

Search for the rare decays $B \rightarrow \pi \ell^+ \ell^-$ and $B^0 \rightarrow \eta \ell^+ \ell^-$

J. P. Lees,¹ V. Poireau,¹ V. Tisserand,¹ E. Grauges,² A. Palano,^{3a,3b} G. Eigen,⁴ B. Stugu,⁴ D. N. Brown,⁵ L. T. Kerth,⁵ Yu. G. Kolomensky,⁵ G. Lynch,⁵ H. Koch,⁶ T. Schroeder,⁶ C. Hearty,⁷ T. S. Mattison,⁷ J. A. McKenna,⁷ R. Y. So,⁷ A. Khan,⁸ V. E. Blinov,⁹ A. R. Buzykaev,⁹ V. P. Druzhinin,⁹ V. B. Golubev,⁹ E. A. Kravchenko,⁹ A. P. Onuchin,⁹ S. I. Serednyakov,⁹ Yu. I. Skovpen,⁹ E. P. Solodov,⁹ K. Yu. Todyshev,⁹ A. N. Yushkov,⁹ D. Kirkby,¹⁰ A. J. Lankford,¹⁰ M. Mandelkern,¹⁰ B. Dey,¹¹ J. W. Gary,¹¹ O. Long,¹¹ G. M. Vitug,¹¹ C. Campagnari,¹² M. Franco Sevilla,¹² T. M. Hong,¹² D. Kovalskyi,¹² J. D. Richman,¹² C. A. West,¹² A. M. Eisner,¹³ W. S. Lockman,¹³ A. J. Martinez,¹³ B. A. Schumm,¹³ A. Seiden,¹³ D. S. Chao,¹⁴ C. H. Cheng,¹⁴ B. Echenard,¹⁴ K. T. Flood,¹⁴ D. G. Hitlin,¹⁴ P. Ongmongkolkul,¹⁴ F. C. Porter,¹⁴ R. Andreassen,¹⁵ Z. Huard,¹⁵ B. T. Meadows,¹⁵ M. D. Sokoloff,¹⁵ L. Sun,¹⁵ P. C. Bloom,¹⁶ W. T. Ford,¹⁶ A. Gaz,¹⁶ U. Nauenberg,¹⁶ J. G. Smith,¹⁶ S. R. Wagner,¹⁶ R. Ayad,^{17,*} W. H. Toki,¹⁷ B. Spaan,¹⁸ K. R. Schubert,¹⁹ R. Schwierz,¹⁹ D. Bernard,²⁰ M. Verderi,²⁰ S. Playfer,²¹ D. Bettoni,^{22a} C. Bozzi,^{22a} R. Calabrese,^{22a,22b} G. Cibinetto,^{22a,22b} E. Fioravanti,^{22a,22b} I. Garzia,^{22a,22b} E. Luppi,^{22a,22b} L. Piemontese,^{22a} V. Santoro,^{22a} R. Baldini-Ferroli,²³ A. Calcaterra,²³ R. de Sangro,²³ G. Finocchiaro,²³ S. Martellotti,²³ P. Patteri,²³ I. M. Peruzzi,^{23,†} M. Piccolo,²³ M. Rama,²³ A. Zallo,²³ R. Contri,^{24a,24b} E. Guido,^{24a,24b} M. Lo Vetere,^{24a,24b} M. R. Monge,^{24a,24b} S. Passaggio,^{24a} C. Patrignani,^{24a,24b} E. Robutti,^{24a} B. Bhuyan,²⁵ V. Prasad,²⁵ M. Morii,²⁶ A. Adametz,²⁷ U. Uwer,²⁷ H. M. Lacker,²⁸ P. D. Dauncey,²⁹ U. Mallik,³⁰ C. Chen,³¹ J. Cochran,³¹ W. T. Meyer,³¹ S. Prell,³¹ A. E. Rubin,³¹ A. V. Gritsan,³² N. Arnaud,³³ M. Davier,³³ D. Derkach,³³ G. Grosdidier,³³ F. Le Diberder,³³ A. M. Lutz,³³ B. Malaescu,³³ P. Roudeau,³³ A. Stocchi,³³ G. Wormser,³³ D. J. Lange,³⁴ D. M. Wright,³⁴ J. P. Coleman,³⁵ J. R. Fry,³⁵ E. Gabathuler,³⁵ D. E. Hutchcroft,³⁵ D. J. Payne,³⁵ C. Touramanis,³⁵ A. J. Bevan,³⁶ F. Di Lodovico,³⁶ R. Sacco,³⁶ G. Cowan,³⁷ J. Bougher,³⁸ D. N. Brown,³⁸ C. L. Davis,³⁸ A. G. Denig,³⁹ M. Fritsch,³⁹ W. Gradl,³⁹ K. Griessinger,³⁹ A. Hafner,³⁹ E. Prencipe,³⁹ R. J. Barlow,^{40,‡} G. D. Lafferty,⁴⁰ E. Behn,⁴¹ R. Cenci,⁴¹ B. Hamilton,⁴¹ A. Jawahery,⁴¹ D. A. Roberts,⁴¹ R. Cowan,⁴² D. Dujmic,⁴² G. Sciolla,⁴² R. Cheaib,⁴³ P. M. Patel,^{43,§} S. H. Robertson,⁴³ P. Biassoni,^{44a,44b} N. Neri,^{44a} F. Palombo,^{44a,44b} L. Cremaldi,⁴⁵ R. Godang,^{45,||} P. Sonnek,⁴⁵ D. J. Summers,⁴⁵ X. Nguyen,⁴⁶ M. Simard,⁴⁶ P. Taras,⁴⁶ G. De Nardo,^{47a,47b} D. Monorchio,^{47a,47b} G. Onorato,^{47a,47b} C. Sciacca,^{47a,47b} M. Martinelli,⁴⁸ G. Raven,⁴⁸ C. P. Jessop,⁴⁹ J. M. LoSecco,⁴⁹ K. Honscheid,⁵⁰ R. Kass,⁵⁰ J. Brau,⁵¹ R. Frey,⁵¹ N. B. Sinev,⁵¹ D. Strom,⁵¹ E. Torrence,⁵¹ E. Feltresi,^{52a,52b} M. Margoni,^{52a,52b} M. Morandin,^{52a} M. Posocco,^{52a} M. Rotondo,^{52a} G. Simi,^{52a} F. Simonetto,^{52a,52b} R. Stroili,^{52a,52b} S. Akar,⁵³ E. Ben-Haim,⁵³ M. Bomben,⁵³ G. R. Bonneaud,⁵³ H. Briand,⁵³ G. Calderini,⁵³ J. Chauveau,⁵³ Ph. Leruste,⁵³ G. Marchiori,⁵³ J. Ocariz,⁵³ S. Sitt,⁵³ M. Biasini,^{54a,54b} E. Manoni,^{54a} S. Pacetti,^{54a,54b} A. Rossi,^{54a,54b} C. Angelini,^{55a,55b} G. Batignani,^{55a,55b} S. Bettarini,^{55a,55b} M. Carpinelli,^{55a,55b,¶} G. Casarosa,^{55a,55b} A. Cervelli,^{55a,55b} F. Forti,^{55a,55b} M. A. Giorgi,^{55a,55b} A. Lusiani,^{55a,55c} B. Oberhof,^{55a,55b} E. Paoloni,^{55a,55b} A. Perez,^{55a} G. Rizzo,^{55a,55b} J. J. Walsh,^{55a} D. Lopes Pegna,⁵⁶ J. Olsen,⁵⁶ A. J. S. Smith,⁵⁶ R. Faccini,^{57a,57b} F. Ferrarotto,^{57a} F. Ferroni,^{57a,57b} M. Gaspero,^{57a,57b} L. Li Gioi,^{57a} G. Piredda,^{57a} C. Büniger,⁵⁸ O. Grünberg,⁵⁸ T. Hartmann,⁵⁸ T. Leddig,⁵⁸ C. Voß,⁵⁸ R. Waldi,⁵⁸ T. Adye,⁵⁹ E. O. Olaiya,⁵⁹ F. F. Wilson,⁵⁹ S. Emery,⁶⁰ G. Hamel de Monchenault,⁶⁰ G. Vasseur,⁶⁰ Ch. Yèche,⁶⁰ F. Anulli,⁶¹ D. Aston,⁶¹ D. J. Bard,⁶¹ J. F. Benitez,⁶¹ C. Cartaro,⁶¹ M. R. Convery,⁶¹ J. Dorfan,⁶¹ G. P. Dubois-Felsmann,⁶¹ W. Dunwoodie,⁶¹ M. Ebert,⁶¹ R. C. Field,⁶¹ B. G. Fulsom,⁶¹ A. M. Gabareen,⁶¹ M. T. Graham,⁶¹ C. Hast,⁶¹ W. R. Innes,⁶¹ P. Kim,⁶¹ M. L. Kocian,⁶¹ D. W. G. S. Leith,⁶¹ P. Lewis,⁶¹ D. Lindemann,⁶¹ B. Lindquist,⁶¹ S. Luitz,⁶¹ V. Luth,⁶¹ H. L. Lynch,⁶¹ D. B. MacFarlane,⁶¹ D. R. Muller,⁶¹ H. Neal,⁶¹ S. Nelson,⁶¹ M. Perl,⁶¹ T. Pulliam,⁶¹ B. N. Ratcliff,⁶¹ A. Roodman,⁶¹ A. A. Salnikov,⁶¹ R. H. Schindler,⁶¹ A. Snyder,⁶¹ D. Su,⁶¹ M. K. Sullivan,⁶¹ J. Va'vra,⁶¹ A. P. Wagner,⁶¹ W. F. Wang,⁶¹ W. J. Wisniewski,⁶¹ M. Wittgen,⁶¹ D. H. Wright,⁶¹ H. W. Wulsin,⁶¹ V. Ziegler,⁶¹ W. Park,⁶² M. V. Purohit,⁶² R. M. White,⁶² J. R. Wilson,⁶² A. Randle-Conde,⁶³ S. J. Sekula,⁶³ M. Bellis,⁶⁴ P. R. Burchat,⁶⁴ T. S. Miyashita,⁶⁴ E. M. T. Puccio,⁶⁴ M. S. Alam,⁶⁵ J. A. Ernst,⁶⁵ R. Gorodeisky,⁶⁶ N. Guttman,⁶⁶ D. R. Peimer,⁶⁶ A. Soffer,⁶⁶ S. M. Spanier,⁶⁷ J. L. Ritchie,⁶⁸ A. M. Ruland,⁶⁸ R. F. Schwitters,⁶⁸ B. C. Wray,⁶⁸ J. M. Izen,⁶⁹ X. C. Lou,⁶⁹ F. Bianchi,^{70a,70b} F. De Mori,^{70a,70b} A. Filippi,^{70a} D. Gamba,^{70a,70b} S. Zambito,^{70a,70b} L. Lanceri,^{71a,71b} L. Vitale,^{71a,71b} F. Martinez-Vidal,⁷² A. Oyanguren,⁷² P. Villanueva-Perez,⁷² H. Ahmed,⁷³ J. Albert,⁷³ Sw. Banerjee,⁷³ F. U. Bernlochner,⁷³ H. H. F. Choi,⁷³ G. J. King,⁷³ R. Kowalewski,⁷³ M. J. Lewczuk,⁷³ T. Lueck,⁷³ I. M. Nugent,⁷³ J. M. Roney,⁷³ R. J. Sobie,⁷³ N. Tasneem,⁷³ T. J. Gershon,⁷⁴ P. F. Harrison,⁷⁴ T. E. Latham,⁷⁴ H. R. Band,⁷⁵ S. Dasu,⁷⁵ Y. Pan,⁷⁵ R. Prepost,⁷⁵ and S. L. Wu⁷⁵

(BABAR Collaboration)

¹Laboratoire d'Annecy-le-Vieux de Physique des Particules (LAPP), Université de Savoie, CNRS/IN2P3, F-74941 Annecy-Le-Vieux, France

