Measurement of an excess of [¯ over B] → D[superscript (*)]#[superscript -]#[¯ over ν][subscript #] decays and implications for charged Higgs bosons

Measurement of an excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays and implications for charged Higgs bosons

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,[4] L. T. Kerth,[5] Yu. G. Kolomensky,[5] M. Lee,[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,[30] D. Derkach,[30] G. Grosdidier,[30] F. Le Diberder,[30] A. M. Lutz,[30] B. Malaescu,[30] P. Roudeau,[30] A. Stocchi,[30] G. Wormser,[30] 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,[5] 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. Bünger,[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,[51] 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. Wu[75]

(The *BABAR* Collaboration)

[1]*Laboratoire d'Annecy-le-Vieux de Physique des Particules (LAPP), Université de Savoie, CNRS/IN2P3, F-74941 Annecy-Le-Vieux, France*
[2]*Facultat de Fisica, Departament ECM, 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]*Institute of Physics, University of Bergen, N-5007 Bergen, Norway*
[5]*Lawrence Berkeley National Laboratory and University of California, Berkeley, California 94720, USA*
[6]*Institut für Experimentalphysik 1, Ruhr Universität Bochum, D-44780 Bochum, Germany*
[7]*University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1*
[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]*Institute for Particle Physics, University of California at Santa Cruz, 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]*Institut für Kern- und Teilchenphysik, Technische Universität Dresden, 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]*Physikalisches Institut, Universität Heidelberg, Philosophenweg 12, D-69120 Heidelberg, Germany*
[28]*Institut für Physik, Humboldt-Universität zu Berlin, 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, Centre Scientifique d'Orsay, IN2P3/CNRS et Université Paris-Sud 11, 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]*Royal Holloway and Bedford New College, University of London, Egham, Surrey TW20 0EX, United Kingdom*
[38]*University of Louisville, Louisville, Kentucky 40292, USA*
[39]*Institut für Kernphysik, Johannes Gutenberg-Universität Mainz, D-55099 Mainz, Germany*
[40]*University of Manchester, Manchester M13 9PL, United Kingdom*
[41]*University of Maryland, College Park, Maryland 20742, USA*
[42]*Massachusetts Institute of Technology, Laboratory for Nuclear Science, Cambridge, Massachusetts 02139, USA*
[43]*McGill University, Montréal, Québec, Canada H3A 2T8*
[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]*Physique des Particules, Université de Montréal, Montréal, Québec, Canada H3C 3J7*
[47a]*INFN Sezione di Napoli, I-80126 Napoli, Italy*
[47b]*Dipartimento di Scienze Fisiche, Università di Napoli Federico II, I-80126 Napoli, Italy*
[48]*National Institute for Nuclear Physics and High Energy Physics, NIKHEF, NL-1009 DB Amsterdam, 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]Centre de Saclay, Irfu, SPP, CEA, 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]School of Physics and Astronomy, Tel Aviv University, 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, Canada V8W 3P6
[74]Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom
[75]University of Wisconsin, Madison, Wisconsin 53706, USA
(Received 5 March 2013; published 31 October 2013)

Based on the full *BABAR* data sample, we report improved measurements of the ratios $\mathcal{R}(D) = \mathcal{B}(\overline{B} \to D\tau^-\overline{\nu}_\tau)/\mathcal{B}(\overline{B} \to D\ell^-\overline{\nu}_\ell)$ and $\mathcal{R}(D^*) = \mathcal{B}(\overline{B} \to D^*\tau^-\overline{\nu}_\tau)/\mathcal{B}(\overline{B} \to D^*\ell^-\overline{\nu}_\ell)$, where ℓ refers to either an electron or muon. These ratios are sensitive to new physics contributions in the form of a charged Higgs boson. We measure $\mathcal{R}(D) = 0.440 \pm 0.058 \pm 0.042$ and $\mathcal{R}(D^*) = 0.332 \pm 0.024 \pm 0.018$, which exceed the standard model expectations by 2.0σ and 2.7σ, respectively. Taken together, the results disagree with these expectations at the 3.4σ level. This excess cannot be explained by a charged Higgs boson in the type II two-Higgs-doublet model. Kinematic distributions presented here exclude large portions of the more general type III two-Higgs-doublet model, but there are solutions within this model compatible with the results.

DOI: 10.1103/PhysRevD.88.072012 PACS numbers: 13.20.He, 14.40.Nd, 14.80.Da

I. INTRODUCTION

In the standard model (SM), semileptonic decays of B mesons proceed via first-order electroweak interactions and are mediated by the W boson [1–3]. Decays involving electrons and muons are expected to be insensitive to non-SM contributions and therefore have been the bases of the determination of the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements $|V_{cb}|$ and $|V_{ub}|$ [4]. Decays involving the higher-mass τ lepton provide additional information on SM processes and are sensitive to additional amplitudes, such as those involving an intermediate charged Higgs boson [5–9]. Thus, they offer an excellent opportunity to search for this and other non-SM contributions.

Over the past two decades, the development of heavy-quark effective theory (HQET) and precise measurements of $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays [10] at the B factories [11,12] have greatly improved our understanding of exclusive semileptonic decays. The relative rates

[*]Deceased.
[†]Present address: the University of Tabuk, Tabuk 71491, Saudi Arabia.
[‡]Also at Università di Perugia, Dipartimento di Fisica, Perugia, Italy.
[§]Present address: the University of Huddersfield, Huddersfield HD1 3DH, United Kingdom.
[‖]Present address: University of South Alabama, Mobile, Alabama 36688, USA.
[¶]Also at Università di Sassari, Sassari, Italy.
[**]Present address: Universidad Técnica Federico Santa Maria, Valparaiso, Chile 2390123.

$$\mathcal{R}(D) = \frac{\mathcal{B}(\bar{B} \to D\tau^- \bar{\nu}_\tau)}{\mathcal{B}(\bar{B} \to D\ell^- \bar{\nu}_\ell)},$$
$$\mathcal{R}(D^*) = \frac{\mathcal{B}(\bar{B} \to D^*\tau^- \bar{\nu}_\tau)}{\mathcal{B}(\bar{B} \to D^*\ell^- \bar{\nu}_\ell)} \quad (1)$$

are independent of the CKM element $|V_{cb}|$ and also, to a large extent, of the parametrization of the hadronic matrix elements. SM expectations [9] for the ratios $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ have uncertainties of less than 6% and 2%, respectively. Calculations [5–9] based on two-Higgs-doublet models predict a substantial impact on the ratio $\mathcal{R}(D)$, and a smaller effect on $\mathcal{R}(D^*)$ due to the spin of the D^* meson.

The decay $\bar{B} \to D^*\tau^- \bar{\nu}_\tau$ was first observed in 2007 by the Belle Collaboration [13]. Since then, both *BABAR* and Belle have published improved measurements, and have found evidence for $\bar{B} \to D\tau^- \bar{\nu}_\tau$ decays [14–16]. Up to now, the measured values for $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ have consistently exceeded the SM expectations, though the significance of the excess is low due to the large statistical uncertainties.

We recently presented an update of the earlier measurement [14] based on the full *BABAR* data sample [17]. This update included improvements to the event reconstruction that increased the signal efficiency by more than a factor of 3. In the following, we describe the analysis in greater detail, present the distributions of some important kinematic variables, and expand the interpretation of the results.

We choose to reconstruct only the purely leptonic decays of the τ lepton, $\tau^- \to e^- \bar{\nu}_e \nu_\tau$ and $\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$, so that $\bar{B} \to D^{(*)}\tau^- \bar{\nu}_\tau$ and $\bar{B} \to D^{(*)}\ell^- \bar{\nu}_\ell$ decays are identified by the same particles in the final state. This leads to the cancellation of various detection efficiencies and the reduction of related uncertainties on the ratios $\mathcal{R}(D^{(*)})$.

Candidate events originating from $\Upsilon(4S) \to B\bar{B}$ decays are selected by fully reconstructing the hadronic decay of one of the B mesons ($B_{\rm tag}$), and identifying the semileptonic decay of the other B by a charm meson (charged or neutral D or D^* meson), a charged lepton (either e or μ) and the missing momentum and energy in the whole event.

Yields for the signal decays $\bar{B} \to D^{(*)}\tau^- \bar{\nu}_\tau$ and the normalization decays $\bar{B} \to D^{(*)}\ell^- \bar{\nu}_\ell$ are extracted by an unbinned maximum-likelihood fit to the two-dimensional distributions of the invariant mass of the undetected particles $m^2_{\rm miss} = p^2_{\rm miss} = (p_{e^+e^-} - p_{B_{\rm tag}} - p_{D^{(*)}} - p_\ell)^2$ (where $p_{e^+e^-}$, $p_{B_{\rm tag}}$, $p_{D^{(*)}}$, and p_ℓ refer to the four-momenta of the colliding beams, the $B_{\rm tag}$, the $D^{(*)}$, and the charged lepton, respectively) versus the lepton three-momentum in the B rest frame, $|\boldsymbol{p}^*_\ell|$. The $m^2_{\rm miss}$ distribution for decays with a single missing neutrino peaks at zero, whereas signal events, which have three missing neutrinos, have a broad $m^2_{\rm miss}$ distribution that extends to about 9 GeV^2. The observed lepton in signal events is a secondary particle from the τ decay, so its $|\boldsymbol{p}^*_\ell|$ spectrum is softer than for primary leptons in normalization decays.

The principal sources of background originate from $B\bar{B}$ decays and from continuum events, i.e., $e^+e^- \to f\bar{f}(\gamma)$ pair production, where $f = u, d, s, c, \tau$. The yields and distributions of these two background sources are derived from selected data control samples. The background decays that are most difficult to separate from signal decays come from semileptonic decays to higher-mass, excited charm mesons, since they can produce similar $m^2_{\rm miss}$ and $|\boldsymbol{p}^*_\ell|$ values to signal decays and their branching fractions and decay properties are not well known. Thus, their impact on the signal yield is examined in detail.

The choice of the selection criteria and fit configuration are based on samples of simulated and data events. To avoid bias in the determination of the signal yield, the signal region was blinded for data until the analysis procedure was settled.

II. THEORY OF $\bar{B} \to D^{(*)}\tau^- \bar{\nu}_\tau$ DECAYS

A. Standard model

Given that leptons are not affected by quantum chromodynamic (QCD) interactions (see Fig. 1), the matrix element of $\bar{B} \to D^{(*)}\tau^- \bar{\nu}_\tau$ decays can be factorized in the form [5]

$$\mathcal{M}^{\lambda_\tau}_{\lambda_{D^{(*)}}}(q^2, \theta_\tau) = \frac{G_F V_{cb}}{\sqrt{2}} \sum_{\lambda_W} \eta_{\lambda_W} L^{\lambda_\tau}_{\lambda_W}(q^2, \theta_\tau) H^{\lambda_{D^{(*)}}}_{\lambda_W}(q^2), \quad (2)$$

where $L^{\lambda_\tau}_{\lambda_W}$ and $H^{\lambda_{D^{(*)}}}_{\lambda_W}$ are the leptonic and hadronic currents defined as

$$L^{\lambda_\tau}_{\lambda_W}(q^2, \theta_\tau) \equiv \epsilon_\mu(\lambda_W) \langle \tau \bar{\nu}_\tau | \bar{\tau} \gamma^\mu (1 - \gamma_5) \nu_\tau | 0 \rangle, \quad (3)$$

$$H^{\lambda_{D^{(*)}}}_{\lambda_W}(q^2) \equiv \epsilon^*_\mu(\lambda_W) \langle D^{(*)} | \bar{c} \gamma^\mu (1 - \gamma_5) b | \bar{B} \rangle. \quad (4)$$

Here, ϵ are polarization vectors, the indices λ refer to the helicities of the W, $D^{(*)}$, and τ, $q = p_B - p_{D^{(*)}}$ is the four-momentum of the virtual W, and θ_τ is the angle between the τ and the $D^{(*)}$ three-momenta measured in the rest frame of the virtual W. The metric factor η in Eq. (2) is $\eta_{\{\pm,0,s\}} = \{1, 1, -1\}$, where $\lambda_W = \pm, 0$, and s refer to the four helicity states of the virtual W boson (s is the scalar state which, of course, has helicity 0).

The leptonic currents can be calculated analytically with the standard framework of electroweak interactions. In the rest frame of the virtual W (W^*), they take the form [18]:

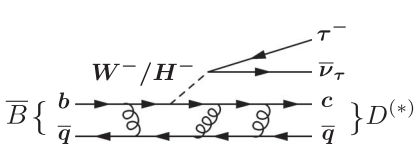

FIG. 1. Parton level diagram for $\bar{B} \to D^{(*)}\tau^- \bar{\nu}_\tau$ decays. The gluon lines illustrate the QCD interactions that affect the hadronic part of the amplitude.

$$L_{\pm}^{-} = -2\sqrt{q^2}vd_{\pm}, \qquad L_{\pm}^{+} = \mp\sqrt{2}m_\tau v d_0, \tag{5}$$

$$L_0^{-} = -2\sqrt{q^2}vd_0, \qquad L_0^{+} = \sqrt{2}m_\tau v(d_+ - d_-), \tag{6}$$

$$L_s^{-} = 0, \qquad L_s^{+} = -2m_\tau v, \tag{7}$$

with

$$v = \sqrt{1 - \frac{m_\tau^2}{q^2}}, \qquad d_\pm = \frac{1 \pm \cos\theta_\tau}{\sqrt{2}}, \qquad d_0 = \sin\theta_\tau. \tag{8}$$

Due to the nonperturbative nature of the QCD interaction at this energy scale, the hadronic currents cannot be calculated analytically. They are expressed in terms of form factors (FF) as functions of q^2 (see Secs. II A 1 and II A 2).

The differential decay rate, integrated over angles, is derived from Eqs. (2) and (5)–(7) [2]:

$$\frac{d\Gamma_\tau}{dq^2} = \frac{G_F^2|V_{cb}|^2|\boldsymbol{p}^*_{D^{(*)}}|q^2}{96\pi^3 m_B^2}\left(1 - \frac{m_\tau^2}{q^2}\right)^2\Big[(|H_+|^2 + |H_-|^2 + |H_0|^2)\left(1 + \frac{m_\tau^2}{2q^2}\right) + \frac{3m_\tau^2}{2q^2}|H_s|^2\Big], \tag{9}$$

where $|\boldsymbol{p}^*_{D^{(*)}}|$ is the three-momentum of the $D^{(*)}$ meson in the B rest frame. For simplicity, the helicities of the $D^{(*)}$ meson and the q^2 dependence of the hadron helicity amplitudes $H_{\pm,0,s}$ have been omitted. The assignment is unambiguous because in $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decays, $H_\pm$ only receive contributions from $\lambda_{D^*} = \pm$, while $H_{0,s}$ require $\lambda_{D^*} = 0$. In $\bar{B} \to D\tau^-\bar{\nu}_\tau$ decays, only $\lambda_D = s$ is possible, which implies $H_\pm = 0$.

1. Form factor parametrization of $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decays

Four independent FFs, V, A_0, A_1, and A_2, describe the nonperturbative QCD interactions in $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decays. Based on the FF convention of Ref. [9], the hadronic currents take the following form:

$$H_\pm(q^2) = (m_B + m_{D^*})A_1(q^2) \mp \frac{2m_B}{m_B + m_{D^*}}|\boldsymbol{p}^*_{D^*}|V(q^2),$$

$$H_0(q^2) = \frac{-1}{2m_{D^*}\sqrt{q^2}}\Big[\frac{4m_B^2|\boldsymbol{p}^*_{D^*}|^2}{m_B + m_{D^*}}A_2(q^2) - (m_B^2 - m_{D^*}^2 - q^2)(m_B + m_{D^*})A_1(q^2)\Big],$$

$$H_s(q^2) = \frac{2m_B|\boldsymbol{p}^*_{D^*}|}{\sqrt{q^2}}A_0(q^2). \tag{10}$$

In this analysis, we use an HQET-based parametrizations for the FFs that is expressed in terms of the scalar product of the B and D^* four-velocities

$$w \equiv v_B \cdot v_{D^*} = \frac{m_B^2 + m_{D^*}^2 - q^2}{2m_{D^*}m_B}. \tag{11}$$

Its minimum value $w_{\min} = 1$ corresponds to $q^2_{\max} = (m_B - m_{D^*})^2$. The maximum value is obtained for the lowest possible value of q^2, which is the square of the mass of the lepton. Thus, $w_{\max} = 1.35$ for $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decays and $w_{\max} = 1.51$ for $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ decays.

In this framework, the FFs are usually expressed in terms of a universal form factor $h_{A_1}(w)$ and ratios $R_i(w)$:

$$A_1(w) = \frac{w+1}{2}r_{D^*}h_{A_1}(w), \qquad A_0(w) = \frac{R_0(w)}{r_{D^*}}h_{A_1}(w),$$

$$A_2(w) = \frac{R_2(w)}{r_{D^*}}h_{A_1}(w), \qquad V(w) = \frac{R_1(w)}{r_{D^*}}h_{A_1}(w),$$

where $r_{D^*} = 2\sqrt{m_B m_{D^*}}/(m_B + m_{D^*})$. Using dispersion relations and analyticity constraints [9,19], the universal FF and the ratios can be expressed in terms of just five parameters:

$$h_{A_1}(w) = h_{A_1}(1)[1 - 8\rho^2_{D^*}z(w) + (53\rho^2_{D^*} - 15)z(w)^2 - (231\rho^2_{D^*} - 91)z(w)^3],$$

$$R_1(w) = R_1(1) - 0.12(w-1) + 0.05(w-1)^2,$$

$$R_2(w) = R_2(1) + 0.11(w-1) - 0.06(w-1)^2,$$

$$R_0(w) = R_0(1) - 0.11(w-1) + 0.01(w-1)^2.$$

Here, $z(w) = (\sqrt{w+1} - \sqrt{2})/(\sqrt{w+1} + \sqrt{2})$. The factor $h_{A_1}(1)$ only affects the overall normalization, so it cancels in the ratio $\mathcal{R}(D^*)$.

Three of the remaining four FF parameters, $R_1(1)$, $R_2(1)$, and $\rho^2_{D^*}$, have been measured in analyses of $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ decays. The most recent averages by the Heavy Flavor Averaging Group (HFAG) [4] and their correlations C are

$$\rho^2_{D^*} = 1.207 \pm 0.028, \qquad C(\rho^2_{D^*}, R_1(1)) = 0.566,$$

$$R_1(1) = 1.401 \pm 0.033, \qquad C(\rho^2_{D^*}, R_2(1)) = -0.807,$$

$$R_2(1) = 0.854 \pm 0.020, \qquad C(R_1(1), R_2(1)) = -0.758.$$

$R_0(w)$ affects the decay rate only via the scalar hadronic amplitude $H_s(q^2)$. The corresponding leptonic amplitude $L_s(q^2, \theta_\tau)$ is helicity suppressed, i.e., its rate is proportional to the mass of the lepton (Eq. (6)). As a result, $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ decays are not sensitive to this FF, and $R_0(w)$ has not been measured. We therefore rely on a theoretical estimate, $R_0(1) = 1.14 \pm 0.07$, based on HQET [9].

2. Form factor parametrization of $\bar{B} \to D\tau^-\bar{\nu}_\tau$ decays

The nonperturbative QCD interactions in $\bar{B} \to D\tau^-\bar{\nu}_\tau$ decays are described by two independent FFs, referred to as V_1 and S_1 [8]. The helicity amplitudes take the form:

$$H_0(w) = \sqrt{m_B m_D}\frac{m_B + m_D}{\sqrt{q^2(w)}}\sqrt{w^2 - 1}V_1(w), \tag{12}$$

$$H_s(w) = \sqrt{m_B m_D}\frac{m_B - m_D}{\sqrt{q^2(w)}}(w+1)S_1(w). \quad (13)$$

The amplitudes corresponding to the helicities $\lambda_W = \pm$ vanish because the D meson has spin 0. For this decay mode, the variable w is defined as in Eq. (11), except that the D^* meson mass is replaced by the D meson mass m_D.

Taking into account dispersion relations [19], V_1 can be expressed as

$$V_1(w) = V_1(1) \times [1 - 8\rho_D^2 z(w) + (51\rho_D^2 - 10)z(w)^2 - (252\rho_D^2 - 84)z(w)^3], \quad (14)$$

where $V_1(1)$ and ρ_D^2 are FF parameters. The normalization $V_1(1)$ cancels in the ratio $\mathcal{R}(D)$. Based on $\bar{B} \to D\ell^-\bar{\nu}_\ell$ decays, the average value of the shape parameter is $\rho_D^2 = 1.186 \pm 0.055$ [4]. As for $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decays, the scalar hadronic amplitude is helicity suppressed and as a result, $S_1(w)$ cannot be measured with $\bar{B} \to D\ell^-\bar{\nu}_\ell$ decays. We use instead the following estimate based on HQET [8]:

$$S_1(w) = V_1(w)\{1 + \Delta[-0.019 + 0.041(w-1) - 0.015(w-1)^2]\}, \quad (15)$$

with $\Delta = 1 \pm 1$.

We have employed this FF parametrization to generate $\bar{B} \to D\tau^-\bar{\nu}_\tau$ and $\bar{B} \to D\ell^-\bar{\nu}_\ell$ decays, as described in Sec. III C 2. Though we used the same FF definitions and parameters, we found a difference of 1% between the value of $\mathcal{R}(D)$ that we obtained by integrating Eq. (9) and the value quoted in Ref. [8].

On the other hand, if we adopt the FF parameters of Ref. [20], we perfectly reproduce the $\mathcal{R}(D)$ predictions presented there. The translation of the FF parametrization of Ref. [20] into standard hadronic amplitudes is not straightforward, so we do not use these FFs in the Monte Carlo simulation. Since both parametrizations yield essentially identical q^2 spectra, they are equivalent with respect to Monte Carlo generation, which is not sensitive to differences in normalization.

