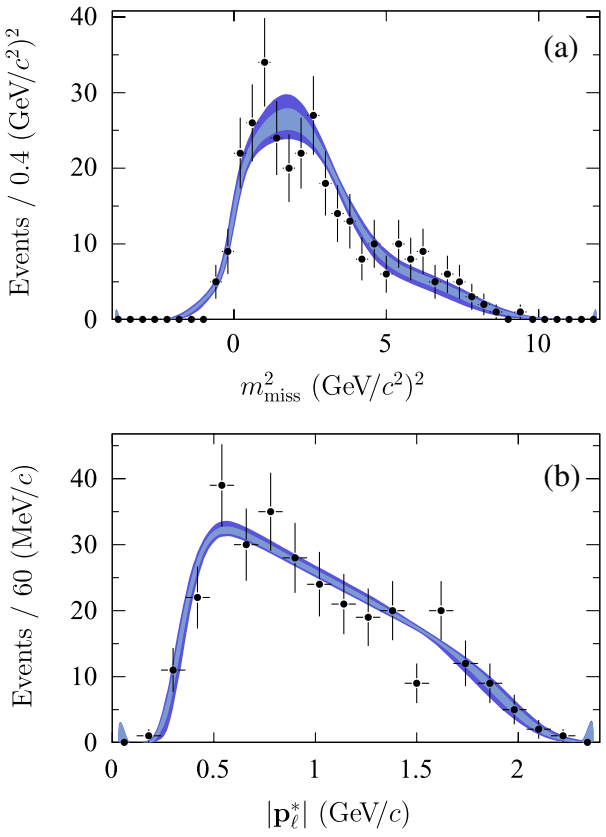

FIG. 16 (color online). Combinatorial background modeling variation for the four signal channels, showing the projections onto (a) $m^2_{\rm miss}$, and (b) $|\mathbf{p}^*_\ell|$. In both figures, the MC sample without reweighting is shown as data points, while the light and dark shaded regions show the 1σ and 2σ envelopes of the ensemble of reweighted PDF shapes.

TABLE IX. $\bar{B} \to X_c \ell^- \bar{\nu}_\ell$ branching fractions used in the D^{**} modeling systematic study. The first line, $c\ell\nu$, represents the inclusive semileptonic branching fraction. For the six lines representing the D, D^*, and D^{**} resonant states, the distribution of these branching fractions is taken to be Gaussian with the given mean and width. For the last four lines, representing the nonresonant D^{**} states, the ranges of variation are not shown in this table; their distribution is determined by the inclusive rate and the other exclusive modes, as described in the text. The generated branching fractions, $\mathcal{B}_{\rm gen}$, are the same for charged and neutral B mesons. All numbers are in %.

Mode	$\mathcal{B}_{\rm gen}$	$\mathcal{B}(\bar{B}^0)$		$\mathcal{B}(B^-)$	
		μ	σ	μ	σ
$c\ell\nu$	10.4	10.17	0	10.9	0
D	2.10	2.14	0.14	2.29	0.16
D^*	5.6	5.54	0.25	5.94	0.24
D_1	0.56	0.47	0.08	0.58	0.06
D_2^*	0.37	0.35	0.07	0.46	0.08
D_0^*	0.20	0.46	0.09	0.45	0.09
D_1'	0.37	0.85	0.20	0.83	0.20
$D^*\pi^0$	0.1	0.03	—	0.029	—
$D\pi^0$	0.3	0.09	—	0.088	—
$D^*\pi^\pm$	0.2	0.06	—	0.058	—
$D\pi^\pm$	0.6	0.18	—	0.175	—

ization modes simultaneously and generate new PDFs, cross feed constraints, and relative efficiencies.

5. $m^2_{\rm miss}$ *Tail modeling*

Studies in the two kinematic control samples show acceptable overall agreement between data and simulation for the $m^2_{\rm miss}$ resolution (see Fig. 8(d)), but suggest that the simulation may underestimate the ratio of the number of events in the large $m^2_{\rm miss}$ tail region to the number of events near $m^2_{\rm miss} = 0$. We estimate that this tail component of the resolution may be underestimated by up to 10%. We study systematic effects related to this by reweighting events at large $m^2_{\rm miss}$, greater than 1 $(\mathrm{GeV}/c^2)^2$, up by 10%, modifying the PDF shapes for $\bar{B} \to D\ell^-\bar{\nu}_\ell$ and $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$. We perform a fit with these modified PDFs and take the difference from the nominal fit as a systematic uncertainty.

