Search for $D^0-\bar{D}^0$ mixing using semileptonic decay modes

B. Aubert,1 R. Barate,1 D. Boutigny,1 F. Coudert,1† J.-M. Gaillard,1 A. Hicheur,1 Y. Karyotakis,1 J. P. Lees,1† V. Tisserand,1 J. A. Nash,1 M. B. Nikolich,1 G. P. Taylor,1 M. J. Charles,1 G. J. Grenier,1 U. Mallik,1 J. Cochran,1 E. Robutti,1 A. Santroni,1 S. Tosi,1 S. Bailey,1 G. Brandenburg,1 K. S. Chaisanguanthum,1 M. Morii,1 E. Won,1 G. Cibinetto,1 E. Luppi,1 M. Negrini,1 L. Piemontese,1 A. Sarti,1 E. Treadwell,1 F. Anulli,2 R. Baldini-Ferroli,2 U. Nauenberg,2 A. Olivas,2 P. Rankin,2 J. G. Smith,2 J. Zhang,2 L. Zhang,2 A. Chen,3 J. L. Harton,3 A. Soffer,3 A. Dvoretskii,4 D. G. Hitlin,4 I. Narsky,4 T. Piatenko,4 F. C. Porter,4 A. Ryd,4 A. Samuel,4 S. Y. Yang,4 S. Jayatilleke,4 T. W. Beck,5 A. M. Eisner,5 C. A. Heusch,5 J. Kroseberg,5 W. S. Lockman,5 G. Nesom,5 T. Schalk,5 B. A. Schumm,5 B. Aubert,1 R. Barate,1 D. Boutigny,1 F. Coudert,1† J.-M. Gaillard,1 A. Hicheur,1 Y. Karyotakis,1 J. P. Lees,1† V. Tisserand,1 J. A. Nash,1 M. B. Nikolich,1 G. P. Taylor,1 M. J. Charles,1 G. J. Grenier,1 U. Mallik,1 J. Cochran,1 E. Robutti,1 A. Santroni,1 S. Tosi,1 S. Bailey,1 G. Brandenburg,1 K. S. Chaisanguanthum,1 M. Morii,1 E. Won,1 G. Cibinetto,1 E. Luppi,1 M. Negrini,1 L. Piemontese,1 A. Sarti,1 E. Treadwell,1 F. Anulli,2 R. Baldini-Ferroli,2 U. Nauenberg,2 A. Olivas,2 P. Rankin,2 J. G. Smith,2 J. Zhang,2 L. Zhang,2 A. Chen,3 J. L. Harton,3 A. Soffer,3 A. Dvoretskii,4 D. G. Hitlin,4 I. Narsky,4 T. Piatenko,4 F. C. Porter,4 A. Ryd,4 A. Samuel,4 S. Y. Yang,4 S. Jayatilleke,4 T. W. Beck,5 A. M. Eisner,5 C. A. Heusch,5 J. Kroseberg,5 W. S. Lockman,5 G. Nesom,5 T. Schalk,5 B. A. Schumm,5
SEARCH FOR $D^0$-$\bar{D}^0$ MIXING USING SEMILEPTONIC \ldots

