Search for Rare or Forbidden Decays of the $D^0$ Meson

J. P. Lees,1 V. Poireau,1 E. Grauges,2 A. Palano,3 G. Eigen,4 D. N. Brown,5 Yu. G. Kolomensky,5 M. Fritsch,5 H. Koch,6 T. Schroeder,8 R. Cheaib,7b C. Hearty,7a,7b T. S. Mattisson,7b J. A. McKenna,7b R. Y. So,7b V. E. Blinov,8a,8b,c A. R. Buzyaev,8a V. P. Druzhinin,8a,b V. B. Golubev,8a,b E. A. Kozyrev,8a,b E. A. Kravchenko,8a,b A. P. Onuchin,8a S. I. Serednyakov,8a,b Yu. I. Skovpen,8a,b E. P. Solodov,8a,b K. Yu. Todyshev,8a,b A. J. Lankford,9 B. Dey,10 J. W. Gary,10 O. Long,10 A. M. Eissner,11 W. S. Lockman,11 W. Panduro Vazquez,11 D. S. Chao,11 C. H. Cheng,12 B. Echenard,12 K. T. Flood,12 D. G. Hitlin,12 J. Kim,12 Y. Li,12 T. S. Misajishita,12 P. Ongmongkolkul,12 F. C. Porter,12 M. Röhrken,12 Z. Huard,13 B. T. Meadows,13 B. G. Pushpawela,13 M. D. Sokoloff,13 L. Sun,13 J. M. LoSecco,41 K. Honscheid,42 R. Kass,42 A. Gaz,43a M. Margoni,43a,43b M. Posocco,43a G. Simi,43a,43b F. Simonetto,43a,43b E. Luppi,16a,16b V. Santoro,16a C. Chen,23 J. Cochran,23 S. Prell,23 A. V. Gritsan,24 N. Arnaud,25 M. Davier,25 A. M. Lutz,25 G. Wormser,25 D. J. Lannon,26 D. M. Wright,26 J. P. Coleman,27 E. Gabathuler,27 D. E. Hutchcroft,27 D. J. Payne,27 C. Touramanis,27 A. J. Bevan,28 R. F. Wilson,28 R. Sacco,28 G. Cowan,29 Sw. Banerjee,30 C. L. Davis,30 M. K. Sullivan,52 J. Va...
We present a search for nine lepton-number-violating and three lepton-flavor-violating neutral charm decays of the type $D^0 \rightarrow h^- h^+ \ell^+\ell^−$ and $D^0 \rightarrow h^- h^+ \ell^\pm\ell^\mp$, where $h$ and $h'$ represent a $K$ or $\pi$ meson and $\ell$ and $\ell'$ an electron or muon. The analysis is based on $468 \text{ fb}^{-1}$ of $e^+e^−$ annihilation data collected at or close to the $\Upsilon(4S)$ resonance with the BABAR detector at the SLAC National Accelerator Laboratory. No significant signal is observed for any of the twelve modes, and we establish 90% confidence level upper limits on the branching fractions in the range $(1.0−30.6) \times 10^{-7}$. The limits are between 1 and 3 orders of magnitude more stringent than previous measurements.

DOI: 10.1103/PhysRevLett.124.071802

Lepton-flavor-violating and lepton-number-violating neutral charm decays can be used to investigate physics beyond the standard model (SM) of particle physics. A potential set of decays for study are of the form $D^0 \rightarrow h^- h^+ \ell^+\ell^−$ and $D^0 \rightarrow h^- h^+ \ell^\pm\ell^\mp$, where $h$ and $h'$ represent a $K$ or $\pi$ meson and $\ell$ and $\ell'$ an electron or muon [1].