6. π^0 *Efficiency and Cross feed Constraints*

While the systematic uncertainties due to detector efficiencies (described in more detail in Sec. IX B 3) are primarily multiplicative, the efficiencies for π^0 reconstruction have a large impact on the feed-down efficiencies and therefore the fit yields. This effect can be enhanced by the fact that the feed-down constraints are defined as the ratio of the number of events reconstructed in the $D\ell^-$ channel to that in the $D^*\ell^-$ channel, which move in opposite directions as the π^0 efficiency is varied.

We generate an ensemble of fits by varying the π^0 efficiency within its uncertainty, 3.0% per π^0. The resulting changes in the feed-down constraints for both signal and background modes are propagated through the signal fit to estimate the resulting systematic uncertainties.

7. $D^{**}\ell^-\bar{\nu}_\ell$ *Feed-down*

We assign an additional systematic uncertainty on $D^{**}\ell^-\bar{\nu}_\ell$ feed-down rates due to the fact that the π^0 mesons involved in feed-down processes typically have low momentum, while the 3.0% systematic uncertainty mentioned above is derived from a control sample with a broad spectrum. Since we float the constraints describing $D^* \Rightarrow D$ feed-down in the fit, D^* feed-down processes are insensitive to systematic effects due to the π^0 efficiency at low momentum. The D^{**} feed-down constraints, however, are taken from simulation and can therefore be affected.

We compare the fitted values of the $D^* \Rightarrow D$ feed-down rates to the simulation to estimate that the efficiency for low-momentum π^0 mesons is correctly modeled to within 10%. We generate an ensemble of fits in which we vary the π^0 reconstruction efficiency $\pm 10\%$ for π^0 mesons with momentum less than 300 MeV/c. We generate new PDFs and feed-down constraints which we propagate through the signal fit to estimate the systematic uncertainties.

8. $\bar{B} \to D^{**}\tau^-\bar{\nu}_\tau$ *Abundance*

We vary the fraction of $\bar{B} \to D^{**}\tau^-\bar{\nu}_\tau$ events in the D^{**} samples by generating random numbers from a Gaussian distribution with mean 1.0 and width 0.3, equivalent to a $\pm 30\%$ variation. For each test, we generate new PDFs and cross feed constraints to estimate the systematic uncertainty.

B. Multiplicative systematic uncertainties

1. Monte Carlo statistics

The dominant multiplicative systematic uncertainty is due to limited Monte Carlo statistics. The various MC samples are independent of one another, so that there is no cancellation between the signal and normalization.

2. Bremsstrahlung and final-state radiation

Based on a control sample of identified electrons and studies in MC samples, we estimate the uncertainty on reconstruction efficiency due to bremsstrahlung and final-state radiation effects to be 2.1%. This uncertainty applies to both signal and normalization modes, however, and so the effect on the relative efficiency is expected to cancel. The fractions of events in which a photon is radiated are nearly the same between signal and normalization modes, within statistical precision of 10%; we therefore treat the uncertainty between signal and normalization modes as 90% correlated to calculate the final systematic uncertainty.

TABLE X. Results from fits to data: the signal yield ($N_{\rm sig}$), the yield of normalization $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ events ($N_{\rm norm}$), the relative systematic error due to the fit yields [$(\Delta R/R)_{\rm fit}$], the relative systematic error due to the efficiency ratios [$(\Delta R/R)_\varepsilon$], the branching-fraction ratio (R), the absolute branching fraction ($\mathcal{B}$), and the total and statistical signal significances ($\sigma_{\rm tot}$ and $\sigma_{\rm stat}$). The first two errors on R and $\mathcal{B}$ are statistical and systematic, respectively; the third error on $\mathcal{B}$ represents the uncertainty on the normalization mode. The last two rows show the results of the fit with the $B^- - \bar{B}^0$ constraint applied, where $\mathcal{B}$ is expressed for the $\bar{B}^0$. The statistical correlation between $R(D)$ and $R(D^*)$ in this fit is -0.51.