PHYSICAL REVIEW D 70 091102

24University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom
25Dipartimento di Fisica and INFN, Università di Ferrara, I-44100 Ferrara, Italy
26Florida A&M University, Tallahassee, Florida 32307, USA
27Laboratori Nazionali di Frascati dell’INFN, I-00044 Frascati, Italy
28Dipartimento di Fisica and INFN, Università di Genova, I-16146 Genova, Italy
29Harvard University, Cambridge, Massachusetts 02138, USA
30Physikalisches Institut, Universität Heidelberg, Philosophenweg 12, D-69120 Heidelberg, Germany
31Imperial College London, London, SW7 2AZ, United Kingdom
32University of Iowa, Iowa City, Iowa 52242, USA
33Iowa State University, Ames, Iowa 50011-3160, USA
34Dipartimento di Fisica and INFN, Università di Perugia, I-06100 Perugia, Italy
35Laboratoire de l’Accélérateur Linéaire, F-91898 Orsay, France
36Lawrence Livermore National Laboratory, Livermore, California 94550, USA
37University of Liverpool, Liverpool L69 7E, United Kingdom
38Queen Mary, University of London, E1 4NS, United Kingdom
39Royal Holloway, University of London, and Bedford New College, Egham, Surrey TW20 0EX, United Kingdom
40University of Louisville, Louisville, Kentucky 40292, USA
41University of Manchester, Manchester M13 9PL, United Kingdom
42University of Maryland, College Park, Maryland 20742, USA
43University of Massachusetts, Amherst, Massachusetts 01003, USA
44Laboratory for Nuclear Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
45McGill University, Montréal, Quebec, Canada H3A 2T8
46Dipartimento di Fisica and INFN, Università di Milano, I-20133 Milano, Italy
47University of Mississippi, University, Mississippi 38667, USA
48Laboratoire René J. A. Lévesque, Université de Montréal, Montréal, Quebec, Canada H3C 3J7
49Mount Holyoke College, South Hadley, Massachusetts 01075, USA
50Dipartimento di Scienze Fisiche and INFN, Università di Napoli Federico II, I-80126, Napoli, Italy
51NIKHEF, National Institute for Nuclear Physics and High Energy Physics, NL-1009 DB Amsterdam, The Netherlands
52University of Notre Dame, Notre Dame, Indiana 46556, USA
53Ohio State University, Columbus, Ohio 43210, USA
54University of Oregon, Eugene, Oregon 97403, USA
55Dipartimento di Fisica and INFN, Università di Padova, I-35131 Padova, Italy
56Laboratoire de Physique Nucléaire et de Hautes Energies, Universités Paris VI et VII, F-75252 Paris, France
57Dipartimento di Elettronica and INFN, Università di Pavia, I-27100 Pavia, Italy
58University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA
59Dipartimento di Fisica, Università di Pisa, Scuola Normale Superiore and INFN, I-56127 Pisa, Italy
60Prairie View A&M University, Prairie View, Texas 77446, USA
61Princeton University, Princeton, New Jersey 08544, USA
62Dipartimento di Fisica and INFN, Università di Roma La Sapienza, I-00185 Roma, Italy
63Universität Rostock, D-18051 Rostock, Germany
64Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, United Kingdom
65DSM/Dapnia, CEA/Saclay, F-91191 Gif-sur-Yvette, France
66University of South Carolina, Columbia, South Carolina 29208, USA
67Stanford Linear Accelerator Center, Stanford, California 94309, USA
68Stanford University, Stanford, California 94305-4060, USA
69State University of New York, Albany, New York 12222, USA
70University of Tennessee, Knoxville, Tennessee 37996, USA
71University of Texas at Austin, Austin, Texas 78712, USA
72University of Texas at Dallas, Richardson, Texas 75083, USA
73Dipartimento di Fisica Sperimentale and INFN, Università di Torino, I-10125 Torino, Italy
74Dipartimento di Fisica and INFN, Università di Trieste, I-34127 Trieste, Italy
75Vanderbilt University, Nashville, Tennessee 37235, USA
76University of Victoria, Victoria, British Columbia, Canada V8W 3P6
77University of Wisconsin, Madison, Wisconsin 53706, USA
78Yale University, New Haven, Connecticut 06511, USA

(Received 15 August 2004; published 10 November 2004)

*Now at Department of Physics, University of Warwick, Coventry, United Kingdom.
†Also with Università della Basilicata, Potenza, Italy.
‡Also with IFIC, Instituto de Física Corpuscular, CSIC-Universidad de Valencia, Valencia, Spain.
§Deceased.
Based on an 87-fb$^{-1}$ data set collected by the BABAR detector at the PEP-II asymmetric-energy $B$ Factory, a search for $D^{0}$-$\bar{D}^{0}$ mixing has been made using the semileptonic decay modes $D^{\ast+}\rightarrow\pi^{+}D^{0}$, $D^{0}\rightarrow K^{(*)}\ell\nu$ (c.c.). The use of these modes allows unambiguous flavor tagging and a combined fit of the $D^{0}$ decay time and $D^{+}$-$\bar{D}^{0}$ mass difference ($\Delta M$) distributions. The high-statistics sample of unmixed semileptonic $D^{0}$ decays is used to model the $\Delta M$ distribution and time dependence of mixed events directly from the data. Neural networks are used to select events and reconstruct the $D^{0}$. A result consistent with no charm mixing has been obtained, $R_{\text{mix}} = 0.0023 \pm 0.0012 \pm 0.0004$. This corresponds to an upper limit of $R_{\text{mix}} \leq 0.0042$ (90% C.L.).