3. SM calculation of $\mathcal{R}(D^{(*)})$ and q^2 spectrum

We determine the SM predictions for the ratios $\mathcal{R}(D^{(*)})$ integrating the expression for the differential decay rate (Eq. (9)) as follows:

$$\mathcal{R}(D^{(*)}) \equiv \frac{\mathcal{B}(B \to D^{(*)}\tau\nu)}{\mathcal{B}(B \to D^{(*)}\ell\nu)} = \frac{\int_{m_\tau^2}^{q^2_{\rm max}} \frac{d\Gamma_\tau}{dq^2}\, dq^2}{\int_{m_\ell^2}^{q^2_{\rm max}} \frac{d\Gamma_\ell}{dq^2}\, dq^2}, \quad (16)$$

with $q^2_{\rm max} = (m_B - m_{D^{(*)}})^2$.

The uncertainty of this calculation is determined by generating one million random sets of values for all the FF parameters assuming Gaussian distributions for the uncertainties and including their correlations. We calculate $\mathcal{R}(D^{(*)})$ with each set of values and assign the root mean square (RMS) of its distribution as the uncertainty.

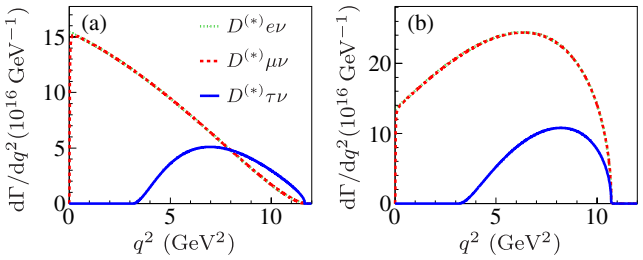

FIG. 2 (color online). Predicted q^2 spectra for (a) $\bar{B} \to D\tau^-\bar{\nu}_\tau$ and $\bar{B} \to D\ell^-\bar{\nu}_\ell$ decays for $V_1(1)V_{cb} = 0.0427$ and (b) $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ and $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ decays for $h_{A_1}(1)V_{cb} = 0.0359$ [4].

We apply this procedure for B^0 and B^- decays, and for $\ell = e$ and μ, and average the four results to arrive at the following predictions,

$$\mathcal{R}(D)_{\rm SM} = 0.297 \pm 0.017, \quad (17)$$

$$\mathcal{R}(D^*)_{\rm SM} = 0.252 \pm 0.003. \quad (18)$$

Additional uncertainties that have not been taken into account could contribute at the percent level. For instance, some electromagnetic corrections could affect $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ and $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays differently [9]. The experimental uncertainty on $\mathcal{R}(D^{(*)})$ is expected to be considerably larger.

The q^2 spectra for $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays in Fig. 2 clearly show the threshold at $q^2_{\rm min} = m_\tau^2$, while for $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays $q^2_{\rm min} \sim 0$. We take advantage of this difference in the signal selection by imposing $q^2 > 4\ \mathrm{GeV}^2$. The spectra for $\ell = e$ and μ are almost identical, except for $q^2 < m_\mu^2 = 0.011\ \mathrm{GeV}^2$.

B. Two-Higgs-doublet models

As we noted in the introduction, $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays are potentially sensitive to new physics (NP) processes. In this paper, we study their sensitivity to new scalar particles such as charged Higgs bosons. The effective Hamiltonian of the $b \to c\ell^-\bar{\nu}_\ell$ transition in the context of general spin-0 contributions [21] is

$$\mathcal{H}_{\rm eff} = \frac{4G_F V_{cb}}{\sqrt{2}}[(\bar{c}\gamma_\mu P_L b)(\bar{\tau}\gamma^\mu P_L \nu_\tau) + S_L(\bar{c}P_L b) \times (\bar{\tau}P_L\nu_\tau) + S_R(\bar{c}P_R b)(\bar{\tau}P_L\nu_\tau)], \quad (19)$$

where S_L and S_R are independent complex parameters, and $P_{L,R} \equiv (1 \mp \gamma_5)/2$. For $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays and large Higgs boson masses, this Hamiltonian describes the most general type of two-Higgs-doublet model (2HDM), sometimes referred to as 2HDM of type III.

The contributions of charged Higgs bosons to $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays can be encapsulated in the scalar helicity amplitude in the following way [5,20]:

$$H_s^{\mathrm{2HDM}} \approx H_s^{\mathrm{SM}} \times \left(1 + (S_R \pm S_L)\frac{q^2}{m_\tau(m_b \mp m_c)}\right). \quad (20)$$

Here, m_b and m_c are the masses of the b and c quarks at the b-quark mass scale, $m_b = 4.20 \pm 0.07$ GeV and $m_c = 1.25 \pm 0.09$ GeV [22]. The top sign in Eq. (20) applies to $\bar{B} \to D\tau^-\bar{\nu}_\tau$ decays and the bottom sign applies to $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decays.

The dependence of the ratios $\mathcal{R}(D^{(*)})$ on the parameters S_R and S_L can be studied by substituting H_s^{2HDM} for H_s^{SM} in Eq. (9). Given that charged Higgs bosons are not expected to contribute significantly to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays, $\mathcal{R}(D^{(*)})_{\mathrm{2HDM}}$ can be described by the following parabolas

$$\mathcal{R}(D) = \mathcal{R}(D)_{\mathrm{SM}} + A'_D \mathrm{Re}(S_R + S_L) + B'_D |S_R + S_L|^2, \quad (21)$$

$$\mathcal{R}(D^*) = \mathcal{R}(D^*)_{\mathrm{SM}} + A'_{D^*} \mathrm{Re}(S_R - S_L) + B'_{D^*} |S_R - S_L|^2. \quad (22)$$

The sign difference arises because $\bar{B} \to D\tau^-\bar{\nu}_\tau$ decays probe scalar operators, while $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decays are sensitive to pseudoscalar operators.

Of particular interest is the 2HDM of type II, which describes the Higgs sector of the minimal supersymmetric model at tree level. In this model, one of the two Higgs doublets couples to up-type quarks, while the other doublet couples to down-type quarks and leptons. The type II 2HDM corresponds to the subset of the type III 2HDM parameter space for which $S_R = -m_b m_\tau \tan^2\beta/m_{H^\pm}^2$ and $S_L = 0$. Here, $\tan\beta$ is the ratio of the vacuum expectation values of the two Higgs doublets and $m_{H^\pm}$ is the mass of the charged Higgs.

In the context of the 2HDM type II, the NP dependence of $\mathcal{R}(D^{(*)})$ is expressed as

$$\mathcal{R}(D^{(*)})_{\mathrm{2HDM}} = \mathcal{R}(D^{(*)})_{\mathrm{SM}} + A_{D^{(*)}}\frac{\tan^2\beta}{m_{H^+}^2} + B_{D^{(*)}}\frac{\tan^4\beta}{m_{H^+}^4}. \quad (23)$$

This expression is accurate to 1% for $m_{H^\pm}$ larger than 15 GeV. The region for $m_{H^\pm} \leq 15$ GeV has already been excluded by $B \to X_s\gamma$ measurements [23].

Table I lists the values of $A_{D^{(*)}}$ and $B_{D^{(*)}}$, which are determined by averaging over B^0 and B^- decays. The uncertainty estimation includes the uncertainties on the mass ratio m_c/m_b and the FF parameters, as well as their correlations. The parameters $A'_{D^{(*)}}$ and $B'_{D^{(*)}}$ in Eqs. (21) and (22) are related to $A_{D^{(*)}}$ and $B_{D^{(*)}}$ by

$$A'_{D^{(*)}} = \frac{-A_{D^{(*)}}}{m_\tau m_b}, \qquad B'_{D^{(*)}} = \frac{B_{D^{(*)}}}{m_\tau^2 m_b^2}.$$

Due to the destructive interference between the SM and 2HDM amplitudes in Eq. (20), charged Higgs contributions depress the ratios $\mathcal{R}(D^{(*)})$ for low values of $\tan\beta/m_{H^\pm}$. For larger values of $\tan\beta/m_{H^\pm}$, the Higgs contributions dominate and $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ increase rapidly. As the coefficients of Table I show, the 2HDM impact is expected to be larger for $\mathcal{R}(D)$ than for $\mathcal{R}(D^*)$. This is because charged Higgs contributions only affect the scalar amplitude H_s^{2HDM}, but $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decays also receive contributions from $H_\pm$, diluting the effect on the total rate.

TABLE I. Dependence of $\mathcal{R}(D^{(*)})$ on $\tan\beta/m_{H^\pm}$ in the 2HDM according to Eq. (23) for $\bar{B} \to D\tau^-\bar{\nu}_\tau$ and $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decays: the values of $\mathcal{R}(D^{(*)})$, the parameters A and B with their uncertainties, and correlations C.

	$\bar{B} \to D\tau^-\bar{\nu}_\tau$	$\bar{B} \to D^*\tau^-\bar{\nu}_\tau$
$\mathcal{R}(D^{(*)})_{\mathrm{SM}}$	0.297 ± 0.017	0.252 ± 0.003
$A_{D^{(*)}}$ (GeV^2)	-3.25 ± 0.32	-0.230 ± 0.029
$B_{D^{(*)}}$ (GeV^4)	16.9 ± 2.0	0.643 ± 0.085
$C(\mathcal{R}(D^{(*)})_{\mathrm{SM}}, A_{D^{(*)}})$	-0.928	-0.946
$C(\mathcal{R}(D^{(*)})_{\mathrm{SM}}, B_{D^{(*)}})$	0.789	0.904
$C(A_{D^{(*)}}, B_{D^{(*)}})$	-0.957	-0.985

Figure 3 shows the impact of the 2HDM on the q^2 spectrum. Given that the B and D mesons have spin $J = 0$, the SM decays $B \to DW^* \to D\tau\nu$ proceed via P-wave for $J_{W^*} = 1$, and via S-wave for $J_{W^*} = 0$. For the P-wave decay, which accounts for about 96% of the total amplitude, the decay rate receives an additional factor $|\boldsymbol{p}_D^*|^2$, which suppresses the q^2 spectrum at high values. Since charged Higgs bosons have $J_H = 0$, their contributions proceed via S-wave, and, thus, have a larger average q^2 than the SM contributions. As a result, for low values of $\tan\beta/m_{H^\pm}$ where the negative interference depresses H_s^{2HDM}, the q^2 spectrum shifts to lower values. For large values of $\tan\beta/m_{H^\pm}$, the Higgs contributions dominate the decay rate and the average q^2 significantly exceeds that of the SM.

The situation is different for $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decays because the D^* meson has spin $J_{D^*} = 1$. The SM decays can proceed via S, P, or D-waves, while the decay via an

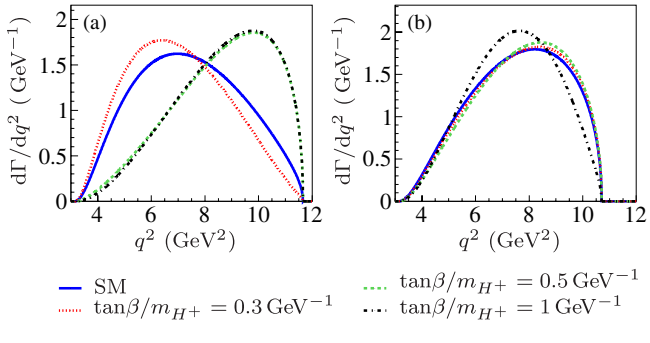

FIG. 3 (color online). Predicted q^2 distributions for (a) $\bar{B} \to D\tau^-\bar{\nu}_\tau$ and (b) $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decays for different values of $\tan\beta/m_{H^\pm}$. All curves are normalized to unit area.

intermediate Higgs boson must proceed via P-wave, suppressing the rate at high q^2.

When searching for charged Higgs contributions, it is important to account for the changes in the q^2 spectrum. This distribution has a significant impact on the analysis due to the close relation between q^2 and $m^2_{\rm miss}$, one of the fit variables.

Charged Higgs contributions also affect the $|\boldsymbol{p}^*_\ell|$ distribution. Given the spin 0 of the Higgs boson and the positive helicity (right-handedness) of the antineutrino, the decays $H^- \to \tau^- \bar{\nu}_\tau$ always produce τ^- leptons with positive helicities ($\lambda_\tau = +$). As a result, the fraction of right-handed τ^- leptons produced in $\bar{B} \to D\tau^- \bar{\nu}_\tau$ and $\bar{B} \to D^*\tau^- \bar{\nu}_\tau$ decays changes from 67% and 23% in the SM, respectively [2], to close to 100% when the 2HDM contributions dominate.

The lepton spectrum of polarized $\tau^\pm \to \ell^\pm \nu_\ell \nu_\tau$ decays is well known [24]. For τ^- leptons with $\lambda_{\tau^-} = -$, the ℓ^- is emitted preferentially in the τ^- direction, while the opposite is true for positive helicities. In the B rest frame, leptons of a certain momentum in the τ^- rest frame have larger momentum if they are emitted in the direction of the τ^- momentum than in the opposite direction. As a result, the $|\boldsymbol{p}^*_\ell|$ spectrum for SM decays is harder than for Higgs dominated decays. For low values of $\tan\beta/m_{H^\pm}$ for which the destructive interference depresses the $\bar{B} \to D^{(*)}\tau^- \bar{\nu}_\tau$ rate, the proportion of left-handed τ^- leptons increases, and therefore, the $|\boldsymbol{p}^*_\ell|$ spectrum is harder than in the SM.

III. DATA SAMPLE, DETECTOR AND SIMULATION

A. Data sample

This analysis is based on the full data sample recorded with the *BABAR* detector [25] at the PEP-II energy-asymmetric e^+e^- storage rings [26]. It operated at a center-of-mass (c.m.) energy of 10.58 GeV, equal to the mass of the $\Upsilon(4S)$ resonance. This resonance decays almost exclusively to $B\bar{B}$ pairs. The collected data sample of 471 million $\Upsilon(4S) \to B\bar{B}$ events (on-peak data), corresponds to an integrated luminosity of 426 fb^{-1} [27]. To study continuum background, an additional sample of 40 fb^{-1} (off-peak data) was recorded approximately 40 MeV below the $\Upsilon(4S)$ resonance, i.e., below the threshold for $B\bar{B}$ production.

B. The *BABAR* detector and single particle reconstruction

The *BABAR* detector and event reconstruction have been described in detail elsewhere [25]. The momentum and angles of charged particles were measured in a tracking system consisting of a 5-layer, double-sided silicon-strip detector (SVT) and a 40-layer, small-cell drift chamber (DCH) filled with a helium-isobutane gas mixture. Charged particles of different masses were distinguished by their ionization energy loss in the tracking devices and by a ring-imaging Cerenkov detector (DIRC). A finely segmented CsI(Tl) calorimeter (EMC) measured the energy and position of electromagnetic showers generated by electrons and photons. The EMC was surrounded by a superconducting solenoid providing a 1.5-T magnetic field and by a segmented flux return with a hexagonal barrel section and two endcaps. The steel of the flux return was instrumented (IFR) with resistive plate chambers and limited streamer tubes to detect particles penetrating the magnet coil and steel.

Within the polar angle acceptance of the SVT and DCH ($0.4 < \theta_{lab} < 2.6$) the efficiency for the reconstruction of charged particles exceeds 99% for momenta above 1 GeV. For low momentum pions, especially from $D^{*+} \to D^0\pi^+$ decays, the efficiency drops to about 90% at 0.4 GeV and to 50% at 0.1 GeV.

The electron and muon identification efficiencies and the probabilities to misidentify a pion, a kaon, or a proton as an electron or a muon are measured as a function of the laboratory momentum and angles using high-purity data samples.

Electrons are separated from charged hadrons primarily on the basis of the ratio of the energy deposited in the EMC to the track momentum. A special algorithm has been developed to identify photons from bremsstrahlung in the inner detector, and to correct the electron momentum for the energy loss. Within the polar angle acceptance, the average electron efficiency for laboratory momenta above 0.5 GeV is 97%, largely independent of momentum. The average pion misidentification rate is less than 0.5%.

Muon identification relies on a new multivariate algorithm that significantly increases the reconstruction efficiency at low muon momenta, $|\boldsymbol{p}_\mu| < 1$ GeV. This algorithm combines information on the measured DCH track, the track segments in the IFR, and the energy deposited in the EMC. The average muon efficiency is close to 90% independent of momentum, except in the forward endcap, where it decreases for laboratory momenta below 1 GeV. The average pion misidentification rate is about 2% above 1.2 GeV, rising at lower momenta and reaching a maximum of 9% at 0.8 GeV.

By choosing a fairly loose selection of charged leptons and taking advantage of improved PID algorithms, we increased the lepton efficiencies by 6% for electrons and 50% for muons compared to the previous *BABAR* analysis [14].

Charged kaons are identified up to 4 GeV on the basis of information from the DIRC, SVT, and DCH. The efficiency exceeds 80% over most of the momentum range and varies with polar angle. The probability that a pion is misidentified as a kaon is close to 2%, varying by about 1% as a function of momentum and polar angle.

The decays $K^0_S \to \pi^+\pi^-$ are reconstructed as pairs of tracks of opposite charge originating from a displaced

vertex. The invariant mass of the pair $m_{\pi\pi}$ is required to be in the range $m_{\pi\pi} \in [0.491, 0.506]$ GeV. No attempt is made to identify interactions of K_L^0 in the EMC or IFR.

To remove beam-generated background in the EMC and electronic noise, photon candidates are required to have a minimum energy of 30 MeV and a shower shape that is consistent with that of an electromagnetic shower. Neutral pions are reconstructed from pairs of photon candidates with an invariant mass in the range $m_{\gamma\gamma} \in [120, 150]$ MeV.

C. Monte Carlo simulation

1. Simulated samples

This analysis relies on Monte Carlo (MC) techniques to simulate the production and decay of continuum and $B\bar{B}$ events. The simulation is based on the EvtGen generator [28]. The $q\bar{q}$ fragmentation is performed by Jetset [29], and the detector response by Geant4 [30]. Radiative effects such as bremsstrahlung in the detector material and initial-state and final-state radiation [31] are included.

We derive predictions for the distributions and efficiencies of the signal and backgrounds from the simulation. The size of the simulated sample of generic $B\bar{B}$ events exceeds that of the $B\bar{B}$ data sample by about a factor of ten, while the sample for $q\bar{q}$ events corresponds to twice the size of the off-peak data sample. We assume that the $\Upsilon(4S)$ resonance decays exclusively to $B\bar{B}$ pairs and use recent measurements of branching fractions [12] for all produced particles. The impact of their uncertainties on the final results is assessed as a systematic uncertainty.

Information extracted from studies of selected data control samples is used to improve the accuracy of the simulation. Specifically, we reweight simulated events to account for small differences observed in comparisons of data and simulation (Sec. V).

2. Implementation of the form factor parametrizations

For reasons of simplicity, the simulation of $\bar{B} \to D\ell^-\bar{\nu}_\ell$ and $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays is based on the ISGW2 model [32], and $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ decays are generated using an HQET-based parametrization [33]. A change to a different FF parametrization is implemented by reweighting the generated events with the weights

$$w_{\rm HQET}(q^2, \theta_i) = \left(\frac{\mathcal{M}(q^2, \theta_i)_{\rm HQET}}{\mathcal{M}(q^2, \theta_i)_{\rm MC}}\right)^2 \times \frac{\mathcal{B}_{\rm MC}}{\mathcal{B}_{\rm HQET}}. \quad (24)$$

Here, $\mathcal{M}(q^2, \theta_i)_{\rm HQET}$ refers to the matrix element for the FF parametrizations described in Secs. II A 1 and II A 2, and $\mathcal{M}(q^2, \theta_i)_{\rm MC}$ is the matrix element employed in the MC generation. The matrix element of decays involving the scalar D meson depends on one angular variable, the lepton helicity angle θ_ℓ, with $\ell = e, \mu, \tau$. In addition to θ_ℓ, the matrix element of decays involving the vector meson D^* is sensitive to two additional angular variables describing the D^* decay. The ratio of the branching fractions $\mathcal{B}_{\rm MC}/\mathcal{B}_{\rm HQET}$ ensures that the sum of all weights equals the number of generated events.

In the SM, this reweighting results in a small shift of the q^2 distribution to higher values, while the changes in the helicity angle θ_τ and the τ polarization are negligible. Therefore, the distributions of the secondary charged lepton are not affected.

In the presence of a charged Higgs boson, however, the τ polarization can change substantially, affecting the momentum of the secondary lepton ℓ originating from the $\tau \to \ell\nu_\ell\nu_\tau$ decays. We account for the potential presence of a charged Higgs of 2HDM type II by reweighting the simulation with the following weights,

$$w_{\rm 2HDM}(q^2, \theta_i, |\boldsymbol{p}_\ell^*|) = \left(\frac{\mathcal{M}(q^2, \theta_i)_{\rm 2HDM}}{\mathcal{M}(q^2, \theta_i)_{\rm MC}}\right)^2 \times \frac{\Gamma(|\boldsymbol{p}_\ell^*|)_{\rm 2HDM}}{\Gamma(|\boldsymbol{p}_\ell^*|)_{\rm MC}} \times \frac{\mathcal{B}_{\rm MC}}{\mathcal{B}_{\rm 2HDM}}, \quad (25)$$

where θ_i refers again to the angular variables. The second factor represents the ratio of the $|\boldsymbol{p}_\ell^*|$ distributions $\Gamma(|\boldsymbol{p}_\ell^*|)$ in the 2HDM parametrization and in the MC simulation. This factorization is necessary because in the MC generation the polarization is handled in a probabilistic manner, so it cannot be corrected on an event-per-event basis. It is only applicable if $|\boldsymbol{p}_\ell^*|$ is uncorrelated with q^2 and the angular variables, which is largely the case. In some regions of phase space, the 2HDM weights have a much larger dispersion than the weights applied in the SM reweighting, leading to larger statistical uncertainties for the simulation of the Higgs boson contributions.