Mode	$N_{\rm sig}$	$N_{\rm norm}$	$(\Delta R/R)_{\rm fit}$ [%]	$(\Delta R/R)_\varepsilon$ [%]	R [%]	$\mathcal{B}$ [%]	$\sigma_{\rm tot}$ ($\sigma_{\rm stat}$)
$B^- \to D^0\tau^-\bar{\nu}_\tau$	35.6 ± 19.4	347.9 ± 23.1	15.5	1.6	$31.4 \pm 17.0 \pm 4.9$	$0.67 \pm 0.37 \pm 0.11 \pm 0.07$	1.8 (1.8)
$B^- \to D^{*0}\tau^-\bar{\nu}_\tau$	92.2 ± 19.6	1629.9 ± 63.6	9.7	1.5	$34.6 \pm 7.3 \pm 3.4$	$2.25 \pm 0.48 \pm 0.22 \pm 0.17$	5.3 (5.8)
$\bar{B}^0 \to D^+\tau^-\bar{\nu}_\tau$	23.3 ± 7.8	150.2 ± 13.3	13.9	1.8	$48.9 \pm 16.5 \pm 6.9$	$1.04 \pm 0.35 \pm 0.15 \pm 0.10$	3.3 (3.6)
$\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$	15.5 ± 7.2	482.3 ± 25.5	3.6	1.4	$20.7 \pm 9.5 \pm 0.8$	$1.11 \pm 0.51 \pm 0.04 \pm 0.04$	2.7 (2.7)
$B \to D\tau^-\bar{\nu}_\tau$	66.9 ± 18.9	497.8 ± 26.4	12.4	1.4	$41.6 \pm 11.7 \pm 5.2$	$0.86 \pm 0.24 \pm 0.11 \pm 0.06$	3.6 (4.0)
$B \to D^*\tau^-\bar{\nu}_\tau$	101.4 ± 19.1	2111.5 ± 68.1	5.8	1.3	$29.7 \pm 5.6 \pm 1.8$	$1.62 \pm 0.31 \pm 0.10 \pm 0.05$	6.2 (6.5)

3. Detector Efficiencies

We estimate systematic uncertainties related to the detector efficiencies—track and neutral reconstruction and charged particle identification—by studying these efficiencies in several control samples in both data and simulation. We correct the MC efficiencies to match those seen in the data, and we take the statistical precision of these studies as an estimate of the systematic uncertainty on absolute efficiencies.

Since we normalize our signals to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$, we calculate systematic uncertainties on the relative efficiency, treating uncertainties on the signal and normalization modes as correlated. The degree of correlation, and therefore, the degree to which the uncertainty cancels, is determined by the kinematics of the two samples. For most of the final-state particles, the kinematic distributions are very similar between signal and normalization modes and so the systematic uncertainty cancels almost entirely. For the charged leptons, however, the momentum spectra are very different between signal and normalization (see Fig. 2), and so the associated systematic uncertainty is larger.

4. Hadronic daughter branching fractions

We reconstruct both signal and normalization modes using the same set of final states, so uncertainties due to the branching fractions of these states very nearly cancel. [The $D^{(*)}$ momentum spectra are slightly different between signal and normalization modes, so this cancellation is not perfect.] We take the uncertainty on each of the reconstructed D^*, D, K^0_S, and π^0 decay modes from [16] and propagate each of these uncertainties through to the relative efficiency, using the relative abundance of each decay chain in the signal and normalization MC samples to determine the correlation and the degree of cancellation.

5. Leptonic τ branching fraction

The τ branching fraction $\mathcal{B}(\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau)$ appears only in the denominator of Eq. (18) and therefore contributes a 0.2% systematic uncertainty on all modes [16] without cancellation.

X. RESULTS

Table X summarizes the results from two fits, one in which all four signal yields can vary independently, and the second $B^- - \bar{B}^0$ constrained fit with $R(D^+) = R(D^0)$ and $R(D^{*+}) = R(D^{*0})$. We observe approximately 67 $\bar{B} \to D\tau^-\bar{\nu}_\tau$ and 101 $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ signal events in this $B^- - \bar{B}^0$-constrained fit, corresponding to signal branching-fraction ratios of $R(D) = (41.6 \pm 11.7 \pm 5.2)\%$ and $R(D^*) = (29.7 \pm 5.6 \pm 1.8)\%$, where the first error is statistical and the second systematic. Normalizing these to known $\bar{B}^0$ branching fractions,[8] we obtain $\mathcal{B}(\bar{B} \to D\tau^-\bar{\nu}_\tau) = (0.86 \pm 0.24 \pm 0.11 \pm 0.06)\%$ and $\mathcal{B}(\bar{B} \to D^*\tau^-\bar{\nu}_\tau) = (1.62 \pm 0.31 \pm 0.10 \pm 0.05)\%$, where the third error is from that on the normalization branching fraction.

