Search for D^0-$\overline{D^0}$ mixing and doubly-Cabibbo-suppressed decays of the D^0 in hadronic final states

E. M. Aitala,[9] S. Amato,[1] J. C. Anjos,[1] J. A. Appel,[5] D. Ashery,[15] S. Banerjee,[5] I. Bediaga,[1] G. Blaylock,[8] S. B. Bracker,[16] P. R. Burchat,[14] R. A. Burnstein,[6] T. Carter,[5] H. S. Carvalho,[1] N. K. Copty,[13] L. M. Cremaldi,[9] C. Darling,[19] A. Fernandez,[12] P. Gagnon,[2] C. Gobel,[1] K. Gounder,[9] A. M. Halling,[5] G. Herrera,[4] G. Hurvits,[15] C. James,[5] P. A. Kasper,[6] S. Kwan,[5] D. C. Langs,[11] J. Leslie,[2] B. Lundberg,[5] S. MayTal-Beck,[15] B. Meadows,[3] J. R. T. de Mello Neto,[1] R. H. Milburn,[17] J. M. de Miranda,[1] A. Napier,[17] A. Nguyen,[7] A. B. d'Oliveira,[3,12] K. O'Shaughnessy,[2] K. C. Peng,[6] L. P. Perera,[3] M. V. Purohit,[13] B. Quinn,[9] S. Radeztsky,[18] A. Rafatian,[9] N. W. Reay,[7] J. J. Reidy,[9] A. C. dos Reis,[1] H. A. Rubin,[6] A. K. S. Santha,[3] A. F. S. Santoro,[1] A. J. Schwartz,[11] M. Sheaff,[18] R. A. Sidwell,[7] A. J. Slaughter,[19] M. D. Sokoloff,[3] N. R. Stanton,[7] K. Stenson,[18] D. J. Summers,[9] S. Takach,[19] K. Thorne,[5] A. K. Tripathi,[10] S. Watanabe,[18] R. Weiss-Babai,[15] J. Wiener,[11] N. Witchey,[7] E. Wolin,[19] D. Yi,[9] S. Yoshida,[7] R. Zaliznyak,[14] and C. Zhang[7]

(Fermilab E791 Collaboration)

[1]*Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil*
[2]*University of California, Santa Cruz, California 95064*
[3]*University of Cincinnati, Cincinnati, Ohio 45221*
[4]*CINVESTAV, Mexico City, Mexico*
[5]*Fermilab, Batavia, Illinois 60510*
[6]*Illinois Institute of Technology, Chicago, Illinois 60616*
[7]*Kansas State University, Manhattan, Kansas 66506*
[8]*University of Massachusetts, Amherst, Massachusetts 01003*
[9]*University of Mississippi, University, Mississippi 38677*
[10]*The Ohio State University, Columbus, Ohio 43210*
[11]*Princeton University, Princeton, New Jersey 08544*
[12]*Universidad Autonoma de Puebla, Puebla, Mexico*
[13]*University of South Carolina, Columbia, South Carolina 29208*
[14]*Stanford University, Stanford, California 94305*
[15]*Tel Aviv University, Tel Aviv, Israel*
[16]*317 Belsize Drive, Toronto, Ontario, Canada M4S 1M7*
[17]*Tufts University, Medford, Massachusetts 02155*
[18]*University of Wisconsin, Madison, Wisconsin 53706*
[19]*Yale University, New Haven, Connecticut 06511*

(Received 29 August 1996; revised manuscript received 27 August 1997; published 8 December 1997)