The $D^0 \rightarrow h^- h^+ \ell^+\ell^−$ decay modes with two opposite-charge, different-flavor leptons in the final state are lepton-flavor-violating (LFV). They are essentially prohibited in the SM because they can occur only through lepton mixing [2]. The $D^0 \rightarrow h^- h^+ \ell^+\ell^−$ decay modes with two same-charge leptons are both lepton-flavor violating and lepton-number violating (LNV) and are forbidden in the SM in low-energy collisions or decays. However, LNV processes can occur in extremely high-energy or high-density interactions [3].

Lepton-number violation is a necessary condition for leptogenesis as an explanation of the baryon asymmetry of the Universe [4]. If neutrinos are Majorana fermions, the neutrino and antineutrino are the same particle, and some LNV processes become possible [5]. Many models beyond the SM allow lepton-number violation. Most models have made predictions for, or used constraints from, three-body decays of the form $D \rightarrow M\ell\ell'$ or $B \rightarrow M\ell\ell'$, where $M$ is a meson [6–12]. However, some models that consider LFV and LNV four-body charm decays predict branching fractions up to $\mathcal{O}(10^{-6})$ to $\mathcal{O}(10^{-5})$, approaching those accessible with current data [11–13].

The branching fractions $B(D^0 \rightarrow h^- h^+ \mu^+\mu^-)$ and $B(D^0 \rightarrow K^- \pi^+ e^+\mu^-)$ have recently been determined to be $\mathcal{O}(10^{-7})$ to $\mathcal{O}(10^{-8})$ [14–16], compatible with SM predictions [17,18]. The most stringent existing upper limits on the branching fractions for the LFV and LNV four-body decays of the type $D^0 \rightarrow h^- h^+ \ell^\pm\ell^\mp$ and $D^0 \rightarrow h^- h^+ \ell^\pm\ell^\mp$ are in the range $(0.3−55.3) \times 10^{-5}$ at the 90% confidence level (C.L.) [19–21]. For the LFV decays $D^0 \rightarrow V\ell^\pm\ell^−$, where $V$ is an intermediate resonance such as a $\rho$ or $\phi$ meson decaying to $h^- h^+$, the 90% C.L. limits are in the range $(3.4−118) \times 10^{-5}$ [19,20,22]. Searches for Majorana neutrinos in $D^+_s \rightarrow \pi^- \mu^+\mu^-$ decays have placed upper limits on the branching fractions as low as $2.2 \times 10^{-8}$ at the 90% C.L. [23].

In this report we present a search for nine $D^0 \rightarrow h^- h^+ \ell^+\ell^−$ LFV decays, with data recorded with the BABAR detector at the PEP-II asymmetric-energy $e^+e^-$ collider operated at the SLAC National Accelerator Laboratory. The data sample corresponds to $424 \text{ fb}^{-1}$ of $e^+e^−$ collisions collected at the center-of-mass (c.m.) energy of the $\Upsilon(4S)$ resonance (on peak) and an additional $44 \text{ fb}^{-1}$ of data collected $40 \text{ MeV}$ below the $\Upsilon(4S)$ resonance (off peak) [24]. The branching fractions for signal modes with zero, one, or two kaons in the final state are measured relative to the normalization decays $D^0 \rightarrow \pi^\pm \pi^\mp \pi^\mp \pi^\mp$, $D^0 \rightarrow K^- \pi^- \pi^0 \pi^0$, and $D^0 \rightarrow K^- K^- \pi^+ \pi^−$, respectively. The $D^0$ mesons are identified from the decay $D^+ \rightarrow D^0 \ell^\mp$ produced in $e^-e^-\rightarrow c\bar{c}$ events.

The BABAR detector is described in detail in Refs. [25,26]. Charged particles are reconstructed as tracks with a five-layer silicon vertex detector and a 40-layer drift chamber inside a 1.5 T solenoidal magnet. An electromagnetic calorimeter comprised of 6580 CsI(Tl) crystals is used to identify electrons and photons. A ring-imaging Cherenkov detector is used to identify charged hadrons and to provide additional lepton
identification information. Muons are identified with an instrumented magnetic-flux return.