- ²*Departament ECM, Facultat de Física, Universitat de Barcelona, E-08028 Barcelona, Spain*
- ^{3a}*INFN Sezione di Bari, I-70126 Bari, Italy*
- ^{3b}*Dipartimento 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*Ruhr Universität Bochum, Institut für Experimentalphysik I, D-44780 Bochum, Germany*
- ⁷*University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada*
- ⁸*Brunel University, Uxbridge, Middlesex UB8 3PH, United Kingdom*
- ⁹*Budker Institute of Nuclear Physics SB RAS, Novosibirsk 630090, Russia*
- ¹⁰*University of California at Irvine, Irvine, California 92697, USA*
- ¹¹*University of California at Riverside, Riverside, California 92521, USA*
- ¹²*University of California at Santa Barbara, Santa Barbara, California 93106, USA*
- ¹³*University of California at Santa Cruz, Institute for Particle Physics, Santa Cruz, California 95064, USA*
- ¹⁴*California Institute of Technology, Pasadena, California 91125, USA*
- ¹⁵*University of Cincinnati, Cincinnati, Ohio 45221, USA*
- ¹⁶*University of Colorado, Boulder, Colorado 80309, USA*
- ¹⁷*Colorado State University, Fort Collins, Colorado 80523, USA*
- ¹⁸*Fakultät Physik, Technische Universität Dortmund, D-44221 Dortmund, Germany*
- ¹⁹*Technische Universität Dresden, Institut für Kern- und Teilchenphysik, D-01062 Dresden, Germany*
- ²⁰*Laboratoire Leprince-Ringuet, Ecole Polytechnique, CNRS/IN2P3, F-91128 Palaiseau, France*
- ²¹*University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom*
- ^{22a}*INFN Sezione di Ferrara, I-44122 Ferrara, Italy*
- ^{22b}*Dipartimento di Fisica e Scienze della Terra, Università di Ferrara, I-44122 Ferrara, Italy*
- ²³*INFN Laboratori Nazionali di Frascati, I-00044 Frascati, Italy*
- ^{24a}*INFN Sezione di Genova, I-16146 Genova, Italy*
- ^{24b}*Dipartimento di Fisica, Università di Genova, I-16146 Genova, Italy*
- ²⁵*Indian Institute of Technology Guwahati, Guwahati, Assam 781 039, India*
- ²⁶*Harvard University, Cambridge, Massachusetts 02138, USA*
- ²⁷*Universität Heidelberg, Physikalisches Institut, Philosophenweg 12, D-69120 Heidelberg, Germany*
- ²⁸*Humboldt-Universität zu Berlin, Institut für Physik, Newtonstrasse 15, D-12489 Berlin, Germany*
- ²⁹*Imperial College London, London SW7 2AZ, United Kingdom*
- ³⁰*University of Iowa, Iowa City, Iowa 52242, USA*
- ³¹*Iowa State University, Ames, Iowa 50011-3160, USA*
- ³²*Johns Hopkins University, Baltimore, Maryland 21218, USA*
- ³³*Laboratoire de l'Accélérateur Linéaire, IN2P3/CNRS et Université Paris-Sud 11, Centre Scientifique d'Orsay, B. P. 34, F-91898 Orsay Cedex, France*
- ³⁴*Lawrence Livermore National Laboratory, Livermore, California 94550, USA*
- ³⁵*University of Liverpool, Liverpool L69 7ZE, United Kingdom*
- ³⁶*Queen Mary, University of London, London E1 4NS, United Kingdom*
- ³⁷*University of London, Royal Holloway and Bedford New College, Egham, Surrey TW20 0EX, United Kingdom*
- ³⁸*University of Louisville, Louisville, Kentucky 40292, USA*
- ³⁹*Johannes Gutenberg-Universität Mainz, Institut für Kernphysik, D-55099 Mainz, Germany*
- ⁴⁰*University of Manchester, Manchester M13 9PL, United Kingdom*
- ⁴¹*University of Maryland, College Park, Maryland 20742, USA*
- ⁴²*Laboratory for Nuclear Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
- ⁴³*McGill University, Montréal, Québec H3A 2T8, Canada*
- ^{44a}*INFN Sezione di Milano, I-20133 Milano, Italy*
- ^{44b}*Dipartimento di Fisica, Università di Milano, I-20133 Milano, Italy*
- ⁴⁵*University of Mississippi, University, Mississippi 38677, USA*
- ⁴⁶*Université de Montréal, Physique des Particules, Montréal, Québec H3C 3J7, Canada*
- ^{47a}*INFN Sezione di Napoli, I-80126 Napoli, Italy*
- ^{47b}*Dipartimento di Scienze Fisiche, Università di Napoli Federico II, I-80126 Napoli, Italy*
- ⁴⁸*NIKHEF, National Institute for Nuclear Physics and High Energy Physics, NL-1009 DB Amsterdam, The Netherlands*
- ⁴⁹*University of Notre Dame, Notre Dame, Indiana 46556, USA*
- ⁵⁰*Ohio State University, Columbus, Ohio 43210, USA*
- ⁵¹*University of Oregon, Eugene, Oregon 97403, USA*
- ^{52a}*INFN Sezione di Padova, I-35131 Padova, Italy*
- ^{52b}*Dipartimento di Fisica, Università di Padova, I-35131 Padova, Italy*
- ⁵³*Laboratoire de Physique Nucléaire et de Hautes Energies, IN2P3/CNRS, Université Pierre et Marie Curie-Paris6, Université Denis Diderot-Paris7, F-75252 Paris, France*
- ^{54a}*INFN Sezione di Perugia, I-06100 Perugia, Italy*

- ^{54b}*Dipartimento di Fisica, Università di Perugia, I-06100 Perugia, Italy*
^{55a}*INFN Sezione di Pisa, I-56127 Pisa, Italy*
^{55b}*Dipartimento di Fisica, Università di Pisa, I-56127 Pisa, Italy*
^{55c}*Scuola Normale Superiore di Pisa, I-56127 Pisa, Italy*
⁵⁶*Princeton University, Princeton, New Jersey 08544, USA*
^{57a}*INFN Sezione di Roma, I-00185 Roma, Italy*
^{57b}*Dipartimento di Fisica, Università di Roma La Sapienza, I-00185 Roma, Italy*
⁵⁸*Universität Rostock, D-18051 Rostock, Germany*
⁵⁹*Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, United Kingdom*
⁶⁰*CEA, Irfu, SPP, Centre de Saclay, F-91191 Gif-sur-Yvette, France*
⁶¹*SLAC National Accelerator Laboratory, Stanford, California 94309 USA*
⁶²*University of South Carolina, Columbia, South Carolina 29208, USA*
⁶³*Southern Methodist University, Dallas, Texas 75275, USA*
⁶⁴*Stanford University, Stanford, California 94305-4060, USA*
⁶⁵*State University of New York, Albany, New York 12222, USA*
⁶⁶*Tel Aviv University, School of Physics and Astronomy, Tel Aviv 69978, Israel*
⁶⁷*University of Tennessee, Knoxville, Tennessee 37996, USA*
⁶⁸*University of Texas at Austin, Austin, Texas 78712, USA*
⁶⁹*University of Texas at Dallas, Richardson, Texas 75083, USA*
^{70a}*INFN Sezione di Torino, I-10125 Torino, Italy*
^{70b}*Dipartimento di Fisica Sperimentale, Università di Torino, I-10125 Torino, Italy*
^{71a}*INFN Sezione di Trieste, I-34127 Trieste, Italy*
^{71b}*Dipartimento di Fisica, Università di Trieste, I-34127 Trieste, Italy*
⁷²*IFIC, Universitat de Valencia-CSIC, E-46071 Valencia, Spain*
⁷³*University of Victoria, Victoria, British Columbia V8W 3P6, Canada*
⁷⁴*Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom*
⁷⁵*University of Wisconsin, Madison, Wisconsin 53706, USA*
(Received 31 March 2013; published 27 August 2013)

We present the results of a search for the rare flavor-changing neutral-current decays $B \rightarrow \pi \ell^+ \ell^-$ ($\pi = \pi^\pm, \pi^0$ and $\ell = e, \mu$) and $B^0 \rightarrow \eta \ell^+ \ell^-$ using a sample of $e^+ e^- \rightarrow Y(4S) \rightarrow B\bar{B}$ decays corresponding to 428 fb^{-1} of integrated luminosity collected by the BABAR detector. No significant signal is observed, and we set an upper limit on the isospin and lepton-flavor averaged branching fraction of $\mathcal{B}(B \rightarrow \pi \ell^+ \ell^-) < 5.9 \times 10^{-8}$ and a lepton-flavor averaged upper limit of $\mathcal{B}(B^0 \rightarrow \eta \ell^+ \ell^-) < 6.4 \times 10^{-8}$, both at the 90% confidence level. We also report 90% confidence level branching fraction upper limits for the individual modes $B^+ \rightarrow \pi^+ e^+ e^-$, $B^0 \rightarrow \pi^0 e^+ e^-$, $B^+ \rightarrow \pi^+ \mu^+ \mu^-$, $B^0 \rightarrow \pi^0 \mu^+ \mu^-$, $B^0 \rightarrow \eta e^+ e^-$, and $B^0 \rightarrow \eta \mu^+ \mu^-$.

DOI: [10.1103/PhysRevD.88.032012](https://doi.org/10.1103/PhysRevD.88.032012)

PACS numbers: 13.20.He, 13.60.Hb

I. INTRODUCTION

In the standard model (SM), the decays $B \rightarrow \pi \ell^+ \ell^-$ ($\pi = \pi^\pm, \pi^0$ and $\ell = e, \mu$) and $B^0 \rightarrow \eta \ell^+ \ell^-$ proceed through the quark-level flavor-changing neutral-current process $b \rightarrow d \ell^+ \ell^-$. Since all flavor-changing neutral-current processes are forbidden at tree level in the SM, the lowest order diagrams representing these transitions

must involve loops. For $b \rightarrow d \ell^+ \ell^-$, these are the electroweak semileptonic penguin diagrams [Fig. 1(a)] and the $W^+ W^-$ box diagrams [Fig. 1(b)]. The $b \rightarrow d \ell^+ \ell^-$ transition is similar to $b \rightarrow s \ell^+ \ell^-$, but its rate is suppressed by the ratio $|V_{td}/V_{ts}|^2 \approx 0.04$ where V_{td} and V_{ts} are elements of the Cabibbo-Kobayashi-Maskawa quark mixing matrix [1,2]. The predicted branching fractions for the $B^+ \rightarrow \pi^+ \ell^+ \ell^-$ decay modes lie in the range of $(1.4\text{--}3.3) \times 10^{-8}$, when the dilepton mass regions near the J/ψ and $\psi(2S)$ are excluded in order to remove decays that proceed through the intermediate charmonium resonances. The largest source of uncertainty in these predictions arises from knowledge of the $B \rightarrow \pi$ form factors [3–5]. These branching fractions imply that 5–15 events occur for each $B \rightarrow \pi \ell^+ \ell^-$ decay channel in the BABAR data set (471 million $B\bar{B}$ pairs). The predicted $B^0 \rightarrow \eta \ell^+ \ell^-$ branching fractions lie in the range $(2.5\text{--}3.7) \times 10^{-8}$ where again the dominant

*Present address: University of Tabuk, Tabuk 71491, Saudi Arabia.

†Also at Dipartimento di Fisica, Università di Perugia, Perugia, Italy.

‡Present address: University of Huddersfield, Huddersfield HD1 3DH, United Kingdom.

§Deceased.

||Present address: University of South Alabama, Mobile, Alabama 36688, USA.

¶Also at Università di Sassari, Sassari, Italy.

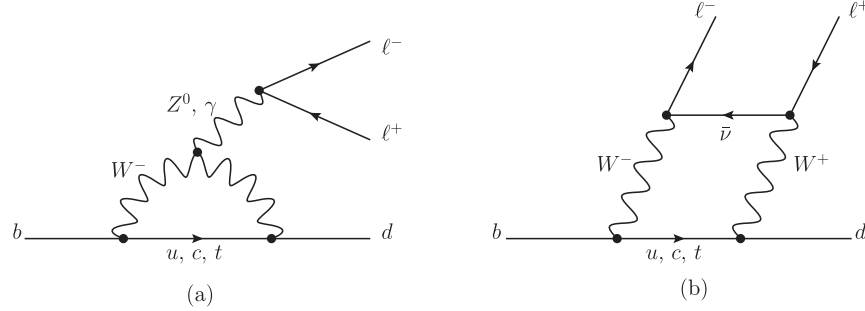


FIG. 1. Lowest order Feynman diagrams describing the quark level $b \rightarrow d\ell^+\ell^-$ transition in B meson decay: (a) electroweak penguin diagrams and (b) W^+W^- box diagrams.

uncertainty is due to lack of knowledge of the $B^0 \rightarrow \eta$ form factors [6].

Many extensions of the SM predict the existence of new, heavy particles which couple to the SM fermions and bosons. The $b \rightarrow d\ell^+\ell^-$ and $b \rightarrow s\ell^+\ell^-$ decays provide a promising avenue in which to search for new physics (NP). Amplitudes from these NP contributions can interfere with those from the SM, altering physical observables (e.g., decay rates, CP , isospin, and forward-backward asymmetries) from the SM predictions [3,4,7–10]. Measurements in the $\pi\ell^+\ell^-$ and $\eta\ell^+\ell^-$ systems complement and provide independent searches of NP from those in the $K^{(*)}\ell^+\ell^-$ channels [11–17], as physics beyond the SM may have nontrivial flavor couplings [18]. The measurement of observables as a function of the square of the invariant dilepton mass $q^2 = m_{\ell\ell}^2$ for exclusive $b \rightarrow d\ell^+\ell^-$ decay modes allows for more thorough tests of SM predictions and deeper probes for NP but is currently not possible due to the size of the data set.