In the standard model (SM), $D^{0}$-$\bar{D}^{0}$ mixing typically proceeds through box diagram loops involving down-type quarks. This results in very effective Glashow-Iliopoulos-Maiani and Cabibbo-Kobayashi-Maskawa suppressions of the mixing rate. The expected rate, relative to the unmixed rate, of $D^{0}$-$\bar{D}^{0}$ mixing through SM box [1] and dipenguin [2] diagrams is $O(10^{-3} - 10^{-10})$, many orders of magnitude below the current experimental sensitivity of $O(10^{-3})$ [3]. Possible enhancements to the SM mixing rate involve nonperturbative effects, and predictions range over many orders of magnitude [4] approximately bounded by the box diagram rate and the experimental sensitivity. New physics charm mixing predictions span the same large range [4] and, therefore, the presence of a mixing signal alone would not be a clear indication of new physics. However, the current experimental bounds on charm mixing already constrain new physics models.

Charm mixing is generally characterized by two dimensionless parameters, $x \equiv \Delta m/\Gamma$ and $y \equiv \Delta \Gamma/2\Gamma$, where $\Delta m = m_{2} - m_{1}$ ($\Delta \Gamma = \Gamma_{2} - \Gamma_{1}$) is the mass (width) difference between the two neutral $D$ mass eigenstates and $\Gamma$ is the average width. If either $x$ or $y$ is nonzero, then $D^{0}$-$\bar{D}^{0}$ mixing will occur.

The time evolution of a neutral $D$ meson depends on the type of final state into which it decays, since doubly Cabibbo-suppressed (DCS) decays to the same final states as mixed events allow for interference between the mixing and DCS amplitudes [5]. Semileptonic decays, for which there are no DCS amplitudes, have a particularly simple time dependence,

$$T_{\text{mix}}(t) \equiv T_{\text{unmix}}(t) \left( \frac{x^{2} + y^{2}}{4} \right) \left( \frac{t}{\tau_{D^{0}}} \right)^{2},$$

(1)

where $t$ is the proper time of the $D^{0}$ decay, $T_{\text{unmix}}(t) \propto e^{-t/\tau_{D^{0}}}$, and the approximation is valid in the limit of small mixing rates. Sensitivity to $x$ and $y$ individually is lost with semileptonic final states. The time-integrated mixing rate $R_{\text{mix}}$ relative to the unmixed rate is

$$R_{\text{mix}} = \frac{x^{2} + y^{2}}{2}.$$

(2)

We present a measurement of $R_{\text{mix}}$ using an 87-fb$^{-1}$ data sample collected on and just below the $Y(4S)$ resonance with the BABAR detector [6] at the PEP-II asymmetric-energy $e^{+}e^{-}$ storage ring. A full BABAR detector Monte Carlo simulation based on GEANT4 [7] is used to study background sources and to develop methods to reconstruct neutral $D$ mesons decaying to semileptonic final states.

Charged tracks are measured in a five-layer silicon vertex tracker (SVT) surrounded by a cylindrical wire drift chamber (DCH). Charged hadrons are identified through $dE/dx$ measurements in the tracking volume and in a ring-imaging Cherenkov detector (DIRC) surrounding the DCH. Identified charged $K$ purities averaged across all momenta are $\sim 99\%$, with $\sim 80\%$ reconstruction efficiency. Relative charge asymmetries between $K^{+}$ and $K^{-}$ in both misidentification rate and reconstruction efficiency are at most a few percent across all momenta and contribute negligibly to the total error on the mixing rate. Charge conjugation is implied unless explicitly stated otherwise.

Electromagnetic showers from electrons and photons are detected by CsI crystals arrayed in an electromagnetic calorimeter (EMC) located between the DIRC and the superconducting solenoid. Electrons are differentiated from other charged tracks by the intrinsic differences between hadronic, muonic, and electronic energy depositions in the EMC. As with charged $K$ candidates, there are relative charge asymmetries of at most a few percent in both misidentification rate and reconstruction efficiency between positrons and electrons, which have a negligible contribution to the total error on the mixing rate. Electrons and positrons arising from gamma conversions are eliminated from the pool of tracks used to reconstruct $D^{0}$ candidates.