We present results of a search for D^0-$\overline{D^0}$ mixing and doubly-Cabibbo-suppressed decays of the D^0 in Fermilab experiment E791, a fixed-target charm hadroproduction experiment. We look for evidence of mixing in the decay chain $D^* \to \pi D \to \pi(K\pi$ or $K\pi\pi\pi)$. If the charge of the pion from the D^* decay is the same as the charge of the kaon from the D decay (a ''wrong-sign'' event), mixing may have occurred. Mixing can be distinguished from other sources of wrong-sign events (such as doubly-Cabibbo-suppressed decays) by analyzing the distribution of decay times. We see no evidence of mixing. Allowing for CP violation in the interference between DCS and mixing amplitudes our fitted ratio for mixed to unmixed decay rates is $r_{mix} = (0.39^{+0.36}_{-0.32} \pm 0.16)\%$. This corresponds to a 90% C.L. upper limit of $r_{mix} < 0.85\%$. The sensitivity of this result is comparable to that of previous measurements, but the assumptions made in fitting the data are notably more general. We present results from many fits to our data under various assumptions. If we assume $r_{mix} = 0$, we find a two-sigma wrong-sign enhancement in the $K\pi$ mode which we ascribe to doubly Cabibbo-suppressed decays. The ratios of doubly Cabibbo-suppressed decays to Cabibbo-favored decays are $r_{dcs}(K\pi) = (0.68^{+0.34}_{-0.33} \pm 0.07)\%$ and $r_{dcs}(K\pi\pi\pi) = (0.25^{+0.36}_{-0.34} \pm 0.03)\%$. [S0556-2821(98)01103-5]

PACS number(s): 13.25.Ft, 12.15.Ff, 14.40.Lb

I. INTRODUCTION

The standard model predicts a rate for D^0-$\overline{D^0}$ mixing which is many orders of magnitude below the reach of present experiments. Typical calculations [1] give r_{mix}, the ratio of mixed to unmixed decay rates, in the range 10^{-10}–10^{-7}. In contrast, various extensions to the standard model [2] allow a mixing rate close to the current experimental sensitivity of 10^{-3}–10^{-2}. Consequently, a discovery of D^0-$\overline{D^0}$ mixing at currently measurable levels would be inconsistent with the standard model, and would provide a clear signal for new physics.

Experimentally, mixing is identified by a change in the charm quantum number of the neutral D meson between its production and decay. In the analysis presented in this paper, the charm of the produced D is determined from the decay

$D^{*+}\to D^0\pi^+$ (or $D^{*-}\to\overline{D^0}\pi^-$), where the charge of the pion indicates whether a D^0 or a $\overline{D^0}$ was produced. D decays are reconstructed in four all-charged hadronic decay modes $D\to K^-\pi^+$, $D\to K^+\pi^-$, $D\to K^-\pi^-\pi^+\pi^+$ or $D\to K^+\pi^+\pi^-\pi^-$. (Hereafter, we will omit the charge superscripts from the final states where context allows.) Possible evidence for mixing comes from the detection of a meson produced as a D^0 ($\overline{D^0}$) decaying to a ''wrong-sign'' final state which contains a K^+ (K^-), with the kaon charge opposite to that expected for unmixed decays.

Fermilab experiment E691 [3] has previously used this technique to set what is currently the strictest upper limit on mixing, $r_{mix}<0.37\%$, albeit with specific assumptions which we will address in this paper. Fermilab experiment E615 obtained a limit of $r_{mix}<0.56\%$ by looking for same-sign muon pairs in π-tungsten interactions, based on a specific model for charm production [4]. Evidence for wrong-sign decays has been presented by the CLEO Collaboration [6], which measures the ratio of wrong-sign to right-sign decays to be $(0.77\pm0.25\pm0.25)\%$ for the $K\pi$ final state. However, the CLEO experiment was unable to distinguish between mixing and doubly-Cabibbo-suppressed decays, which also produce wrong-sign events. Recently we have reported on a search for mixing using semileptonic decays of the D^0 ($\overline{D^0}$) which do not have a doubly-Cabibbo-suppressed background [5]. We found that $r_{mix}<0.50\%$.

It is possible to distinguish doubly-Cabibbo-suppressed (DCS) and mixing contributions to the wrong-sign rate by studying the distribution of D decay times. In the limit of small mixing, the rate for wrong-sign decays takes the form

$$\Gamma[D^0(t)\to f]=\frac{e^{-\Gamma t}}{4}|\langle f|H|\overline{D^0}\rangle_{CF}|^2\left|\frac{q}{p}\right|^2\times\left[4|\lambda|^2+\left((\Delta M)^2+\frac{(\Delta\Gamma)^2}{4}\right)t^2+[2Re(\lambda)\Delta\Gamma+4Im(\lambda)\Delta M]t\right],\qquad(1)$$