Only one $b \rightarrow d\ell^+\ell^-$ decay has been observed to date, with LHCb measuring the $B^+ \rightarrow \pi^+\mu^+\mu^-$ branching fraction to be $(2.3 \pm 0.6 \pm 0.1) \times 10^{-8}$ [19]. Both *BABAR* [20] and *Belle* [21] have performed searches for $B \rightarrow \pi\ell^+\ell^-$ decays but have observed no significant signal. For the $\pi\ell^+\ell^-$ modes, the smallest upper limits from the B factories lie within an order of magnitude of the SM predictions [3–5] and are beginning to exclude portions of the NP parameter space. No previous searches for $B^0 \rightarrow \eta\ell^+\ell^-$ have been reported. Observation of $b \rightarrow d\ell^+\ell^-$ decays at the B factories is currently limited by the size of the available data sets. Additionally, for $B^+ \rightarrow \pi^+\ell^+\ell^-$, background from $B^+ \rightarrow K^+\ell^+\ell^-$ decays where the kaon is misidentified as a pion must be treated carefully as $K^+\ell^+\ell^-$ can appear very signal-like and occurs at a rate approximately 25 times the expected $B^+ \rightarrow \pi^+\ell^+\ell^-$ rate.

In this article we report on our study of the $B \rightarrow \pi\ell^+\ell^-$ and $B^0 \rightarrow \eta\ell^+\ell^-$ decays using the full *BABAR* data set, presenting branching fraction upper limits for the modes $B^+ \rightarrow \pi^+e^+e^-$, $B^0 \rightarrow \pi^0e^+e^-$, $B^+ \rightarrow \pi^+\mu^+\mu^-$, $B^0 \rightarrow \pi^0\mu^+\mu^-$, $B^0 \rightarrow \eta e^+e^-$, and $B^0 \rightarrow \eta\mu^+\mu^-$. Charge conjugation is implied throughout unless specified

otherwise. We also present upper limits for the lepton-flavor averaged modes $B^+ \rightarrow \pi^+\ell^+\ell^-$, $B^0 \rightarrow \pi^0\ell^+\ell^-$, and $B^0 \rightarrow \eta\ell^+\ell^-$, where we constrain the e^+e^- and $\mu^+\mu^-$ branching fractions to be equal; for the isospin averaged modes $B \rightarrow \pi e^+e^-$ and $B \rightarrow \pi\mu^+\mu^-$, where the $B^+ \rightarrow \pi^+\ell^+\ell^-$ decay rate is constrained to be twice the $B^0 \rightarrow \pi^0\ell^+\ell^-$ decay rate; and for the isospin and lepton-flavor averaged mode $B \rightarrow \pi\ell^+\ell^-$. For the lepton-flavor averaged measurements, we neglect differences in available phase space due to the difference between the electron and muon masses. The branching fractions are based on signal yields that are extracted through an unbinned, extended maximum likelihood fit to two kinematic variables. All selection criteria are determined before the fit was performed on data, i.e., the analysis is performed “blind.”

II. *BABAR* DETECTOR, SIMULATION, AND DATA SETS

The results of this analysis are based upon a sample of $e^+e^- \rightarrow Y(4S) \rightarrow B\bar{B}$ interactions provided by the PEP-II asymmetric-energy storage rings and collected by the *BABAR* detector located at SLAC National Accelerator Laboratory. The *BABAR* data sample corresponds to an integrated luminosity of 428 fb^{-1} [22] containing 471 million $B\bar{B}$ decays. This is the full data set collected at the $Y(4S)$ resonance. A detailed description of the *BABAR* detector can be found elsewhere [23]. Charged particle momenta are measured with a five-layer, double-sided silicon vertex tracker and a 40-layer drift chamber operated in proportional mode. These two tracking systems are immersed in the 1.5 T magnetic field of a superconducting solenoid. A ring-imaging Cherenkov detector with fused silica radiators, aided by ionization loss dE/dx measurements from the tracking system, provides identification of charged particles. Electromagnetic showers from electrons and photons are detected with an electromagnetic calorimeter (EMC) constructed from a finely segmented array of thallium-doped CsI scintillating crystals. The steel flux return of the solenoid was initially instrumented with resistive plate chambers and functions primarily to identify

muons. For the later data taking periods, the resistive plate chambers of the steel flux return of the solenoid were replaced with limited streamer tubes and brass to increase absorption.

The *BABAR* Monte Carlo (MC) simulation utilizes the GEANT4 package [24] for detector simulation, and EvtGen [25] and Jetset7.4 [26] for $B\bar{B}$ and $e^+e^- \rightarrow q\bar{q}$ ($q = u, d, s, c$) decays, respectively. The $B\bar{B}$ and continuum ($e^+e^- \rightarrow q\bar{q}$, $q = u, d, s, c$) MC samples correspond to an effective luminosity of about ten times the data sample collected at the $Y(4S)$ resonance. Simulated $B \rightarrow \pi\ell^+\ell^-$ signal decay samples are generated according to the form-factor model of Ref. [27], with the Wilson coefficients taken from Refs. [28–30], and the decay amplitudes calculated in Ref. [10]. The $B^0 \rightarrow \eta\ell^+\ell^-$ signal MC samples utilize the same kinematics, Wilson coefficients, and form-factor model as the $\pi\ell^+\ell^-$ modes. The effects of the choice of form-factor model and the values of the Wilson coefficients are considered as sources of systematic uncertainty in the signal efficiency. We also make use of simulated $B \rightarrow K^{(*)}\ell^+\ell^-$, $B \rightarrow J/\psi X$, and $B \rightarrow \psi(2S)X$ samples. Signal efficiencies, as well as parameters of the fit model, are determined from signal and $K^{(*)}\ell^+\ell^-$ MC data sets. The $B \rightarrow J/\psi X$ and $B \rightarrow \psi(2S)X$ MC samples allow us to study background from these decays and also serve as the data sets from which we fix the parameters of the fit model used in the $B \rightarrow J/\psi\pi/\eta$ fit validation, as described later.

III. EVENT RECONSTRUCTION AND CANDIDATE SELECTION

Event reconstruction begins by building dilepton candidates from two leptons (e^+e^- or $\mu^+\mu^-$). Leptons are selected as charged tracks with momenta in the laboratory reference (lab) frame greater than 300 MeV/c. Loose particle identification (PID) requirements are placed upon the two leptons. More stringent PID requirements are applied later, and the optimization of these selection criteria is discussed in Sec. V. The lepton pair is fit to a common vertex to form a dilepton candidate with the requirement that $m_{\ell\ell} < 5.0$ GeV/c² [31]. We also place a loose constraint on the χ^2 probability of the vertex fit by requiring it to be greater than 10^{-10} . For electrons, we apply an algorithm which associates photons with electron candidates in an attempt to recover energy lost through bremsstrahlung, allowing at most one photon to be associated with each electron. The photon trajectory is required to lie within a small cone of opening angle 0.035 rad about the initial momentum vector of the electron, and the photon energy in the lab frame must be greater than 30 MeV. Additionally, we suppress background from photon conversions by requiring that the invariant mass of the electron (or positron) paired with any other oppositely charged track in the event be greater than 30 MeV/c².

Charged pion candidates are charged tracks passing pion PID requirements which retain approximately 90%–95% of charged pions and only 2%–5% of charged kaons. We reconstruct π^0 candidates from two photons with invariant diphoton mass $m_{\gamma\gamma}$ lying in the range $115 < m_{\gamma\gamma} < 150$ MeV/c². A minimum value of 50 MeV is required for the lab energy of each photon. Photons are detected as EMC clusters not associated with a charged track. The clusters are also required to have a lateral shower profile consistent with originating from a photon. We reconstruct η as $\eta \rightarrow \gamma\gamma$ ($\eta_{\gamma\gamma}$) and $\eta \rightarrow \pi^+\pi^-\pi^0$ ($\eta_{3\pi}$), which constitute 39.3% and 22.7% of the η branching fraction, respectively. As in the case of the π^0 , we require the $\eta_{\gamma\gamma}$ photon daughters to have energy greater than 50 MeV in the lab frame. Additionally, the photon energy asymmetry $A_\gamma = |E_{1,\gamma} - E_{2,\gamma}|/(E_{1,\gamma} + E_{2,\gamma})$ must be less than 0.8, where $E_{1,\gamma}$ and $E_{2,\gamma}$ are the energies of the photons in the lab frame. The invariant diphoton mass must lie in the range $500 < m_{\gamma\gamma} < 575$ MeV/c². For $\eta_{3\pi}$, the pion candidates are fit to a common vertex to form an η candidate. In the fit the η candidate mass is constrained to the nominal η mass, while the invariant three-pion mass is required to lie in the range $535 < m_{3\pi} < 565$ MeV/c².

B candidates are reconstructed from a hadron candidate (π^+ , π^0 , $\eta_{\gamma\gamma}$, or $\eta_{3\pi}$) and a dilepton candidate (e^+e^- or $\mu^+\mu^-$). The hadron and dilepton candidates are fit to a common vertex, and the entire decay chain is refit. We make use of two kinematic, Lorentz-invariant quantities, m_{ES} and ΔE , defined as

$$m_{\text{ES}} = \sqrt{(s/2 + \vec{p}_B \cdot \vec{p}_0)^2/E_0^2 - p_B^2} \quad (1)$$

$$\Delta E = (2q_B q_0 - s)/2\sqrt{s}, \quad (2)$$

where $\sqrt{s} = 2E_{\text{beam}}^*$ is the total energy of the e^+e^- system in the c.m. frame, q_B and $q_0 = (E_0, \vec{p}_0)$ are the four-vectors representing the momentum of the B candidate and of the e^+e^- system, respectively, and \vec{p}_B is the three-momentum of the B candidate. In the c.m. frame, these expressions simplify to

$$m_{\text{ES}} = \sqrt{E_{\text{beam}}^{*2} - p_B^{*2}} \quad (3)$$

$$\Delta E = E_B^* - E_{\text{beam}}^*, \quad (4)$$

where the asterisk indicates evaluation in the c.m. frame. These variables make use of precisely measured beam quantities. All B candidates are required to have $m_{\text{ES}} > 5.1$ GeV/c² and $-300 < \Delta E < 250$ MeV. The distributions of these two variables are later fit to extract the $\pi\ell^+\ell^-$ and $\eta\ell^+\ell^-$ branching fractions.

A large background is present from $B \rightarrow J/\psi X$ and $B \rightarrow \psi(2S)X$ decays where J/ψ and $\psi(2S)$ decay to $\ell^+\ell^-$. Here X represents a hadronic state, typically π , η , ρ , or $K^{(*)}$. These events are removed from our data sample

by rejecting any event with a value of $m_{\ell\ell}$ consistent with originating from a J/ψ or $\psi(2S)$ decay. The rejected J/ψ events are useful as they provide a control sample which can be used to test the fit model. We also use these samples to estimate systematic uncertainties and to correct for differences between data and MC selection efficiencies. For the electron modes we reject events in the following regions about the J/ψ mass: $2.90 < m_{\ell\ell} < 3.20$ GeV/ c^2 , or $m_{ee} < 2.90$ GeV/ c^2 and $\Delta E < m_{ee}c^2 - 2.875$ GeV. For the muon modes the region is $3.00 < m_{\mu\mu} < 3.20$ GeV/ c^2 , or $m_{\mu\mu} < 3.00$ GeV/ c^2 and $\Delta E < 1.11m_{\mu\mu}c^2 - 3.31$ GeV. The rejection region about the $\psi(2S)$ mass is the same for electrons and muons: $3.60 < m_{\ell\ell} < 3.75$ GeV/ c^2 , or $m_{\ell\ell} < 3.60$ GeV/ c^2 and $\Delta E < m_{\ell\ell}c^2 - 3.525$ GeV. Introducing ΔE dependence into the region boundaries allows us to account for some of the effects of track mismeasurement and energy lost due to bremsstrahlung.

The largest source of background comes from random combinations of particles from continuum events or semileptonic B and D decays in $B\bar{B}$ events. Continuum events tend to be jetlike as the $q\bar{q}$ pair is produced back-to-back with relatively large momentum in the c.m. frame. The topology of $B\bar{B}$ decays is more isotropic as the B mesons are produced nearly at rest in the $Y(4S)$ rest frame. Semileptonic decays are characterized by the presence of a neutrino, e.g., missing energy in the event and nonzero total transverse momentum of the event. Due to the differences in these two background types we train separate artificial neural networks (NNs) to reject each of them. By selecting inputs to the NNs which are independent of the final state we are able to train only one NN for each lepton flavor. We do not train separate NNs for π^+ , π^0 , and η . This increases the size of the training samples, improving the performance of the NNs. We train four NNs: one to reject $B\bar{B}$ background in the e^+e^- modes, one to reject $B\bar{B}$ background in the $\mu^+\mu^-$ modes, one to reject continuum background in the e^+e^- modes, and one to reject continuum background in the $\mu^+\mu^-$ modes.