Table X also gives the significances of the signal yields. The statistical significance is determined from $\sqrt{2\Delta(\ln\mathcal{L})}$, where $\Delta(\ln\mathcal{L})$ is the change in log-likelihood between the nominal fit and the no-signal hypothesis. The total significances are determined by including the systematic uncertainties on the fit yields in quadrature with the statistical errors. In the $B^- - \bar{B}^0$-constrained fit, the signal significances are 3.6σ and 6.2σ for $R(D)$ and $R(D^*)$, respectively.

The statistical correlation between $R(D)$ and $R(D^*)$ is -0.51 in the $B^- - \bar{B}^0$-constrained fit. This correlation is due to the fact that most of the events at large $m^2_{\rm miss}$ are either $\bar{B} \to D\tau^-\bar{\nu}_\tau$ or $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ signal events, and increasing either of the two signal yields in the fit necessarily decreases the other. The systematic uncertanties have a correlation of -0.03 between $R(D)$ and $R(D^*)$; most of the systematic uncertainties have large negative

[8] We use [16] to normalize the four individual branching fractions. For the $B^- - \bar{B}^0$-constrained measurement, we use our own averages of the values in [16]: $\mathcal{B}(\bar{B}^0 \to D^+\ell^-\bar{\nu}_\ell) = (2.07 \pm 0.14)\%$ and $\mathcal{B}(\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell) = (5.46 \pm 0.18)\%$.

correlations for the same reason that the statistical uncertainty does, but the combinatorial background uncertainty affects both signal yields in a coherent manner and so contributes a large positive correlation. The sum of the two branching fractions, taking all correlations into account, is $\mathcal{B}(\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau) = (2.48 \pm 0.28 \pm 0.15 \pm 0.08)\%$.

Figures 17 and 18 show the observed q^2 distributions in the four signal channels in the low and high $m^2_{\rm miss}$ regions, respectively. The histograms in these figures are taken from MC samples of the various components, with each component scaled to match the yield in the $B^- - \bar{B}^0$-constrained fit; since q^2 is not a fit variable, we cannot show a projection of a continuous PDF as was done in Figs. 10–14. As before, we observe good agreement between the data and the expectation from simulation, in both the low and high $m^2_{\rm miss}$ regions. Since the q^2 distribution is highly dependent on the form factor model, we note that the CLN model describes both normalization and signal events within the available statistics.

Table XI summarizes the results of several cross-checks, including splitting up the sample according to lepton flavor, lepton charge, and data-taking period. We have done these checks by performing "cut-and-count" analyses, both in the data and in simulated event samples. In all cases, the results in data are consistent with our expectations from simulation. The first row in this table shows the fraction of events with muon candidates in data and simu-