Monte Carlo (MC) simulation is used to investigate sources of background contamination, evaluate selection efficiencies, cross-check the selection procedure, and for studies of systematic effects. The signal and normalization channels are simulated with the EVTGEN package [27]. We generate the signal channel decays uniformly throughout the four-body phase space, while the normalization modes include two-body and three-body intermediate resonances, as well as nonresonant decays. We also generate $e^+e^- \rightarrow q\bar{q}$ ($q = u, d, s, c$), dimuon, Bhabha elastic $e^+e^-$ scattering, $B\bar{B}$ background, and two-photon events [28,29]. Final-state radiation is generated using PHOTOS [30]. The detector response is simulated with GEANT4 [31,32].

In order to optimize the event reconstruction, candidate selection criteria, multivariate analysis training, and fit procedure, a rectangular area in the $m(D^0) \times \Delta m = m(D^{+}) - m(D^0)$ plane is defined, where $m(D^+)$ and $m(D^0)$ are the reconstructed masses of the $D^+$ and $D^0$ candidates, respectively. This optimization region is kept hidden (blinded) in data until the analysis steps are finalized. The blinded region is approximately 3 times the width of the $\Delta m$ and $m(D^0)$ resolutions. The $\Delta m$ region is $0.1447 < \Delta m < 0.1462$ GeV/$c^2$ for all modes. The $m(D^0)$ signal peak distribution is asymmetric due to bremsstrahlung emission. The upper bound $m(D^0)$ on the blinded region is 1.874 GeV/$c^2$ for all modes, and the lower bound is 1.848, 1.852, and 1.856 GeV/$c^2$ for modes with two, one, or no electrons, respectively.

Events are required to contain at least five charged tracks. Particle identification (PID) criteria are applied to all the charged tracks to identify kaons, pions, electrons, and muons [26,33]. For modes with two kaons in the final state, the PID requirement on the kaons is relaxed compared to the single-kaon modes. This increases the reconstruction efficiency for the modes with two kaons, with little increase in backgrounds or misidentified candidates. The PID efficiency depends on the track momentum, and is in the range 0.96–0.99 for electrons, 0.60–0.95 for muons, and 0.90–0.98 for kaons and pions. The misidentification probability is typically less than 0.03 for all selection criteria, except for the pion selection criteria, where the muon misidentification rate can be as high as 0.35 at low momentum.

Candidate $D^0$ mesons are formed from four charged tracks reconstructed with the appropriate mass hypotheses for the signal and normalization decays. The four tracks must form a good-quality vertex with a $\chi^2$ probability for the vertex fit greater than 0.005. A bremsstrahlung energy recovery algorithm is applied to electrons [16]. The invariant mass of any $e^+e^-$ pair is required to be greater than 0.1 GeV/$c^2$. For the normalization modes, the reconstructed $D^0$ meson mass is required to be in the range $1.81 < m(D^0) < 1.91$ GeV/$c^2$, while for the signal modes, $m(D^0)$ must be in the blinded $m(D^0)$ range defined above.

The candidate $D^{++}$ is formed by combining the $D^0$ candidate with a charged pion with a momentum in the laboratory frame greater than 0.1 GeV/$c$. For the normalization mode $D^0 \rightarrow K^-\pi^+\pi^+\pi^-$, this pion is required to have a charge opposite to that of the final-state kaon. A vertex fit is performed with the $D^0$ mass constrained to its known value [20] and the requirement that the $D^0$ meson and the pion originate from the PEP-II interaction region. The $\chi^2$ probability of the fit is required to be greater than 0.005. For signal modes with two kaons, the mass difference $\Delta m$ is required to be $0.141 < \Delta m < 0.201$ GeV/$c^2$. Signal modes with fewer than two kaons have almost no candidates beyond $\Delta m = 0.149$ GeV/$c^2$, and the range for these modes is restricted to $0.141 < \Delta m < 0.149$ GeV/$c^2$.