The signal training samples were assembled from equal portions of correctly reconstructed $B^+ \rightarrow \pi^+\ell^+\ell^-$, $B^0 \rightarrow \pi^0\ell^+\ell^-$, $B^+ \rightarrow \rho^+\ell^+\ell^-$, $B^0 \rightarrow \rho^0\ell^+\ell^-$, $B^0 \rightarrow \eta_{\gamma\gamma}\ell^+\ell^-$, and $B^0 \rightarrow \omega\ell^+\ell^-$ MC events. The size of the training samples, particularly the background training samples, that could be formed from events reconstructed as one of our signal modes was a limiting factor in the performance of the NNs. To increase the available statistics the other events from the $\rho\ell^+\ell^-$ and $\omega\ell^+\ell^-$ modes were added to the training samples. The ρ^+ (ρ^0) was reconstructed as $\pi^+\pi^0$ ($\pi^+\pi^-$) with two-pion invariant mass $m_{\pi\pi}$ in the range $0.455 < m_{\pi\pi} < 1.095$ GeV/ c^2 ($0.475 < m_{\pi\pi} < 1.075$ GeV/ c^2). The ω was reconstructed as $\pi^+\pi^-\pi^0$ and required to have a three-pion invariant mass lying within 50 MeV/ c^2 of the nominal ω mass. No $\eta_{3\pi}\ell^+\ell^-$ events were used in the training due to very

low statistics for the background $B\bar{B}$ and continuum samples for these modes. For background we combined the MC data sets, either $B\bar{B}$ or continuum depending upon the classifier to be trained, from the six modes listed above and randomly select events from this data set to form the training sample. The performances of the NNs trained with samples from several different $b \rightarrow d\ell^+\ell^-$ (global NNs) modes were compared with NNs trained specifically for each of our four $B \rightarrow \pi\ell^+\ell^-$ modes (single mode NNs). The background rejection of the global NNs at a fixed signal efficiency was found to be similar to that of the single mode NNs.

The input variables to the continuum NNs are related mostly to event topology and include the ratios of Fox-Wolfram moments [32]; the cosine of the polar angle of the thrust axis [33] of the event; the cosine of the polar angle of the thrust axis of the rest-of-the-event (ROE), which consists of all particles in the event not associated with the signal B candidate; the momentum weighted polynomials L_i^j [34] computed using tracks and EMC clusters; the cosine of the polar angle of the B candidate momentum; and the χ^2 probability of the B candidate vertex fit. The input variables to the $B\bar{B}$ NNs reflect the effort to reject background from semileptonic B and D decays and include m_{ES} and ΔE constructed from the ROE, total momentum of the event transverse to the beam, missing energy in the event, momentum of the ROE transverse to the beam direction, momentum of the ROE transverse to the thrust axis of the event, cosine of the polar angle of the B candidate momentum, and χ^2 probability of the B candidate and dilepton candidate vertex fits. The NN outputs show only weak correlation with the fit variables m_{ES} and ΔE .

Figure 2, as representative of the several neural networks, shows the output of the $e^+e^- B\bar{B}$ NN for a sample of signal and $B\bar{B}$ background $\pi^+e^+e^-$ MC events. Also shown is the output of the $\mu^+\mu^-$ continuum NN for a sample of signal and continuum background $\pi^0\mu^+\mu^-$ MC events. Requirements on the NN outputs are optimized for each of our eight modes to produce the lowest branching fraction upper limit. A description of the optimization procedure is given in Sec. V.

Due to their similarity to signal, $B \rightarrow K^{(*)}\ell^+\ell^-$ decays constitute a background that mimics signal by peaking in either one or both m_{ES} and ΔE . The $b \rightarrow s\ell^+\ell^-$ transition occurs at a rate approximately 25 times greater than the SM $b \rightarrow d\ell^+\ell^-$ rate, and due to particle misidentification and event misreconstruction, its contribution is expected to be of the same order as the $\pi\ell^+\ell^-$ signal in the BABAR data sample. In the charged pion modes, $B^+ \rightarrow K^+\ell^+\ell^-$ peaks in m_{ES} as $\pi^+\ell^+\ell^-$ signal but in ΔE near -70 MeV due to the misidentification of the kaon as a pion. There are also contributions from $B^0 \rightarrow K_S^0\ell^+\ell^-$, where one of the pions from the K_S^0 decay is missed, and from $B \rightarrow K^*(\rightarrow K^+\pi)\ell^+\ell^-$, where the pion from the K^* decay is

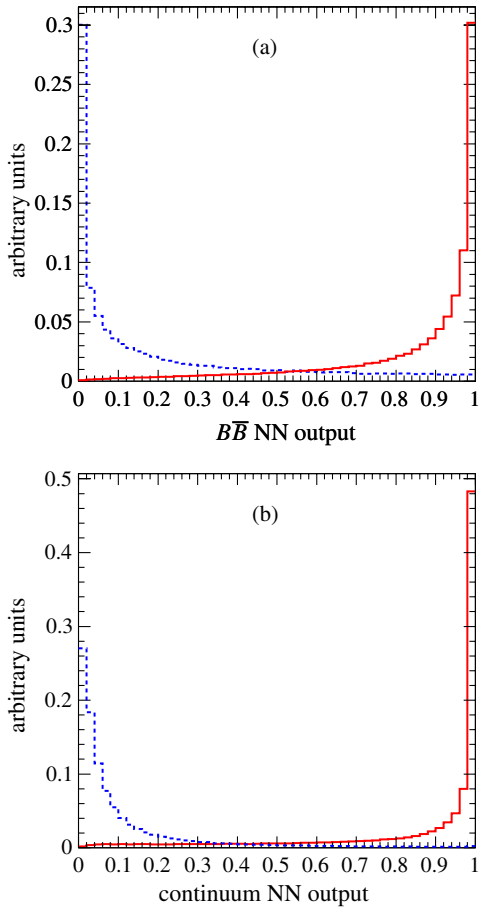


FIG. 2 (color online). Outputs of (a) the $e^+e^- B\bar{B}$ neural network for a sample of $B^+ \rightarrow \pi^+ e^+ e^-$ signal (solid red) and $B\bar{B}$ background (dashed blue) MC events and (b) the $\mu^+\mu^-$ continuum neural network for a sample of $B^0 \rightarrow \pi^0 \mu^+ \mu^-$ signal (solid red) and continuum background (dashed blue) MC events. For both (a) and (b) the signal and background distributions are normalized to equal areas.

missed. In the case of $K_S^0 \ell^+ \ell^-$, the remaining pion and the two leptons are reconstructed as $\pi^+ \ell^+ \ell^-$. For $K^*(\rightarrow K^+ \pi) \ell^+ \ell^-$, the K^+ is misidentified as π^+ and reconstructed with the two leptons as $\pi^+ \ell^+ \ell^-$. In both cases, the decays peak in m_{ES} like signal but at $\Delta E < -140$ MeV due to the missing pion. For $K^* \ell^+ \ell^-$ the ΔE peak occurs at even lower values due to the kaon misidentification. For $B^0 \rightarrow \pi^0 \ell^+ \ell^-$, there is a similar background from $B^0 \rightarrow K_S^0(\rightarrow \pi^0 \pi^0) \ell^+ \ell^-$ decays where one π^0 is reconstructed along with the lepton pair as $\pi^0 \ell^+ \ell^-$. These events produce a peak in m_{ES} in the same location as $\pi^0 \ell^+ \ell^-$ signal but peak at smaller values of ΔE due to the missing pion from the decay. In all three cases [$K^+ \ell^+ \ell^-$ in $\pi^+ \ell^+ \ell^-$, $K_S^0(\rightarrow \pi^+ \pi^-) \ell^+ \ell^-$ and $K^*(\rightarrow K^+ \pi) \ell^+ \ell^-$ in $\pi^+ \ell^+ \ell^-$, and $K_S^0(\rightarrow \pi^0 \pi^0) \ell^+ \ell^-$ in $\pi^0 \ell^+ \ell^-$] we include a separate component in the fit model to account for the corresponding contribution.

For $\pi e^+ e^-$ and $\eta_{\gamma\gamma} e^+ e^-$, there is an additional background that originates from two-photon events, given by

the process $e^+ e^- \rightarrow e^+ e^- \gamma \gamma \rightarrow (e^+ e^-) q \bar{q}$ where q is a u , d , or s quark. The background is characterized by a small transverse momentum of the pion and a large lepton-lepton opening angle $\theta_{\ell^+ \ell^-}$. There is also a correlation between the polar angles of the electron, θ_{e^-} and of the positron, θ_{e^+} . The e^- tends to be in the forward direction while the e^+ tends to be in the backward direction, consistent with the $e^+ e^-$ beam particles scattering into the detector. Events of this type are rejected using the following requirements. For $\pi^+ e^+ e^-$, $\pi^0 e^+ e^-$, and $\eta_{\gamma\gamma} e^+ e^-$ we require $p_{\text{had}}^* > 750$ MeV/ c and $N_{\text{trk}} > 4$ where p_{had}^* is the hadron momentum in the c.m. frame and N_{trk} is the number of charged tracks in the event. Additionally, for $\pi^+ e^+ e^-$ we require $E_{1,\text{neut}} < 1.75$ GeV, $\cos \theta_{\ell^+ \ell^-} > -0.95$, and $\theta_{e^-} > (0.57\theta_{e^+} - 0.7)$ rad where $E_{1,\text{neut}}$ is the energy of the highest energy neutral cluster in the event in the lab frame. Similarly, $\pi^0 e^+ e^-$ candidates must satisfy $\theta_{e^-} > (0.64\theta_{e^+} - 0.8)$ rad, and $\eta_{\gamma\gamma} e^+ e^-$ candidates are required to have $\theta_{e^-} > (0.6\theta_{e^+} - 0.55)$ rad and $\cos \theta_{\ell^+ \ell^-} > -0.95$. These criteria were determined by maximizing the quantity $\varepsilon/\sqrt{N_{\text{SB}}}$, where ε is the signal efficiency and N_{SB} is the number of events lying in the sideband region $5.225 < m_{\text{ES}} < 5.26$ GeV/ c^2 in data. We assume that the two-photon background in the m_{ES} sideband occurs similarly to the two-photon background in the region $m_{\text{ES}} > 5.26$ GeV/ c^2 . The optimization was carried out with all other selection criteria applied, including those on the NN outputs.

To guard against possible background from $B \rightarrow D\pi$ and $B \rightarrow D\eta$ decays where $D \rightarrow K\pi$, $\pi\pi$, or $\eta\pi$ and the kaon or pions are misidentified as leptons, we assign the lepton candidates either a kaon or pion mass and discard any event with a combination of $\mu^+ \mu^-$, $\mu^\pm \pi$, or $\mu^\pm \eta$ with invariant mass in the range (1.83–1.89) GeV/ c^2 . The probability of misidentifying a hadron as an electron is negligible, and this requirement is therefore only applied to the $\mu^+ \mu^-$ modes.

Hadronic decays such as $B^+ \rightarrow \pi^+ \pi^- \pi^+$, where two pions are misidentified as muons, peak in both m_{ES} and ΔE similarly to signal due to the relatively small difference between the pion and muon masses. This hadronic peaking background is modeled by a component in the fit. A dedicated data control sample is used to determine its normalization and shape. This sample is constructed from events where one lepton candidate passes the muon identification requirements but the other does not. The events in these samples are weighted with particle misidentification probabilities determined from control samples in *BABAR* data. Studies of MC samples indicate that this background is consequential only for the $\pi\mu^+\mu^-$ modes.

After applying all selection criteria there are sometimes multiple candidates within a given mode remaining in an event. This occurs for approximately 20%–25% (35%–40%) of $\pi^+ e^+ e^-$ and $\pi^0 e^+ e^-$ ($\eta_{\gamma\gamma} e^+ e^-$ and $\eta_{3\pi} e^+ e^-$) candidates and 5%–10% (25%–30%) of

$\pi^+ \mu^+ \mu^-$ and $\pi^0 \mu^+ \mu^-$ ($\eta_{\gamma\gamma} \mu^+ \mu^-$ and $\eta_{3\pi} \mu^+ \mu^-$) candidates. There tend to be more events containing multiple candidates in the $e^+ e^-$ modes due to the bremsstrahlung recovery. For instance, there may be multiple candidates arising from the same $\pi^+ e^+ e^-$ combination where the bremsstrahlung photons associated with the e^+ or e^- are different.