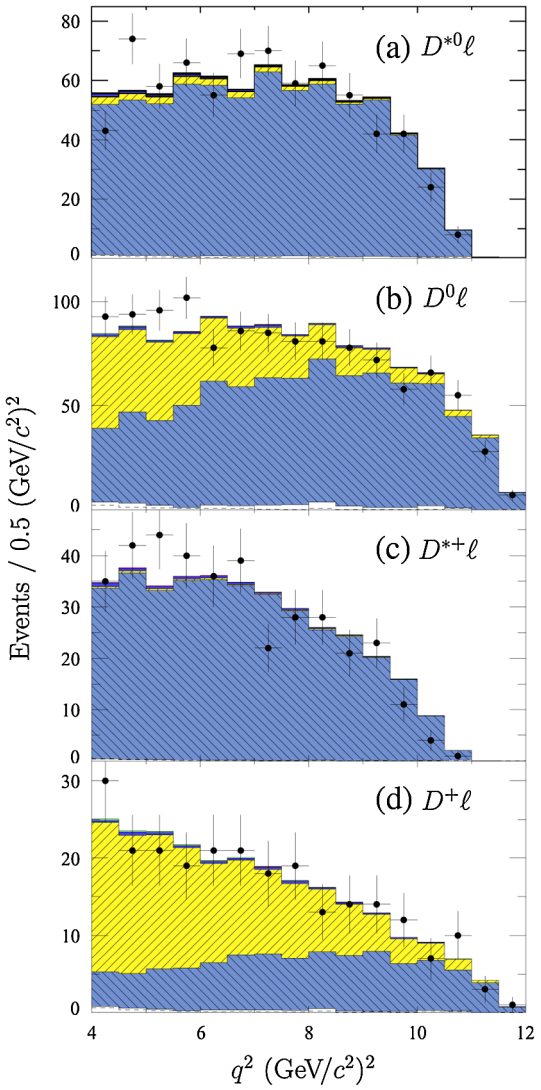

FIG. 17 (color online). q^2 distributions of events in the four final states $D^{*0}\ell^-$, $D^0\ell^-$, $D^{*+}\ell^-$, and $D^+\ell^-$, shown in the normalization region, $m^2_{\rm miss} < 1$ (GeV/c^2)^2. The data are shown as points with error bars. The shaded histograms are taken from MC samples with normalizations from the fit to data. The components are shaded as in Fig. 10.

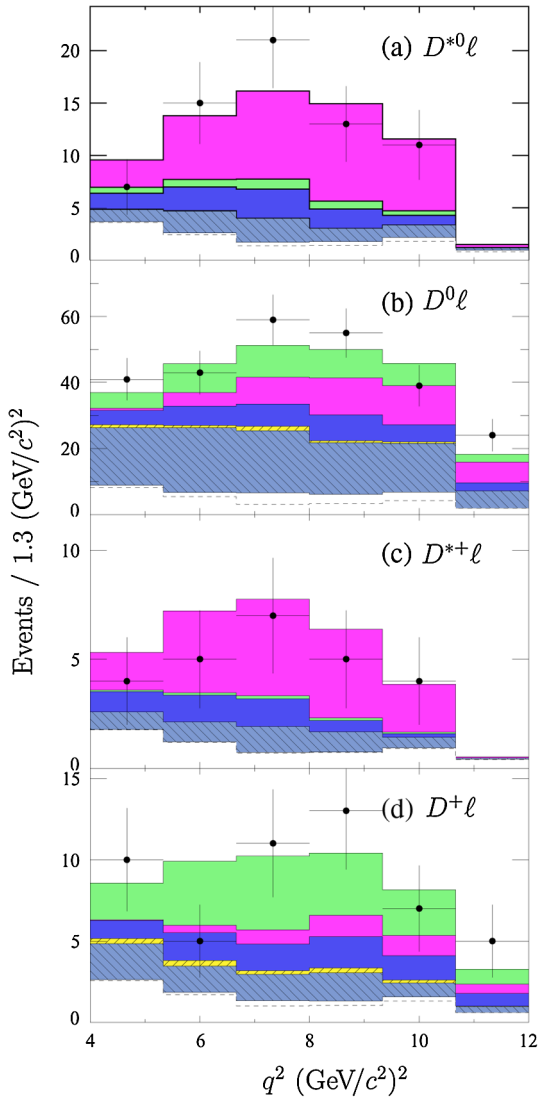

FIG. 18 (color online). q^2 distributions of events in the four final states $D^{*0}\ell^-$, $D^0\ell^-$, $D^{*+}\ell^-$, and $D^+\ell^-$, shown in the signal region, $m^2_{\rm miss} > 1$ (GeV/c^2)^2. The data are shown as points with error bars. The shaded histograms are taken from MC samples with normalizations from the fit to data. The components are shaded as in Fig. 10.

TABLE XI. Cross-check studies, splitting the data according to lepton flavor, lepton charge, and running period. The first row shows the fraction of events with muon candidates for both data and MC samples, for both the full-event sample and for the signal-sensitive region $m^2_{\rm miss} > 1\ ({\rm GeV}/c^2)^2$. The second row shows fractions of events with positively charged lepton candidates, and the third row shows the fractions of events recorded in Run 4. In all cases, the data are consistent with the simulation and with expectations.