Neutral $D$ candidates are selected by reconstructing the decay chain $D^{\ast+} \rightarrow \pi^{+}D^{0}$, $D^{0} \rightarrow K^{(*)}\ell\nu$. There are no essential differences for this analysis between the $K$ and $K^{*}$ modes, either theoretically or empirically, and thus no attempt is made to reconstruct the $K^{*}$—its charged $K$ daughter is treated as if it was a direct daughter of the $D^{0}$. The charge of the pion daughter of the charged $D^{\ast}$ identifies the production flavor of the neutral $D$, while the charge of the electron identifies the decay flavor. These charges are the same for unmixed decays and different for mixed decays, denoted as right-sign (RS) and wrong-sign (WS) decays, respectively.
The measured mixing rate is parametrized as $R_{\text{mix}} = n_{\text{WS}}/n_{\text{RS}}$. We fit the number of RS (WS) signal decays, $n_{\text{RS}}$ ($n_{\text{WS}}$), with a likelihood combining the $D^+D^0$ mass difference ($\Delta M$) distribution with the unmixed and mixed [Eq. (1)] decay time distributions, respectively. The high-statistics sample of RS signal events is used to model the $\Delta M$ distribution and the time dependence of mixed events directly from the data. To avoid potential bias, we perform a blind analysis in which the event selection criteria, and procedures for fitting the data and extracting an upper limit (UL), are determined prior to examining the WS signal region $\Delta M$ and decay time distributions in the data.

Identified $K$ and $e$ candidates of opposite charges are combined to create neutral $D$ candidate decay vertices. Only candidates with vertex fit probability $>0.01$ and mass $<1.82$ GeV/$c^2$ are retained. The value of the mass cut is chosen to exclude all hadronic two-body $D^0$ decays. The average PEP-II interaction point (IP), measured on a run-to-run basis using Bhabha and $\mu^+\mu^-$ events, is taken as the production point of $D^0$ candidates. The transverse $D^0$ decay time is used due to the relative narrowness of, and small errors on, the IP in the transverse ($r$-$\phi$) plane, as opposed to the IP's longitudinal ($z$) profile. Because of the very small transverse distance traveled by $B$ mesons prior to decay, there are no significant differences between the decays of charmed parents arising from the continuum (prompt events) and those coming from the decay of $B$ mesons, other than the harder $D^0$ momentum spectrum from prompt events.

The pions from $D^+D^0$ decays are relatively soft tracks with $p_{\pi} < 450$ MeV/$c$, where the asterisk denotes a parameter measured in the Y(4S) center-of-mass frame. Charged tracks identified as either a charged $K$ or $e$ candidate are not considered as $\pi$ candidates. To reject poorly reconstructed tracks, $\pi$ candidates are required to have six or more SVT hits, with $\geq 2r$-$\phi$ layer hits, $\geq 2z$ layer hits, at least one hit on the inner three $r$-$\phi$ layers, and at least one hit on the inner three $z$ layers. Pion candidates are also refit with the IP as a constraint and accepted only if the refit probability is greater than 0.01.

The individual $K$, $e$, and $\pi$ tracks, the $K$-$e$ vertex, and the event thrust are used to reconstruct the three orthogonal components of the $D^0$ momentum vector with three JETNET 3.4 [8] neural networks (NN’s). Each NN has two hidden layers, and is trained and tested with a large sample [$O(10^5)$] of simulated signal events. The following vector inputs to the NN’s are used: the momentum of the event, $p_e$ ($p_K$), and the event thrust vector $T$ (calculated using charged and neutral candidates except the $K$ and $e$ candidates). In addition, four opening angles among the above parameters are used: $\theta^*(p_K, T)$, $\theta^*(p_e, p_K)$, $\theta^*(p_\pi, T)$, and $\theta^*(p_\pi, p_e)$. The vector inputs are separated into their respective orthogonal components (azimuthal/polar angles and magnitude) and, additionally using the four scalar inputs, separate NN’s for $p_e^*(D^0)$, $p_K^*(D^0)$, and $|p^*(D^0)|$ are trained with the seven input parameters.

The residuals between the true $p^*(D^0)$ direction for simulated signal events and the NN output are unbiased and Gaussian distributed with $\sigma \sim 110$ mrad. The distribution of momentum magnitude residuals shows an error of $\sigma_{p}/p \sim 10\%$ for a typical signal event.