After the application of the $D^{++}$ vertex fit, the $D^0$ candidate momentum in the c.m. system $p^*$ is required to be greater than 2.4 GeV/$c$. This removes most sources of combinatorial background and also charm hadrons produced in $B$ decays, which are limited to $p^* \lesssim 2.2$ GeV/$c$ [34]. Remaining backgrounds are mainly radiative Bhabha scattering, initial-state radiation, and two-photon events, which are all rich in electrons and positrons. We suppress these backgrounds by requiring that the PID signatures of the hadron candidates be inconsistent with the electron hypothesis.

Hadronic $D^0$ decays with large branching fractions, where one or more charged tracks are misidentified as leptons, will usually have reconstructed $D^0$ masses well away from the known $D^0$ mass [20]. To ensure rejection of this type of background, the $D^0$ candidate is also reconstructed assuming the kaon or pion mass hypothesis for the lepton candidates. If the resulting $D^0$ candidate mass is within 20 MeV/$c^2$ of the known $D^0$ mass, and if $|\Delta m| < 2$ MeV/$c^2$, the event is discarded. After these criteria are applied, the background from these hadronic decays is negligible.

Two particular sources of background are semileptonic charm decays in which a charged hadron is misidentified as a lepton; and charm decays in which the final state contains a neutral particle or more than four charged tracks. In both cases, tracks can be selected from elsewhere in the event to form a $D^0$ candidate. To reject these backgrounds, a multivariate selection based on a Fisher discriminant is applied [35]. The discriminant uses nine input observables: the momenta of the four tracks used to form the $D^0$ candidate; the thrust and sphericity of the $D^{++}$ candidate [36]; the angle between the $D^{++}$ meson candidate sphericity axis and the sphericity axis defined by the charged particles in the rest of the event (ROE); the angle between the $D^{++}$ meson candidate thrust axis and the thrust axis defined by the charged particles in the ROE; and the second Fox-Wolfram moment [37] calculated from the entire event using both charged and neutral particles. The input
observables are determined in the laboratory frame after the application of the $D^{+}$ vertex fit. The discriminant is trained and tested using MC for the signal modes; for the background, data outside the optimization region, together with $e^+e^-\rightarrow c\bar{c}$ MC samples, are used. The training is performed independently for each signal mode. A requirement on the Fisher discriminant output is chosen such that approximately 90% of the simulation signal candidates are accepted. Depending on the signal mode, this rejects 30% to 50% of the background in data.

The cross feed to one signal mode from the other eleven is estimated from MC samples to be $\lesssim 0.5\%$ in all cases, assuming equal branching fractions for all signal modes. The cross feed to a specific normalization mode from the other two normalization modes is predicted from simulation to be $\lesssim 0.7\%$, where the branching fractions are taken from Ref. [20]. Multiple candidates occur in 4.5% to 7.1% of simulated signal events and in 2.4% to 4.4% of the normalization events in data. If two or more candidates are found in an event, the one with the highest $D^{+}$ vertex $\chi^2$ probability is selected. After the application of all selection criteria and corrections for small differences between data and MC simulation in tracking and PID performance derived from high purity control samples [26], the reconstruction efficiency $\epsilon_{\text{sig}}$ for the simulated signal decays is between 3.2% and 6.2%, depending on the mode. For the normalization decays, the reconstruction efficiency $\epsilon_{\text{norm}}$ is between 19.2% and 24.7%. The difference between $\epsilon_{\text{sig}}$ and $\epsilon_{\text{norm}}$ is mainly due to the momentum dependence of the lepton PID [26].