To choose the best candidate we construct a ratio \mathcal{L}_R from the $B\bar{B}$ and continuum NN classifier output distributions of the signal and background samples. The ratio \mathcal{L}_R is defined as

$$\mathcal{L}_R(x, y) = \frac{\mathcal{P}_{B\bar{B}}^{\text{sig}}(x) + \mathcal{P}_{\text{cont}}^{\text{sig}}(y)}{(\mathcal{P}_{B\bar{B}}^{\text{sig}}(x) + \mathcal{P}_{\text{cont}}^{\text{sig}}(y)) + (\mathcal{P}_{B\bar{B}}^{\text{bkg}}(x) + \mathcal{P}_{\text{cont}}^{\text{bkg}}(y))}, \quad (5)$$

where $\mathcal{P}_{B\bar{B}}^{\text{sig}}(x)$ [$\mathcal{P}_{\text{cont}}^{\text{sig}}(y)$] is the probability that a signal candidate has a $B\bar{B}$ (continuum) NN output value of x (y). The quantities $\mathcal{P}_{B\bar{B}}^{\text{bkg}}(x)$ and $\mathcal{P}_{\text{cont}}^{\text{bkg}}(y)$ are defined analogously for background events. Signal-like candidates have values of \mathcal{L}_R near 1 while more backgroundlike candidates have values near 0. If multiple candidates are present in an event, we choose the candidate with the greatest value of \mathcal{L}_R as the best candidate. For events containing multiple candidates, this procedure chooses the correct candidate approximately 90%–95% of the time for $\pi \ell^+ \ell^-$ and 75%–80% of the time for $\eta \ell^+ \ell^-$. The ratio \mathcal{L}_R is used only to select a best candidate.

IV. BRANCHING FRACTION MEASUREMENT AND UPPER LIMIT CALCULATION

Branching fractions are extracted through an unbinned extended maximum likelihood fit to m_{ES} and ΔE with the fit region defined as $m_{\text{ES}} > 5.225 \text{ GeV}/c^2$ and $-300 < \Delta E < 250 \text{ MeV}$. The probability density functions (PDFs) in the fit model contain several components corresponding to the different contributions in the data set. To model the various components, we use a combination of products of one-dimensional parametric PDFs, two-dimensional histograms, and two-dimensional nonparametric shapes determined by a Gaussian kernel density estimation algorithm (KEYS PDF) [35]. For components that are described by the product of one-dimensional PDFs, we are allowed to use such a model because m_{ES} and ΔE are uncorrelated for these components.

A. $B^+ \rightarrow \pi^+ \ell^+ \ell^-$

The $\pi^+ \ell^+ \ell^-$ fit model involves four components: signal, $K^+ \ell^+ \ell^-$ background, $K_S^0/K^* \ell^+ \ell^-$ background, and combinatoric background. There is an additional component in $B^+ \rightarrow \pi^+ \mu^+ \mu^-$ representing the $B^+ \rightarrow \pi^+ \pi^+ \pi^-$ hadronic peaking background. The $K^+ \ell^+ \ell^-$ background arises from decays where the kaon is misidentified as a pion. The K^+ misidentification rate is such that

the $K^+ \ell^+ \ell^-$ background in $\pi^+ \ell^+ \ell^-$ is approximately the same size as the expected SM $\pi^+ \ell^+ \ell^-$ signal. Since the K^+ misidentification probability is well measured, it is possible to measure this background contribution directly from our data. This is done by simultaneously fitting two data samples, comprised by the $B^+ \rightarrow \pi^+ \ell^+ \ell^-$ candidates and the $B^+ \rightarrow K^+ \ell^+ \ell^-$ candidates in our data set. The K^+ misidentification background to $B^+ \rightarrow K^+ \ell^+ \ell^-$ is included in the fit at a level fixed to the $B^+ \rightarrow K^+ \ell^+ \ell^-$ yield using the known misidentification probability (which depends on the momentum of the kaon). The $B^+ \rightarrow K^+ \ell^+ \ell^-$ branching fraction that is measured from the simultaneous fit of the $B^+ \rightarrow \pi^+ \ell^+ \ell^-$ and $B^+ \rightarrow K^+ \ell^+ \ell^-$ data samples provides an additional validation of our procedure, since this branching fraction has been previously measured [36].

The $K^+ \ell^+ \ell^-$ sample is selected in exactly the same way as the $\pi^+ \ell^+ \ell^-$ sample except the charged pion identification requirements are reversed and the J/ψ and $\psi(2S)$ rejection window includes the following regions: $m_{ee} > 3.20 \text{ GeV}/c^2$ and $1.11m_{ee}c^2 - 3.67 < \Delta E < m_{ee}c^2 - 2.875 \text{ GeV}$ for $\pi^+ e^+ e^-$ surrounding the J/ψ mass, $m_{\mu\mu} > 3.20 \text{ GeV}/c^2$ and $1.11m_{\mu\mu}c^2 - 3.614 < \Delta E < m_{\mu\mu}c^2 - 2.925 \text{ GeV}$ for $\pi^+ \mu^+ \mu^-$ surrounding the J/ψ mass, and $m_{\ell\ell} > 3.75 \text{ GeV}/c^2$ and $1.11m_{\ell\ell}c^2 - 4.305 < \Delta E < m_{\ell\ell}c^2 - 3.525 \text{ GeV}$ for both modes surrounding the $\psi(2S)$ mass. Also, the ΔE window is $-200 < \Delta E < 250 \text{ MeV}$ for $K^+ e^+ e^-$ and $-100 < \Delta E < 250 \text{ MeV}$ for $K^+ \mu^+ \mu^-$.

The $\pi^+ \ell^+ \ell^-$ and $K^+ \ell^+ \ell^-$ background m_{ES} and ΔE distributions are modeled by products of one-dimensional PDFs. The $\pi^+ \ell^+ \ell^-$ signal and $K^+ \ell^+ \ell^-$ background m_{ES} distributions are described by a Crystal Ball function [37]. The $\pi^+ e^+ e^-$ ΔE signal distribution is modeled by the sum of a Crystal Ball function and a Gaussian which share a common mean, while the $\pi^+ \mu^+ \mu^-$ signal and both the $K^+ e^+ e^-$ and $K^+ \mu^+ \mu^-$ ΔE distributions are modeled by a modified Gaussian with tail parameters whose functional form is given by

$$f(\Delta E) = \exp \left[-\frac{(\Delta E - \mu)^2}{2\sigma_{L,R}\alpha_{L,R} + \alpha_{L,R}(\Delta E - \mu)} \right], \quad (6)$$

where σ_L and α_L (σ_R and α_R) are the width and tail parameters used when $\Delta E < \mu$ ($\Delta E > \mu$), respectively. A two-dimensional histogram models the contribution from $B \rightarrow K_S^0/K^* \ell^+ \ell^-$ decays. Combinatoric background is described by the product of an ARGUS function [38] in m_{ES} with endpoint fixed to $5.29 \text{ GeV}/c^2$ and a second-order polynomial in ΔE . The $\pi^+ \mu^+ \mu^-$ hadronic peaking background component is modeled by a two-dimensional KEYS PDF [35].

The PDF fit to the $K^+ \ell^+ \ell^-$ sample contains a similar set of components. Signal $K^+ \ell^+ \ell^-$ distributions are modeled by the product of a Crystal Ball function in m_{ES} and the line shape of Eq. (6) in ΔE . The contribution from other

$b \rightarrow s\ell^+\ell^-$ decays is dominated by $B \rightarrow K^*(K^+\pi)\ell^+\ell^-$ where the pion is lost. We use a two-dimensional histogram to model this background. Combinatoric background is modeled by the product of an ARGUS distribution in m_{ES} and by an exponential function for $K^+e^+e^-$ and a second-order polynomial for $K^+\mu^+\mu^-$ in ΔE . A KEYS PDF models the hadronic peaking background in $K^+\mu^+\mu^-$.

In both the $\pi^+\ell^+\ell^-$ and $K^+\ell^+\ell^-$ PDFs, the signal and combinatoric background yields float along with the shapes of the combinatoric background PDFs. The $K^+\ell^+\ell^-$ background yield in the $\pi^+\ell^+\ell^-$ sample is constrained so that the $B^+ \rightarrow K^+\ell^+\ell^-$ branching fractions measured in the $\pi^+\ell^+\ell^-$ and $K^+\ell^+\ell^-$ samples are equal. All fixed shapes and yields are determined from exclusive MC samples except for the hadronic peaking background which uses a data control sample. Normalizations of the $K_S^0/K^*\ell^+\ell^-$ component of the $\pi^+\ell^+\ell^-$ PDF and of $K^*\ell^+\ell^-$ component in the $K^+\ell^+\ell^-$ PDF are fixed from efficiencies determined from MC samples and world average branching fractions [36].

B. $B^0 \rightarrow \pi^0\ell^+\ell^-$

The $B^0 \rightarrow \pi^0\ell^+\ell^-$ signal distribution is modeled by the product of a Crystal Ball function in m_{ES} and by the line shape given in Eq. (6) in ΔE . Background from $B^0 \rightarrow K_S^0(\rightarrow \pi^0\pi^0)\ell^+\ell^-$ decays is modeled by a two-dimensional histogram. The product of an ARGUS shape in m_{ES} with an exponential function in ΔE models the combinatoric background distribution. As in the $\pi^+\mu^+\mu^-$ and $K^+\mu^+\mu^-$ PDFs, there is an additional component in the $\pi^0\mu^+\mu^-$ fit model devoted to hadronic peaking background which is described by a KEYS PDF.

In the fit, only the signal $\pi^0\ell^+\ell^-$ and combinatoric background yields along with the shape of the combinatoric background PDF float. The signal and $K_S^0(\rightarrow \pi^0\pi^0)\ell^+\ell^-$ shapes are determined from fits to MC samples, and the $K_S^0(\rightarrow \pi^0\pi^0)\ell^+\ell^-$ normalization comes from efficiencies taken from MC samples and world average branching fractions [36]. The shape and normalization of the peaking hadronic component are determined from a data control sample.

C. $B^0 \rightarrow \eta\ell^+\ell^-$

The $\eta\ell^+\ell^-$ fit model is simple, consisting of only three components, and is the same for all four $\eta\ell^+\ell^-$ channels. The signal component is modeled by the product of a Crystal Ball function in m_{ES} and the line shape of Eq. (4) in ΔE . We include a component for events containing a signal decay where the signal B is incorrectly reconstructed, which we refer to as self-cross-feed. In these events the signal decay is typically reconstructed as a combination of particles from the B decaying to our signal mode and the other B . In most self-cross-feed events the dilepton pair is correctly reconstructed, and the hadron is

misreconstructed. The self-cross-feed contribution is represented by a two-dimensional histogram, and its normalization is a fixed fraction of the signal yield with the fraction determined from signal MC. The self-cross-feed-to-signal ratio varies from 0.1 to 0.15 for the $\eta_{\gamma\gamma}$ channels to 0.25–0.3 for the $\eta_{3\pi}$ channels. Combinatoric background is described by the product of an ARGUS function in m_{ES} and an exponential function in ΔE . From studies of MC samples, we find no indication of potential peaking background contributions from $b \rightarrow s\ell^+\ell^-$ decays or any other sources. The $\eta_{\gamma\gamma}\ell^+\ell^-$ yield and the $\eta_{3\pi}\ell^+\ell^-$ yield are constrained in the fit to be consistent with the same $B^0 \rightarrow \eta\ell^+\ell^-$ branching fraction. The signal yield, combinatoric background yield, ARGUS slope and exponential argument float in the fit. All other parameters are fixed from MC samples.

D. Lepton-flavor averaged and isospin averaged fits

In addition to branching fraction measurements and upper limits for the $B \rightarrow \pi\ell^+\ell^-$ and $B^0 \rightarrow \eta\ell^+\ell^-$ modes we also present lepton-flavor averaged, isospin averaged, and lepton-flavor and isospin averaged results. The lepton-flavor averaged measurement of $\mathcal{B}(B^+ \rightarrow \pi^+\ell^+\ell^-)$ is the branching fraction obtained from a simultaneous fit to the $\pi^+e^+e^-$ and $\pi^+\mu^+\mu^-$ samples subject to the constraint $\mathcal{B}(B^+ \rightarrow \pi^+e^+e^-) = \mathcal{B}(B^+ \rightarrow \pi^+\mu^+\mu^-)$. Here we have neglected the difference between the electron and muon masses. The measurements of $\mathcal{B}(B^0 \rightarrow \pi^0\ell^+\ell^-)$ and $\mathcal{B}(B^0 \rightarrow \eta\ell^+\ell^-)$ are subject to a similar set of constraints and are determined in an analogous way. The isospin averaged branching fraction $\mathcal{B}(B \rightarrow \pi e^+e^-)$ is the measured value of $\mathcal{B}(B^+ \rightarrow \pi^+e^+e^-)$ after simultaneously fitting the $\pi^+e^+e^-$ and $\pi^0e^+e^-$ samples subject to the constraint $\mathcal{B}(B^+ \rightarrow \pi^+e^+e^-) = (\tau_{B^0}/2\tau_{B^+})\mathcal{B}(B^0 \rightarrow \pi^0e^+e^-)$ where τ_{B^0} and τ_{B^+} are the mean lifetimes of the neutral and charged B mesons, respectively [36]. An analogous expression is applied for the $\mathcal{B}(B \rightarrow \pi\mu^+\mu^-)$ measurement. The lepton-flavor and isospin averaged measurement of $\mathcal{B}(B \rightarrow \pi\ell^+\ell^-)$ is the value of $\mathcal{B}(B^+ \rightarrow \pi^+\ell^+\ell^-)$ determined from a simultaneous fit to all four samples subject to both the lepton flavor and isospin constraints listed above.