3. Simulation of $\bar{B} \to D^{**}(\tau^-/\ell^-)\bar{\nu}$ decays

By D^{**} we refer to excited charm resonances heavier than the D^* meson. We include in the simulation the $\bar{B} \to D^{**}\tau^-\bar{\nu}_\tau$ and $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ decays that involve the four D^{**} states with $L = 1$ that have been measured [4]. This simulation takes into account their helicities [34] and the following decay modes: D_0^*, $D_2^* \to D\pi$ and D_1', D_1, $D_2^* \to D^*\pi$. Three-body decays $D^{**} \to D^{(*)}\pi\pi$ are not included in the nominal fit for lack of reliable measurements.

To estimate the rate of $B \to D^{**}\tau\nu_\tau$ decays, we rely on ratios of the available phase space Φ,

$$\mathcal{R}(D^{**}) \equiv \frac{\mathcal{B}(\bar{B} \to D^{**}\tau^-\bar{\nu}_\tau)}{\mathcal{B}(\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell)} \approx \frac{\Phi(\bar{B} \to D^{**}\tau^-\bar{\nu}_\tau)}{\Phi(\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell)}. \quad (26)$$

The value of this ratio depends on the mass of the D^{**} state involved in the $\bar{B} \to D^{**}(\tau^-/\ell^-)\bar{\nu}$ decay. We use the largest of the four possible choices, $\mathcal{R}(D^{**}) = 0.18$.

Possible contributions from nonresonant $\bar{B} \to D^{(*)}\pi(\pi)\ell^-\bar{\nu}_\ell$ decays and semileptonic decays involving higher-mass excited charm mesons are not included in the nominal fit, and will be treated as a systematic uncertainty.

IV. EVENT SELECTION

The event selection proceeds in two steps. First, we select $B\bar{B}$ events in which one of the B mesons, the $B_{\rm tag}$, is fully reconstructed in a hadronic decay, while the other B meson decays semileptonically. To increase the event selection efficiency compared to earlier analyses, we have added more decay chains to the $B_{\rm tag}$ selection and have chosen a looser charged lepton selection. This leads to significantly higher backgrounds, primarily combinatorial background from $B\bar{B}$ and continuum events, and charge-crossfeed events. Charge-crossfeed events are $\bar{B} \to D^{(*)}(\tau^-/\ell^-)\bar{\nu}$ decays in which the charge of the reconstructed $B_{\rm tag}$ and $D^{(*)}$ mesons are wrong, primarily because of an incorrectly assigned low-momentum $\pi^{\pm}$.

Semileptonic decays to higher mass charm mesons have a signature similar to that of signal events and their composition is not well measured. This background is fitted in selected control samples that are enriched with these decays.

As the second step in the event selection, we introduce kinematic criteria that increase the fraction of selected signal events with respect to normalization and background decays. We also apply a multivariate algorithm to further improve the signal-to-background ratio.

A. Selection of events with a $B_{\rm tag}$ and a semileptonic B decay

$\Upsilon(4S) \to B\bar{B}$ events are tagged by the hadronic decay of one of the B mesons. We use a semiexclusive algorithm which includes additional $B_{\rm tag}$ decay chains and enhances the efficiency by a factor of 2 compared to the earlier version employed by *BABAR* [14]. We look for decays of the type $B_{\rm tag} \to SX^{\pm}$, where S refers to a *seed* meson and $X^{\pm}$ is a charged state comprising of up to five hadrons, pions or kaons, among them up to two neutral mesons, π^0 or K^0_S. The seed mesons, D, D^*, D_s, D_s^*, and J/ψ, are reconstructed in 56 decay modes. As a result, the $B_{\rm tag}$ is reconstructed in 1,680 different decay chains, which are further subdivided into 2,968 kinematic modes.

To isolate the true tag decays from combinatorial background, we use two kinematic variables: the energy substituted mass $m_{\rm ES} = \sqrt{E^2_{\rm beam} - \mathbf{p}^2_{\rm tag}}$ and the energy difference $\Delta E = E_{\rm tag} - E_{\rm beam}$. Here $\mathbf{p}_{\rm tag}$ and $E_{\rm tag}$ refer to the c.m. momentum and energy of the $B_{\rm tag}$, and $E_{\rm beam}$ is the c.m. energy of a single beam particle. These variables make optimum use of the precisely known energies of the colliding beams. For correctly reconstructed B decays, the $m_{\rm ES}$ distribution is centered at the B-meson mass with a resolution of 2.5 MeV, while ΔE is centered at zero with a resolution of 18 MeV which is dominated by the detector resolution. We require $m_{\rm ES} > 5.27$ GeV and $|\Delta E| < 0.072$ GeV.

For each $B_{\rm tag}$ candidate in a selected event, we look for the signature of the semileptonic decay of the second B meson, a D or D^* meson and a charged lepton ℓ. 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. We require all charged particles to be associated with the $B_{\rm tag}D^{(*)}\ell$ candidate, but we allow for any number of additional photons in the event.

The laboratory momentum of the electron or muon is required to exceed 300 MeV or 200 MeV, respectively. For D mesons, we reconstruct the following decay modes: $D^0 \to K^-\pi^+$, K^-K^+, $K^-\pi^+\pi^0$, $K^-\pi^+\pi^-\pi^+$, $K^0_S\pi^+\pi^-$, and $D^+ \to K^-\pi^+\pi^+$, $K^-\pi^+\pi^+\pi^0$, $K^0_S\pi^+$, $K^0_S\pi^+\pi^+\pi^-$, $K^0_S\pi^+\pi^0$, $K^0_SK^+$, with $K^0_S \to \pi^+\pi^-$. The reconstructed invariant mass of D candidates is required to be consistent with the nominal D mass to within four standard deviations (σ). The combined reconstructed branching fractions are 35.8% and 27.3% for D^0 and D^+, respectively. We identify D^* mesons by their decays $D^{*+} \to D^0\pi^+$, $D^+\pi^0$, and $D^{*0} \to D^0\pi^0$, $D^0\gamma$. For these decays, the c.m. momentum of the pion or the c.m. energy of the photon are required to be less than 400 MeV. Furthermore, the mass difference $\Delta m = m(D^*) - m(D)$ is required to differ by less than 4σ from the expected value [12].

To further reduce the combinatorial background, we perform a kinematic fit to the event, constraining tracks of secondary charged particles to the appropriate B, $D^{(*)}$, or K^0_S decay vertices. The fit also constrains the reconstructed masses of the D, D^*, and K^0_S mesons to their nominal values. The vertex of the $\Upsilon(4S) \to B\bar{B}$ decay has to be compatible with a beam-beam interaction. Candidates for which this fit does not converge are rejected. The $m^2_{\rm miss}$ resolution improves by about 25% and becomes more symmetric for the remaining candidates.

To select a single $B\bar{B}$ candidate, we determine $E_{\rm extra} = \sum_i E^{\gamma}_i$, the sum of the energies of all photons that are not associated with the reconstructed $B\bar{B}$ pair. We only include photons of more than 50 MeV, thereby eliminating about 99% of the beam-generated background. We retain the candidate with the lowest value of $E_{\rm extra}$, and if more than one candidate survives, we select the one with the smallest $|\Delta E|$. This procedure preferentially selects $D^*\ell$ candidates over $D\ell$ candidates. Thus, we reduce the fraction of misreconstructed events with a $D^* \to D(\pi/\gamma)$ decay for which the pion or photon is not properly assigned to the D^* meson.

As a consequence of the rather loose lepton selection criteria and the addition of decay modes with multiple neutral pions and K^0_S for the $B_{\rm tag}$ selection, the number of $B_{\rm tag}D^{(*)}\ell$ candidates per event is very large. To address this problem, we identify the $B_{\rm tag}$ decay modes that contribute primarily to the combinatorial background. Specifically, we determine for each of the 2,968 kinematic modes $R_{\rm tc}$,

the fraction of events for which all charged particles in the $B_{\rm tag}$ final state are correctly reconstructed and associated with the tag decay. This assessment is based on a large sample of simulated $B\bar{B}$ events equivalent to 700 fb^{-1}. We observe that for decay chains with low multiplicity final states and no neutral hadrons the signal-to-background ratio (S/B) is very high. For instance, for the $B^-_{\rm tag} \to J/\psi(\to \mu^+\mu^-)K^-$ decay, we obtain $S/B = 316/79$, whereas for the decay $B^0_{\rm tag} \to D^-(\to K^0_S\pi^-)\pi^+\pi^+\pi^-\pi^-$ this ratio is $S/B = 20/145$. For this decay mode, typically 3.5 of the 8 $B_{\rm tag}$ final state particles are incorrectly associated with the second B decay in the event or otherwise misidentified. Based on this study, we only retain $B_{\rm tag}$ decay chains with $R_{\rm tc} > 0.3$. With this criterion, we remove 2100 $B_{\rm tag}$ kinematic modes, eliminate 2/3 of the combinatorial background, and retain 85% of the signal $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays. Thanks to this procedure, the average number of candidates per event before single candidate selection is reduced to 1.8 for the $D^0\ell$ and $D^+\ell$ samples, and 3.1 and 4.8 for the $D^{*0}\ell$ and $D^{*+}\ell$ samples, respectively.

B. Selection of the $D^{(*)}\pi^0\ell$ control samples

To constrain the $\bar{B} \to D^{**}(\tau^-/\ell^-)\bar{\nu}$ background, we select four $D^{(*)}\pi^0\ell$ control samples, identical to the $D^{(*)}\ell$ samples except for an additional reconstructed π^0. The π^0 is selected in the mass range $m_{\gamma\gamma} \in [120, 150]$ MeV. Decays of the form $B \to D^{(*)}\pi\ell\nu$ peak at $m^2_{\rm miss} = 0$ in these samples. As a result, we can extract their yields together with the signal and normalization yields by fitting the $D^{(*)}\ell$ and $D^{(*)}\pi^0\ell$ samples simultaneously.

More than half of the events in these control samples originate from continuum $e^+e^- \to q\bar{q}(\gamma)$ events. Since the fragmentation of light quarks leads to a two-jet event topology, this background is very effectively suppressed by the requirement $|\cos\Delta\theta_{\rm thrust}| < 0.8$, where $\Delta\theta_{\rm thrust}$ is the angle between the thrust axes of the $B_{\rm tag}$ and of the rest of the event. Since B mesons originating from $\Upsilon(4S)$ decays are produced just above threshold, their final state particles are emitted almost isotropically, and, therefore, the $\cos\Delta\theta_{\rm thrust}$ distribution is uniform. As a result, the loss of $\bar{B} \to D^{**}(\tau^-/\ell^-)\bar{\nu}$ decays due to this restriction is significantly smaller than the amount of continuum events rejected.

C. Optimization of the signal selection

We introduce criteria that discriminate signal from background, and also differentiate between signal $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ and $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays. For semileptonic decays the minimum momentum transfer is largely determined by the mass of the charged lepton. For decays involving τ leptons, $q^2_{\rm min} = m^2_\tau \simeq 3.16$ GeV^2. Thus the selection $q^2 > 4$ GeV^2 retains 98% of the $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays and rejects more than 30% of the $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays. The event sample with $q^2 < 4$ GeV^2 is dominated by $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ and serves as a very clean data sample for comparisons with the MC simulation. To reject background from hadronic B decays in which a pion is misidentified as muon, we require $|\boldsymbol{p}_{\rm miss}| > 200$ MeV, where $|\boldsymbol{p}_{\rm miss}|$ is the missing momentum in the c.m. frame.

To further improve the separation of well-reconstructed signal and normalization decays from various backgrounds, we employ a boosted decision tree (BDT) multivariate method [35]. This method relies on simple classifiers which determine signal and background regions by using binary selections on various input distributions. For each of the four $D^{(*)}\ell$ samples, we train a BDT to select signal and normalization events and reject

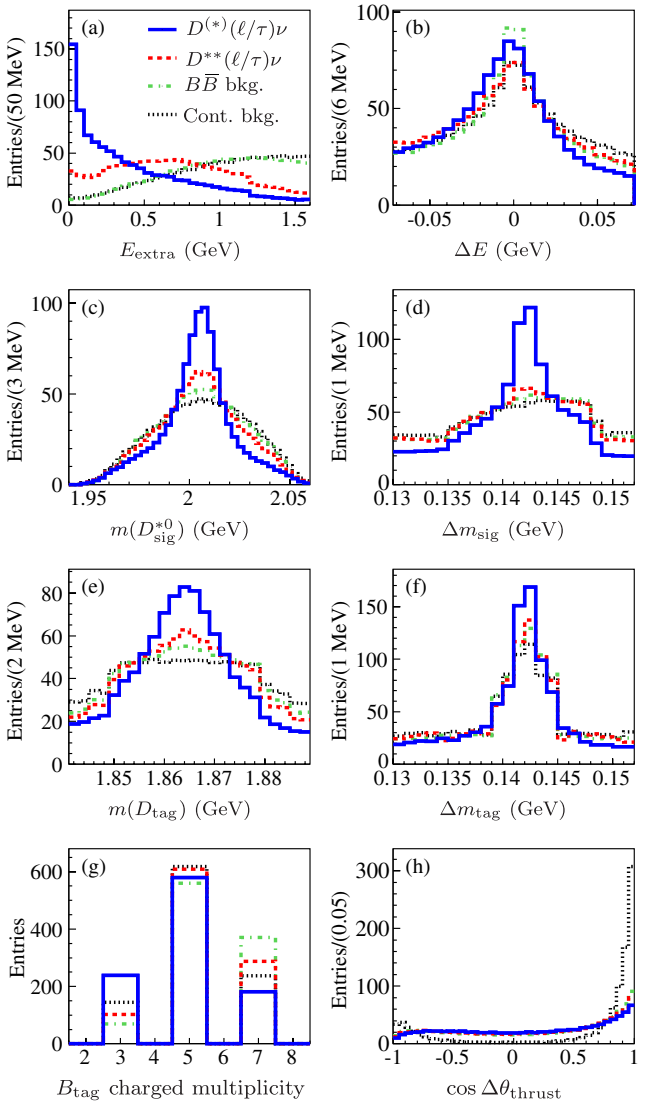

FIG. 4 (color online). Input variables for the BDT selector trained on the $D^{*0}\ell$ sample. Histograms are normalized to 1000 entries.

$\bar{B} \to D^{**}(\tau^-/\ell^-)\bar{\nu}$ and charge cross-feed backgrounds. Each BDT selector relies on the simulated distributions of the following variables: (a) $E_{\rm extra}$; (b) ΔE; (c) the reconstructed mass of the signal $D^{(*)}$ meson; (d) the mass difference for the reconstructed signal D^*: $\Delta m = m(D\pi) - m(D)$; (e) the reconstructed mass of the seed meson of the $B_{\rm tag}$; (f) the mass difference for a D^* originating from the $B_{\rm tag}$, $\Delta m_{\rm tag} = m(D_{\rm tag}\pi) - m(D_{\rm tag})$; (g) the charged particle multiplicity of the $B_{\rm tag}$ candidate; and (h) $\cos\Delta\theta_{\rm thrust}$. The input distributions for one of the BDT selectors are shown in Fig. 4. For the $D^{(*)}\pi^0\ell$ samples, we use similar BDT selectors that are trained to reject continuum, $D^{(*)}(\ell/\tau)\nu$, and other $B\bar{B}$ background. After the BDT requirements are applied, the fraction of events attributed to signal in the $m^2_{\rm miss} > 1.5$ GeV^2 region, which excludes most of the normalization decays, increases from 2% to 39%. The background remaining in that region is composed of normalization events (10%), continuum (19%), $D^{**}(\ell/\tau)\nu$ events (13%), and other $B\bar{B}$ events (19%), primarily from $B \to D^{(*)}D_s^{(*)+}$ decays with $D_s^+ \to \tau^+\nu_\tau$.

V. CORRECTION AND VALIDATION OF THE MC SIMULATION

The simulation of the full reconstruction of high-multiplicity events, including the veto of events with extra tracks or higher values of $E_{\rm extra}$ is a rather challenging task. To validate the simulation, we compare simulated distributions with data control samples, and, when necessary, correct the MC simulations for the observed differences. The figures shown in this section combine events from all four channels ($D^0\ell$, $D^{*0}\ell$, $D^+\ell$, and $D^{*+}\ell$); the observed differences are similar in the individual samples.

The control samples are selected to have little or no contamination from signal decays. Specifically we select,

(i) *Continuum events:* off-peak data.

(ii) *Normalization decays:* $q^2 \le 4$ GeV^2.

(iii) *Combinatorial $B\bar{B}$ and continuum backgrounds:* $5.20 < m_{\rm ES} < 5.26$ GeV.

(iv) *Incorrectly reconstructed events:* events in three $E_{\rm extra}$ intervals, *high* ($1.2 < E_{\rm extra} < 2.4$ GeV), *intermediate* ($0.5 < E_{\rm extra} < 1.2$ GeV), and *low* ($E_{\rm extra} < 0.5$ GeV for events that fail the BDT selection). N. B. the BDT selection results in the elimination of all events with $E_{\rm extra} > 0.4$ GeV.

The off-peak data sample confirms the $m^2_{\rm miss}$ distribution of simulated continuum events, but shows discrepancies in the $|\boldsymbol{p}^*_\ell|$ spectrum and overall normalization of the simulation [Fig. 5(a)]. These features are also observed in other control samples, such as on-peak data with high $E_{\rm extra}$ [Fig. 5(b)]. We correct the simulated $|\boldsymbol{p}^*_\ell|$ spectrum and yield of the continuum contribution by reweighting it to match off-peak data, on an event-by-event basis. After this correction, the $|\boldsymbol{p}^*_\ell|$ distributions of the expected backgrounds agree well in independent control samples down to low lepton momenta where the misidentification rates are significant [Fig. 5(c)]. We observe that in the high $E_{\rm extra}$

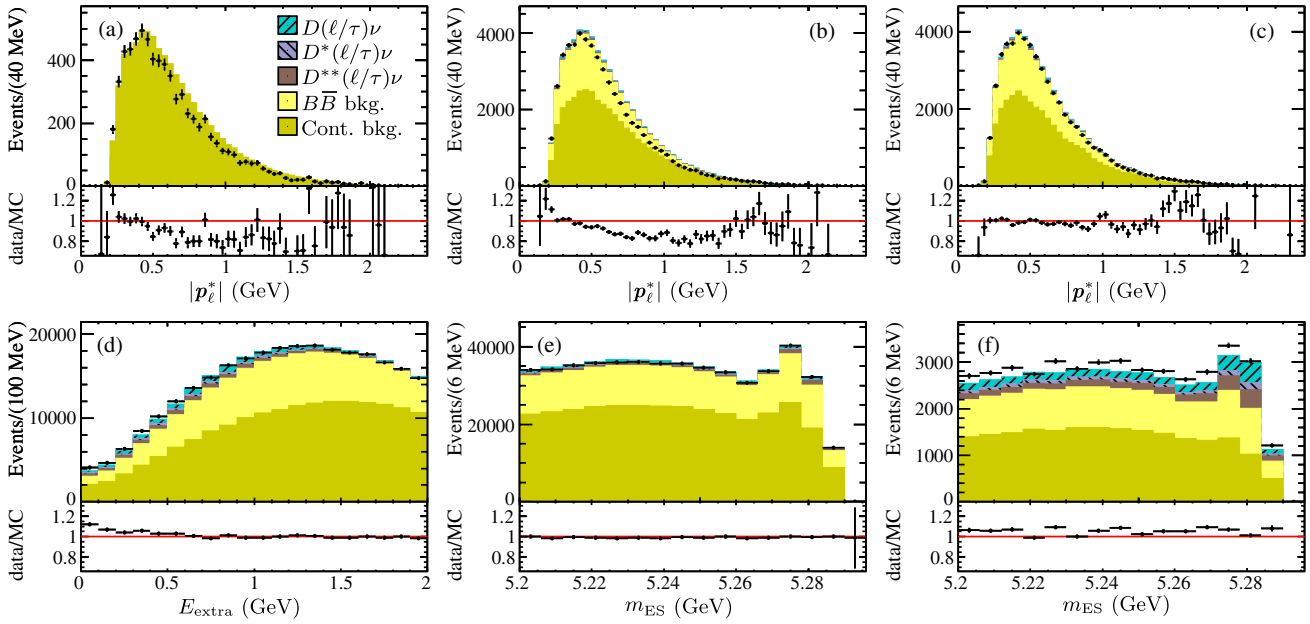

FIG. 5 (color online). Comparison of data control samples (data points) with MC simulated samples (histograms) of the $|\boldsymbol{p}^*_\ell|$ distributions for (a) off-peak data prior to $|\boldsymbol{p}^*_\ell|$ reweighting, (b) the intermediate $E_{\rm extra}$ sample prior to $|\boldsymbol{p}^*_\ell|$ reweighting, and (c) the intermediate $E_{\rm extra}$ sample after $|\boldsymbol{p}^*_\ell|$ reweighting; (d) the $E_{\rm extra}$ distribution for the combinatorial background; and the $m_{\rm ES}$ distributions for (e) the intermediate $E_{\rm extra}$ sample, and (f) the low $E_{\rm extra}$ sample. The results are shown for the four $D^{(*)}\ell$ samples combined.

region, the simulation exceeds data yield by $(1.3 \pm 0.5)\%$. This small excess is corrected by decreasing the expected $B\bar{B}$ background yield by $(4.3 \pm 1.9)\%$. After this correction, the simulation provides accurate yield predictions for the backgrounds at intermediate and high $E_{\rm extra}$. For instance, the ratio of the expected to observed yield of events with $m^2_{\rm miss} > 1.5\ {\rm GeV}^2$ is 0.998 ± 0.006. The $m^2_{\rm miss}$ distributions of the continuum and $B\bar{B}$ backgrounds are described well in all control samples.