where

$$\lambda\equiv\frac{p}{q}\frac{\langle f|H|D^0\rangle_{DCS}}{\langle f|H|\overline{D^0}\rangle_{CF}},\qquad(2)$$

and p and q describe the relationship between the charm eigenstates $|D^0\rangle$ and $|\overline{D^0}\rangle$ and the mass eigenstates $|D_{1,2}\rangle$:

$$|D_1\rangle=p|D^0\rangle+q|\overline{D^0}\rangle,$$

$$|D_2\rangle=p|D^0\rangle-q|\overline{D^0}\rangle.\qquad(3)$$

The amplitude $\langle f|H|D^0\rangle_{DCS}$ represents the DCS decay of the D^0 while $\langle f|H|\overline{D^0}\rangle_{CF}$ is the Cabibbo-favored counterpart for the decay of $\overline{D^0}$. The parameters ΔM and $\Delta\Gamma$ describe the differences in mass and width of the two physical states. The term proportional to $|\lambda|^2$ in Eq. (1) describes the contribution from DCS amplitudes, the term proportional to t^2 describes the lowest-order contribution from mixing, and the term proportional to t represents the interference between mixing and DCS amplitudes. We can apply this formula to the measured time distribution of wrong-sign decays to determine the separate contributions from DCS and mixing amplitudes.

In the study that follows, we examine a sample of about 9100 reconstructed, tagged D^0 decays to look for wrong-sign decays, using the different time distributions to separate the DCS and mixing contributions in our search. As we shall see, there are no significant wrong-sign signals in our data, which leads us to set restrictions on the ratio of wrong-sign to right-sign rates. The most likely fit (in the possible standard model scenarios) will be presented first. Afterwards, we will determine the effects of relaxing all constraints and of additional constraints (absence of DCS-mixing interference, no mixing at all) which investigate interesting physics cases or are necessary to compare with previously published results.

II. EFFECTS OF *CP* VIOLATION

Equation (1) describes the rate for D^0 to decay to a wrong-sign final state f. Within the context of some new physics models, it is possible that the rate for $\overline{D^0}$ to decay to $\overline{f}$ is not the same, and that CP is violated to a significant extent. Thus, it is important to allow for the possibility of CP violation. This results in the most conservative upper limit on wrong-sign decays. The analysis presented here is the first experimental study to allow for the possibility of CP violation. (For recent discussions of the role of CP violation in D^0-$\overline{D^0}$ mixing see [7, 8].)

Formally, the conjugate equation is

$$\Gamma[\overline{D^0}(t)\to\overline{f}]=\frac{e^{-\Gamma t}}{4}|\langle\overline{f}|H|D^0\rangle_{CF}{}^2|\left|\frac{p}{q}\right|^2\times\left[4|\overline{\lambda}|^2+\left((\Delta M)^2+\frac{(\Delta\Gamma)^2}{4}\right)t^2+[2Re(\overline{\lambda})\Delta\Gamma+4Im(\overline{\lambda})\Delta M]t\right],\qquad(4)$$

with

$$\overline{\lambda}\equiv\frac{q}{p}\frac{\langle\overline{f}|H|\overline{D^0}\rangle_{DCS}}{\langle\overline{f}|H|D^0\rangle_{CF}}.\qquad(5)$$

In principle, CP violation can arise through a difference between Eqs. (1) and (4) in any one of the three terms. Any term in Eq. (1) can differ from its charge conjugate in Eq. (4) as a result of the interference of two or more contributing amplitudes which have nonzero relative phases of both the CP-conserving and CP-violating type.

Inequality of the two constant terms (i.e., $(|q/p|)^2|\lambda|^2\neq(|p/q|)^2|\overline{\lambda}|^2$, but see comment in [9]) is referred to as *direct CP* violation. This could be significant if two or more comparable DCS amplitudes contribute with different CP-conserving and CP-violating phases. However, the standard model contribution (which is expected to dominate) provides only one weak, CP-violating phase. Direct CP violation is therefore likely to be small. Similarly, the two charge conjugate terms proportional to t^2 will be the same

unless there are two or more mixing amplitudes with relative CP-violating and CP-conserving phases. On the contrary, most models suggest that if mixing occurs at all, it is likely to be dominated by a single CP-violating phase. Therefore, the most plausible constraint involving CP violation restricts CP violation to the interference term. We will explore this possibility, as well as the more general case without CP restrictions, in the study of our data which follows.