E. Upper limit calculation

We set upper limits on the branching fractions following a method which utilizes the profile likelihood. Upper limits at the α confidence level (CL) are set by scanning the profile likelihood λ as a function of the signal branching fraction to determine where $-2 \ln \lambda$ changes by α percentile of a χ^2 random variable with 1 degree of freedom. For $\alpha = 0.9$ we look for a change in $-2 \ln \lambda$ of 1.642. If the measured branching fraction is negative, we begin our scan from zero rather than the minimum [39]. This is a conservative approach that always produces physical, i.e., non-negative, upper limits, even in the case of low

statistics. Systematic uncertainties are incorporated into the limit by convolving the profile likelihood with a Gaussian distribution whose width is equal to the total systematic uncertainty.

V. OPTIMIZATION OF SELECTION

We simultaneously optimize the selection criteria for the two NN outputs and the PID selection criteria for the charged pions and leptons. *BABAR* employs algorithms which use outputs from one or more multivariate classifiers to identify charged particle species. A few (3–6) standard selections on the outputs of these algorithms are used to identify particles of a given species with different efficiencies. Greater identification efficiencies typically imply greater misidentification rates. Due to this tradeoff, it is not clear *a priori* which selection is best for a particular analysis. Therefore, for each charged particle type (e^- , μ^- , π^+) we optimize the PID requirements for the leptons and pions along with the NN output criteria.

For the optimization we assume that $B \rightarrow \pi\ell^+\ell^-$ and $B^0 \rightarrow \eta\ell^+\ell^-$ occur near the center of the branching fraction ranges expected in the SM. Under this assumption, no statistically significant signal is expected, and the selection is optimized to produce the smallest branching fraction upper limit. We divide the $B\bar{B}$ and continuum NN output space into a grid and generate 2,500 parametrically simulated data sets per grid point according to our fit model. Each simulated data set is fit, and a branching fraction upper limit is calculated. The figure of merit (FOM) for each point is the average branching fraction upper limit determined from the 2,500 data sets, and we take the combination of PID and NN output selection producing the smallest FOM as optimal.

The results of the optimization show that the upper limits are rather insensitive to the PID selection. Also, in the two-dimensional NN output space, there is a region about the optimal selection where the FOM changes slowly, giving confidence that our optimization procedure is robust because the expected limits do not depend critically on the NN selection requirements.

The e^+e^- ($\mu^+\mu^-$) modes use the same electron (muon) selection, while more efficient charged pion selection is favored for $\pi^+e^+e^-$ and $\eta_{3\pi}e^+e^-$ than $\pi^+\mu^+\mu^-$ and $\eta_{3\pi}\mu^+\mu^-$. Tighter selection is favored on the continuum NN output than the $B\bar{B}$ NN output. The optimization favors looser requirements for the $\eta\ell^+\ell^-$ modes as the size of the background in these channels is much smaller than for $\pi\ell^+\ell^-$.

VI. FIT VALIDATION

We validate our fit methodology in three ways: (1) generating an ensemble of data sets from our fit model and fitting them with the same model (“pure pseudoexperiments”), (2) generating and fitting an ensemble of data sets

with signal events from the *BABAR* MC simulation embedded into the data set (“embedded pseudoexperiments”), and (3) extracting $B \rightarrow J/\psi\eta$ and $B^0 \rightarrow J/\psi\eta$ branching fractions from the *BABAR* data sample.

From our studies of both pure and embedded pseudoexperiments, we find no significant source of bias in our fit. Distributions of branching fractions and their errors obtained from fits to these data sets are consistent with expectations.

Measuring the $B^+ \rightarrow J/\psi\pi^+$, $B^0 \rightarrow J/\psi\pi^0$, $B^0 \rightarrow J/\psi\eta$, and $B^+ \rightarrow J/\psi K^+$ branching fractions in the control sample of vetoed charmonium events allows us to validate our fit methodology on data. We employ the same fit model to extract these branching fractions as we do for the $\pi^+\ell^+\ell^-$, $\pi^0\ell^+\ell^-$, $K^+\ell^+\ell^-$, and $\eta\ell^+\ell^-$ branching fractions. Fixed shape parameters and yields are determined through fits to exclusive MC samples. We find that all measurements are in good agreement with world averages [36].

VII. SYSTEMATIC UNCERTAINTIES

Systematic uncertainties are included in the branching fraction upper limit calculation by convolving the profile likelihood with a Gaussian whose width is equal to the total systematic uncertainty. The systematic uncertainties are divided into “multiplicative” uncertainties, which scale with the true value of the branching fraction, and “additive” uncertainties, which are added to the true value of the branching fraction, independent of its value.

A. Multiplicative uncertainties

We list the sources of multiplicative systematic uncertainty below and their assigned values for each of the $\pi\ell^+\ell^-$ and $\eta\ell^+\ell^-$ signal modes in Table I.

The systematic uncertainty in the measured number of $B\bar{B}$ pairs is estimated to be 0.6% [40].

The difference between the π^0 reconstruction efficiency in data and MC has been studied in $\tau^+\tau^-$ decays where one τ decays via the channel $\tau^\pm \rightarrow e^\pm\nu\bar{\nu}$ and the other τ decays via $\tau^\pm \rightarrow \pi^\pm\nu$ or $\tau^\pm \rightarrow \rho^\pm\nu$ with ρ^\pm reconstructed as $\pi^\pm\pi^0$. The $\tau^\pm \rightarrow \rho^\pm\nu$ yields are roughly proportional to the product of the π^\pm and π^0 reconstruction efficiencies, while the $\tau^\pm \rightarrow \pi^\pm\nu$ yields are proportional to the π^\pm reconstruction efficiencies. A correction proportional to the ratio of the $\tau \rightarrow \rho\nu$ to $\tau \rightarrow \pi\nu$ yields in the data and MC samples is applied to better reproduce the data reconstruction efficiency in MC simulation. The uncertainty due to this correction is estimated as 3.0% per π^0 . We take the uncertainty in the $\eta_{\gamma\gamma}$ reconstruction efficiency associated with this correction to also be 3.0% per $\eta_{\gamma\gamma}$.

A correction to the MC tracking efficiency was developed from the study of $\tau^+\tau^-$ decays where one τ has a single charged daughter (1-prong decays) allowing the event to be identified as a $\tau^+\tau^-$ event and the other τ

TABLE I. Multiplicative systematic uncertainties for the $\pi\ell^+\ell^-$ and $\eta\ell^+\ell^-$ modes. The lepton and π^\pm PID and NN output selection efficiency correction uncertainties are determined using $J/\psi K^{(*)}$ control samples, while the tracking and $\pi^0/\eta_{\gamma\gamma}$ efficiency correction and B counting uncertainties are taken from dedicated *BABAR* studies. The total uncertainty is the sum in quadrature of the individual uncertainties.

	$\pi^+e^+e^-$	$\pi^0e^+e^-$	$\pi^+\mu^+\mu^-$	$\pi^0\mu^+\mu^-$	$\eta_{\gamma\gamma}e^+e^-$	$\eta_{3\pi}e^+e^-$	$\eta_{\gamma\gamma}\mu^+\mu^-$	$\eta_{3\pi}\mu^+\mu^-$
$N_{B\bar{B}}$ (%)	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
$\pi^0/\eta_{\gamma\gamma}$ efficiency	...	3.0%	...	3.0%	3.0%	3.0%	3.0%	3.0%
Tracking efficiency (%)	0.9	0.6	0.9	0.6	0.6	1.2	0.6	1.2
Lepton PID (%)	1.3	1.3	1.4	1.4	1.4	1.4	1.5	1.5
π^\pm PID	2.5%	...	3.5%	2.3%	...	3.7%
NN cut (%)	1.4	1.3	1.5	1.5	1.4	1.3	1.5	1.5
Wilson coefficient (%)	2.7	2.3	1.0	1.9	3.1	3.1	0.3	0.9
FF model (%)	9.1	7.7	0.7	7.1	3.4	1.3	0.2	1.6
Total (%)	10.1	8.8	4.4	8.2	5.9	5.6	3.8	5.7

has three charged daughters (3-prong decays). By measuring the event yields where the 3-prong τ has either two or three tracks reconstructed, the track reconstruction efficiency can be measured. This efficiency can be used to correct the MC to match the efficiency measured in data. The systematic uncertainty associated with this correction is estimated to be 0.3% per charged track taken to be 100% correlated among tracks in the event.

We correct for the difference between the lepton PID selection efficiencies in data and MC by measuring the $B^+ \rightarrow J/\psi K^+$ yields in data and $J/\psi K^+$ MC control samples with and without the PID selection requirements applied to both leptons. The ratios of the yields are used to correct the lepton particle identification selection efficiency derived from MC to match data. The error on the correction is taken as the associated systematic uncertainty, which ranges from 1.3% to 1.5%. The available statistics in the samples used to calculate the correction determine the size of the error which is associated with it.

In an analogous procedure, we correct for the difference between the charged pion PID selection efficiency obtained by measuring signal yields in high statistics $B^0 \rightarrow J/\psi K^{*0}(\rightarrow K^-\pi^+)$ data and exclusive MC control samples with and without pion PID selection criteria applied. A correction is derived, and the error on the correction is taken as the associated systematic uncertainty. These uncertainties are approximately 2.5% and 3.5% for e^+e^- and $\mu^+\mu^-$ modes, respectively.

The high statistics of the $B^+ \rightarrow J/\psi K^+$ data and MC control samples are again exploited to derive a correction for the NN output selection efficiency on MC. The $J/\psi K^+$ signal yields were measured with only the $B\bar{B}$ NN output selection applied, only the continuum NN output selection applied, and with both selections applied. The error on the correction is taken as the associated systematic uncertainty and ranges from 1.3% to 1.5%.

We conservatively vary the Wilson coefficients C_7 , C_9 , and C_{10} from their nominal values of -0.313 , 4.344 , and

-4.669 , respectively, by a factor of ± 2 (e.g., C_7 is varied to -0.157 and -0.616) and generate new simulated samples with all possible combinations of the varied Wilson coefficients. For each varied sample we apply the full event selection and calculate the efficiency for that set of Wilson coefficients, taking the largest relative difference between the varied Wilson coefficient efficiencies and the efficiency of our default model as the associated systematic uncertainty.

Simulated MC samples using several different form-factor models were generated. Ultimately, we compare the efficiency from our default model with the efficiency calculated from the “set 2” and “set 4” form-factor models of Ref. [41]. The maximum relative difference between our default model efficiency and the efficiency obtained with the “set 2” and “set 4” form-factor models is taken as the associated systematic uncertainty. There is a large variation in this uncertainty from one mode to another. The source of this effect is due in part to the correlation between the NN output and q^2 . Selection on the NN outputs is mode dependent and therefore changes the q^2 dependence of the efficiency. Also, for the modes $\pi^+e^+e^-$, $\pi^0e^+e^-$, and $\eta_{\gamma\gamma}e^+e^-$ we require that the hadron momentum be greater than 750 MeV/c in the c.m. frame. The hadron momentum is highly correlated with q^2 . Removing events with small hadron momentum also removes events with large q^2 . Differences at large q^2 between the differential branching fractions calculated using the default and alternative models lead to greater sensitivity to the choice of form-factor model and therefore larger uncertainties associated with the choice of model. For $\pi^+e^+e^-$, $\pi^0e^+e^-$, and $\pi^0\mu^+\mu^-$, the uncertainties associated with the form-factor models dominate the multiplicative uncertainty, while for $\pi^+\mu^+\mu^-$ the uncertainty is 0.7%.

The uncertainty in the efficiency due to the size of the simulated MC samples is less than 0.1% and is negligible.

B. Additive uncertainties

We consider the following sources of additive systematic uncertainty with their values given in Table II.

The fixed parameters of the $\pi\ell^+\ell^-$, $\eta\ell^+\ell^-$, and $K^+\ell^+\ell^-$ signal and the $K^+\ell^+\ell^-$ background PDFs are varied individually within the errors obtained from fits to exclusive MC samples, and the data sample is refit. For simultaneous fits we additionally vary the efficiencies within their uncertainties, and for $\eta\ell^+\ell^-$ we vary the self-cross-feed-to-signal ratio by $\pm 10\%$. The size of the variation is arbitrary but conservative enough since the number of expected self-cross-feed events is at most 0.15. The difference between the branching fraction from this fit and that from the nominal fit is taken as the associated systematic uncertainty. We take the largest change in the branching fraction as the systematic uncertainty associated with each fixed quantity. The uncertainties from individual variations are summed in quadrature.