The region of low $E_{\rm extra}$, which includes the signal region, is more difficult to model, primarily due to low energy photons and K^0_L mesons interacting in the EMC. Figure 5(d) shows that the data in the $m_{\rm ES}$ sideband agree well with the combinatorial background predictions for $E_{\rm extra} > 0.5$ GeV, but are underestimated for low $E_{\rm extra}$. This, and small differences in the other BDT input distributions, result in a underestimation of the combinatorial background when the BDT requirements are applied. Based on the $5.20 < m_{\rm ES} < 5.26$ GeV sideband, we find scale factors of 1.099 ± 0.019 and 1.047 ± 0.034 for the combinatorial background in the $D\ell$ and $D^*\ell$ samples, respectively. The uncertainties are given by the statistics of the data and simulated samples. The ratio of the observed to the expected number of events is independent of $m_{\rm ES}$ [Figs. 5(e) and 5(f)], so we apply these corrections to the continuum and $B\bar{B}$ backgrounds in the signal region. The same correction is applied to $\bar{B} \to D^{**}(\tau^-/\ell^-)\bar{\nu}$ events, which cannot be easily isolated, because their simulated $E_{\rm extra}$ distributions are very similar to those of combinatorial background. These corrections affect the fixed $B\bar{B}$ and continuum yields in the fit, as well as the relative efficiency of $\bar{B} \to D^{**}(\tau^-/\ell^-)\bar{\nu}$ events in the $D^{(*)}\ell$ and $D^{(*)}\pi^0\ell$ samples. As a result, these corrections are the source of the dominant systematic uncertainties.

Relying on the $q^2 \leq 4\ {\rm GeV}^2$ control sample, where $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays account for 96% of the events, we correct the $E_{\rm extra}$ distribution and an 8.5% overestimation of the simulated normalization events. We apply the same correction to simulated signal events which are expected to have a similar $E_{\rm extra}$ distribution. This procedure does not affect the relative efficiency of signal to normalization events, so it has a very small impact on the $\mathcal{R}(D^{(*)})$ measurements.

We use the same $q^2 \leq 4\ {\rm GeV}^2$ control sample to compare and validate the $|\boldsymbol{p}^*_\ell|$ distributions of $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ events. We observe that the $m^2_{\rm miss}$ resolution of the narrow peaks at $m^2_{\rm miss} = 0$ is slightly underestimated by the simulation. This effect is corrected by convolving the simulated distributions with a Gaussian resolution function, for which the width is adjusted by iteration.

VI. FIT PROCEDURE AND RESULTS

A. Overview

We extract the signal and normalization yields from an extended, unbinned maximum-likelihood fit to two-dimensional $m^2_{\rm miss}$-$|\boldsymbol{p}^*_\ell|$ distributions. The fit is performed simultaneously to the four $D^{(*)}\ell$ samples and the four $D^{(*)}\pi^0\ell$ samples. The distribution of each $D^{(*)}\ell$ and $D^{(*)}\pi^0\ell$ sample is fit to the sum of eight or six contributions, respectively. Each of the $4 \times 8 + 4 \times 6 = 56$ contributions is described by a probability density function (PDF). Their relative scale factor determines the number of events from each source. Tables II and III summarize the contributions to the fit for the four $D^{(*)}\ell$ samples and the four $D^{(*)}\pi^0\ell$ samples. These tables also list the relative yield for each contribution as estimated from MC simulation (for SM signal), and specify whether the yield is free, fixed, or constrained in the fit.

We introduce the following notation to uniquely identify each contribution to the fit: source $\Rightarrow$ sample. For instance, $D^{*0}\tau\nu \Rightarrow D^{*0}\ell$ refers to signal $D^{*0}\tau\nu$ decays that are correctly reconstructed in the $D^{*0}\ell$ sample, while $D^{*0}\tau\nu \Rightarrow D^0\ell$ refers to the same decays, but incorrectly reconstructed in the $D^0\ell$ sample. We refer to the latter as

TABLE II. Contributions to the four $D^{(*)}\ell$ samples. The expected relative abundance of events in each data sample is represented by $f_{\rm exp}$. The columns labeled Yield indicate whether the contribution is free in the fit, fixed, or linked to another component through a cross-feed constraint. The charged cross-feed components, marked with Fix./It., are fixed in the fit, but updated in the iterative process.

	$D^0\ell$		$D^{*0}\ell$		$D^+\ell$		$D^{*+}\ell$	
Source	$f_{\rm exp}$ (%)	Yield	$f_{\rm exp}$ (%)	Yield	$f_{\rm exp}$ (%)	Yield	$f_{\rm exp}$ (%)	Yield
$D^{(*)}\tau\nu$ signal	2.6	Free	4.9	Free	4.3	Free	5.0	Free
$D^{(*)}\tau\nu$ signal feed-down/up	2.8	$D^{*0}\ell$	0.4	$D^0\ell$	1.8	$D^{*+}\ell$	0.1	$D^+\ell$
$D^{(*)}\ell\nu$ normalization	24.5	Free	80.7	Free	37.3	Free	88.0	Free
$D^{(*)}\ell\nu$ norm. feed-down/up	53.5	Free	2.7	$D^0\ell$	35.0	Free	0.3	$D^+\ell$
$D^{**}(\ell/\tau)\nu$ background	4.3	$D^0\pi^0\ell$	3.6	$D^{*0}\pi^0\ell$	6.6	$D^+\pi^0\ell$	3.0	$D^{*+}\pi^0\ell$
Cross-feed background	3.8	Fix./It.	1.3	Fix./It.	2.1	Fix./It.	0.4	Fix./It.
$B\bar{B}$ background	4.1	Fixed	3.7	Fixed	7.1	Fixed	2.8	Fixed
Continuum background	4.4	Fixed	2.6	Fixed	5.9	Fixed	0.5	Fixed

TABLE III. Contributions to the four $D^{(*)}\pi^0\ell$ samples. The expected relative abundance of events in each data sample is represented by $f_{\rm exp}$. The columns labeled Yield indicate whether the contribution is free in the fit, fixed, or linked to another component through a cross-feed constraint. The $D(\ell/\tau)\nu$ components are linked to the $D^*(\ell/\tau)\nu$ components, and the cross-feed constraint is updated in the iteration. The charged cross-feed components, marked with Fix./It., are fixed in the fit, but updated in the iterative process.

	$D^0\pi^0\ell$		$D^{*0}\pi^0\ell$		$D^+\pi^0\ell$		$D^{*+}\pi^0\ell$	
Source	$f_{\rm exp}$ (%)	Yield	$f_{\rm exp}$ (%)	Yield	$f_{\rm exp}$ (%)	Yield	$f_{\rm exp}$ (%)	Yield
$D^{**}(\ell/\tau)\nu$ background	20.1	Free	16.4	Free	19.9	Free	22.1	Free
$D^*(\ell/\tau)\nu$ feed-up	19.1	Free	20.6	Free	10.0	Free	25.2	Free
$D(\ell/\tau)\nu$ feed-up	6.4	$D^0\pi^0\ell$	2.3	$D^{*0}\pi^0\ell$	4.7	$D^+\pi^0\ell$	0.8	$D^{*+}\pi^0\ell$
Cross-feed background	4.9	Fix./It.	3.6	Fix./It.	4.4	Fix./It.	2.5	Fix./It.
$B\bar{B}$ background	28.4	Free	36.4	Free	38.7	Free	37.4	Free
Continuum background	21.0	Fixed	20.8	Fixed	22.2	Fixed	12.0	Fixed

feed-down. Contributions of the form $D(\tau/\ell)\nu \Rightarrow D^*(\tau/\ell)$ and $D^{(*)}(\tau/\ell)\nu \Rightarrow D^{**}(\tau/\ell)$ are referred to as feed-up.

The contributions from the continuum, $B\bar{B}$, and cross-feed backgrounds, with the exception of $B\bar{B}$ background in the $D^{(*)}\pi^0\ell$ samples, are fixed to the yields determined by MC simulation after small adjustments based on data control regions. The yields of the remaining 36 contributions are determined in the fit. Some of these contributions share the same source and therefore the ratio of their yields is constrained to the expected value, e.g., $D^{*0}\tau\nu \Rightarrow D^{*0}\ell$ and $D^{*0}\tau\nu \Rightarrow D^0\ell$. Of special importance are the constraints linking the $D^{**}(\ell/\tau)\nu$ yields in the $D^{(*)}\ell$ samples ($N_{D^{**}\Rightarrow D^{(*)}}$) to the yields in the $D^{(*)}\pi^0\ell$ samples ($N_{D^{**}\Rightarrow D^{(*)}\pi^0}$),

$$f_{D^{**}} = \frac{N_{D^{**}\Rightarrow D^{(*)}}}{N_{D^{**}\Rightarrow D^{(*)}\pi^0}} = \frac{\varepsilon_{D^{**}\Rightarrow D^{(*)}}}{\varepsilon_{D^{**}\Rightarrow D^{(*)}\pi^0}}. \quad (27)$$

Given that these constraints share the same source, $f_{D^{**}}$ is equivalent to the ratio of the $D^{**}(\ell/\tau)\nu$ reconstruction efficiencies for the two samples.

Taking into account the constraints imposed on event yields from a common source, there are 22 free parameters in the standard fit, as listed in Table IV. In addition, we perform a fit in which we impose the isospin relations $\mathcal{R}(D^0) = \mathcal{R}(D^+) \equiv \mathcal{R}(D)$ and $\mathcal{R}(D^{*0}) = \mathcal{R}(D^{*+}) \equiv \mathcal{R}(D^*)$. We choose not to impose isospin relations for the $D^{(*)}\pi^0\ell$ samples. Consequently, this fit has a total of 17 free parameters.

TABLE IV. Number of free parameters in the isospin-unconstrained ($N_{\rm un}$) and constrained ($N_{\rm cons}$) fits.

Sample	Contribution	$N_{\rm un}$	$N_{\rm cons}$
$D^{(*)}\ell$	$D^{(*)}\tau\nu$ signal	4	2
$D^{(*)}\ell$	$D^{(*)}\ell\nu$ normalization	4	2
$D^{(*)}\ell$	$D^*\ell\nu$ norm. feed-down	2	1
$D^{(*)}\pi^0\ell$	$D^{**}(\ell/\tau)\nu$ background	4	4
$D^{(*)}\pi^0\ell$	$D^{(*)}\ell\nu$ norm. feed-up	4	4
$D^{(*)}\pi^0\ell$	$B\bar{B}$ background	4	4

The following inputs are updated by iterating the fit:

(i) The eight $D^{(*)}(\ell/\tau)\nu \Rightarrow D^{(*)}\pi^0\ell$ PDFs are recalculated taking into account the fitted $D^{(*)}\ell\nu$ and $D^{(*)}\tau\nu$ contributions to the $D^{(*)}\ell$ samples.

(ii) The fixed charge cross-feed yields are updated based on the deviation of the fitted $D^{(*)}\ell\nu$ yields from the expected values.

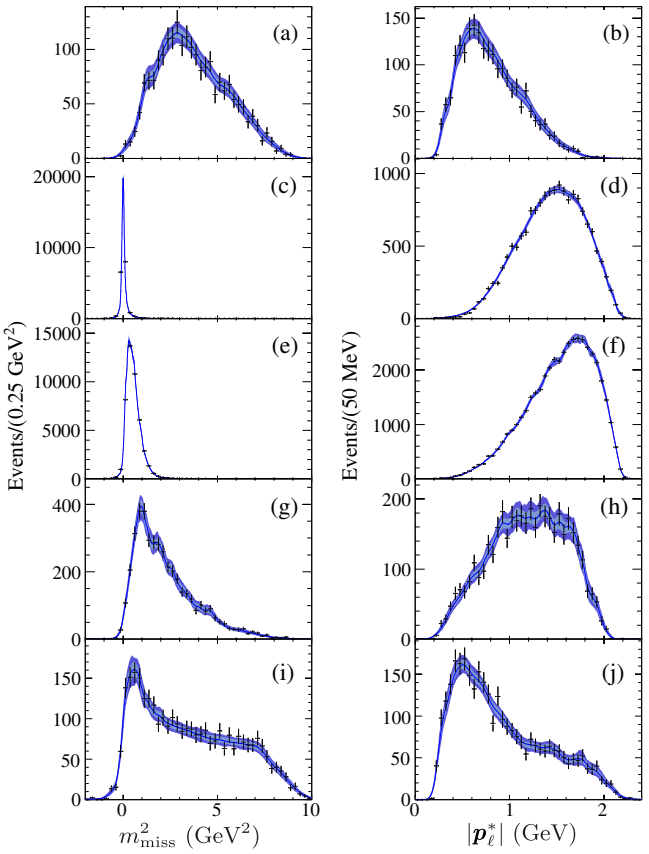

FIG. 6 (color online). Projections of the simulated $m^2_{\rm miss}$ and $|\boldsymbol{p}^*_\ell|$ distributions and the PDFs for the following contributions to the $D^0\ell$ sample: (a), (b) $D^0\tau\nu$; (c), (d) $D^0\ell\nu$; (e), (f) $D^{*0}\ell\nu$; (g), (h) $D^{**}(\ell/\tau)\nu$, and (i), (j) $B\bar{B}$ background. The light and dark blue (gray) bands mark the 1σ and 2σ envelopes of the variations of the PDF projections due to their statistical uncertainty.

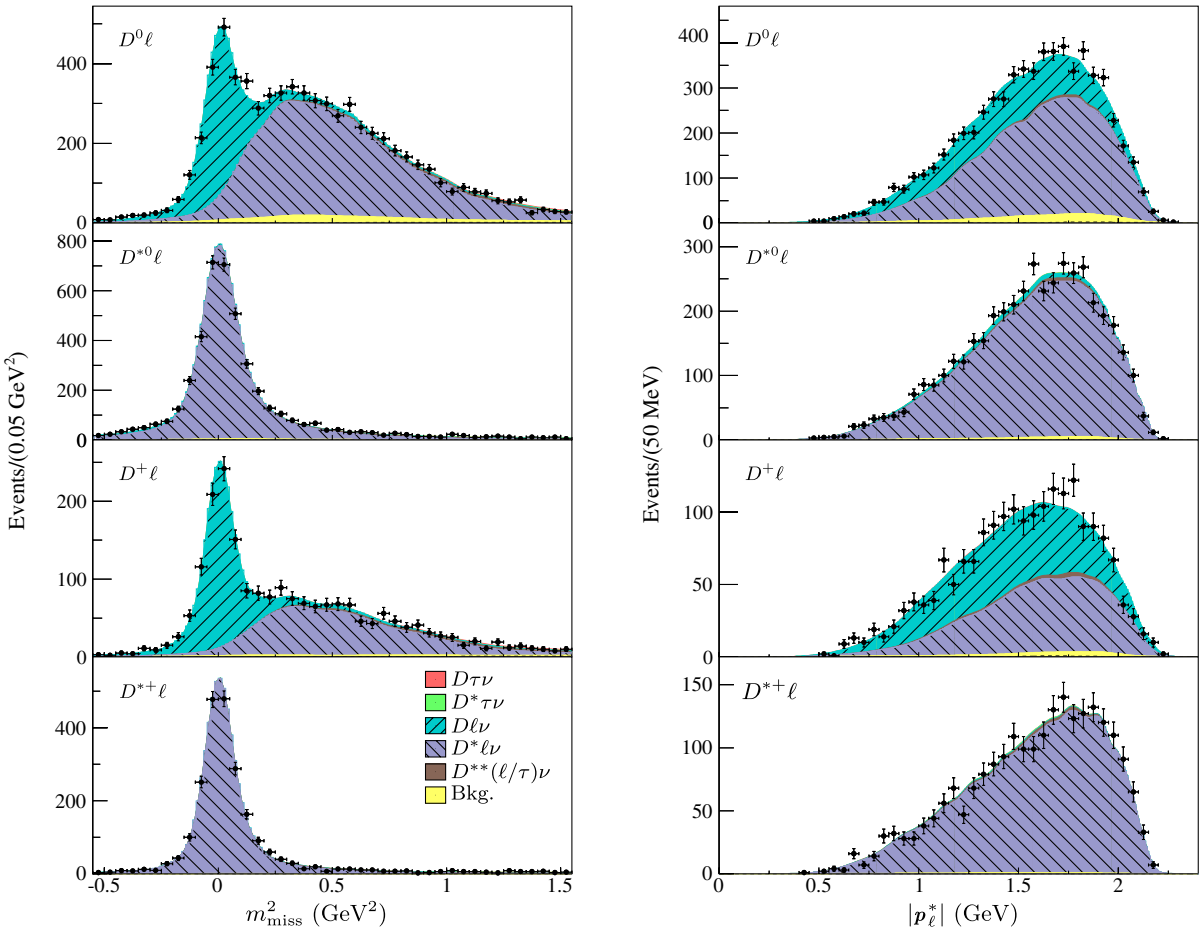

FIG. 7 (color online). Comparison of the $m^2_{\rm miss}$ and $|\boldsymbol{p}^*_\ell|$ distributions of the $D^{(*)}\ell$ samples (data points) with the projections of the results of the isospin-unconstrained fit (stacked colored distributions). The $|\boldsymbol{p}^*_\ell|$ distributions show the normalization-enriched region with $m^2_{\rm miss} < 1$ GeV^2, thus excluding most of the signal events in these samples.

(iii) The continuum, $B\bar{B}$, and $D^{**}(\ell/\tau)\nu$ background corrections are recalculated. They have a slight dependence on the fitted $D^{(*)}\ell\nu$ events because some of these events extend into the $m_{\rm ES}$ sideband.

(iv) The correction to the $m^2_{\rm miss}$ resolution of the normalization contributions is readjusted.

(v) The two feed-down constraints for $D^*\tau\nu$ are updated using the fitted feed-down constraints for the normalization contributions in the following way:

$$\left.\frac{N_{D^*\tau\nu\Rightarrow D\ell}}{N_{D^*\tau\nu\Rightarrow D^*\ell}}\right|_{\rm Iter} = \left.\frac{N_{D^*\tau\nu\Rightarrow D\ell}}{N_{D^*\tau\nu\Rightarrow D^*\ell}}\right|_{\rm MC} \times \left.\frac{N_{D^*\ell\nu\Rightarrow D\ell}}{N_{D^*\ell\nu\Rightarrow D^*\ell}}\right|_{\rm Fit} \left.\frac{N_{D^*\ell\nu\Rightarrow D^*\ell}}{N_{D^*\ell\nu\Rightarrow D\ell}}\right|_{\rm MC}. \quad (28)$$

The iterations continue until the change on the values of $\mathcal{R}(D^{(*)})$ is less than 0.01%. The update of the feed-down rates has a significant impact on the fits to the D^0 and D^+ samples because of the large signal feed-down. The other iterative updates have only a marginal impact.

B. Probability density functions and validation

The fit relies on 56 PDFs, which are derived from MC samples of continuum and $B\bar{B}$ events equivalent to 2 and 9 times the size of the data sample, respectively. The two-dimensional $m^2_{\rm miss}$-$|\boldsymbol{p}^*_\ell|$ distributions for each of the 56 contributions to the fit are described by smooth nonparametric kernel estimators [36]. These estimators enter a two-dimensional Gaussian function centered at the $m^2_{\rm miss}$ and $|\boldsymbol{p}^*_\ell|$ values of each simulated event. The width of the Gaussian function determines the smoothness of the PDF. We find the optimum level of global smoothing with a cross-validation algorithm [37]. For PDFs that have variations in shape that require more than one level of smoothing, we combine estimators with different Gaussian widths in up to four areas in the $m^2_{\rm miss}$-$|\boldsymbol{p}^*_\ell|$ space. For instance, we use different levels of smoothing in the $D^{*0}\ell\nu \Rightarrow D^{*0}\ell$ contribution for the narrow peak at $m^2_{\rm miss} = 0$ and the smooth $m^2_{\rm miss}$ tail that extends up to 7 GeV^2. Figure 6 shows one-dimensional projections of five two-dimensional PDFs. The bands indicate the statistical uncertainty on the PDFs estimated with a bootstrap algorithm [37].

The $m^2_{\rm miss}$ distributions of signal and normalization are very distinct due to the different number of neutrinos in the final state. The $m^2_{\rm miss}$ distributions of the backgrounds resemble those of the signal, and therefore these contributions to the fit are either fixed or constrained by the $D^{(*)}\pi^0\ell$ samples.

To validate the PDFs and the fit procedure, we divide the large sample of simulated $B\bar{B}$ events into two: sample A with about 3.3×10^9 $B\bar{B}$ events, and sample B with 9.4×10^8 $B\bar{B}$ events. We determine the PDFs with sample A, and create histograms by integrating the PDFs in bins of their $m^2_{\rm miss}$ and $|\boldsymbol{p}^*_\ell|$ projections. We compare the resulting histograms with the events in sample A, and derive a χ^2 based on the statistical significance of the difference for each bin. The distribution of the corresponding p values for these PDFs is uniform, as expected for an unbiased estimation. As another test, we extract the signal and normalization yields from fits to the events of sample B, using the PDFs obtained from sample A. Again, the results are compatible with an unbiased fit. Furthermore, we validate the fit procedure based on a large number of pseudoexperiments generated from these PDFs. Fits to these samples also show no bias in the extracted signal and normalization yields.

C. Fit results

Figures 7 and 8 show the $m^2_{\rm miss}$ and $|\boldsymbol{p}^*_\ell|$ projections of the fits to the $D^{(*)}\ell$ samples. In Fig. 7, the $|\boldsymbol{p}^*_\ell|$ projections do not include events with $m^2_{\rm miss} > 1$ GeV^2, i.e., most of the signal events. In Fig. 8, the vertical scale is enlarged and the horizontal axis is extended for the $m^2_{\rm miss}$ projection to reveal the signal and background contributions. The $|\boldsymbol{p}^*_\ell|$ projections emphasize the signal events by excluding events with $m^2_{\rm miss} < 1$ GeV^2. Both figures demonstrate that the fit describes the data well and the observed differences are consistent with the statistical and systematic uncertainties on the PDFs and the background contributions.