The signal mode branching fraction $B_{\text{sig}}$ is determined relative to that of the normalization decay using

$$B_{\text{sig}} = \frac{N_{\text{sig}} \epsilon_{\text{norm}} L_{\text{norm}}}{N_{\text{norm}} \epsilon_{\text{sig}} L_{\text{sig}}} B_{\text{norm}},$$

where $B_{\text{norm}}$ is the normalization mode branching fraction [20], and $N_{\text{sig}}$ and $N_{\text{norm}}$ are the fitted yields of the signal and normalization mode decays, respectively. The symbols $L_{\text{sig}}$ and $L_{\text{norm}}$ represent the integrated luminosities of the data samples used for the signal ($468.2 \pm 2.0$ fb$^{-1}$) and the normalization decays ($39.3 \pm 0.2$ fb$^{-1}$), respectively [24]. For the signal modes, we use the on-peak and off-peak data samples, while the normalization modes use only a subset of the off-peak data.

Each normalization mode yield $N_{\text{norm}}$ is extracted by performing an extended two-dimensional unbinned maximum likelihood (ML) fit [38] to the observables $\Delta m$ and $m(D^0)$ in the range $0.141 < \Delta m < 0.201$ GeV/c$^2$ and $1.81 < m(D^0) < 1.91$ GeV/c$^2$. The measured $\Delta m$ and $m(D^0)$ values are not correlated and are treated as independent observables in the fits. The probability density functions (PDFs) in the fits depend on the normalization mode and use sums of multiple Crujiff [16] and Crystal Ball [39] functions in both $\Delta m$ and $m(D^0)$. The functions for each observable use a common mean. The background is modeled with an ARGUS threshold function [40] for $\Delta m$ and a Chebyshev polynomial for $m(D^0)$. The ARGUS end point parameter is fixed to the kinematic threshold for a $D^{+} \rightarrow D^0 \pi^+$ decay. All other PDF parameters, together with the normalization mode and background yields, are allowed to vary in the fit. The fitted yields and reconstruction efficiencies for the normalization modes are given in Table I.

| Decay mode | $N_{\text{norm}}$ (candidates) | Systematic (%) | $\epsilon_{\text{norm}}$ (%) |
|------------|-------------------------------|---------------|-----------------|
| $D^0 \rightarrow K^-\pi^+\pi^+\pi^-$ | 260870 ± 520 | 4.7 | 20.1 ± 0.2 |
| $D^0 \rightarrow K^-K^-\pi^+\pi^-$ | 8480 ± 110 | 6.6 | 19.2 ± 0.2 |
| $D^0 \rightarrow \pi^-\pi^+\pi^+\pi^-$ | 28470 ± 220 | 6.8 | 24.7 ± 0.2 |

Each signal mode yield $N_{\text{sig}}$ is extracted by performing the ML fit with the single observable $\Delta m$ in the range $0.141 < \Delta m < 0.201$ GeV/c$^2$ for signal modes with two kaons and $0.141 < \Delta m < 0.149$ GeV/c$^2$ for all other signal modes. The signal PDF is a Crujiff function with parameters obtained by fitting the signal MC. The background is modeled with an ARGUS function with an end point that is set to the same value that is used for the normalization modes. The signal PDF parameters and the end point parameter are fixed in the fit. All other background parameters and the signal and background yields are allowed to vary. Figures 1 and 2 show the results of the fits to the $\Delta m$ distributions for the twelve signal modes.

We test the performance of the ML fit for the normalization modes by generating ensembles of MC samples from the normalization and background PDF distributions. The mean numbers of normalization and background candidates used in the ensembles are taken from the fits to the data. The numbers of background and normalization mode candidates are allowed to fluctuate according to a Poisson distribution and all background and normalization mode PDF parameters are allowed to vary. No significant biases are observed in fitted yields of the normalization modes. The same procedure is repeated for the ML fit to signal modes, with ensembles of MC samples generated from the background PDF distributions only, assuming a signal yield of zero. The signal PDF parameters are fixed to the values used for the fits to the data but the signal yield is allowed to vary. The biases in the fitted signal yields are less than $\pm 0.2$ for all modes, and these are subtracted from the fitted yields before calculating the signal branching fractions.