Nonparametric PDFs include the two-dimensional histograms and KEYS shapes. We vary the binning of the two-dimensional PDFs, increasing and decreasing them by a factor of 2. The data sample is refit, and we take the largest change in the branching fraction as the associated systematic uncertainty. For the KEYS shapes, we increase and decrease the width of the Gaussian kernel used to generate the shapes and take the largest change in the branching fraction as the associated systematic uncertainty. If there are multiple nonparametric PDFs, we add their associated uncertainties in quadrature.

The hadronic peaking background yields are fixed from the control sample of hadronic decays and are varied within their statistical uncertainties. The data sample is refit, and we take the largest change in the branching fraction as the associated systematic uncertainty.

We fix the $K_S^0/K^*\ell^+\ell^-$ yield in the $\pi^+\ell^+\ell^-$ PDF, the $K^*\ell^+\ell^-$ yield in the $K^+\ell^+\ell^-$ PDF, and the $K_S^0\ell^+\ell^-$ yield in the $\pi^0\ell^+\ell^-$ PDF. These values are determined from efficiencies taken from exclusive MC samples and the current world average branching fractions for these modes [36]. We vary the yields according to the errors on their branching fractions and refit the data sample. The change in the branching fraction from its nominal value is taken as the associated systematic uncertainty. In Table II, these uncertainties are classified as ‘‘nonhadronic peaking bkg yields.’’

TABLE II. Additive systematic uncertainties for the $B \rightarrow \pi\ell^+\ell^-$ and $B^0 \rightarrow \eta\ell^+\ell^-$ channels. The total uncertainty is the sum in quadrature of the individual uncertainties. All uncertainties are given in units of 10^{-8} .

Mode	$\pi^+e^+e^-$	$\pi^0e^+e^-$	$\pi^+\mu^+\mu^-$	$\pi^0\mu^+\mu^-$	ηe^+e^-	$\eta\mu^+\mu^-$
Fixed parameters	1.9	0.2	0.6	0.3	0.6	0.4
Nonparametric shapes	<0.1	<0.1	0.7	0.5	0.1	0.1
Hadronic peaking background yields	<0.1	<0.1
Nonhadronic peaking background yields	<0.1	<0.1
Total	1.9	0.2	0.9	0.6	0.6	0.4

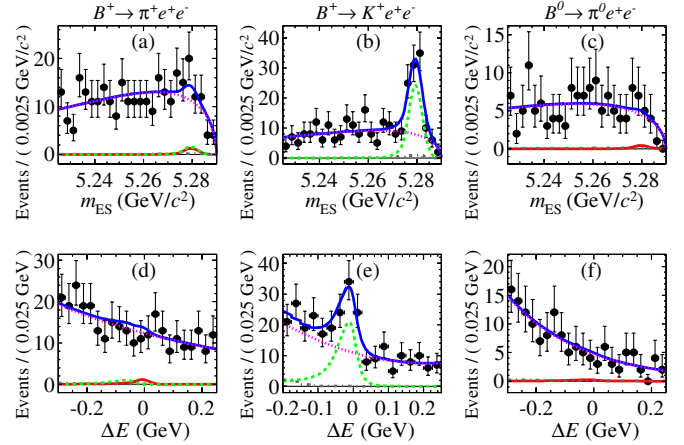


FIG. 3 (color online). Fit projections of m_{ES} (top) and ΔE (bottom) for the isospin averaged $B \rightarrow \pi e^+e^-$ fit to the $\pi^+e^+e^-$ [(a) and (d)], $K^+e^+e^-$ [(b) and (e)], and $\pi^0e^+e^-$ [(c) and (f)] data sets. Points with error bars represent data. The curves are dotted (magenta) for combinatoric background, dotted-dashed (gray) for $K^*/K_S^0\ell^+\ell^-$ background, dashed (green) for $K^+\ell^+\ell^-$ signal and background, and solid (red) for $\pi\ell^+\ell^-$ signal. The solid blue curve represents the total fit function.

VIII. RESULTS

We extract branching fractions by fitting the data set with the fit model described in Sec. IV. Projections of the PDFs and data sets in m_{ES} and ΔE are shown for the isospin averaged $B \rightarrow \pi e^+e^-$ fit, the lepton-flavor averaged $B^0 \rightarrow \eta\ell^+\ell^-$ fit, and the isospin and lepton-flavor averaged $B \rightarrow \pi\ell^+\ell^-$ fit in Figs. 3–5 respectively. Figure 6 shows $-2 \ln \lambda$ as a function of the branching fraction for the $\pi\ell^+\ell^-$, $\pi^+\ell^+\ell^-$, $\pi^0\ell^+\ell^-$, $\eta\ell^+\ell^-$, ηe^+e^- , and $\eta\mu^+\mu^-$ measurements. Branching fraction measurements and upper limits at 90% CL are given in Table II for each mode.

As a cross-check, we measure the $B^+ \rightarrow K^+\ell^+\ell^-$ branching fractions and find them consistent with the current world averages [36].

We set upper limits on the lepton-flavor averaged branching fractions of

$$\mathcal{B}(B^+ \rightarrow \pi^+\ell^+\ell^-) < 6.6 \times 10^{-8}, \quad (7)$$

$$\mathcal{B}(B^0 \rightarrow \pi^0\ell^+\ell^-) < 5.3 \times 10^{-8}, \quad (8)$$

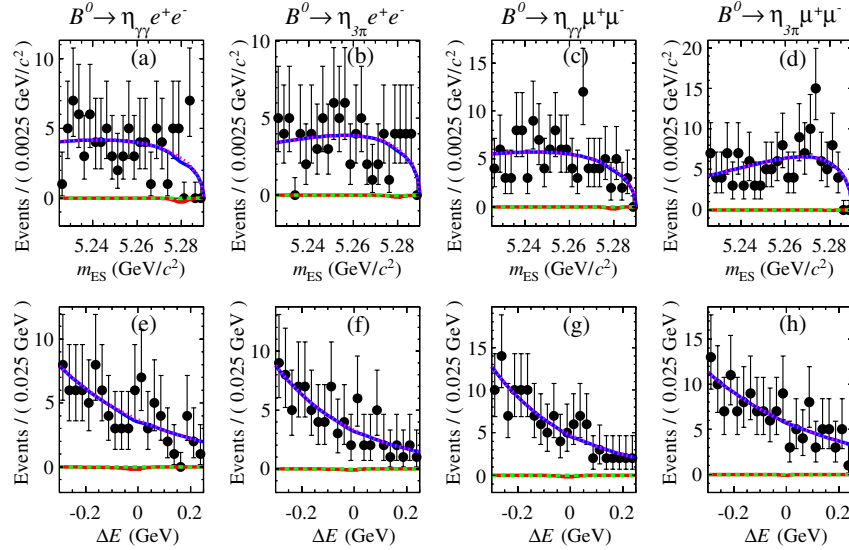


FIG. 4 (color online). Fit projections of m_{ES} (top) and ΔE (bottom) for the lepton-flavor averaged fit to the $\eta_{\gamma\gamma}e^+e^-$ [(a) and (e)], $\eta_{3\pi}e^+e^-$ [(b) and (f)], $\eta_{\gamma\gamma}\mu^+\mu^-$ [(c) and (g)], and $\eta_{3\pi}\mu^+\mu^-$ [(d) and (h)] samples. Points with error bars represent data. The curves are dotted (magenta) for combinatoric background, dashed (green) for self-cross-feed, and solid (red) for $\eta\ell^+\ell^-$ signal. The solid blue curve represents the total fit function.

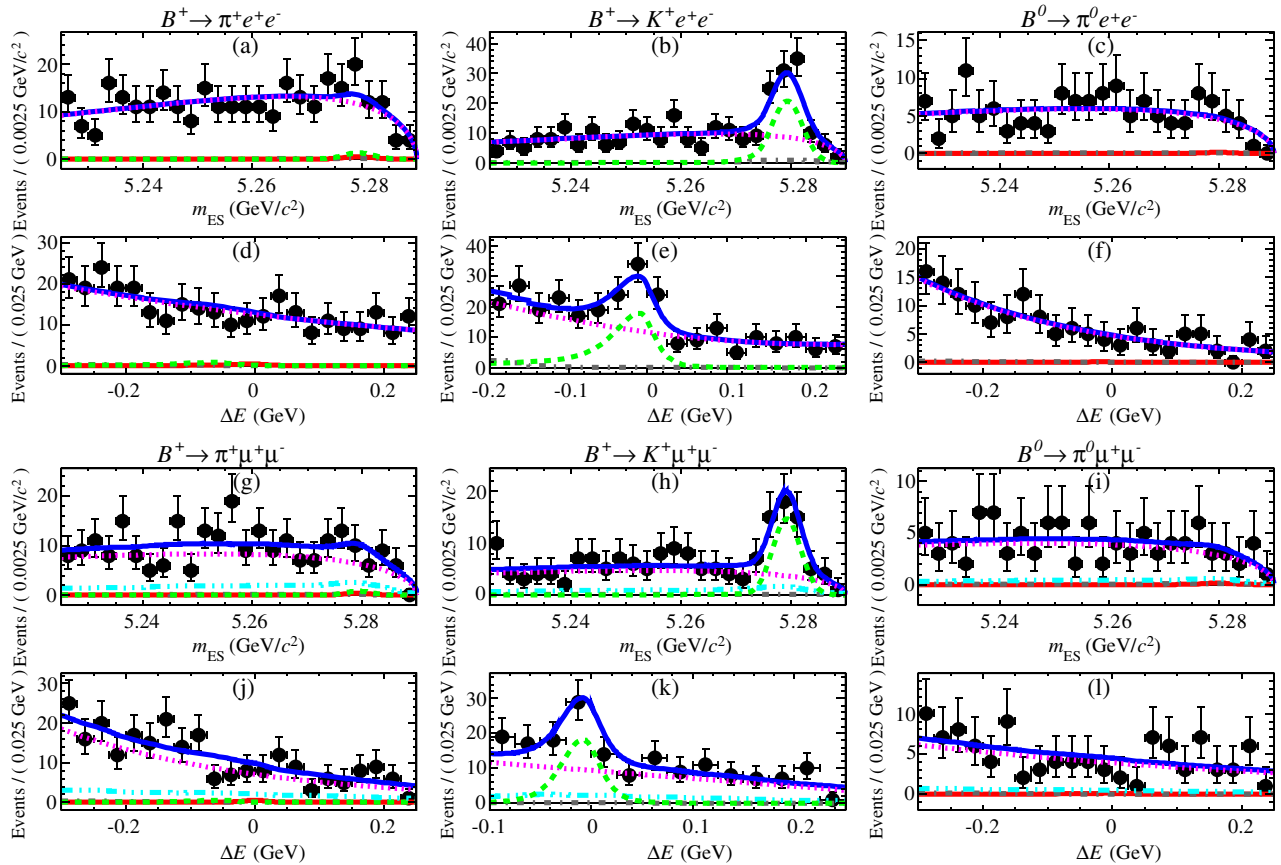


FIG. 5 (color online). Fit projections of m_{ES} and ΔE for the lepton-flavor and isospin averaged $\pi\ell^+\ell^-$ fit to the $\pi^+e^+e^-$ [(a) and (d)], $K^+e^+e^-$ [(b) and (e)], $\pi^0e^+e^-$ [(c) and (f)], $\pi^+\mu^+\mu^-$ [(g) and (j)], $K^+\mu^+\mu^-$ [(h) and (k)], and $\pi^0\mu^+\mu^-$ [(i) and (l)] data sets. Points with error bars represent data. The curves are dotted (magenta) for combinatoric background, triple dotted-dashed (cyan) for hadronic peaking background in the $\mu^+\mu^-$ modes, dotted-dashed (gray) for $K^*/K_S^0\ell^+\ell^-$ background, dashed (green) for $K^+\ell^+\ell^-$ signal and background, and solid (red) for $\pi\ell^+\ell^-$ signal. The solid blue curve represents the total fit function.