Figure 9 shows the $m^2_{\rm miss}$ and $|\boldsymbol{p}^*_\ell|$ projections of the fit to the four $D^{(*)}\pi^0\ell$ samples. The narrow $m^2_{\rm miss}$ peak is

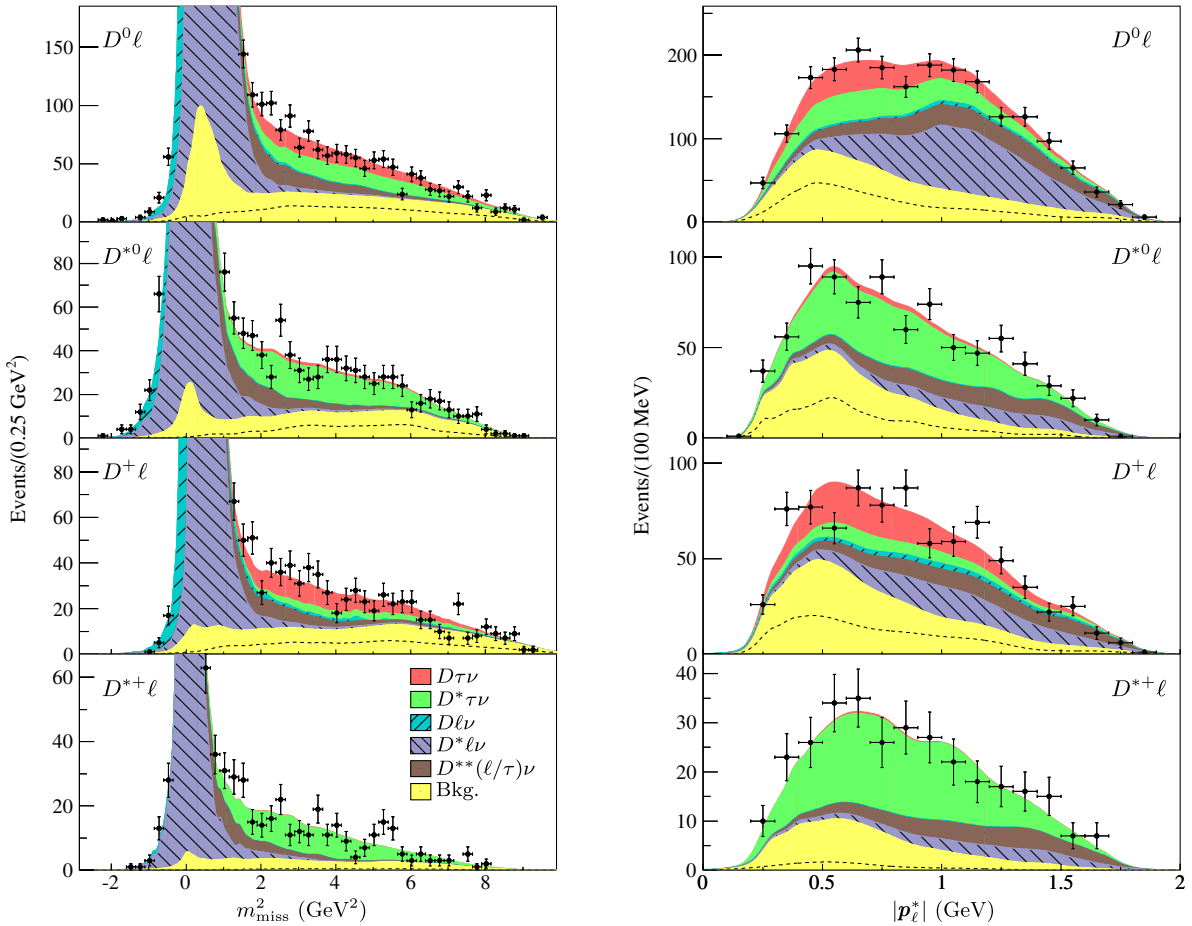

FIG. 8 (color online). Comparison of the $m^2_{\rm miss}$ and $|\boldsymbol{p}^*_\ell|$ distributions of the $D^{(*)}\ell$ samples (data points) with the projections of the results of the isospin-unconstrained fit (stacked colored distributions). The region above the dashed line of the background component corresponds to $B\bar{B}$ background and the region below corresponds to continuum. The peak at $m^2_{\rm miss} = 0$ in the background component is due to charge cross-feed events. The $|\boldsymbol{p}^*_\ell|$ distributions show the signal-enriched region with $m^2_{\rm miss} \geq 1$ GeV^2, thus excluding most of the normalization events in these samples.

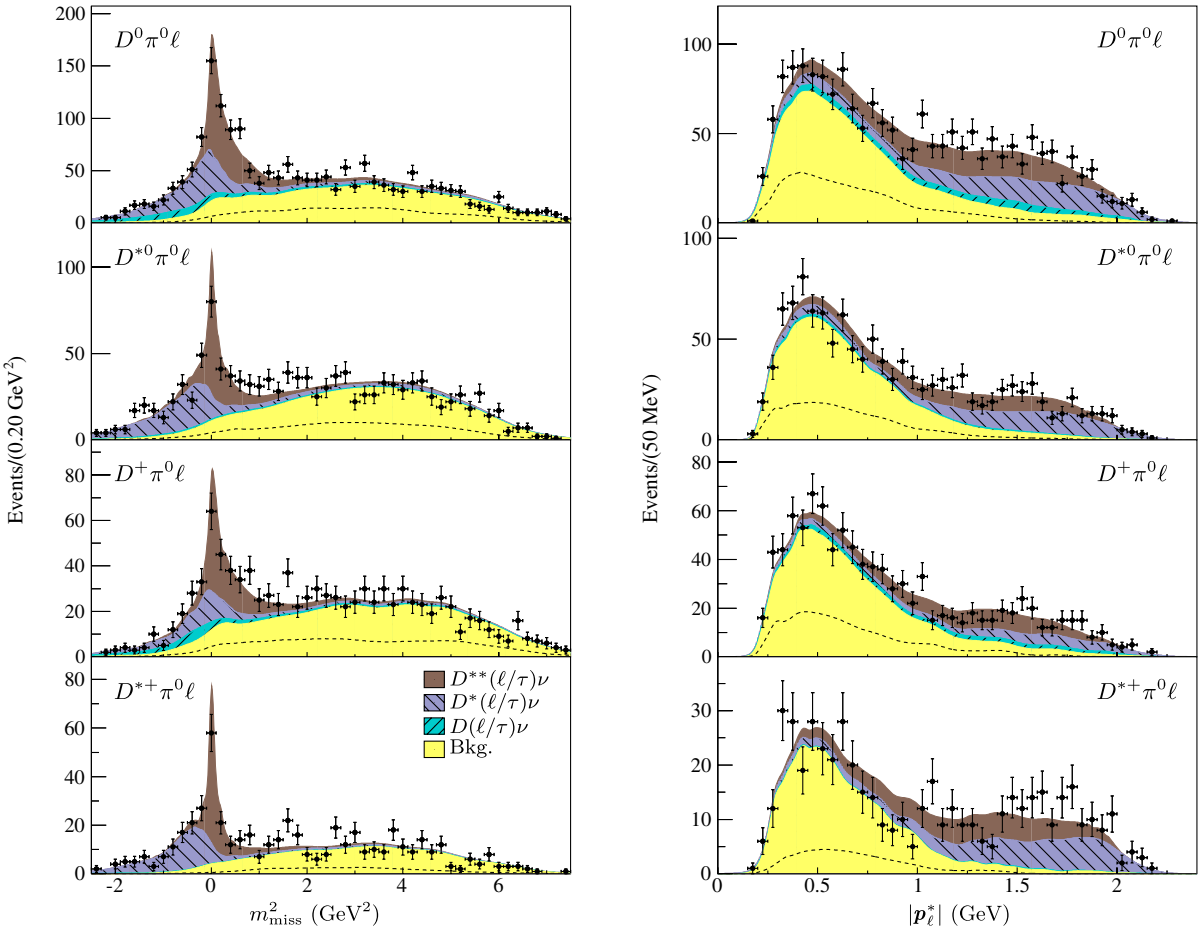

FIG. 9 (color online). Comparison of the $m^2_{\rm miss}$ and $|\boldsymbol{p}^*_\ell|$ distributions of the $D^{(*)}\pi^0\ell$ samples (data points) with the projections of the results of the isospin-unconstrained fit (stacked colored distributions). The region above the dashed line of the background component corresponds to $B\bar{B}$ background and the region below corresponds to continuum.

described well by the fit. It tightly constrains contributions from $B \to D^{(*)}\pi\ell\nu$ decays, including the nonresonant $D^{(*)}\pi$ states as well as decays of D^{**} states, narrow or wide. There appears to be a small excess of events in the data for $1 < m^2_{\rm miss} < 2$ GeV^2. This might be an indication for an underestimation of the $D^{**}(\ell/\tau)\nu$ background. The impact of this effect is assessed as a systematic uncertainty.

The fit determines, for each signal decay mode, the number of signal events in the data sample, $N_{\rm sig}$, and the corresponding number of normalization events, $N_{\rm norm}$. We derive the ratios of branching fractions as

$$\mathcal{R}(D^{(*)}) = \frac{N_{\rm sig}}{N_{\rm norm}} \frac{\varepsilon_{\rm norm}}{\varepsilon_{\rm sig}}, \tag{29}$$

where $\varepsilon_{\rm sig}/\varepsilon_{\rm norm}$ is the ratio of efficiencies (including the $\tau^\pm$ branching fractions) taken from MC simulation. These relative efficiencies are larger for $\mathcal{R}(D)$ than for $\mathcal{R}(D^*)$, because the $q^2 > 4$ GeV^2 requirement rejects a larger fraction of $\bar{B} \to D\ell^-\bar{\nu}_\ell$ decays than of $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ decays, while keeping almost 100% of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays.

The results of the fits in terms of the number of events, the efficiency ratios, and $\mathcal{R}(D^{(*)})$ are listed in Table VIII, for both the standard and the isospin-constrained fits. Due to the large signal feed-down, there are significant negative correlations between the fits to the $D\ell$ and $D^*\ell$ samples. The statistical correlations are -0.59 for $\mathcal{R}(D^0)$ and $\mathcal{R}(D^{*0})$, -0.23 for $\mathcal{R}(D^+)$ and $\mathcal{R}(D^{*+})$, and -0.45 for $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$.

VII. SYSTEMATIC UNCERTAINTIES

Table V lists the systematic uncertainties considered in this analysis, as well as their correlations in the measurements of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$. We distinguish two kinds of uncertainties that affect the measurement of $\mathcal{R}(D^{(*)})$: *additive* uncertainties which impact the signal and background yields and thereby the significance of the results, and *multiplicative* uncertainties that affect the $\varepsilon_{\rm sig}/\varepsilon_{\rm norm}$ ratios and, thus, do not change the significance. The limited size of the simulated signal and background samples impact both additive and multiplicative uncertainties.

TABLE V. Systematic uncertainties and correlations on $\mathcal{R}(D^{(*)})$ for the isospin-unconstrained (columns 1–4 and 7–8) and isospin-constrained (columns 5–6 and 9) fits. The total uncertainties and correlations are calculated based on Eq. (30).

	Fractional uncertainty (%)						Correlation		
Source of uncertainty	$\mathcal{R}(D^0)$	$\mathcal{R}(D^{*0})$	$\mathcal{R}(D^+)$	$\mathcal{R}(D^{*+})$	$\mathcal{R}(D)$	$\mathcal{R}(D^*)$	D^0/D^{*0}	D^+/D^{*+}	D/D^*
Additive uncertainties									
PDFs									
MC statistics	6.5	2.9	5.7	2.7	4.4	2.0	−0.70	−0.34	−0.56
$\bar{B} \to D^{(*)}(\tau^-/\ell^-)\bar{\nu}$ FFs	0.3	0.2	0.2	0.1	0.2	0.2	−0.52	−0.13	−0.35
$D^{**} \to D^{(*)}(\pi^0/\pi^\pm)$	0.7	0.5	0.7	0.5	0.7	0.5	0.22	0.40	0.53
$\mathcal{B}(\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell)$	1.0	0.4	1.0	0.4	0.8	0.3	−0.63	−0.68	−0.58
$\mathcal{B}(\bar{B} \to D^{**}\tau^-\bar{\nu}_\tau)$	1.2	2.0	2.1	1.6	1.8	1.7	1.00	1.00	1.00
$D^{**} \to D^{(*)}\pi\pi$	2.1	2.6	2.1	2.6	2.1	2.6	0.22	0.40	0.53
Cross-feed constraints									
MC statistics	2.6	0.9	2.1	0.9	2.4	1.5	0.02	−0.02	−0.16
$f_{D^{**}}$	6.2	2.6	5.3	1.8	5.0	2.0	0.22	0.40	0.53
Feed-up/feed-down	1.9	0.5	1.6	0.2	1.3	0.4	0.29	0.51	0.47
Isospin constraints	···	···	···	···	1.2	0.3	···	···	−0.60
Fixed backgrounds									
MC statistics	4.3	2.3	4.3	1.8	3.1	1.5	−0.48	−0.05	−0.30
Efficiency corrections	4.8	3.0	4.5	2.3	3.9	2.3	−0.53	0.20	−0.28
Multiplicative uncertainties									
MC statistics	2.3	1.4	3.0	2.2	1.8	1.2	0.00	0.00	0.00
$\bar{B} \to D^{(*)}(\tau^-/\ell^-)\bar{\nu}$ FFs	1.6	0.4	1.6	0.3	1.6	0.4	0.00	0.00	0.00
Lepton PID	0.9	0.9	0.9	0.8	0.9	0.9	1.00	1.00	1.00
$\pi^0/\pi^\pm$ from $D^* \to D\pi$	0.1	0.1	0.0	0.0	0.1	0.1	1.00	1.00	1.00
Detection/Reconstruction	0.7	0.7	0.7	0.7	0.7	0.7	1.00	1.00	1.00
$\mathcal{B}(\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau)$	0.2	0.2	0.2	0.2	0.2	0.2	1.00	1.00	1.00
Total syst. uncertainty	12.2	6.7	11.4	6.0	9.6	5.6	−0.21	0.10	0.05
Total stat. uncertainty	19.2	9.8	18.0	11.0	13.1	7.1	−0.59	−0.23	−0.45
Total uncertainty	22.8	11.9	21.3	12.5	16.2	9.0	−0.48	−0.15	−0.27

A. Additive uncertainties

Additive uncertainties affect the results of the fit. To assess their impact, we vary the source of uncertainty 1000 times following a given distribution, and repeat the fit for each variation. We adopt as the uncertainty the standard deviation of the distribution of the resulting $R(D^{(*)})$ values. From this ensemble of fits, we also estimate the correlation between the uncertainties of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$.

1. PDF estimation

MC statistics: We employ a bootstrap algorithm [37] to estimate the uncertainty due to the limited size of the simulated event samples on which we base the 56 PDFs. We generate 1000 samples of simulated events by sampling the original MC sample with replacement [38]. The PDFs are recalculated with each bootstrapped sample, and the fit is repeated for each set of PDFs. Figure 6 shows the 1σ and 2σ bands for the projections of five selected PDFs. The impact on the final result is 4.4% for $\mathcal{R}(D)$ and 2.0% for $\mathcal{R}(D^*)$.

Form factors for $\bar{B} \to D^{(*)}(\tau^-/\ell^-)\bar{\nu}$: We estimate the impact on the signal and normalization PDFs due to the uncertainties on the FF parameters, ρ_D^2, Δ, $\rho_{D^*}^2$, $R_0(1)$, $R_1(1)$, and $R_2(1)$, taking into account their uncertainties and correlations. We recalculate the $D^{(*)}\tau\nu$ and $D^{(*)}\ell\nu$ PDFs with each set of 1000 Gaussian variations of the parameter values, and repeat the fit with each set of PDFs to determine the impact on $\mathcal{R}(D^{(*)})$.

$D^{**} \to D^{(*)}(\pi^0/\pi^\pm)$ *fraction:* The simulation of $D^{**}(\ell/\tau)\nu$ decays only includes the two-body decays $D^{**} \to D^{(*)}\pi$ of the four $L = 1$ charm meson states. The ratio of $D^{**} \to D^{(*)}\pi^0$ decays to $D^{**} \to D^{(*)}\pi^\pm$ decays which is fixed by isospin relations has a significant impact on the PDFs, because $D^{**} \to D^{(*)}\pi^0$ decays result in a sharply peaked $m_{\rm miss}^2$ distribution for the $D^{(*)}\pi^0\ell$ samples. The measured uncertainty on the π^0 detection efficiency is 3%. We assume a 4% uncertainty to the probability that a low momentum charged pion from $D^{**} \to D^{(*)}\pi^\pm$ decays is misassigned to the $B_{\rm tag}$ decay. Combining these two uncertainties, we arrive at an uncertainty on the relative proportion of the two-body decays of D^{**} of 5%. We repeat the fit increasing and decreasing this ratio by 5%, and adopt the largest variation of the isospin-constrained fit results as the systematic uncertainty.

$\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ *branching fractions:* Since decays to the four D^{**} states are combined in the $\bar{B} \to D^{(*)}(\tau^-/\ell^-)\bar{\nu}$ samples, the PDFs depend on the relative $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ branching fractions for the four $L = 1$ states [4]. The impact of the branching fraction uncertainties is assessed by recalculating the $\bar{B} \to D^{(*)}(\tau^-/\ell^-)\bar{\nu}$ PDFs and adopting the variation of the fit results from the ensemble of PDFs as the uncertainty.

$\bar{B} \to D^{**}(\tau^-/\ell^-)\bar{\nu}$ *branching fractions:* As noted above, the sharp peak in the $m^2_{\rm miss}$ distribution of the $D^{(*)}\pi^0\ell$ samples constrains contributions from $B \to D^{(*)}\pi\ell\nu$ decays. Events with additional unreconstructed particles contribute to the tail of the $m^2_{\rm miss}$ distribution and, thus, are more difficult to separate from other backgrounds and signal events. This is the case for $\bar{B} \to D^{**}\tau^-\bar{\nu}_\tau$ decays, which are combined with $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ decays in the $D^{**}(\ell/\tau)\nu$ PDFs with the relative proportion $\mathcal{R}(D^{**})_{\rm PS} = 0.18$. This value has been derived from the ratio of the available phase space. The same estimate applied to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays results in $\mathcal{R}(D)_{\rm PS} = 0.279$ and $\mathcal{R}(D^*)_{\rm PS} = 0.251$, values that are 58% and 32% smaller than the measured values. Taking this comparison as guidance for the error on $\mathcal{R}(D^{**})$, we increase $\mathcal{R}(D^{**})$ by 50%, recalculate the $D^{**}(\ell/\tau)\nu$ PDFs, and repeat the fit. As a result, the values of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ decrease by 1.8% and 1.7%, respectively. The impact is relatively small, because $\bar{B} \to D^{**}\tau^-\bar{\nu}_\tau$ contributions are small with respect to signal decays, which have much higher reconstruction efficiencies.

Unmeasured $B \to D^{**}(\to D^{(*)}\pi\pi)\ell\nu_\ell$ *decays:* To assess the impact of other potential $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ contributions, we modify the standard fit by adding an additional component. Out of the four contributions listed in Table VI, the three-body decays of the D^{**} states with $L = 1$ give the best agreement in the fits to the $D^{(*)}\pi^0\ell$ samples. For this decay chain, the $m^2_{\rm miss}$ distribution has a long tail due to an additional undetected pion. This could account for some of the observed excess at $1 < m^2_{\rm miss} < 2$ GeV^2 in Fig. 9. We assign the observed change in $\mathcal{R}(D^{(*)})$ as a systematic uncertainty.

2. Cross-feed constraints

MC statistics: Constraints on the efficiency ratios that link contributions from the same source are taken from MC simulation. The impact of their statistical uncertainty is assessed by varying the simulated event yields assuming Poisson errors.

TABLE VI. Additional $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ decays and the MC model implemented for their decays. The fourth decay mode refers to three-body decay of the four $L = 1$ D^{**} states.

Decay	Decay model
Non-resonant $B \to D^{(*)}\pi\ell\nu_\ell$	Goity-Roberts [39]
Non-resonant $B \to D^{(*)}\pi\pi\ell\nu_\ell$	Phase Space
$B \to D^{(*)}\eta\ell\nu_\ell$	Phase Space
$B \to D^{**}(\to D^{(*)}\pi\pi)\ell\nu_\ell$	ISGW2 [32]

The ratios $f_{D^{**}}$: We assess the uncertainty on $f_{D^{**}}$, the constraints linking the $D^{**}(\ell/\tau)\nu$ yields in the $D^{(*)}\ell$ and $D^{(*)}\pi^0\ell$ samples, by estimating the relative efficiencies of the selection criteria that differ in the two samples. The main differences in the selection of these samples are due to differences in the $D^{(*)}\ell$ and $D^{(*)}\pi^0\ell$ BDTs.

In the $D^{(*)}\ell$ samples, we observed that differences between data and simulation cause a 5%–10% underestimation of the continuum and $B\bar{B}$ backgrounds after the BDT requirements are applied. Since the $D^{**}(\ell/\tau)\nu$ contributions have similar $E_{\rm extra}$ distributions, and these distributions are the key inputs to the BDTs, we applied the same 5%-10% corrections to these contributions. We conservatively assign 100% of this correction as the systematic uncertainty on the $D^{**}(\ell/\tau)\nu$ efficiency in the $D^{(*)}\ell$ samples.

Since $\bar{B} \to D^{**}(\tau^-/\ell^-)\bar{\nu}$ decays are difficult to isolate in samples other than the $D^{(*)}\pi^0\ell$ control samples, we estimate the uncertainty on the $D^{**}(\ell/\tau)\nu$ efficiency due to the $D^{(*)}\pi^0\ell$ BDT selection by relying on the observed data-MC difference of the BDT selection efficiency for the $D^{(*)}\ell\nu$ sample. We assign the full 8.5% overestimate of the $D^{(*)}\ell\nu$ contribution as the systematic uncertainty on the $D^{**}(\ell/\tau)\nu$ efficiency in the $D^{(*)}\pi^0\ell$ samples.

The $f_{D^{**}}$ constraints also depend on the relative branching fractions of the four $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ decays that are combined in the $D^{**}(\ell/\tau)\nu$ contributions. We estimate their impact on $f_{D^{**}}$ from the branching fraction variations observed in the evaluation of the PDF uncertainty. The largest standard deviation for the four $f_{D^{**}}$ distributions is 1.8%.