The $D^+D^0$ mass difference is calculated using the $D^0$ candidate and the tagging pion candidate. The signal $\Delta M$ distribution peaks $\sim3$ MeV/$c^2$ below the Particle Data Group (PDG) [9] value of 145.4 MeV$/c^2$ and has a rms width of $\sim2.2$ MeV/$c^2$. The shift in the peak introduces no bias into the measurement of the mixing rate and is due to the use of tagging pion information in reconstructing the $D^0$ momentum. A $\sim80$ MeV/$c^2$ wide $\Delta M$ sideband is retained from 160 to 240 MeV/$c^2$ in order to characterize the level of background events.

The transverse momentum of a $D^0$ candidate, and the projections of the IP and $K$-$e$ vertex loci on the $r$-$\phi$ plane, are used to calculate the candidate's proper decay time. The error on the decay time, calculated using only the errors on the IP and $K$-$e$ vertex, is typically 0.6$\tau_{D^0}$, where $\tau_{D^0}$ is the nominal mean $D^0$ lifetime. The contribution of the $p^*(D^0)$ estimator to the total decay time uncertainty is negligible. Poorly reconstructed events with calculated decay time uncertainties greater than 2$\tau_{D^0}$ are discarded. Only events with decay times between $-12\tau_{D^0}$ and $16\tau_{D^0}$ are retained.

In addition to the above criteria, events are selected using a neural network trained to distinguish prompt charm signal from background events. The event selector NN uses a five-element input vector: $p_{e}^*$, $p_{K}^*$, $|T|$, $\theta^*(p_K, T)$, and $\theta^*(p_e, p_K)$. It has a single hidden layer of nine nodes and is also constructed using JETNET 3.4. A cut is made on the event selector NN output such that the statistical sensitivity to mixed charm events is optimized for the 87-fb$^{-1}$ data set used.

Events with multiple candidates meeting all of the above selection criteria but which share tracks are discarded, leading to a loss of $\sim10\%$ of signal events passing all other cuts. There are no differences in RS and WS reconstruction efficiencies in the final event selection.

Backgrounds predominantly come from prompt charm events, with minor contributions from $uds$ and $b\bar{b}$ events. The backgrounds from misidentified charged particle species are negligible. Nearly all background events come from $D^0$ and $D^+$ semileptonic decays to final states including both a charged $K$ and an $e$ that are combined with a random $\pi^+$, and truly random combinatorics in which the $K$ and $e$ do not share a common charm parent. The decay time distributions of background charm parents differ somewhat from those of well-reconstructed charm parents because the $p^*(D^0)$ estimator returns a correct momentum estimate only for signal events. The
decay time probability density functions (PDF’s) for these charm backgrounds are thus not fit from the data but, instead, are derived from simulated events. The decay time distribution for truly random (i.e., zero-lifetime) events is fit in the data using a sum of three Gaussians.

The $\Delta M$ backgrounds qualitatively fall into two categories—those which peak under the signal $\Delta M$ distribution and those which do not. Both RS and WS events have small peaking $\Delta M$ backgrounds. In the WS sample, the peaking component arises from semielectronic $D^+\pi^-$ decays with a $K^+$ daughter—these events are fit using the $D^+$ decay time PDF and are thus differentiated from mixed signal events. In the RS sample, the fraction of background events sharing the RS unmixed signal PDF is estimated from fits to simulated data sets and a correction of $\sim 4\%$ is made to the fit number of RS signal decays to obtain the actual number. The RS $D^+$ $\Delta M$ background peaks well away from the $\Delta M$ signal region and its shape is fixed from simulated events. The shape is floated as a systematic check and there is a negligible effect on the final value of $R_{\text{mix}}$.

Nonpeaking combinatoric $\Delta M$ backgrounds are modeled from the data by combining $K^-$-$e$ vertex and $\pi$ candidates from different events (“event mixing”). Events are selected from the off-resonance data such that a $K^-$-$e$ vertex drawn from one event may be combined with a $\pi$ candidate from any other event. Approximately 70,000 off-resonance data events are used to generate $O(10^6)$ RS and WS pseudoevents passing the final event selection criteria. A Kolmogorov-Smirnov test [10] shows that the event-mixed RS and WS $\Delta M$ distributions have a $\sim 92\%$ probability of sharing the same underlying parent distribution. The shape of the PDF associated with the nonpeaking combinatoric $\Delta M$ background is taken from the combined event-mixed RS and WS $\Delta M$ distributions and denoted by $H(\Delta M)$.