III. DESCRIPTION OF E791

We report the results of a search for D^0-$\overline{D^0}$ mixing and DCS decays using hadronic decays found in data from our experiment, Fermilab E791. We collected approximately 2×10^{10} hadronic interactions in the 1991–1992 fixed-target run using the TPL spectrometer [10] with a 500 GeV/c π^- beam. There were five foil targets: one 0.5-mm-thick Pt foil followed by four 1.6-mm-thick diamond foils with 15-mm center-to-center separations. This arrangement allowed us to greatly reduce secondary interaction backgrounds by selecting only charm candidates which decayed in air.

The target region was preceded by 6 planes of silicon microstrip detectors and 8 proportional wire chambers (PWC's) used for beam tracking and was followed by 17 additional planes of silicon microstrip detectors for measuring tracks produced at and downstream of the primary vertex. The track momenta and slopes were also measured in the downstream spectrometer which had two magnets, 35 planes of drift chambers, and two PWC's. Two threshold Čerenkov counters provided π/K separation in the 6–60 GeV/c momentum range [11].

The mixing analysis in this paper relies heavily on track reconstruction, which begins by using hits in the silicon detector and folds in additional information from the downstream devices. The tracking efficiency is approximately 80% for particles with momenta greater than 30 GeV/c and drops to around 60% for particle momenta of 10 GeV/c. The mean number of reconstructed tracks used to fit the primary vertex is 7. After reconstruction, events with evidence of multiple vertices were kept for further analysis. The list of reconstructed vertices is used in the selection criteria described below.

We determined our production (primary) and decay vertex resolutions by comparing reconstructed and true vertex positions using our Monte Carlo detector simulation. The transverse resolutions quoted below are one-dimensional values. Longitudinal and transverse position resolutions for the primary vertex are 350 and 6 μm, respectively. For the mean D^0 momentum of 65 GeV/c the longitudinal resolutions for $K\pi$ and $K\pi\pi\pi$ vertices are 320 and 395 μm, respectively, and increase by 33 and 36 μm, respectively, for every 10 GeV/c D^0 momentum. Similarly, for the mean momentum, 65 GeV/c, of the observed D^0's, the transverse resolutions for $K\pi$ and $K\pi\pi\pi$ vertices are 10 and 12 μm respectively and decrease by about 0.5 μm for every 10 GeV/c increase in D^0 momentum.

The kaon and pion identification efficiencies and misidentification probabilities vary with momentum and with the signatures we require in the Čerenkov detector. For typical particle momenta in the range 20–40 GeV/c, the Čerenkov identification efficiency of a kaon is around 58% when the probability for a pion to be misidentified as a kaon is 4%. In the same momentum range the Čerenkov identification efficiency of a pion is around 93% when the probability for a kaon to be misidentified as a pion is 35%.

IV. SELECTION OF DATA SAMPLE

A search for the rare wrong-sign mixing and DCS decays requires selection criteria that emphasize background reduction. We achieve this goal in two stages: initially reconstructing displaced secondary vertices and using a few loose criteria for selecting D^0 decays to reduce the data sample, and then optimizing the data selection with artificial neural networks.