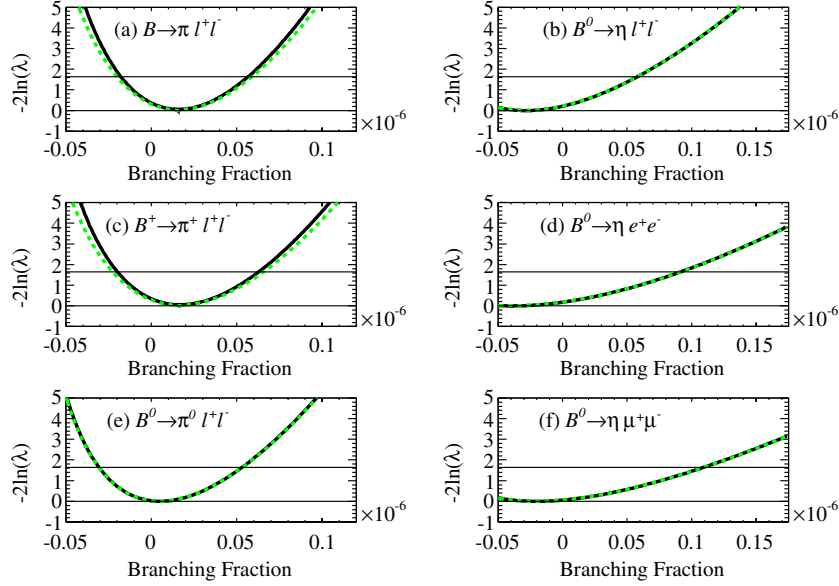


FIG. 6 (color online). The negative logarithm of the profile likelihood as a function of branching fraction for (a) $B \rightarrow \pi \ell^+ \ell^-$, (b) $B^0 \rightarrow \eta \ell^+ \ell^-$, (c) $B^+ \rightarrow \pi^+ \ell^+ \ell^-$, (d) $B^0 \rightarrow \eta e^+ e^-$, (e) $B^0 \rightarrow \pi^0 \ell^+ \ell^-$, and (f) $B^0 \rightarrow \eta \mu^+ \mu^-$. The solid black lines are the negative log likelihood curves including only statistical errors, and the dashed green lines correspond to the same curves convolved with a Gaussian distribution whose width is equal to the total systematic error.

$$\mathcal{B}(B^0 \rightarrow \eta \ell^+ \ell^-) < 6.4 \times 10^{-8}, \quad (9)$$

all at the 90% CL. A lepton-flavor and isospin averaged branching fraction upper limit of

$$\mathcal{B}(B \rightarrow \pi \ell^+ \ell^-) < 5.9 \times 10^{-8} \quad (10)$$

is set at 90% CL. Branching fraction measurements and upper limits at 90% CL for the modes $B^+ \rightarrow \pi^+ e^+ e^-$,

$B^0 \rightarrow \pi^0 e^+ e^-$, $B^+ \rightarrow \pi^+ \mu^+ \mu^-$, $B^0 \rightarrow \pi^0 \mu^+ \mu^-$, $B^0 \rightarrow \eta e^+ e^-$, and $B^0 \rightarrow \eta \mu^+ \mu^-$ are listed in Table III.

In conclusion, we have searched for the rare decays $B \rightarrow \pi \ell^+ \ell^-$ and $B^0 \rightarrow \eta \ell^+ \ell^-$ in a sample of 471 million $B\bar{B}$ decays and observe no statistically significant signal in any of the decay channels studied. We set a lepton-flavor and isospin averaged upper limit at the 90% CL of $\mathcal{B}(B \rightarrow \pi \ell^+ \ell^-) < 5.9 \times 10^{-8}$, within a factor of three of

TABLE III. $B \rightarrow \pi \ell^+ \ell^-$ and $B^0 \rightarrow \eta \ell^+ \ell^-$ corrected efficiencies, yields, branching fractions, and branching fraction upper limits at the 90% CL. The error on the yield is statistical. The first error quoted on the branching fractions is statistical while the second is systematic. Branching fraction upper limits include systematic uncertainties.

Mode	ε	Yield	$\mathcal{B}(10^{-8})$	Upper limit (10^{-8})
$B^+ \rightarrow \pi^+ e^+ e^-$	0.207	$4.2^{+5.7}_{-4.6}$	$4.3^{+5.9}_{-4.3} \pm 2.0$	12.5
$B^0 \rightarrow \pi^0 e^+ e^-$	0.166	$1.0^{+4.2}_{-3.2}$	$1.3^{+5.4}_{-4.1} \pm 0.2$	8.4
$B^0 \rightarrow \eta e^+ e^-$			$-4.0^{+10.0}_{-8.0} \pm 0.6$	10.8
$B^0 \rightarrow \eta_{\gamma\gamma} e^+ e^-$	0.166	$-1.2^{+3.1}_{-2.4}$		
$B^0 \rightarrow \eta_{3\pi} e^+ e^-$	0.111	$-0.5^{+1.2}_{-0.9}$		
$B^+ \rightarrow \pi^+ \mu^+ \mu^-$	0.149	$-0.5^{+3.1}_{-2.3}$	$-0.7^{+4.4}_{-3.2} \pm 0.9$	5.5
$B^0 \rightarrow \pi^0 \mu^+ \mu^-$	0.121	$-0.2^{+3.0}_{-2.0}$	$-0.3^{+5.3}_{-3.6} \pm 0.6$	6.9
$B^0 \rightarrow \eta \mu^+ \mu^-$			$-2.0^{+10.0}_{-6.6} \pm 0.4$	11.2
$B^0 \rightarrow \eta_{\gamma\gamma} \mu^+ \mu^-$	0.104	$-0.4^{+1.9}_{-1.3}$		
$B^0 \rightarrow \eta_{3\pi} \mu^+ \mu^-$	0.063	$-0.1^{+0.6}_{-0.4}$		
$B \rightarrow \pi e^+ e^-$			$4.0^{+5.1}_{-4.3} \pm 1.6$	11.0
$B \rightarrow \pi \mu^+ \mu^-$			$-0.7^{+4.1}_{-3.1} \pm 1.2$	5.0
$B^+ \rightarrow \pi^+ \ell^+ \ell^-$			$1.6^{+3.6}_{-3.0} \pm 1.2$	6.6
$B^0 \rightarrow \pi^0 \ell^+ \ell^-$			$0.5^{+3.7}_{-2.9} \pm 0.3$	5.3
$B^0 \rightarrow \eta \ell^+ \ell^-$			$-2.8^{+6.6}_{-5.2} \pm 0.3$	6.4
$B \rightarrow \pi \ell^+ \ell^-$			$1.6^{+3.2}_{-2.7} \pm 1.0$	5.9

the SM expectation. We also set lepton-flavor averaged upper limits of $\mathcal{B}(B^+ \rightarrow \pi^+ \ell^+ \ell^-) < 6.6 \times 10^{-8}$ and $\mathcal{B}(B^0 \rightarrow \pi^0 \ell^+ \ell^-) < 5.3 \times 10^{-8}$. Branching fraction upper limits at 90% CL have also been calculated for the modes $B^+ \rightarrow \pi^+ e^+ e^-$, $B^0 \rightarrow \pi^0 e^+ e^-$, $B^+ \rightarrow \pi^+ \mu^+ \mu^-$, and $B^0 \rightarrow \pi^0 \mu^+ \mu^-$. Our upper limits on the $B^0 \rightarrow \pi^0 e^+ e^-$, $B^0 \rightarrow \pi^0 \mu^+ \mu^-$, and $B^0 \rightarrow \pi^0 \ell^+ \ell^-$ branching fractions are the lowest upper limits to date. The results presented for the $\pi \ell^+ \ell^-$ modes supersede those of the previous *BABAR* analysis [20]. We have also performed the first search for the decays $B^0 \rightarrow \eta \ell^+ \ell^-$ and set an upper limit on the lepton-flavor averaged branching fraction of $\mathcal{B}(B^0 \rightarrow \eta \ell^+ \ell^-) < 6.4 \times 10^{-8}$ at the 90% CL. Upper limits at 90% CL for the $B^0 \rightarrow \eta e^+ e^-$ and $B^0 \rightarrow \eta \mu^+ \mu^-$ branching fractions have been reported.

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 U.S. Department of Energy and National Science Foundation, the Natural Sciences and Engineering Research Council (Canada), the Commissariat à l'Énergie 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 (Netherlands), the Research Council of Norway, the Ministry of Education and Science of the Russian Federation, Ministerio de Economía y Competitividad (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 (USA).

-
- [1] N. Cabibbo, *Phys. Rev. Lett.* **10**, 531 (1963).
 [2] M. Kobayashi and T. Maskawa, *Prog. Theor. Phys.* **49**, 652 (1973).
 [3] T.M. Aliev and M. Savci, *Phys. Rev. D* **60**, 014005 (1999).
 [4] J.J. Wang, R.M. Wang, Y.G. Xu, and Y.D. Yang, *Phys. Rev. D* **77**, 014017 (2008).
 [5] H.Z. Song, L.X. Lu, and G.R. Lu, *Commun. Theor. Phys.* **50**, 696 (2008).
 [6] G. Erkol and G. Turan, *Eur. Phys. J. C* **28**, 243 (2003).
 [7] M. Beneke, T. Feldmann, and D. Seidel, *Nucl. Phys.* **B612**, 25 (2001).
 [8] M. Beneke, T. Feldmann, and D. Seidel, *Eur. Phys. J. C* **41**, 173 (2005).
 [9] T. Feldmann and J. Matias, *J. High Energy Phys.* **01** (2003) 074.
 [10] A. Ali, P. Ball, L. T. Handoko, and G. Hiller, *Phys. Rev. D* **61**, 074024 (2000).
 [11] J.P. Lees *et al.* (*BABAR* Collaboration), *Phys. Rev. D* **86**, 032012 (2012).
 [12] B. Aubert *et al.* (*BABAR* Collaboration), *Phys. Rev. Lett.* **102**, 091803 (2009).
 [13] T. Altonen *et al.* (CDF Collaboration), *Phys. Rev. Lett.* **108**, 081807 (2012).
 [14] J.T. Wei *et al.* (Belle Collaboration), *Phys. Rev. Lett.* **103**, 171801 (2009).
 [15] R. Aaij *et al.* (LHCb Collaboration), *J. High Energy Phys.* **07** (2012) 133.
 [16] R. Aaij *et al.* (LHCb Collaboration), *J. High Energy Phys.* **02** (2013) 105.
 [17] R. Aaij *et al.* (LHCb Collaboration), *Phys. Rev. Lett.* **110**, 031801 (2013).
 [18] S. Davidson, D. Bailey, and B.A. Campbell, *Z. Phys. C* **61**, 613 (1994).
 [19] R. Aaij *et al.* (LHCb Collaboration), *J. High Energy Phys.* **12** (2012) 125.
 [20] B. Aubert *et al.* (*BABAR* Collaboration), *Phys. Rev. Lett.* **99**, 051801 (2007).
 [21] J.T. Wei *et al.* (Belle Collaboration), *Phys. Rev. D* **78**, 011101 (2008).
 [22] J.P. Lees *et al.* (*BABAR* Collaboration), [arXiv:1301.2703](https://arxiv.org/abs/1301.2703).
 [23] B. Aubert *et al.* (*BABAR* Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **479**, 1 (2002).
 [24] S. Agostinelli *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **506**, 250 (2003).
 [25] D.J. Lange, *Nucl. Instrum. Methods Phys. Res., Sect. A* **462**, 152 (2001).
 [26] T. Sjöstrand, [arXiv:hep-ph/9508391](https://arxiv.org/abs/hep-ph/9508391).
 [27] P. Ball, *Proceedings of the Second Workshop on the CKM Unitarity Triangle*, eConf. C0304052, WG101 (2003).
 [28] A. Ali, E. Lunghi, C. Greub, and G. Hiller, *Phys. Rev. D* **66**, 034002 (2002).
 [29] H.H. Asatryan, H.M. Asatryan, C. Greub, and M. Walker, *Phys. Rev. D* **65**, 074004 (2002).
 [30] C. Bobeth, M. Misiak, and J. Urban, *Nucl. Phys.* **B574**, 291 (2000).
 [31] W.D. Hulsbergen, *Nucl. Instrum. Methods Phys. Res., Sect. A* **552**, 566 (2005).
 [32] G.C. Fox and S. Wolfram, *Nucl. Phys.* **B149**, 413 (1979).
 [33] S. Brandt, C. Peyrou, R. Sosnowski, and A. Wroblewski, *Phys. Lett.* **12**, 57 (1964).
 [34] The momentum weighted polynomials are defined as $L_i^j = \sum_{k=0}^N |\vec{p}_k|^j \cos^i \theta_k$ with the sum running over N particles (charged tracks and EMC clusters) in the event.

The angle θ_k is the angle between \vec{p}_k , the momentum of the k th particle in the event and some predefined direction, e.g., the thrust axis, the ROE thrust axis, or the dilepton momentum.

- [35] K. S. Cranmer, *Comput. Phys. Commun.* **136**, 198 (2001).
- [36] J. Beringer *et al.* (Particle Data Group), *Phys. Rev. D* **86**, 010001 (2012).
- [37] M. J. Oreglia, Ph. D. thesis, Stanford University, 1980.
- [38] H. Albrecht *et al.*, *Phys. Lett. B* **241**, 278 (1990).
- [39] W. A. Rolke, A. M. Lopez, and J. Conrad, *Nucl. Instrum. Methods Phys. Res., Sect. A* **551**, 493 (2005).
- [40] G. D. McGregor, Master's thesis, University of British Columbia, 2008, [arXiv:0812.1954](https://arxiv.org/abs/0812.1954).
- [41] P. Ball and R. Zwicky, *Phys. Rev. D* **71**, 014015 (2005).