	Full sample		High $m^2_{\rm miss}$	
Sample	$f_{\rm data}$ (%)	$f_{\rm MC}$ (%)	$f_{\rm data}$ (%)	$f_{\rm MC}$ (%)
μ	40.0 ± 0.9	40.9 ± 0.4	30.7 ± 2.3	31.9 ± 1.0
ℓ^+	50.2 ± 1.0	49.2 ± 0.4	49.3 ± 2.5	48.9 ± 1.0
Run 4	44.6 ± 1.0	47.6 ± 0.5	46.8 ± 2.5	48.5 ± 1.2

lation, both for the full-event sample and for the signal-sensitive region in $m^2_{\rm miss}$. Electron identification is more efficient than muon ID, which is why the muon fraction in the final sample is less than 50%, and, at lower momenta (which generally correspond to larger $m^2_{\rm miss}$), this efficiency difference is more pronounced; in both cases, however, the muon abundance is well modeled by the simulation. The next row shows the fraction of positively charged lepton candidates (versus negatively charged candidates), and all samples are consistent with the expected 50/50 split. The last row shows the fraction of events recorded during the Run 4 *BABAR* data-taking period; Run 4 had significantly different accelerator background conditions from Runs 1–3, which could affect missing-energy analyses. The fraction of events in the Run 4 subsample is consistent with expectations: Run 4 makes up 47% of the total luminosity.

We also check that the yield of normalization events is consistent with the $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ branching fractions we use above. It is difficult to estimate the absolute efficiency of our $B_{\rm tag}$ reconstruction with sufficient precision to measure the $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ branching fractions directly, which is why we instead base our signal measurement on the relative normalization. We see normalization event yields consistent with the known values to within about 20%, which is what we expect given the uncertainties on our $B_{\rm tag}$ reconstruction efficiency.

We estimate the goodness of fit using an ensemble of simulated experiments. We generate 1000 event samples, using the nominal PDFs for the fit to data and event yields based on the $B^- - \bar{B}^0$-constrained fit to data. We fit each of these samples both with and without the $B^- - \bar{B}^0$ constraints and study the distribution of $-\log\mathcal{L}$ in these fits.

Figure 19 shows the distribution of $-\log\mathcal{L}$ for the two ensembles of fits. In both cases, the value of $-\log\mathcal{L}$ obtained in the fit to data is indicated with an arrow, and, in both cases, this value is found within the central part of the Monte Carlo distribution, indicating a good fit. In the unconstrained fit, 11.7% of the simulated experiments have a value of $-\log\mathcal{L}$ greater than the value observed in data,

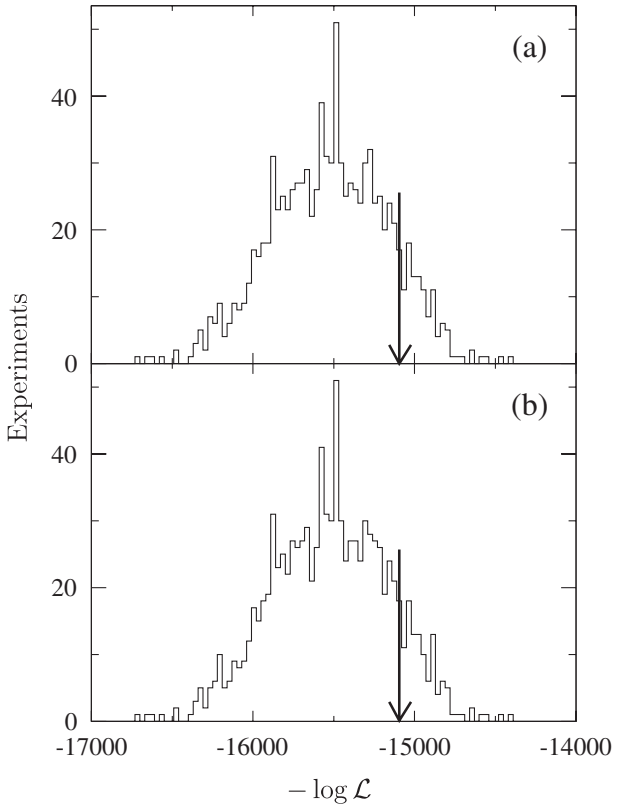

FIG. 19. Distribution of $-\log\mathcal{L}$ from simulated experiments, showing (a) the unconstrained fit and (b) the $B^- - \bar{B}^0$ constrained fit. The observed values of $-\log\mathcal{L}$ in the fit to data are indicated with arrows. The fraction of experiments with $-\log\mathcal{L}$ larger than the observed value is used to estimate the goodness of fit.

corresponding to the probability that we expect to observe a fit as bad, or worse, than the one actually observed. This probability is large, indicating an acceptable goodness of fit. The corresponding probability for the $B^- - \bar{B}^0$ constrained fit is 11.8%, also large.