Initial reduction of the large E791 data set to a manageable size was achieved with the aid of a few simple criteria. Here, we describe the cuts made in these initial stages for the $D^0 \to K\pi$ mode. [When we refer to $K\pi$ or $K\pi\pi\pi$ in this paper, without any explicit signs, we include the charge conjugate states. Otherwise, we indicate a specific final state by explicitly specifying at least the kaon charge or specifying right-sign (RS) or wrong-sign (WS) decays.] Two-prong vertices were used to start the search for D^0 decays. The invariant mass of the two-prong D^0 candidate, assumed to be $K\pi$, was required to be in the range 1.7–2.0 GeV/c^2. The kaon candidate was chosen as the one with the higher probability of being a kaon based on Čerenkov information. To further reduce the contributions from misidentified D^0 decays, $K\pi$ candidates were rejected if the reverse hypothesis (πK) fell within 2σ of the D^0 mass, where σ is the measurement resolution for the D^0 mass. Similarly, to reduce contamination from D^0 decays to K^+K^- and $\pi^+\pi^-$, $K\pi$ candidates were rejected if the K^+K^- or $\pi^+\pi^-$ mass hypotheses fell within 2σ of the D^0 mass. To help ensure that the reconstructed secondary vertex was a true decay vertex, we required that it be separated from the primary vertex by at least 8 standard deviations ($\sigma_{\Delta z}$) in the beam direction (i.e., $\Delta z/\sigma_{\Delta z} > 8$). Two further requirements ensured that the reconstructed D^0 was consistent with originating at the primary vertex. First, the impact parameter of the reconstructed D momentum with respect to the primary vertex, b_p, was required to be less than 60 μm. Second, the component of the reconstructed D momentum perpendicular to the D line of flight (as determined from the primary and secondary vertex positions), p_T^D, was required to be less than 0.35 GeV/c. The K and π decay tracks were required to be well reconstructed in the silicon detectors and drift chambers. Finally, the momentum asymmetry of the K and π as measured in the laboratory frame (i.e., $p_{asy} \equiv |\vec{p}_K - \vec{p}_\pi|/|\vec{p}_K + \vec{p}_\pi|$) was required to be less than 0.65. This reduced the contribution from random track combinations, which tended to be asymmetric.

The cuts for the $D^0 \to K\pi\pi\pi$ case were similar in the initial stages. We used candidates arising from both 4-prong vertices and 3-prong vertices (with an added track). The two vertex samples contributed roughly equal amounts to the signal. We required $\Delta z/\sigma_{\Delta z} > 8$, $b_p < 60$ μm, $p_T^D < 0.5$ GeV/c, and that the decay vertex be outside the target foils. The D^0 candidate mass was required to be in the range 1.7–2.0 GeV/c^2. To eliminate reflections from Cabibbo-favored decays we examined the hypothesis that the kaon was actually

a pion and one of the pions opposite in charge to the kaon was actually a kaon. Since there are two such possibilities, candidate $K\pi\pi\pi$ decays were rejected if either possibility yielded a candidate mass within 2σ of the D^0 mass. Tracks were required to have momenta greater than 0.5 GeV/c and to have greater than 4σ of transverse separation from the primary vertex where σ is the measurement resolution on the separation. Finally, the decay tracks were required to form a vertex that was no more than 2.5 cm downstream of the final target. Beyond that point the silicon detectors and other material in the beam path provided large numbers of secondary interactions.

In the final stage of analysis, we used two-layer feed-forward neural networks to optimize the signal selection [12,13]. Specifically, we chose selection criteria that maximize $S/\sqrt{B}$ where S and B were the signal and the background under the signal for the right-sign decays. A vector, whose components are variables such as the ones just mentioned, was fed into each neural net as the input layer. Each node in the next (hidden) layer computed the sigmoid of the sum of an offset and the inner product of the input vector with a weight vector. The results from this layer in turn formed the input for a single node in the final (output) layer. Thus, the networks effectively combined information from each variable we would otherwise have ''cut'' on and provided a single output value in the range 0–1. This output was monotonically related to the probability that a given candidate was signal and not background.

Since our two major sources of background were false D^0 candidates and real D^0 candidates combined with random pions to produce fake D^* candidates, we used separate neural networks to classify the D^0 and D^* candidates. Although there are only two modes of D^0 decay examined in this work, three D^0 samples were used to train the neural nets: one for the $D^0\to K\pi$ mode and two for the $D^0\to K\pi\pi\pi$ mode. The two separate $D^0\to K\pi\pi\pi$ samples contained candidates from vertices that had either all four or only three of the four tracks.

In order to minimize our dependence on Monte Carlo (MC) calculations, we used D^0 candidates in our real data to train separate neural nets for each of the three samples. We chose D^0 candidates that do not combine with pions to give a D^* candidate. The training sample is thus independent of our mixing sample and ensures that the neural net training was unbiased. A fourth neural net was trained using part of the right-sign $D^{*+}\to D^0\pi^+$ sample to classify D^{*+} candidates. Every net was trained using events in the peak region as ''signal'' and the remaining events as ''background.'' We selected only those events for which the product of the D^0 and D^* net outputs was greater than a certain value rather than making individual ''cuts'' on many variables.