An initial fit to the RS data is used to extract the unmixed signal mean lifetime, shape of the signal $\Delta M$ distribution, and number of RS signal events. The mixed WS signal PDF parameters are taken from this high-statistics fit. The full extended likelihood function is

$$L = \frac{\lambda^N e^{-\lambda}}{N!} \prod_{i=1}^{N} \sum_{j=1}^{m} f_j P_j,$$

where $f_j$ ($n_j$) is the fraction (number) of events assigned to event class $j$, $\lambda = \sum n_j$ is the fit total number of candidates, $P_j$ is the value of the PDF for event class $j$, and the extended product runs over all $N$ events. The PDF’s for RS and WS signal and background classes are detailed below. The shapes of all PDF’s are fit in the data unless otherwise noted. All fits are done as unbinned maximum likelihood fits.

Four classes of events are fit in the RS data: unmixed signal events ($n_{\text{unmix}}$), background $D^0$ decays ($n_{\text{bkgdD^0}}$), RS $D^+$ decays ($n_{\text{RSD+}}$), and completely random three-track combinations ($n_{\text{zero life}}$). The RS signal $\Delta M$ shape is fit to a threshold function with a power-law turn-on and an exponentially decaying tail, while the decay time distribution is fit as an exponential decay with floating decay constant convoluted with a resolution model ($G$) taken from simulated signal events. The RS signal $\Delta M$ and decay time combined PDF is

$$P_{\text{unmix}} = u^2 \exp\left(\sum_{i=1}^{3} \alpha_i u^i\right) \cdot [e^{-t/\tau_{\text{peaking}}} \theta(ct > 0)] \otimes G,$$

where $u = \Delta M - M_{\pi^0}$, and $M_{\pi^0}$ is fixed to the nominal PDG value [9]. The four parameters of the signal $\Delta M$ PDF ($\gamma, \alpha_i, \tau_{\text{peaking}}$), and the unmixed signal population are allowed to float in the RS fit.

The background $D^0$ event class uses the combined PDF

$$P_{\text{bkgdD^0}} = H(\Delta M) \cdot T_{D^0}(t),$$

where $T_{D^0}(t)$ represents the shape of the decay time distribution for this event class derived from simulated events. The RS $D^+$ event class has the combined PDF:

$$P_{\text{RSD+}} = u^2 \exp(\xi_1 u + \xi_2 u^2 + \xi_3 u^3) \cdot T_{D^+}(t),$$

where $\xi$ and $\xi_i$ are fixed from fits to simulated events, and $T_{D^+}(t)$ represents the shape of the decay time distribution for this event class derived from simulated events. Finally, zero-lifetime events use the combined PDF:

$$P_{\text{zero life}} = H(\Delta M) \cdot \sum_{i=1}^{3} f_i \exp\left[-\frac{1}{2} \left(\frac{\mu - c t^i}{\sigma_t}\right)^2\right].$$

with the constraint $f_3 = 1 - f_1 - f_2$. Fifteen parameters are floated in the RS fit: the signal $\Delta M$ shape parameters and mean $D^0$ lifetime; the mean, widths, and relative fractions of the zero-lifetime triple Gaussian; and the population of the four RS event classes.

Figure 1 shows the projections of the RS fit model onto the RS data. After correcting for peaking RS background events as described above, we find $n_{\text{unmix}} = 49620 \pm 265$ events. The fit value for the unmixed $D^0$ mean lifetime is, within its statistical error of $\sim 0.6\%$, consistent with the current PDG [9] value.

Five classes of events are fit in the WS data: mixed signal events ($n_{\text{mix}}$), background $D^0$ decays ($n_{\text{bkgdD^0}}$), peaking WS $D^+$ decays ($n_{\text{pkng D^+}}$), nonpeaking WS $D^+$ decays ($n_{\text{non-pkng D^+}}$), and completely random three-track combinations ($n_{\text{zero life}}$). WS signal events share the sharply peaked $\Delta M$ distribution of the RS signal, but have the modified lifetime distribution given by Eq. (1). The full form of the WS mixed signal PDF is
Mixed events greater than the result of the toy MC—the data-fit NLL values lie well within the range predicted by the toy fits. The WS fit model is also tested for bias and correct error scaling with fits to simulated data sets containing 0, 50, and 100 mixed WS events ($R_{\text{mix}} \sim 0, 0.001, 0.002$, respectively), and no evidence of bias or improperly scaled errors is seen.