By adding the uncertainties on $f_{D^{**}}$ described above in quadrature, we obtain total uncertainties of 13.2% for the D samples, and 10.0% for the D^* samples. Given that there are similarities between the BDT selections applied to the D and D^* samples, we adopt a 50% correlation between their uncertainties. With these uncertainties and correlations, we derive the total impact on the results, 5.0% for $\mathcal{R}(D)$ and 2.0% for $\mathcal{R}(D^*)$.

Feed-down constraints: The feed-down constraints of the signal yields are corrected as part of the iteration of the fit. The uncertainties on these corrections are given by the statistical uncertainty on the ratios of the fitted $D^*\ell\nu \Rightarrow D^*\ell$ and $D^*\ell\nu \Rightarrow D\ell$ yields. They are 2.4% and 4.4% on the $D^{*0}\tau\nu$ and $D^{*+}\tau\nu$ feed-down constraints, respectively.

Feed-up constraints: We estimate the uncertainty on the $D\tau\nu$ and $D\ell\nu$ feed-up constraints as 100% of the corrections on the feed-down constraints. This results in 6.8% on the $D^0(\ell/\tau)\nu$ feed-up and 9.9% on the $D^+(\ell/\tau)\nu$ feed-up. These two effects combined lead to an uncertainty of 1.3% on $\mathcal{R}(D)$ and 0.4% on $\mathcal{R}(D^*)$.

Isospin constraints: In the isospin-constrained fit, we employ five additional constraints to link the signal and

normalization yields of the samples corresponding to B^- and B^0 decays. Since we reweight these contributions with the $q^2 \leq 4$ GeV^2 control sample, the uncertainty on the isospin constraints is given by the statistical uncertainty on the ratios of the $q^2 \leq 4$ GeV^2 yields. This uncertainty is 3.4% in the $D\ell$ samples and 3.6% in the $D^*\ell$ samples. This translates into uncertainties of 1.2% on $\mathcal{R}(D)$ and 0.3% on $\mathcal{R}(D^*)$.

3. Fixed background contributions

MC statistics: The yields of the continuum, $B\bar{B}$, and cross-feed backgrounds are fixed in the fit. The uncertainty due to the limited size of the MC samples is estimated generating Poisson variations of these yields, and repeating the fit with each set of values. A significant part of this uncertainty is due to the continuum yields, since the size of simulated continuum sample is equivalent to only twice the data sample,

Efficiency corrections: To account for the correlations among the various corrections applied to the continuum and $B\bar{B}$ backgrounds, we follow this multistep procedure:

(i) We vary the continuum corrections within their statistical uncertainties of 3%–9%, given by the number of events in the off-peak data control samples.
(ii) The branching fractions of the most abundant decays in the $B\bar{B}$ background are varied within their uncertainties [12].
(iii) The $B\bar{B}$ correction is reestimated in the high $E_{\rm extra}$ control sample, and varied within the statistical uncertainty of 1.9%.
(iv) The BDT bias corrections are reestimated in the $m_{\rm ES}$ sideband, and varied within their statistical uncertainties, 2.1% in the $D\ell$ samples and 3.6% in the $D^*\ell$ samples.
(v) The $B\bar{B}$ background PDFs are recalculated.
(vi) The fit is repeated for each set of PDF and yield variations.

Table VII shows the size of the continuum and $B\bar{B}$ backgrounds and their uncertainties due to the limited size of the MC samples and the various corrections implemented by comparisons with control samples.

B. Multiplicative uncertainties

MC statistics: The relative efficiency $\varepsilon_{\rm sig}/\varepsilon_{\rm norm}$ is estimated as the ratio of expected yields, so the limited size of the MC samples contributes to its uncertainty. We estimate it assuming Poisson errors on the MC yields.

Form factors for $\bar{B} \to D^{(*)}(\tau^-/\ell^-)\bar{\nu}$: The $q^2 > 4$ GeV^2 requirement introduces some dependence on the FF parametrization. This uncertainty is assessed based on the effect of the FF variations calculated for the uncertainty on the PDFs.

$\pi^0/\pi^\pm$ from $D^* \to D\pi$: There is a significant momentum-dependent uncertainty on the reconstruction efficiency of soft pions originating from $D^* \to D\pi$ decays. However, the momentum spectra of soft pions in signal and normalization decays are rather similar, see Fig. 10. As a result, the uncertainty on $\mathcal{R}(D^{(*)})$ is less than 0.1%.

Detection and reconstruction: Given that signal and normalization decays are reconstructed by the same particles in the final state, many of the uncertainties that impact their efficiencies cancel in the ratios $\varepsilon_{\rm sig}/\varepsilon_{\rm norm}$. Uncertainties due to $B_{\rm tag}$-efficiency, final-state radiation, soft-pion reconstruction, and others related to the detector performance contribute less than 1%.

One difference between signal and normalization decays is that τ leptons can decay hadronically. Due to the large misidentification rate of the muon selector, we estimate that in 6.1% of all signal events, the τ does not decay leptonically. The uncertainty on the misidentification rate is less than 10%, so the uncertainty on the signal efficiency due to this effect is 0.6%, included in the lepton ID row of Table V.

$\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau$ branching fraction: We use the world averages $\mathcal{B}(\tau^- \to e^-\bar{\nu}_e\nu_\tau) = (17.83 \pm 0.04)\%$ and $\mathcal{B}(\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau) = (17.41 \pm 0.04)\%$ [12].

C. Correlations

Even though several of the uncertainties listed in Table V have the same source, their impact on $\mathcal{R}(D^{(*)})$ is largely uncorrelated, i.e., the correlation between uncertainties in different rows of Table V is negligible. However, the correlation between the uncertainties on $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ (different columns) is significant, and important for the comparison of these measurements with theoretical predictions.

TABLE VII. Continuum and other $B\bar{B}$ background yields; the first uncertainty is due to MC statistics, the second to efficiency corrections, and σ refers to the total uncertainty.

Sample	Continuum	σ (%)	$B\bar{B}$	σ (%)
$D^0\ell$	$355 \pm 13 \pm 12$	4.9	$330 \pm 6 \pm 17$	5.3
$D^{*0}\ell$	$132 \pm 8 \pm 6$	7.6	$188 \pm 4 \pm 10$	5.9
$D^+\ell$	$157 \pm 9 \pm 6$	6.9	$191 \pm 5 \pm 9$	5.5
$D^{*+}\ell$	$12 \pm 3 \pm 1$	23.6	$72 \pm 3 \pm 4$	6.9

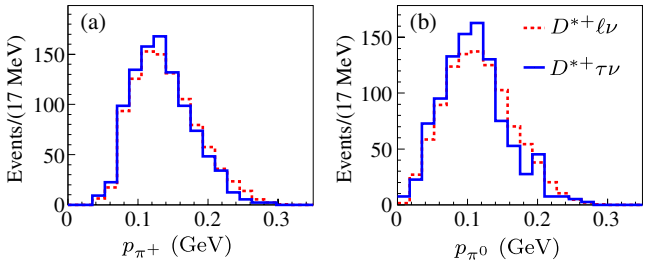

FIG. 10 (color online). Pion momentum in the laboratory from $B \to D^{*+}\ell\nu$ and $B \to D^{*+}\tau\nu$ decays: (a) $D^{*+} \to D^0\pi^+$, and (b) $D^{*+} \to D^+\pi^0$ decays. Histograms are normalized to 1000 entries.

For most of the additive systematic uncertainties, we estimate the correlation from the two-dimensional $\mathcal{R}(D)$-$\mathcal{R}(D^*)$ distribution resulting from the fit variations. This is not possible for the $D^{**} \to D^{(*)}\pi^0/\pi^\pm$ and $D^{**} \to D^{(*)}\pi\pi$ uncertainties. These uncertainties affect the size of the $D^{**}(\ell/\tau)\nu$ background in the $D^{(*)}\ell$ samples in the same way that as $f_{D^{**}}$ does. Thus, we derive their correlations from the $f_{D^{**}}$ correlations. Since the signal and $D^{**}\tau\nu$ PDFs are very similar, we assign a 100% correlation on $\mathcal{B}(\bar{B} \to D^{**}\tau^-\bar{\nu}_\tau)$.

The multiplicative uncertainties on the efficiency due to the MC statistics are uncorrelated. The FFs for $\bar{B} \to D\ell^-\bar{\nu}_\ell$ and $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ decays are measured separately, so their uncertainties are also not correlated. The uncertainty on $\mathcal{B}(\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau)$ affects all channels equally. We assume that the remaining small uncertainties on the efficiencies due to detector effects are 100% correlated as well.

The uncertainties and their correlations are listed in Table V. We combine these correlations ρ_i and the uncertainties by adding their covariance matrices,

$$\sum_i \begin{pmatrix} \sigma_i^2 & \rho_i\sigma_i\sigma_i^* \\ \rho_i\sigma_i\sigma_i^* & \sigma_i^{*2} \end{pmatrix} = \begin{pmatrix} \sigma_{\rm tot}^2 & \rho_{\rm tot}\sigma_{\rm tot}\sigma_{\rm tot}^* \\ \rho_{\rm tot}\sigma_{\rm tot}\sigma_{\rm tot}^* & \sigma_{\rm tot}^{*2} \end{pmatrix}. \tag{30}$$

Here, σ_i and σ_i^* refer to the uncertainties on $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$, respectively.

VIII. STABILITY CHECKS AND KINEMATIC DISTRIBUTIONS

A. Stability tests

We have checked the stability of the fit results for different data subsamples and different levels of background suppression.

To look for possible dependence of the results on the data taking periods, we divide the data sample into four periods corresponding to approximately equal luminosity, and fit each sample separately. The results are presented in Fig. 11. The eight measurements each for $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$, separately for B^+ and B^0, are compared to the isospin-constrained fit results obtained from the complete data sample. Based on the values of χ^2 for 7 degrees of freedom, we conclude that the results of these fits are statistically consistent with the fit to the whole data sample.

A similar test is performed for two samples identified by the final state lepton, an electron or a muon. This test includes the uncertainties on the background corrections that affect the electron and muon samples differently. These uncertainties are statistically dominated and, thus, independent for both samples. The results are presented in the bottom panels of Fig. 11. The χ^2 tests confirm the stability of these measurements within the uncertainties.

To assess the sensitivity of the fit results on the purity of the data sample and the BDT selection, we perform fits for

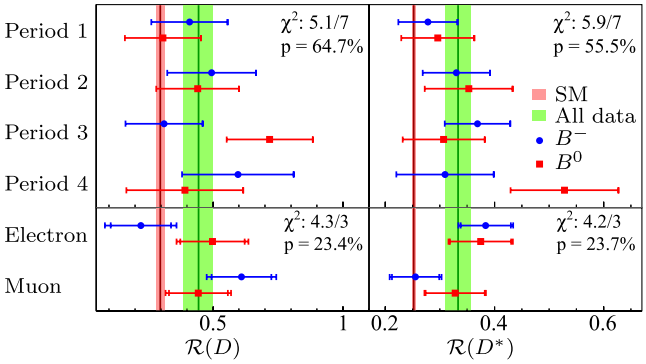

FIG. 11 (color online). Measurements of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ for different data subsamples. Top: for four run periods with statistical uncertainties only. Bottom: for electrons and muons with statistical and uncorrelated systematic uncertainties. The vertical bands labeled "SM" and "All data" mark the SM predictions and the results of the fits to the whole data sample, respectively.

samples selected with different BDT requirements. We identify each sample by the relative number of events in the signal region ($m^2_{\rm miss} > 1$ GeV^2) with respect to the nominal sample, which is labeled as the 100% sample. The ratio of the number of fitted signal events S to the number of background events B varies from $S/B = 1.27$ in the 30% sample, to $S/B = 0.27$ in the 300% sample, while the backgrounds increase by a factor of 18. The BDT bias correction and the PDFs are recalculated for each sample. Figure 12 shows the results of fits to the different samples with tighter and looser BDT requirements. We take into account the large correlations between these nested samples and conclude that the results are stable for the very large variations of the BDT requirements.

B. Gaussian uncertainties

For a maximum likelihood fit with Gaussian uncertainties, the logarithm of the likelihood is described by the parabola $P(Y) = (Y - Y_{\rm fit})^2/2\sigma^2_{\rm fit}$, where $Y_{\rm fit}$ is the fitted yield and $\sigma_{\rm fit}$ is the uncertainty on $Y_{\rm fit}$. Figure 13 compares

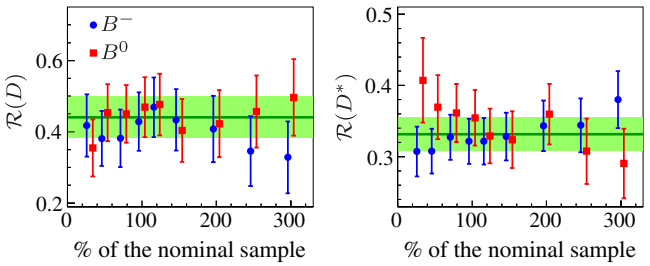

FIG. 12 (color online). Measurements of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ for different BDT requirements, impacting the signal/background ratio. The horizontal bands mark the $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ results for the isospin-constrained fit to the nominal (100%) sample. The data points represent the results of the fits for B^+ and B^0 mesons with their statistical uncertainties.

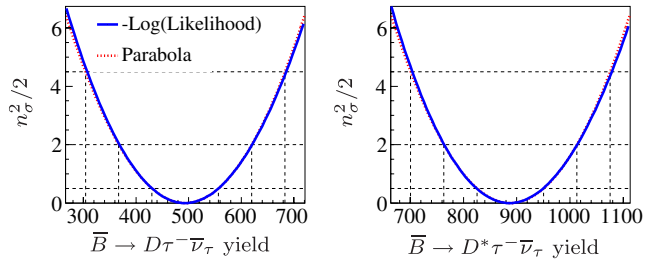

FIG. 13 (color online). Likelihood scan for the two signal yields compared to a parabola. The dashed lines indicate the number of standard deviations (n_σ) away from the fit result.

the likelihood scan of the signal yields for the isospin-constrained fit with the parabola that results from the fitted yields, presented in Table VIII. There is a slight asymmetry in the likelihood function, but good agreement overall. Thus, we conclude that the statistical uncertainty on $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ may be considered Gaussian.

Figure 14 shows the effect on $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ from variations on $f_{D^{**}}$, the largest source of systematic uncertainty. The distributions are well described by a Gaussian function. This is also the case for the other major sources of systematic uncertainty.

C. Kinematic distributions

We further study the results of the fit by comparing the kinematic distributions of data events with the SM expectations. Specifically, we focus on the signal-enriched region with $m^2_{\rm miss} > 1.5$ GeV^2 and scale each component in the simulation by the results of the fits. To compare the data and MC distributions we calculate a χ^2 per degree of freedom which only includes the statistical uncertainty of bins with 8 or more events. The number of degrees of freedom is given by the number of bins minus the number of fitted signal yields.

Figure 15 shows the $E_{\rm extra}$ distribution of events in the $D^{(*)}\ell$ samples. This variable is key in the BDT selection and overall background suppression. There is a clear enhancement of signal events at $E_{\rm extra} = 0$ in all four $D^{(*)}\ell$ samples. The background contributions, which are significantly more uniform in $E_{\rm extra}$ than those of signal, appear to be well reproduced. We conclude that the simulation agrees well with the data distribution.

Figure 16 also shows clear signal enhancements in the $m_{\rm ES}$ and $|\boldsymbol{p}^*_\ell|$ distributions of events in the $m^2_{\rm miss} > 1.5$ GeV^2 region. The data and simulation agree well within the limited statistics.

IX. RESULTS

A. Comparison with SM expectations

Table VIII shows the results of the measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ extracted from the fit without and with isospin constraints linking B^+ and B^0 decays.

The $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ branching fractions are calculated from the measured values of $\mathcal{R}(D^{(*)})$,

$$\mathcal{B}(\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau) = \mathcal{R}(D^{(*)}) \times \mathcal{B}(\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell). \tag{31}$$

For B^-, we use the average branching fractions measured by *BABAR* [40–42],

$$\mathcal{B}(B^- \to D^0\ell^-\bar{\nu}_\ell) = (2.32 \pm 0.03 \pm 0.08)\%,$$
$$\mathcal{B}(B^- \to D^{*0}\ell^-\bar{\nu}_\ell) = (5.31 \pm 0.02 \pm 0.19)\%,$$

and for $\bar{B}^0$, the corresponding branching fractions related by isospin.

We estimate the statistical significance of the measured signal branching fractions as $\Sigma_{\rm stat} = \sqrt{2\Delta(\ln \mathcal{L})}$, where $\Delta(\ln \mathcal{L})$ is the increase in log-likelihood for the nominal fit relative to the no-signal hypothesis. The total significance $\Sigma_{\rm tot}$ is determined as

$$\Sigma_{\rm tot} = \Sigma_{\rm stat}\frac{\sigma_{\rm stat}}{\sqrt{\sigma^2_{\rm stat} + \sigma^2_{\rm asys}}}. \tag{32}$$

In this expression, the statistical significance is scaled by the sum of the statistical uncertainty $\sigma_{\rm stat}$ and the additive systematic uncertainty $\sigma_{\rm asys}$. The significance of the $\bar{B} \to D\tau^-\bar{\nu}_\tau$ signal is 6.8σ, the first such measurement exceeding 5σ.

TABLE VIII. Results of the isospin-unconstrained (top four rows) and isospin-constrained fits (last two rows). The columns show the signal and normalization yields, the ratio of their efficiencies, $\mathcal{R}(D^{(*)})$, the signal branching fractions, and $\Sigma_{\rm stat}$ and $\Sigma_{\rm tot}$, the statistical and total significances of the measured signal yields. Where two uncertainties are given, the first is statistical and the second is systematic. The second and third uncertainties on the branching fractions $\mathcal{B}(\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau)$ correspond to the systematic uncertainties due to $\mathcal{R}(D^{(*)})$ and $\mathcal{B}(\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell)$, respectively. The stated branching fractions for the isospin-constrained fit refer to B^- decays.

Decay	$N_{\rm sig}$	$N_{\rm norm}$	$\varepsilon_{\rm sig}/\varepsilon_{\rm norm}$	$\mathcal{R}(D^{(*)})$	$\mathcal{B}(B \to D^{(*)}\tau\nu)(\%)$	$\Sigma_{\rm stat}$	$\Sigma_{\rm tot}$
$B^- \to D^0\tau^-\bar{\nu}_\tau$	314 ± 60	1995 ± 55	0.367 ± 0.011	$0.429 \pm 0.082 \pm 0.052$	$0.99 \pm 0.19 \pm 0.12 \pm 0.04$	5.5	4.7
$B^- \to D^{*0}\tau^-\bar{\nu}_\tau$	639 ± 62	8766 ± 104	0.227 ± 0.004	$0.322 \pm 0.032 \pm 0.022$	$1.71 \pm 0.17 \pm 0.11 \pm 0.06$	11.3	9.4
$\bar{B}^0 \to D^+\tau^-\bar{\nu}_\tau$	177 ± 31	986 ± 35	0.384 ± 0.014	$0.469 \pm 0.084 \pm 0.053$	$1.01 \pm 0.18 \pm 0.11 \pm 0.04$	6.1	5.2
$\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$	245 ± 27	3186 ± 61	0.217 ± 0.005	$0.355 \pm 0.039 \pm 0.021$	$1.74 \pm 0.19 \pm 0.10 \pm 0.06$	11.6	10.4
$\bar{B} \to D\tau^-\bar{\nu}_\tau$	489 ± 63	2981 ± 65	0.372 ± 0.010	$0.440 \pm 0.058 \pm 0.042$	$1.02 \pm 0.13 \pm 0.10 \pm 0.04$	8.4	6.8
$\bar{B} \to D^*\tau^-\bar{\nu}_\tau$	888 ± 63	11953 ± 122	0.224 ± 0.004	$0.332 \pm 0.024 \pm 0.018$	$1.76 \pm 0.13 \pm 0.10 \pm 0.06$	16.4	13.2

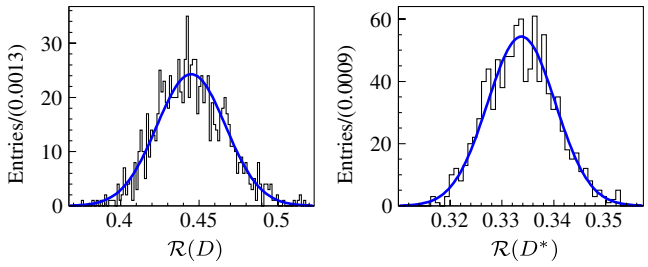

FIG. 14 (color online). Histograms: $\mathcal{R}(D^{(*)})$ distributions resulting from 1000 variations of $f_{D^{**}}$. Solid curves: Gaussian fits to the $\mathcal{R}(D^{(*)})$ distributions.

We compare the measured $\mathcal{R}(D^{(*)})$ to the calculations based on the SM,

$$\begin{aligned}\mathcal{R}(D)_{\rm exp} &= 0.440 \pm 0.072\\ \mathcal{R}(D^*)_{\rm exp} &= 0.332 \pm 0.030,\\ \mathcal{R}(D)_{\rm SM} &= 0.297 \pm 0.017\\ \mathcal{R}(D^*)_{\rm SM} &= 0.252 \pm 0.003,\end{aligned}$$

and observe an excess over the SM predictions for $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ of 2.0σ and 2.7σ, respectively. We combine these two measurements in the following way

$$\chi^2 = (\Delta, \Delta^*)\begin{pmatrix}\sigma^2_{\rm exp}+\sigma^2_{\rm th} & \rho\sigma_{\rm exp}\sigma^*_{\rm exp}\\ \rho\sigma_{\rm exp}\sigma^*_{\rm exp} & \sigma^{*2}_{\rm exp}+\sigma^{*2}_{\rm th}\end{pmatrix}^{-1}\begin{pmatrix}\Delta\\ \Delta^*\end{pmatrix}, \qquad (33)$$

where $\Delta^{(*)} = \mathcal{R}(D^{(*)})_{\rm exp} - \mathcal{R}(D^{(*)})_{\rm th}$, and ρ is the total correlation between the two measurements, $\rho(\mathcal{R}(D), \mathcal{R}(D^*)) = -0.27$. Since the total uncertainty is dominated by the experimental uncertainty, the expression in Eq. (33) is expected to be distributed as a χ^2 distribution for two degrees of freedom. Figure 17 shows this distribution in the $\mathcal{R}(D)$-$\mathcal{R}(D^*)$ plane. The contours are ellipses slightly rotated with respect to the $\mathcal{R}(D)$-$\mathcal{R}(D^*)$ axes, due to the nonzero correlation.