In the $D^0\to K^-\pi^+$ mode, the net was presented with 12 input variables: the p_T of the D^0 relative to the incident pion beam direction, the separation between the secondary and primary vertices (Δz), p_T^D, b_p, the χ^2 per degree of freedom for the secondary vertex fit, the Cerenkov-based probability for the kaon to be a kaon, the momentum asymmetry (p_{asy}), the consistency probability for the secondary vertex to be in a target foil, the track fit χ^2 per degree of freedom for the two tracks and the number of tracking systems traversed by each of the two tracks. The two nets for the $D^0\to K^-\pi^+\pi^+\pi^-$ mode used seven variables, of which the first six variables were the same as the first six just listed for the $D^0\to K^-\pi^+$ mode. The seventh variable was the smallest contribution to the χ^2 of the fit to the primary vertex from any of the four decay tracks. Finally, the D^* neural net was constructed using only five variables: the χ^2 per degree of freedom for the track fit for the pion from the D^{*+} decay (referred to hereafter as the ''bachelor pion''), the number of tracking stations in which the pion is detected, the probability that it is not a fictitious track, its momentum and the χ^2 per degree of freedom for the vertex fit of the D^0 and the bachelor pion. There were three nodes in the hidden layer for the $D^0\to K^-\pi^+$ and D^* nets and four nodes in the hidden layers of each of the $D^0\to K^-\pi^+\pi^+\pi^-$ nets.

Although we considered many variables, we pruned the list down to the variables listed above and also pruned some of the connections to the hidden layer nodes when their contributions to the output were deemed unimportant using a technique called subset reduction, implemented as follows. Nodes in a given layer were viewed as a linear array, one row for each event. The matrix thus formed was subjected to singular value decomposition using QRcp factorization [14]. The ''energy content'' of the nodes was determined by the resulting eigenvalues [15]. Nodes with an ''energy content'' $<1\%$ were deleted.

We also tried other techniques for selecting events, including the more common method of using independent ''cuts'' in each variable and a binary decision tree (BDT) technique [16]. The sensitivity of the neural net technique was about 10% higher than the BDT in the $D^0\to K\pi$ mode and about 30% higher than the BDT in the $D^0\to K\pi\pi\pi$ mode; in turn the BDT was better than the commonly used ''cuts'' technique. One further advantage of the neural net technique was that the output could be used to choose the best candidate in an event, should there be more than one (a rare occurrence). This simplified the statistical analysis in the fits to our data.

The results of our neural net optimizations are shown in Fig. 1 for right-sign and wrong-sign $K\pi$ and $K\pi\pi\pi$ final states. In this figure, we plot the candidate D^0 mass [$m(K\pi)$ or $m(K\pi\pi\pi)$] versus Q, defined as $Q\equiv m(K\pi\pi)-m(K\pi)-m(\pi)$ or $Q\equiv m(K\pi\pi\pi\pi)-m(K\pi\pi\pi)-m(\pi)$. For real D^* decays, Q has a value of about 5.8 MeV. In the right-sign plots (top of Fig. 1), clear signals are apparent over small backgrounds. The bands of events at $m(K\pi)$ or $m(K\pi\pi\pi)\approx 1.865$ GeV/c^2 are due to real D^0 decays combining with random pions in the event to give a false D^* candidate. These bands are more readily seen in the wrong-sign plots (bottom of Fig. 1) where the vertical scale is expanded by a factor of 20. This background, which we will refer to as ''random pion'' background, is the dominant one in our analysis. We will call the remaining broad background visible in the plots the ''false D^0'' background.