To cross-check the normalization procedure, the signal modes in Eq. (1) are replaced with the decay $D^0 \rightarrow K^-\pi^+$, which has a well-known branching fraction [20]. The
$D^0 \to K^-\pi^+$ decay is selected using the same criteria as used for the $D^0 \to K^-\pi^+\pi^+\pi^-$ mode, which is used as the normalization mode for this cross-check. The $D^0 \to K^-\pi^+$ signal yield is $1881950 \pm 1380$ with $\epsilon_{\text{sig}} = (27.4 \pm 0.2)\%$. We determine $B(D^0 \to K^-\pi^+) = (3.98 \pm 0.08 \pm 0.10)\%$, where the uncertainties are statistical and systematic, respectively. This is consistent with the current world average of $(3.89 \pm 0.04)\%$ [20]. Similar compatibility with the $B(D^0 \to K^-\pi^+)$ world average, but with larger uncertainties, is observed when the normalization mode $D^0 \to K^-\pi^+\pi^+\pi^-$ in Eq. (1) is replaced with the decay modes $D^0 \to K^-K^+\pi^+\pi^-$ or $D^0 \to \pi^-\pi^+\pi^+\pi^-$. The main sources of systematic uncertainties in the signal yields are associated with the model parametrizations used in the fits to the signal modes and backgrounds, the fit biases, and the limited MC and data sample sizes available for the optimization of the Fisher discriminants. The uncertainties associated with the fit model parametrizations of the signal modes are estimated by repeating the fits with alternative PDFs. This involves swapping the Cruijff and Crystal Ball functions, using Gaussian functions with different asymmetric widths, and changing the number of functions used. For the background, the order of the polynomials is changed and the ARGUS function is replaced by a second-order polynomial. The fits are also performed with the fixed signal parameters allowed to vary within the statistical uncertainties obtained from fits to the signal MC samples. The systematic uncertainty is taken as half the maximum deviation from the default fit.

Systematic uncertainties that impact the calculation of the branching fractions of the signal modes are due to assumptions made about the distributions of the final-state particles in the signal simulation modeling, the model parametrizations used in the fits to the normalization modes, the normalization mode branching fractions, tracking and PID efficiencies, and luminosity. Since the decay mechanism of the signal modes is unknown, we vary the angular distributions of the simulated final-state particles from the $D^0$ signal decay, where the three angular variables are defined following the prescription of Ref. [41]. We weight the reconstruction efficiencies of the phase-space simulation samples as a function of the angular-variable distributions, trying combinations of $\sin$, $\cos$, $\sin^2$, and $\cos^2$ functions. Half the maximum change in the average reconstruction efficiency is assigned as a systematic uncertainty.
TABLE II. Summary of fitted signal yields with statistical and systematic uncertainties, reconstruction efficiencies, branching fractions with statistical and systematic uncertainties, 90% C.L. branching fraction upper limits (U.L.), and the previous limits [20,21]. The branching fraction systematic uncertainties take into account correlations and cancellations between the signal and normalization modes and include the uncertainties in the normalization mode branching fractions.