XI. CONCLUSIONS

We have presented measurements of the branching fractions for the decays $\bar{B} \to D\tau^-\bar{\nu}_\tau$ and $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$, determined relative to the corresponding decays to light leptons. We measure the branching-fraction ratios for four individual $D^{(*)}$ states, as well as two $B^- - \bar{B}^0$-constrained ratios

$$R(D^0) = (31.4 \pm 17.0 \pm 4.9)\%,$$
$$R(D^{*0}) = (34.6 \pm 7.3 \pm 3.4)\%,$$
$$R(D^+) = (48.9 \pm 16.5 \pm 6.9)\%,$$
$$R(D^{*+}) = (20.7 \pm 9.5 \pm 0.8)\%,$$
$$R(D) = (41.6 \pm 11.7 \pm 5.2)\%,$$
$$R(D^*) = (29.7 \pm 5.6 \pm 1.8)\%,$$

where the first uncertainty is statistical and the second is

systematic. The significances of these signals are 1.8σ, 5.3σ, 3.3σ, 2.7σ, 3.6σ, and 6.2σ, respectively. The statistical and systematic uncertainties on $R(D)$ and $R(D^*)$ have correlations of -0.51 and -0.03, respectively.

From these branching-fraction ratios and known branching fractions of the normalization modes $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$, we derive the absolute branching fractions

$$\mathcal{B}(B^- \to D^0\tau^-\bar{\nu}_\tau) = (0.67 \pm 0.37 \pm 0.11 \pm 0.07)\%,$$
$$\mathcal{B}(B^- \to D^{*0}\tau^-\bar{\nu}_\tau) = (2.25 \pm 0.48 \pm 0.22 \pm 0.17)\%,$$
$$\mathcal{B}(\bar{B}^0 \to D^+\tau^-\bar{\nu}_\tau) = (1.04 \pm 0.35 \pm 0.15 \pm 0.10)\%,$$
$$\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau) = (1.11 \pm 0.51 \pm 0.04 \pm 0.04)\%,$$
$$\mathcal{B}(\bar{B} \to D\tau^-\bar{\nu}_\tau) = (0.86 \pm 0.24 \pm 0.11 \pm 0.06)\%,$$
$$\mathcal{B}(\bar{B} \to D^*\tau^-\bar{\nu}_\tau) = (1.62 \pm 0.31 \pm 0.10 \pm 0.05)\%,$$

where the third uncertainty reflects that of the normalization mode branching fraction.

The measurement of $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)$ is consistent with the Belle result [17]. The branching-fraction ratios $R(D)$ and $R(D^*)$ are about 1σ higher than the SM predictions but, given the uncertainties, there is still room for a sizeable non-SM contribution.

We have also presented distributions of the lepton momentum $|\mathbf{p}^*_\ell|$ and the squared momentum transfer q^2 for $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ events. In all cases, these distributions are consistent with expectations based on the SM and the CLN form factor model with measured form factors.

ACKNOWLEDGMENTS

We are grateful for the extraordinary contributions of our PEP-II colleagues in achieving the excellent luminosity and machine conditions that have made this work possible. The success of this project also relies critically on the expertise and dedication of the computing organizations that support *BABAR*. The collaborating institutions wish to thank SLAC for its support and the kind hospitality extended to them. This work is supported by the US Department of Energy and National Science Foundation, the Natural Sciences and Engineering Research Council (Canada), the Commissariat à l'Energie Atomique and Institut National de Physique Nucléaire et de Physique des Particules (France), the Bundesministerium für Bildung und Forschung and Deutsche Forschungsgemeinschaft (Germany), the Istituto Nazionale di Fisica Nucleare (Italy), the Foundation for Fundamental Research on Matter (The Netherlands), the Research Council of Norway, the Ministry of Education and Science of the Russian Federation, Ministerio de Educación y Ciencia (Spain), and the Science and Technology Facilities Council (United Kingdom). Individuals have received support from the Marie-Curie IEF program (European Union) and the A. P. Sloan Foundation.