The systematic error includes variations of the WS combined mixed signal PDF, random combinatoric $\Delta M$ PDF shape, signal decay time resolution model, and background $D^0$ and $D^+$ decay time PDF’s. By far, the dominant systematic is the statistical precision with which the RS $\Delta M$ PDF is known. This systematic is characterized by sequentially varying by $\pm 1\sigma$ one of the signal $\Delta M$ shape parameters floated in the RS fit.

TABLE I. Systematic variations in the value of $R_{\text{mix}}$.

| Varied parameter | $|\Delta R_{\text{mix}}|$ |
|------------------|------------------|
| Mixed $\Delta M$ PDF | 0.00032 |
| Mixed decay time PDF | 0.00008 |
| Random combinatoric $\Delta M$ shape | 0.00015 |
| Background $D^0$ decay model | 0.00016 |
| Background $D^+$ decay model | 0.00012 |
along with the mean $D^0$ lifetime, and refitting the WS data. The resulting deviations from the central value of $n_{\text{mix}}$ obtained in the nominal fit to the WS data are summed in quadrature, taking into account correlations among the fit parameters. Table I shows the error on $R_{\text{mix}}$ attributable to each source of systematic error listed above. There are no significant effects on $R_{\text{mix}}$ attributable to the choice of vertexing algorithm, IP, $K$, or $e$ particle identification, NN event selector cut, decay time error cut, signal resolution model, or $\Delta M$ sideband cut. Taking the total systematic error as the sum in quadrature of the values in Table I, $\sigma_{R_{\text{mix}}}^{\text{stat}} = 0.0004 = 0.34\sigma_{R_{\text{mix}}}^{\text{stat}}$.

A scan of the change in NLL for $n_{\text{mix}}$ values in the region surrounding the fit minimum is used to calculate upper limits. By construction, the NLL scan includes only the statistical error of the fit—the systematic error is included as a small perturbation on the values of $\Delta$NLL used to establish confidence intervals. The total error is taken as the sum in quadrature of the statistical and systematic errors, $\sigma_{R_{\text{mix}}}^{\text{total}} = \sqrt{1 + 0.34^2} \sigma_{R_{\text{mix}}}^{\text{stat}} = 1.06\sigma_{R_{\text{mix}}}^{\text{stat}}$.

The 95% C.L. UL is taken as the value of $n_{\text{mix}}$ where the NLL value changes from its minimum by $\Delta$NLL = 0.5, which yields $R_{\text{mix}} < 0.0046$ (95% C.L.). The factor of “0.97” in the preceding expression for the UL arises from the fraction of the a posteriori distribution of $n_{\text{mix}}$ lying in the physical region. A similar calculation shows $R_{\text{mix}} < 0.0042$ at the 90% C.L. The relatively small error ($\sim 0.5\%$) on $n_{\text{RS}}$ is negligible and has been ignored.

In summary, we have obtained a value of $R_{\text{mix}}$ consistent with no mixing and set a limit of $R_{\text{mix}} < 0.0042(0.0046)$ at the 90% (95%) C.L. This is slightly lower than the previous best published limit using semileptonic decays of $R_{\text{mix}} < 0.005$ (90% C.L.) set by E791 [11]. Our result can be compared to the current best published UL on $R_{\text{mix}}$ (< 0.0016, 95% C.L.), which comes from the BABAR hadronic $K\pi$ mixing analysis [12]. Had the central value of $n_{\text{mix}}$ been observed here to be zero, the 90% C.L. (95%) UL for the mixing rate would have been 0.0016 (0.0020), comparable to the sensitivity of both the hadronic BABAR result and a recent semileptonic FOCUS analysis [13].

We are grateful for the excellent luminosity and machine conditions provided by our PEP-II colleagues, and for the substantial 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 (U.S.A.), NSERC (Canada), IHEP (China), CEA and CNRS-IN2P3 (France), BMBF and DFG (Germany), INFN (Italy), FOM (The Netherlands), NFR (Norway), MIST (Russia), and PPARC (United Kingdom). Individuals have received support from CONACyT (Mexico), A.P. Sloan Foundation, Research Corporation, and Alexander von Humboldt Foundation.