| Decay mode $D^0 \rightarrow$ | $N_{sig}$ (candidates) | $e_{sig}$ (%) | $B \times 10^{-7}$ | $B$ 90% U.L. ($\times 10^{-7}$) |
|-----------------------------|------------------------|--------------|-------------------|-----------------------------|
| $\pi^+\pi^-e^+e^-$          | $0.22 \pm 0.15 \pm 0.54$ | $4.38 \pm 0.05$ | $0.27 \pm 3.90 \pm 0.67$ | $9.1 \pm 1120$ |
| $\pi^+\mu^+\mu^-$          | $6.69 \pm 4.88 \pm 0.80$ | $4.91 \pm 0.05$ | $7.40 \pm 5.40 \pm 0.91$ | $15.2 \pm 290$ |
| $\pi^+e^+\mu^-$            | $12.42 \pm 5.30 \pm 1.45$ | $4.38 \pm 0.05$ | $15.41 \pm 6.59 \pm 1.85$ | $30.6 \pm 790$ |
| $K^-\pi^+e^+\mu^-$         | $1.37 \pm 6.15 \pm 1.28$ | $4.79 \pm 0.06$ | $1.55 \pm 6.97 \pm 1.45$ | $17.1 \pm 150$ |
| $K^-\pi^+\mu^+\mu^-$       | $0.03 \pm 2.10 \pm 0.40$ | $3.30 \pm 0.05$ | $-0.05 \pm 3.34 \pm 0.64$ | $5.3 \pm 3900$ |
| $K^-\pi^+e^+\mu^+$          | $0.37 \pm 3.96 \pm 2.36$ | $3.48 \pm 0.04$ | $5.84 \pm 5.97 \pm 3.56$ | $21.0 \pm 2180$ |
| $K^-\pi^+e^+\mu^+$          | $2.52 \pm 4.60 \pm 1.35$ | $3.65 \pm 0.05$ | $3.62 \pm 6.61 \pm 1.95$ | $19.0 \pm 5530$ |
| $K^-K^-\mu^+\mu^-$          | $0.30 \pm 1.08 \pm 0.41$ | $3.25 \pm 0.04$ | $0.43 \pm 1.54 \pm 0.58$ | $3.4 \pm 1520$ |
| $K^-K^-e^+e^-$              | $1.09 \pm 1.29 \pm 0.42$ | $6.21 \pm 0.06$ | $-0.81 \pm 0.96 \pm 0.32$ | $1.0 \pm 940$ |
| $K^-K^-\mu^+\mu^-$          | $1.93 \pm 1.92 \pm 0.83$ | $4.63 \pm 0.05$ | $1.93 \pm 1.93 \pm 0.84$ | $5.8 \pm 570$ |
| $K^-K^-e^+\mu^+$            | $4.09 \pm 3.00 \pm 1.59$ | $4.83 \pm 0.05$ | $3.93 \pm 2.89 \pm 1.45$ | $10.0 \pm 1800$ |


dedicated effort from the computing organizations that support BABAR. The collaborating institutions wish to thank SLAC for its support and kind hospitality. This work is supported by DOE and NSF (USA), NSERC (Canada), CEA and CNRS-IN2P3 (France), BMBF and DFG (Germany), INFN (Italy), FOM (Netherlands), NFR (Norway), MES (Russia), MINECO (Spain), STFC (United Kingdom), BSF (USA-Israel). Individuals have received support from the Marie Curie EIF (European Union) and the A.P. Sloan Foundation (USA).

1. Deceased.
2. Present address: Wuhan University, Wuhan 430072, China.
3. Present address: Università di Bologna and INFN Sezione di Bologna, I-47921 Rimini, Italy.
4. Present address: University of Huddersfield, Huddersfield HD1 3DH, United Kingdom.
5. Present address: University of South Alabama, Mobile, Alabama 36688, USA.
6. * Also at Università di Sassari, I-07100 Sassari, Italy.
7. † Also at Gran Sasso Science Institute, I-67100 LAquila, Italy.