[1] J. G. Körner and G. A. Schuler, Phys. Lett. B **231**, 306 (1989); Z. Phys. C **46**, 93 (1990).
[2] A. F. Falk *et al.*, Phys. Lett. B **326**, 145 (1994).
[3] D. S. Hwang and D.-W. Kim, Eur. Phys. J. C **14**, 271 (2000).
[4] B. Grządkowski and W.-S. Hou, Phys. Lett. B **283**, 427 (1992).
[5] M. Tanaka, Z. Phys. C **67**, 321 (1995).
[6] K. Kiers and A. Soni, Phys. Rev. D **56**, 5786 (1997).
[7] H. Itoh, S. Komine, and Y. Okada, Prog. Theor. Phys. **114**, 179 (2005).
[8] C.-H. Chen and C.-Q. Geng, J. High Energy Phys. 10 (2006) 053.
[9] J. F. Kamenik and F. Mescia, Phys. Rev. D **78**, 014003 (2008).
[10] J. E. Duboscq *et al.* (CLEO Collaboration), Phys. Rev. Lett. **76**, 3898 (1996).
[11] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **74**, 092004 (2006).
[12] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **77**, 032002 (2008).
[13] N. Isgur and M. B. Wise, Phys. Lett. B **232**, 113 (1989); **237**, 527 (1990).
[14] U. Nierste, S. Trine, and S. Westhoff, Phys. Rev. D **78**, 015006 (2008).
[15] M. Acciarri *et al.* (L3 Collaboration), Phys. Lett. B **332**, 201 (1994); Z. Phys. C **71**, 379 (1996); P. Abreu *et al.* (DELPHI Collaboration), Phys. Lett. B **496**, 43 (2000); R. Barate *et al.* (ALEPH Collaboration), Eur. Phys. J. C **19**, 213 (2001); G. Abbiendi *et al.* (OPAL Collaboration), Phys. Lett. B **520**, 1 (2001).
[16] W.-M. Yao *et al.* (Particle Data Group), J. Phys. G **33**, 1 (2006).
[17] A. Matyja *et al.* (Belle Collaboration), Phys. Rev. Lett. **99**, 191807 (2007).
[18] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **100**, 021801 (2008).
[19] B. Aubert *et al.* (*BABAR* Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 1 (2002).
[20] D. Lange *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).
[21] E. Richter-Was, Phys. Lett. B **303**, 163 (1993).
[22] S. Agostinelli *et al.* (GEANT Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **506**, 250 (2003).
[23] D. Scora and N. Isgur, Phys. Rev. D **52**, 2783 (1995).
[24] M. Neubert, Phys. Rep. **245**, 259 (1994).
[25] J. L. Goity and W. Roberts, Phys. Rev. D **51**, 3459 (1995).
[26] D. Côté *et al.*, Eur. Phys. J. C **38**, 105 (2004).
[27] I. Caprini, L. Lellouch, and M. Neubert, Nucl. Phys. **B530**, 153 (1998).

[28] E. Barberio *et al.* (HFAG), arXiv:0704.3575.
[29] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **92**, 071802 (2004); **93**, 091802 (2004); Phys. Rev. D **69**, 111103 (2004); **69**, 111104 (2004); Phys. Rev. Lett. **94**, 101801 (2005); **96**, 241802 (2006); **99**, 201801 (2007); Phys. Rev. D **77**, 011104 (2008); **77**, 011107 (2008); **77**, 032007 (2008); **77**, 051103 (2008); Phys. Rev. Lett. **100**, 151802 (2008); **100**, 171802 (2008).
[30] W. D. Hulsburgen, Nucl. Instrum. Methods Phys. Res., Sect. A **552**, 566 (2005).
[31] S. Eidelman *et al.* (Particle Data Group), Phys. Lett. B **592**, 1 (2004) and 2005 partial online update for the 2006 edition. Some of these numbers have since been updated in [16] and some new modes are available, but we estimate these changes to give a negligible improvement in the analysis.
[32] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **73**, 112004 (2006).
[33] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **76**, 051101 (2007); Phys. Rev. Lett. **100**, 231803 (2008); A. Hauke, arXiv:0706.4468.