For the assumption that $\mathcal{R}(D^{(*)})_{\rm th} = \mathcal{R}(D^{(*)})_{\rm SM}$, we obtain $\chi^2 = 14.6$, which corresponds to a probability of 6.9×10^{-4}. This means that the possibility that the measured $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ both agree with the SM predictions is excluded at the 3.4σ level [43]. Recent calculations [7,8,44,45] have resulted in values of $\mathcal{R}(D)_{\rm SM}$ that slightly exceed our estimate. For the largest of those values, the significance of the observed excess decreases to 3.2σ.

B. Search for a charged Higgs

To examine whether the excess in $\mathcal{R}(D^{(*)})$ can be explained by contributions from a charged Higgs boson in the type II 2HDM, we study the dependence of the fit results on $\tan\beta/m_{H^\pm}$.

For 20 values of $\tan\beta/m_{H^\pm}$, equally spaced in the $[0.05, 1.00]$ GeV^{-1} range, we recalculate the eight signal PDFs, accounting for the charged Higgs contributions as described in Sec. II. Figure 18 shows the $m^2_{\rm miss}$ and $|\boldsymbol{p}^*_\ell|$ projections of the $D^0\tau\nu \Rightarrow D^0\ell$ PDF for four values of $\tan\beta/m_{H^\pm}$. The impact of charged Higgs contributions on the $m^2_{\rm miss}$ distribution mirrors those in the q^2 distribution, see Fig. 3, because of the relation

$$m^2_{\rm miss} = (p_{e^+e^-} - p_{B_{\rm tag}} - p_{D^{(*)}} - p_\ell)^2 = (q - p_\ell)^2,$$

The changes in the $|\boldsymbol{p}^*_\ell|$ distribution are due to the change in the τ polarization.

We recalculate the value of the efficiency ratio $\varepsilon_{\rm sig}/\varepsilon_{\rm norm}$ as a function of $\tan\beta/m_{H^\pm}$ (see Fig. 19). The efficiency

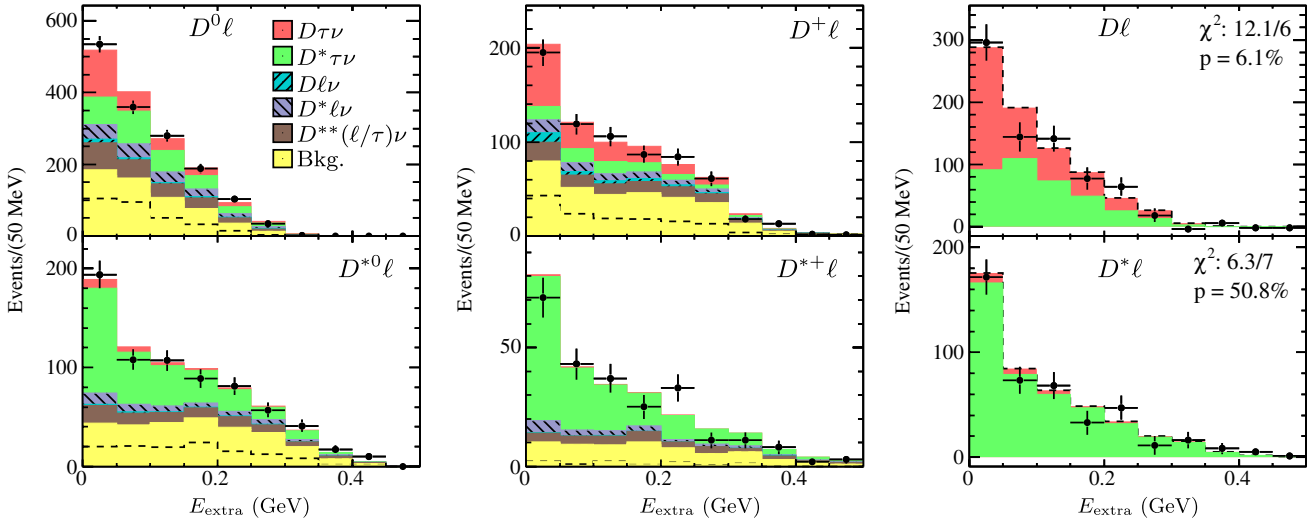

FIG. 15 (color online). $E_{\rm extra}$ distributions for events with $m^2_{\rm miss} > 1.5$ GeV^2 scaled to the results of the isospin-unconstrained (first two columns) and isospin-constrained (last column) fits. The region above the dashed line of the background component corresponds to $B\bar{B}$ background and the region below corresponds to continuum. In the third column, the B^0 and B^+ samples are combined, and the normalization and background events are subtracted.

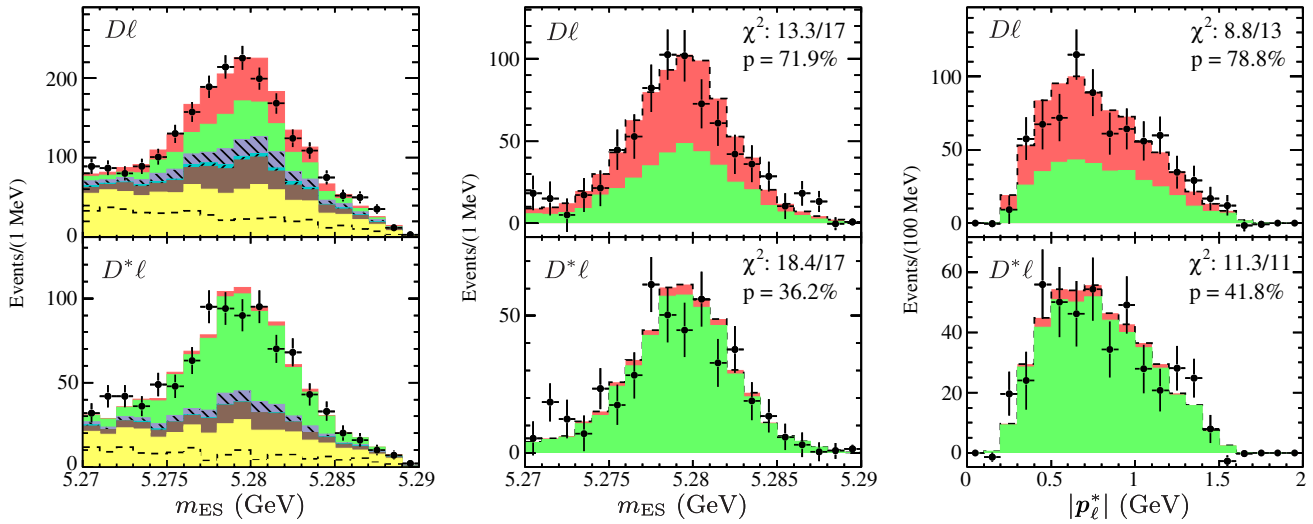

FIG. 16 (color online). $m_{\rm ES}$ distributions before (left) and after (center) subtraction of normalization of background events, and lepton momentum distributions after this subtraction (right) for events with $m^2_{\rm miss} > 1.5\ {\rm GeV}^2$ scaled to the results of the isospin-constrained fit. The B^0 and B^+ samples are combined. See Fig. 15 for a legend.

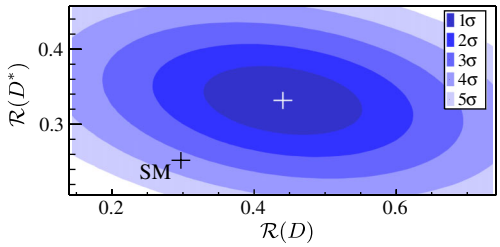

FIG. 17 (color online). Representation of χ^2 [Eq. (33)] in the $\mathcal{R}(D)$-$\mathcal{R}(D^*)$ plane. The white cross corresponds to the measured $\mathcal{R}(D^{(*)})$, and the black cross to the SM predictions. The shaded bands represent one standard deviation each.

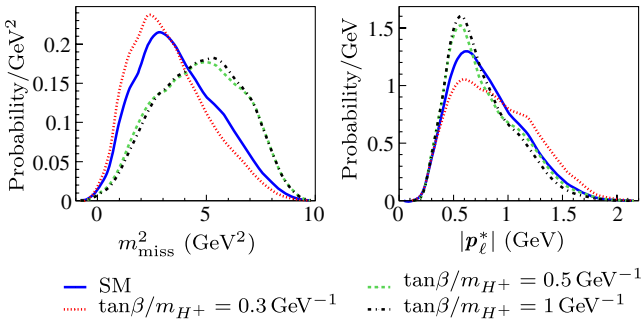

FIG. 18 (color online). $m^2_{\rm miss}$ and $|\boldsymbol{p}^*_\ell|$ projections of the $D^0\tau\nu \Rightarrow D^0\ell$ PDF for various values of $\tan\beta/m_{H^\pm}$.

increases up to 8% for large values of $\tan\beta/m_{H^\pm}$, and, as we noted earlier, its uncertainty increases due to the larger dispersion of the weights in the 2HDM reweighting.

The variation of the fitted signal yields as a function of $\tan\beta/m_{H^\pm}$ is also shown in Fig. 19. The sharp drop in the $\bar{B} \to D\tau^-\bar{\nu}_\tau$ yield at $\tan\beta/m_{H^\pm} \approx 0.4\ {\rm GeV}^{-1}$ is due to the large shift in the $m^2_{\rm miss}$ distribution which occurs when

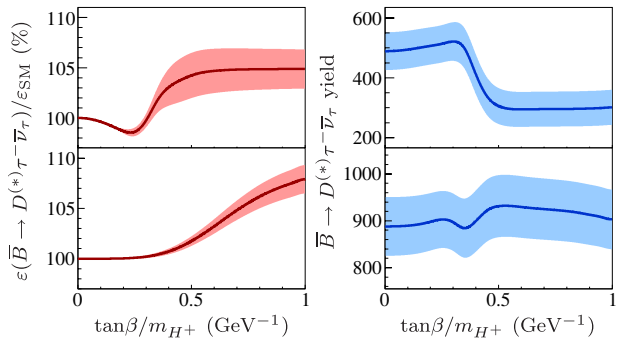

FIG. 19 (color online). Left: Variation of the $\bar{B} \to D\tau^-\bar{\nu}_\tau$ (top) and $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ (bottom) efficiency in the 2HDM with respect to the SM efficiency. The band indicates the increase on statistical uncertainty with respect to the SM value. Right: Variation of the fitted $\bar{B} \to D\tau^-\bar{\nu}_\tau$ (top) and $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ (bottom) yields as a function of $\tan\beta/m_{H^\pm}$. The band indicates the statistical uncertainty of the fit.

the Higgs contribution begins to dominate the total rate. This shift is also reflected in the q^2 distribution and, as we will see in the next section, the data do not support it. The change of the $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ yield, mostly caused by the correlation with the $\bar{B} \to D\tau^-\bar{\nu}_\tau$ sample, is much smaller.

Figure 20 compares the measured values of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ in the context of the type II 2HDM to the theoretical predictions as a function of $\tan\beta/m_{H^\pm}$. The increase in the uncertainty on the signal PDFs and the efficiency ratio as a function of $\tan\beta/m_{H^\pm}$ are taken into account. Other sources of systematic uncertainty are kept constant in relative terms.

The measured values of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ match the predictions of this particular Higgs model for $\tan\beta/m_{H^\pm} = 0.44 \pm 0.02\ {\rm GeV}^{-1}$ and $\tan\beta/m_{H^\pm} = 0.75 \pm 0.04\ {\rm GeV}^{-1}$, respectively. However, the combination of $\mathcal{R}(D)$ and

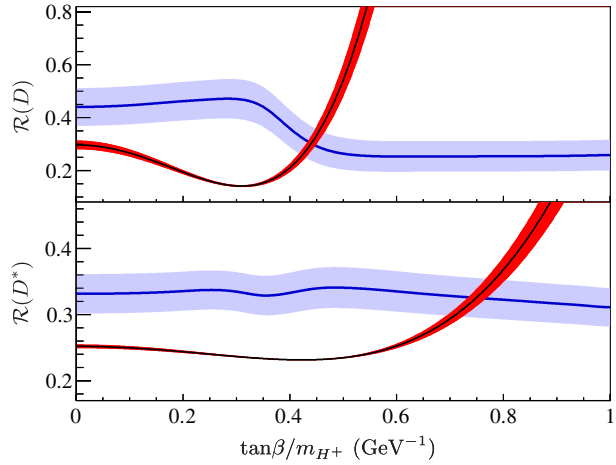

FIG. 20 (color online). Comparison of the results of this analysis (light band, blue) with predictions that include a charged Higgs boson of type II 2HDM (dark band, red). The widths of the two bands represent the uncertainties. The SM corresponds to $\tan\beta/m_{H^\pm} = 0$.

$\mathcal{R}(D^*)$ excludes the type II 2HDM charged Higgs boson at 99.8% confidence level for any value of $\tan\beta/m_{H^\pm}$, as illustrated in Fig. 21. This calculation is only valid for values of $m_{H^\pm}$ greater than 15 GeV [5,8]. The region for $m_{H^\pm} \leq 15$ GeV has already been excluded by $B \to X_s\gamma$ measurements [23], and therefore, the type II 2HDM is excluded in the full $\tan\beta$-$m_{H^\pm}$ parameter space.

As we detailed in Sec. II B, the type II 2HDM is a subset of more general 2HDMs that corresponds to values of $(S_R - S_L, S_R + S_L)$ that lie in the line joining $(-\infty, -\infty)$ and (0, 0) with slope 1. Since the dependence of the measured $\mathcal{R}(D^*)$ on $\tan\beta/m_{H^\pm}$, or, equivalently, on $\mathrm{Re}(S_R - S_L)$, is smaller than the total uncertainties considered, we can extend the measurement of $\mathcal{R}(D^{(*)})$ to the bottom half of the real $(S_R - S_L, S_R + S_L)$ plane by using the values of $\mathcal{R}(D^{(*)})$ obtained with $H_s(S_R \pm S_L)$ for $H_s(-S_R \mp S_L)$.

We also employ this extrapolation in the top half of the $(S_R - S_L, S_R + S_L)$ plane, that is, for $S_R + S_L > 0$. In this

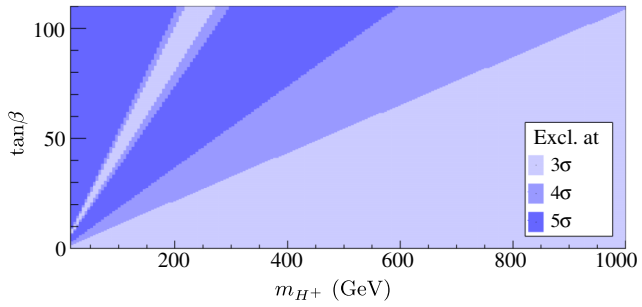

FIG. 21 (color online). Level of disagreement between this measurement of $\mathcal{R}(D^{(*)})$ and the type II 2HDM predictions for all values in the $\tan\beta$-$m_{H^\pm}$ parameter space.

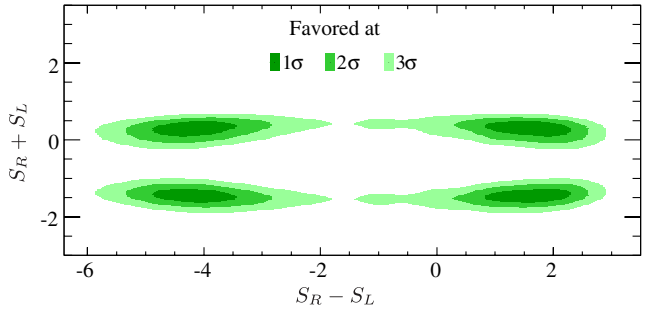

FIG. 22 (color online). Favored regions for real values of the type III 2HDM parameters S_R and S_L given by the measured values of $\mathcal{R}(D^{(*)})$. The bottom two solutions are excluded by the measured q^2 spectra.

case, the extrapolation is only a good approximation when the decay amplitude is dominated either by SM or NP contributions, that is, for small or large values of $|S_R + S_L|$. In the intermediate region, the q^2 spectra first shifts slightly to lower average values, and then moves sharply in the opposite direction. This is reflected in the measured value of $\mathcal{R}(D)$, and corresponds to the small rise up to $\tan\beta/m_{H^\pm} \sim 0.36$ GeV^{-1} ($S_R + S_L \sim -0.97$) in Fig. 20, and the sharp drop in the $0.36 < \tan\beta/m_{H^\pm} < 0.46$ GeV^{-1} region ($-0.97 > S_R + S_L > -1.58$).

For positive values of $S_R + S_L$ the interference between SM and 2HDM contributions is constructive, so the q^2 spectrum never shifts to lower values. By matching the q^2 spectra for positive and negative values of $S_R + S_L$, we can estimate that the drop in the value of $\mathcal{R}(D)$ becomes much more gradual and occurs in the $0.15 < S_R + S_L < 6.05$ region. Based on the extrapolation described above, the measured and expected values of $\mathcal{R}(D)$ match for $S_R + S_L \sim 0.3$. In this region, the NP contributions are small and the approximation is accurate to ~5%.

Figure 22 shows that for real values of S_R and S_L, there are four regions in the type III parameter space that can explain the excess in both $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$. This figure does not include uncertainties due to the extrapolation of the type II 2HDM measurements, which could somewhat affect the top two solutions. In addition, a range of complex values of the parameters are also compatible with this measurement [21,45–47].

C. Study of the q^2 spectra

As shown in Sec. II B, the q^2 spectrum of $\bar{B} \to D\tau^-\bar{\nu}_\tau$ decays could be significantly impacted by charged Higgs contributions. Figure 23 compares the q^2 distribution of background subtracted data, corrected for detector efficiency, with the expectations of three different scenarios. Due to the subtraction of the large $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ feed-down in the $D\ell$ samples, the measured q^2 spectrum of $\bar{B} \to D\tau^-\bar{\nu}_\tau$ decays depends on the signal hypothesis. This dependence is very small, however, because the q^2

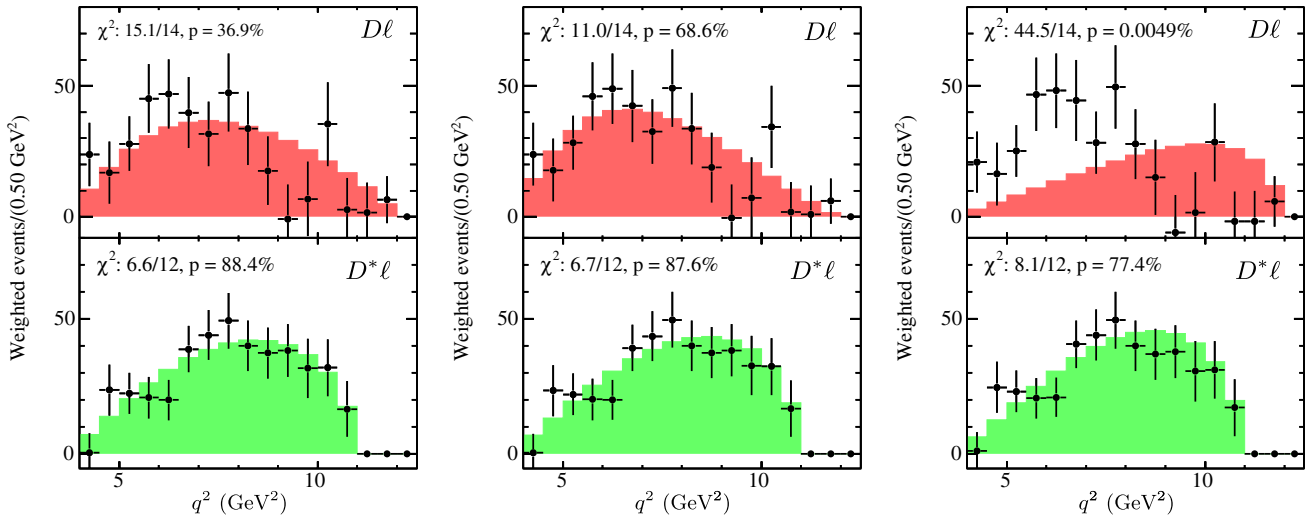

FIG. 23 (color online). Efficiency corrected q^2 distributions for $\bar{B} \to D\tau^-\bar{\nu}_\tau$ (top) and $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ (bottom) events with $m^2_{\rm miss} > 1.5\ {\rm GeV}^2$ scaled to the results of the isospin-constrained fit. Left: SM. Center: $\tan\beta/m_{H^\pm} = 0.30\ {\rm GeV}^{-1}$. Right: $\tan\beta/m_{H^\pm} = 0.45\ {\rm GeV}^{-1}$. The points and the shaded histograms correspond to the measured and expected distributions, respectively. The B^0 and B^+ samples are combined and the normalization and background events are subtracted. The distributions are normalized to the number of detected events. The uncertainty on the data points includes the statistical uncertainties of data and simulation. The values of χ^2 are based on this uncertainty.

spectrum of $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decays is largely independent of $\tan\beta/m_{H^\pm}$.