[1] Charge conjugation is implied throughout.
[2] D. Guadagnoli and K. Lane, Phys. Lett. B 751, 54 (2015).
[3] F. R. Klinkhamer and N. S. Manton, Phys. Rev. D 30, 2212 (1984).
[4] S. Davidson, E. Nardi, and Y. Nir, Phys. Rev. D 466, 105 (2008).
[5] E. Majorana, Nuovo Cimento 14, 171 (1937).
[6] A. Paul, I. I. Bigi, and S. Recksiegel, Phys. Rev. D 43, 114006 (2011).
[7] A. Paul, A. de la Puente, and I. I. Bigi, Phys. Rev. D 90, 014005 (2014).
[8] G. Burdman, E. Golowich, J. A. Hewett, and S. Pakvasa, Phys. Rev. D 66, 014009 (2002).
[9] S. Fajfer and S. Prelovšek, Phys. Rev. D 73, 054026 (2006).
[10] S. Fajfer, N. Košnik, and S. Prelovšek, Phys. Rev. D 76, 074010 (2007).
[11] A. Atre, T. Han, S. Pascoli, and B. Zhang, J. High Energy Phys. 05 (2009) 030.
[12] H. Yuan, T. Wang, G.-L. Wang, W.-L. Ju, and J.-M. Zhang, J. High Energy Phys. 08 (2013) 066.
[13] D. Hai-Rong, F. Feng, and L. Hai-Bo, Chin. Phys. C 39, 013101 (2015).
[14] R. Aaij et al. (LHCb Collaboration), Phys. Lett. B 757, 558 (2016).
[15] R. Aaij et al. (LHCb Collaboration), Phys. Rev. Lett. 119, 181805 (2017).
[16] J. P. Lees et al. (BABAR Collaboration), Phys. Rev. Lett. 122, 081802 (2019).
[17] L. Cappiello, O. Cata, and G. D’Ambrosio, J. High Energy Phys. 04 (2013) 135.
[18] S. de Boer and G. Hiller, Phys. Rev. D 98, 035041 (2018).
[19] E. M. Aitala et al. (E791 Collaboration), Phys. Rev. Lett. 86, 3969 (2001).
[20] M. Tanabashi et al. (Particle Data Group), Phys. Rev. D 98, 030001 (2018).
[21] M. Ablikim et al. (BESIII Collaboration), Phys. Rev. D 99, 112002 (2019).
[22] A. Freyberger et al. (CLEO Collaboration), Phys. Rev. Lett. 76, 3065 (1996).
[23] R. Aaij et al. (LHCb Collaboration), Phys. Lett. B 724, 203 (2013).
[24] J. P. Lees et al. (BABAR Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A 726, 203 (2013).
[25] B. Aubert et al. (BABAR Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A 479, 1 (2002).
[26] B. Aubert et al. (BABAR Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A 729, 615 (2013).
[27] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A 462, 152 (2001).
[28] B. F. L. Ward, S. Jadach, and Z. Was, Nucl. Phys. B, Proc. Suppl. 116, 73 (2003).
[29] T. Sjöstrand, Comput. Phys. Commun. 82, 74 (1994).
[30] P. Golonka and Z. Was, Eur. Phys. J. C 45, 97 (2006).
[31] S. Agostinelli et al. (GEANT 4 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A 506, 250 (2003).
[32] J. Allison, K. Amako, J. Apostolakis, H. Arai, P. Dubois et al. (GEANT 4 Collaboration), IEEE Trans. Nucl. Sci. 53, 270 (2006).
[33] I. Adam et al., Nucl. Instrum. Methods Phys. Res., Sect. A 538, 281 (2005).
[34] B. Aubert et al. (BABAR Collaboration), Phys. Rev. D 69, 111104 (2004).
[35] R. A. Fisher, Ann. Eugen. 7, 179 (1936).
[36] A. De Rujula, J. R. Ellis, E. G. Floratos, and M. K. Gaillard, Nucl. Phys. B138, 387 (1978).
[37] G. C. Fox and S. Wolfram, Nucl. Phys. B149, 413 (1979).
[38] J. P. Lees et al. (BABAR Collaboration), Phys. Rev. D 89, 011102 (2014).
[39] T. Skwarnicki, Ph.D. thesis, Institute of Nuclear Physics, Krakow, DESY-F31-86-02, 1986.
[40] H. Albrecht et al. (ARGUS Collaboration), Phys. Lett. B 241, 278 (1990).
[41] R. Aaij et al. (LHCb Collaboration), Phys. Rev. D 88, 052002 (2013).
[42] T. Allmendinger et al., Nucl. Instrum. Methods Phys. Res., Sect. A 704, 44 (2013).
[43] G. J. Feldman and R. D. Cousins, Phys. Rev. D 57, 3873 (1998).