The measured q^2 spectra agree with the SM expectations within the statistical uncertainties. For $\bar{B} \to D\tau^-\bar{\nu}_\tau$ decays, there might be a small shift to lower values, which is indicated by the increase in the p value for $\tan\beta/m_{H^\pm} = 0.30\ {\rm GeV}^{-1}$. As we showed in Sec. II B, the average q^2 for $\tan\beta/m_{H^\pm} = 0.30\ {\rm GeV}^{-1}$ shifts to lower values because the charged Higgs contribution to $\bar{B} \to D\tau^-\bar{\nu}_\tau$ decays, which always proceeds via an S-wave, interferes destructively with the SM S-wave. As a result, the decay proceeds via an almost pure P-wave and is suppressed at large q^2 by a factor of p_D^2, thus improving the agreement with data. The negative interference suppresses the expected value of $\mathcal{R}(D)$ as well, however, so the region with small $\tan\beta/m_{H^\pm}$ is excluded by the measured $\mathcal{R}(D)$.

The two favored regions in Fig. 22 with $S_R + S_L \sim -1.5$ correspond to $\tan\beta/m_{H^\pm} = 0.45\ {\rm GeV}^{-1}$ for $\bar{B} \to D\tau^-\bar{\nu}_\tau$ decays. However, as we saw in Fig. 3, the charged Higgs contributions dominate $\bar{B} \to D\tau^-\bar{\nu}_\tau$ decays for values of $\tan\beta/m_{H^\pm} > 0.4\ {\rm GeV}^{-1}$ and the q^2 spectrum shifts significantly to larger values. The data do not appear to support this expected shift to larger values of q^2.

To quantify the disagreement between the measured and expected q^2 spectra, we conservatively estimate the systematic uncertainties that impact the distributions shown in Fig. 23 (Appendix). Within these uncertainties, we find the variation that minimizes the χ^2 value of those distributions. Table IX shows that, as expected, the conservative uncertainties give rise to large p values in most cases. However, the p value is only 0.4% for $\bar{B} \to D\tau^-\bar{\nu}_\tau$ decays and $\tan\beta/m_{H^\pm} = 0.45\ {\rm GeV}^{-1}$. Given that this value of $\tan\beta/m_{H^\pm}$ corresponds to $S_R + S_L \sim -1.5$, we exclude the two solutions at the bottom of Fig. 22 with a significance of at least 2.9σ.

The other two solutions corresponding to $S_R + S_L \sim 0.4$ do not impact the q^2 distributions of $\bar{B} \to D\tau^-\bar{\nu}_\tau$ to the same large degree, and, thus, we cannot exclude them with the current level of uncertainty. However, these solutions also shift the q^2 spectra to larger values due to the S-wave contributions from the charged Higgs boson, so the agreement with the measured spectra is worse than in the case of the SM. This is also true for any other solutions corresponding to complex values of S_R and S_L.

On the other hand, contributions to $\bar{B} \to D\tau^-\bar{\nu}_\tau$ decays proceeding via P-wave tend to shift the expected q^2 spectra to lower values. Thus, NP processes with spin 1 could simultaneously explain the excess in $\mathcal{R}(D^{(*)})$ [21,45] and improve the agreement with the measured q^2 distributions.

TABLE IX. Maximum p value for the q^2 distributions in Fig. 23 corresponding to the variations due to the systematic uncertainties.

	$\bar{B} \to D\tau^-\bar{\nu}_\tau$	$\bar{B} \to D^*\tau^-\bar{\nu}_\tau$
SM	83.1%	98.8%
$\tan\beta/m_{H^\pm} = 0.30\ {\rm GeV}^{-1}$	95.7%	98.9%
$\tan\beta/m_{H^\pm} = 0.45\ {\rm GeV}^{-1}$	0.4%	97.9%

X. CONCLUSIONS

In summary, we have measured the ratios $\mathcal{R}(D^{(*)}) = \mathcal{B}(\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell)$ based on the full *BABAR* data sample, resulting in

$$\mathcal{R}(D) = 0.440 \pm 0.058 \pm 0.042,$$
$$\mathcal{R}(D^*) = 0.332 \pm 0.024 \pm 0.018,$$

where the first uncertainty is statistical and the second is systematic. These results supersede the previous *BABAR* measurements [14]. Improvements of the event selection have increased the reconstruction efficiency of signal events by more than a factor of 3, and the overall statistical uncertainty has been reduced by more than a factor of 2.

Table X shows the results of previous $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ analyses. In 2007 and 2010, the Belle collaboration measured the absolute $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ branching fractions which we translate to $\mathcal{R}(D^{(*)})$ with $\mathcal{B}(B^- \to D^0\ell^-\bar{\nu}_\ell) = (2.26 \pm 0.11)\%$ [12] and $\mathcal{B}(B^0 \to D^{*+}\ell^-\bar{\nu}_\ell) = (4.59 \pm 0.26)\%$ [48]. For the translation of $\mathcal{R}(D^*)$, we choose Belle's measurement of the branching fraction, instead of the world average, because of the current large spread of measured values. For Belle 2009, we average the results for B^0 and B^- decays.

The values measured in this analysis are compatible with those measured by the Belle Collaboration, as illustrated in Fig. 24.

The results presented here exceed the SM predictions of $\mathcal{R}(D)_{\rm SM} = 0.297 \pm 0.017$ and $\mathcal{R}(D^*)_{\rm SM} = 0.252 \pm 0.003$ by 2.0σ and 2.7σ, respectively. The combined significance of this disagreement, including the negative correlation between $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$, is 3.4σ. Together with the measurements by the Belle Collaboration, which also exceed the SM expectations, this could be an indication of NP processes affecting $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays.

These results are not compatible with a charged Higgs boson in the type II 2HDM, and, together with $B \to X_s\gamma$ measurements, exclude this model in the full $\tan\beta$-$m_{H^\pm}$ parameter space. More general charged Higgs models, or NP contributions with nonzero spin, are compatible with the measurements presented here.

An analysis of the efficiency corrected q^2 spectra of $\bar{B} \to D\tau^-\bar{\nu}_\tau$ and $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decays shows good agreement with the SM expectations, within the estimated uncertainties. The combination of the measured values of $\mathcal{R}(D^{(*)})$ and the q^2 spectra exclude a significant portion of the type III 2HDM parameter space. Charged Higgs contributions with small scalar terms, $|S_R + S_L| < 1.4$, are compatible with the measured $\mathcal{R}(D^{(*)})$ and q^2 distributions, but NP contributions with spin 1 are favored by data.

TABLE X. Previous measurements of $\mathcal{R}(D^{(*)})$.

Measurement	$\mathcal{R}(D)$	$\mathcal{R}(D^*)$
Belle 2007 [13]	...	$0.44 \pm 0.08 \pm 0.08$
BABAR 2008 [14]	$0.42 \pm 0.12 \pm 0.05$	$0.30 \pm 0.06 \pm 0.02$
Belle 2009 [15]	$0.59 \pm 0.14 \pm 0.08$	$0.47 \pm 0.08 \pm 0.06$
Belle 2010 [16]	$0.34 \pm 0.10 \pm 0.06$	$0.43 \pm 0.06 \pm 0.06$

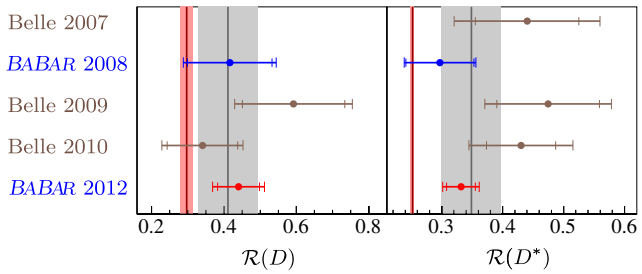

FIG. 24 (color online). Comparison of the previous measurements of $\mathcal{R}(D^{(*)})$ with statistical and total uncertainties (Table X) with this measurement (*BABAR* 2012). The vertical bands represent the average of the previous measurements (light shading) and SM predictions (dark shading), separately for $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$. The widths of the bands represents the uncertainties.

ACKNOWLEDGMENTS

The concept for this analysis is to a large degree based on earlier *BABAR* work and we acknowledge the guidance provided by M. Mazur. The authors consulted with theorists A. Datta, S. Westhoff, S. Fajfer, J. Kamenik, and I. Nišandžić on the calculations of the charged Higgs contributions to the decay rates. 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 relied 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'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 (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).

APPENDIX: SYSTEMATIC UNCERTAINTIES ON THE q^2 SPECTRA

To assess the systematic uncertainty on the measured q^2 distributions of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays, we examine their sensitivity to the estimated contributions from background and normalization events. The q^2 distributions of signal and the various backgrounds are presented in Fig. 25 (left). There is good agreement between the data and the background contributions as derived from the isospin-constrained fit. To further examine the shape of the fixed contributions from $B\bar{B}$ and continuum background, we show two comparisons with data control samples: one for medium values of $E_{\rm extra}$ in the $m_{\rm ES}$ peak regions without

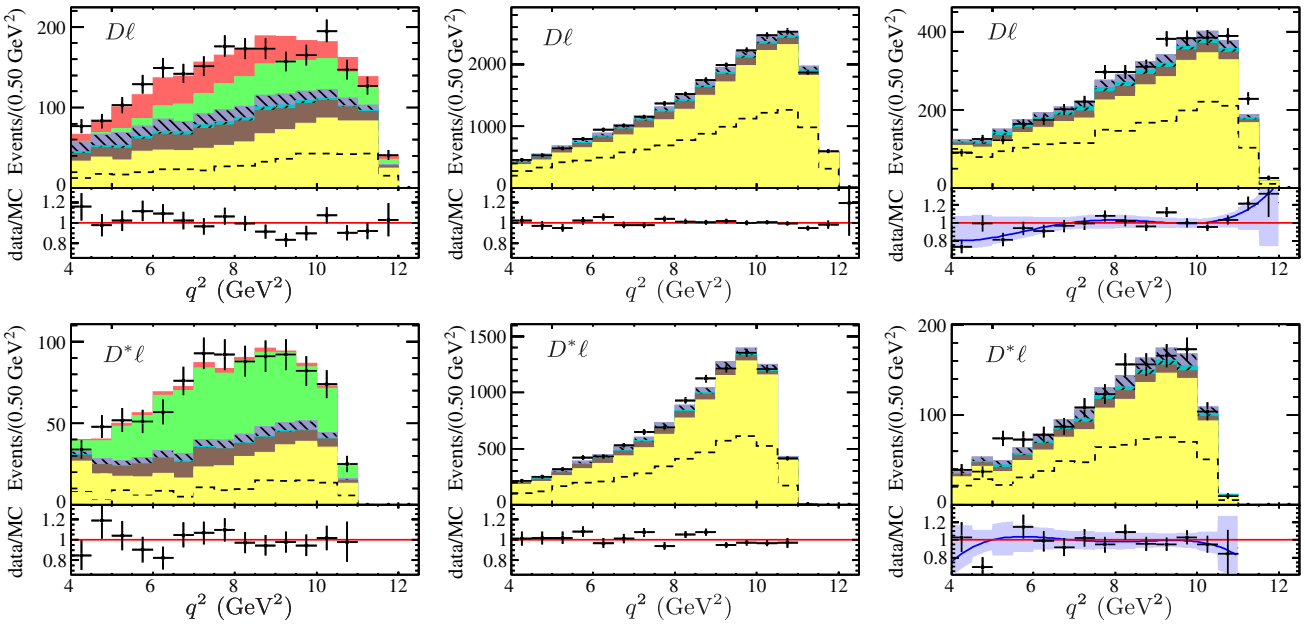

FIG. 25 (color online). Assessment of the uncertainties on the q^2 distributions of background events with $m^2_{\rm miss} > 1.5$ GeV^2. Left: results of the isospin-constrained fit for the SM. Center: sample with $0.5 < E_{\rm extra} < 1.2$ GeV and $5.27 < m_{\rm ES} < 5.29$ GeV. Right: sample satisfying the BDT requirements in the $5.20 < m_{\rm ES} < 5.26$ GeV region. The data/MC plots show a fourth order polynomial fit and the total systematic uncertainty considered. The simulation in the control samples is normalized to the number of events in data. See Fig. 15 for a legend.

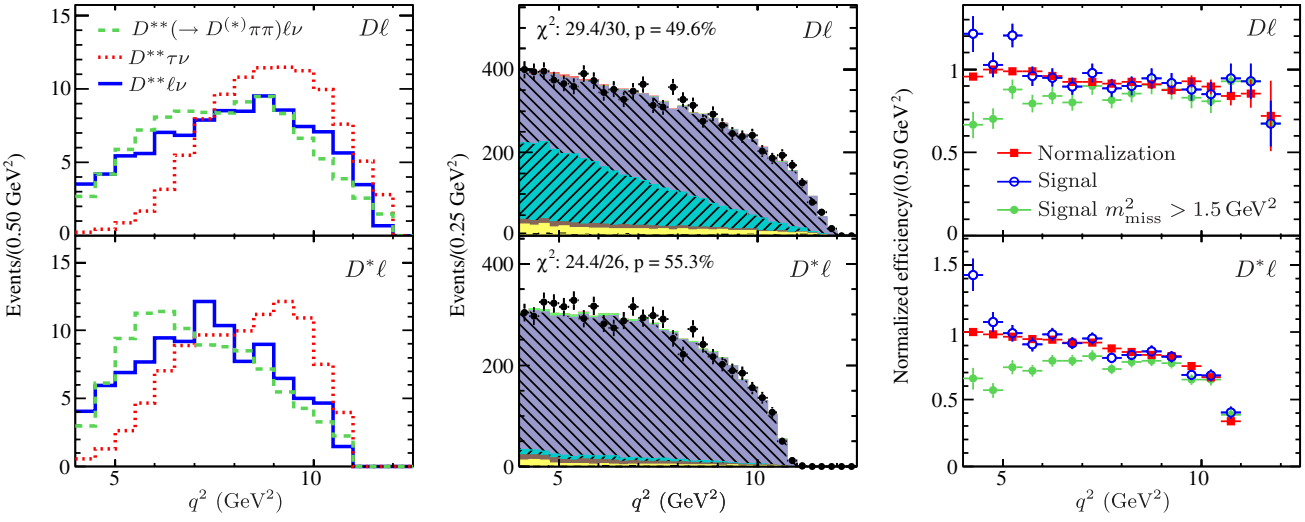

FIG. 26 (color online). Left: q^2 distributions for the different $\bar{B} \to D^{**}(\tau^-/\ell^-)\bar{\nu}$ contributions, all normalized to 100 events. Center: q^2 distributions for events with $m^2_{\rm miss} < 1.5$ GeV^2 scaled to the results of the isospin-constrained fit for the SM. See Fig. 15 for a legend. Right: q^2 dependence of the efficiency. The scale for the efficiency of the normalization decays is chosen so that the maximum value is 1. The efficiency data for the signal are adjusted so that they overlap with the data for normalization decays in the central part of the q^2 range. The signal efficiencies with and without the $m^2_{\rm miss}$ selection have the same scale.

the BDT requirements imposed, and the other for the $m_{\rm ES}$ sidebands with the BDT requirements. While the first sample shows excellent agreement over the full q^2 range, the smaller second sample shows some deviations at low and high q^2. We approximate the deviation of the data from the simulation by a fourth order polynomial, and we adopt this difference plus the statistical uncertainty of each bin as the overall uncertainty of the $B\bar{B}$ and continuum backgrounds. We conservatively consider it uniformly distributed between the limits of the band shown in Fig. 25 and uncorrelated between different bins.

The systematic uncertainty on the shape of the q^2 distribution of $\bar{B} \to D^{**}(\tau^-/\ell^-)\bar{\nu}$ decays is estimated by varying the relative abundance of the contributions shown in Fig. 26. We allow a variation of $\mathcal{R}(D^{**})$, the ratio of $\bar{B} \to D^{**}\tau^-\bar{\nu}_\tau$ decays to $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ decays, between -20% and $+50\%$. We also allow a contribution of up to 30% of $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ decays with the D^{**} decaying into $D^{(*)}\pi^+\pi^-$. In addition, we assume a $\pm 15\%$ variation of the total $\bar{B} \to D^{**}(\tau^-/\ell^-)\bar{\nu}$ yield.

The q^2 spectrum of normalization decays, both well reconstructed and cross-feed $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays, is well described by the simulation, see Fig. 26. Given that the normalization decays are well understood theoretically, we adopt the statistical uncertainty of the simulated distributions as the overall uncertainty of this contribution. Except for $q^2 < 5\ {\rm GeV}^2$, where the rate of signal decays is highly suppressed, the efficiency and detector effects are very similar for signal and normalization. Thus, we also derive the overall uncertainty from the statistical uncertainty of the simulated signal q^2 distributions.

Since it is not feasible to repeat the $m^2_{\rm miss}$-$|\boldsymbol{p}^*_\ell|$ fit for each variation of the background contributions, we adopt the following procedure to account for the impact of these changes on the χ^2: for each of the three q^2 distributions in Fig. 23 and each variation of the background components, we determine the $\bar{B} \to D\tau^-\bar{\nu}_\tau$ and $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ yields by a fit that minimizes the χ^2 of those distributions.

[1] P. Heiliger and L. Sehgal, Phys. Lett. B **229**, 409 (1989).
[2] J. G. Körner and G. A. Schuler, Z. Phys. C **46**, 93 (1990).
[3] D. S. Hwang and D.-W. Kim, Eur. Phys. J. C **14**, 271 (2000).
[4] Y. Amhis *et al.* (Heavy Flavor Averaging Group), arXiv:1207.1158.
[5] M. Tanaka, Z. Phys. C **67**, 321 (1995).
[6] H. Itoh, S. Komine, and Y. Okada, Prog. Theor. Phys. **114**, 179 (2005).
[7] U. Nierste, S. Trine, and S. Westhoff, Phys. Rev. D **78**, 015006 (2008).
[8] M. Tanaka and R. Watanabe, Phys. Rev. D **82**, 034027 (2010).
[9] S. Fajfer, J. F. Kamenik, and I. Nišandžić, Phys. Rev. D **85**, 094025 (2012).
[10] Throughout this paper, ℓ refers only to the light leptons e and μ, $D^{(*)}$ refers to a D or a D^* meson, and charge-conjugate decay modes are implied.
[11] M. Antonelli, D. M. Asner, D. Bauer, T. Becher, M. Beneke, A. J. Bevan, M. Blanke, C. Bloise, M. Bona, and A. Bondar, Phys. Rep. **494**, 197 (2010).
[12] K. Nakamura *et al.* (Particle Data Group), J. Phys. G **37**, 075021 (2010).
[13] A. Matyja *et al.* (Belle Collaboration), Phys. Rev. Lett. **99**, 191807 (2007).
[14] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **100**, 021801 (2008).
[15] I. Adachi *et al.* (Belle Collaboration), arXiv:0910.4301.
[16] A. Bozek *et al.* (Belle Collaboration), Phys. Rev. D **82**, 072005 (2010).
[17] J. Lees *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **109**, 101802 (2012).
[18] K. Hagiwara, A. D. Martin, and M. Wade, Nucl. Phys. **B327**, 569 (1989).
[19] I. Caprini, L. Lellouch, and M. Neubert, Nucl. Phys. **B530**, 153 (1998).
[20] J. F. Kamenik and F. Mescia, Phys. Rev. D **78**, 014003 (2008).
[21] A. Datta, M. Duraisamy, and D. Ghosh, Phys. Rev. D **86**, 034027 (2012).
[22] Z.-Z. Xing, H. Zhang, and S. Zhou, Phys. Rev. D **77**, 113016 (2008).
[23] M. Misiak *et al.*, Phys. Rev. Lett. **98**, 022002 (2007).
[24] Y.-S. Tsai, Phys. Rev. D **4**, 2821 (1971); **13**, 771(E) (1976).
[25] B. Aubert *et al.* (*BABAR* Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 1 (2002).
[26] J. Seeman, in 11th European Particle Accelerator Conference (EPAC 2008), Genoa, Italy, 23-28 June 2008 (to be published).
[27] J. Lees *et al.* (*BABAR* Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **726**, 203 (2013).
[28] D. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).
[29] T. Sjostrand, Comput. Phys. Commun. **82**, 74 (1994).
[30] S. Agostinelli *et al.* (GEANT4), Nucl. Instrum. Methods Phys. Res., Sect. A **506**, 250 (2003).
[31] E. Barberio and Z. Was, Comput. Phys. Commun. **79**, 291 (1994).
[32] D. Scora and N. Isgur, Phys. Rev. D **52**, 2783 (1995).
[33] N. Isgur and M. B. Wise, Phys. Lett. B **237**, 527 (1990).
[34] A. K. Leibovich, Z. Ligeti, I. W. Stewart, and M. B. Wise, Phys. Rev. D **57**, 308 (1998).
[35] P. Speckmayer, A. Hocker, J. Stelzer, and H. Voss, J. Phys. Conf. Ser. **219**, 032057 (2010).

[36] K. S. Cranmer, Comput. Phys. Commun. **136**, 198 (2001).
[37] A. Bowman and A. Azzalini, *Applied Smoothing Techniques for Data Analysis* (Clarendon Press, Oxford, 1997).
[38] B. P. Roe, *Probability and Statistics in Experimental Physics* (Springer, Berlin, 2001), 2nd ed., p. 29.
[39] J. L. Goity and W. Roberts, Phys. Rev. D **51**, 3459 (1995).
[40] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **104**, 011802 (2010).
[41] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **79**, 012002 (2009).
[42] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **77**, 032002 (2008).
[43] In this paper, the significance of an observation with probability p is expressed by the number of standard deviations σ of a one-dimensional Gaussian function for this probability. The shaded bands in Figs. 17, 21, and 22, correspond to p values of 0.683, 0.955, 0.997 and so on.
[44] J. A. Bailey *et al.*, Phys. Rev. Lett. **109**, 071802 (2012).
[45] D. Becirevic, N. Kosnik, and A. Tayduganov, Phys. Lett. B **716**, 208 (2012).
[46] S. Fajfer, J. F. Kamenik, I. Nisandzic, and J. Zupan, Phys. Rev. Lett. **109**, 161801 (2012).
[47] A. Crivellin, C. Greub, and A. Kokulu, Phys. Rev. D **86**, 054014 (2012).
[48] W. Dungel *et al.* (Belle Collaboration), Phys. Rev. D **82**, 112007 (2010).