In the right-sign plot for the $K\pi$ mode, there is a signal of 5643 events above a background of 235 events. In the right-sign plot for the $K\pi\pi\pi$ mode, there is a signal of 3469 events above a background of 146 events. The signals and backgrounds are estimated in a region spanning $\pm 1.75\sigma$ around the peak in Q, for $1.77<m_{D^0}<1.97$ GeV/c^2. The precise region used to estimate S and B is not important for

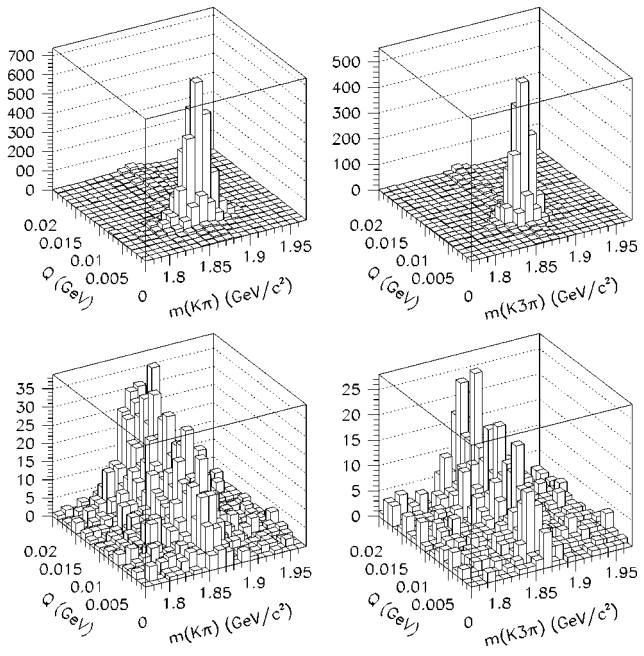

FIG. 1. Plots of Q (defined in the text) versus the candidate D mass for right-sign $D\to K\pi$ (top, left), right-sign $D\to K\pi\pi\pi$ (top, right), wrong-sign $D\to K\pi$ (bottom, left), and wrong-sign $D\to K\pi\pi\pi$ (bottom, right). Clean signals are apparent in both right-sign plots. In all four plots, the bands of events at $m(K\pi), m(K\pi\pi\pi)\approx 1.87$ GeV/c^2 are due to real D decays combining with random pions to give false D^* candidates.

the optimization. The resulting sensitivities $(S/\sqrt{B})$ for the two modes are 368 $(K\pi)$ and 287 $(K\pi\pi\pi)$.

V. FIT TECHNIQUE

The process just described results in eight separate data sets: D^0 or $\overline{D^0}$, decaying to $K\pi$ or $K\pi\pi\pi$, right-sign or wrong-sign decays. Although it is possible for us to fit each data set separately (which we have done as a check), it is useful to combine all eight data sets into a single fit. This allows us to take advantage of the fact that the central values of the D and D^* signals, as well as the $m_{K\pi}$, $m_{K\pi\pi\pi}$ and Q resolutions, are the same for the different data samples. Under these circumstances, most of the parameters of the single fit (which are largely parameters to describe background) remain uncoupled, and in that sense are no different from eight separate fits. Only the signal masses and resolutions are constrained across data samples. Studies of separate fits for the different samples show no significant shifts from the single fit results and have convinced us that these constraints are valid.

Our most general fit includes no constraints beyond those just described, and is summarized in Sec. VII and in Table V. However, as discussed in Sec. II, the most likely scenario is that there is no CP violation in either the DCS or the pure mixing terms of the wrong-sign rates. This leads to three additional constraints, discussed at the end of this section, which then lead to the results of Table II. These results are the main focus of our studies in mixing. Finally, in Sec. VII we perform other fits using additional physical restrictions (no DCS-mixing interference, or no mixing at all) in order to explore other physics hypotheses and to compare with previous measurements. In what follows, we describe the terms of the fit in detail.

As stated, we perform a single unbinned maximum likelihood fit to the data, using the following form for the ln of the likelihood:

$$\ln\mathcal{L}=\sum_i \ln\mathcal{L}_i-\sum_f\left(\frac{1}{2}\ln[2\pi N^f_{pred}]+\frac{[N^f_{pred}-N^f_{obs}]^2}{2N^f_{pred}}\right), \tag{6}$$

where the first sum is over all the D^* candidates, the second sum is over the eight decays used in the analysis, $\mathcal{L}_i$ represents the likelihood for each candidate, and N^f_{obs} is the observed number of candidates for each final state. The argument of the second sum is the logarithm of a normalized Gaussian, and serves to constrain the number of candidates predicted by the fit, N^f_{pred}. There are three contributions to each $\mathcal{L}_i$: signal, random pions with real D^0's, and random pions with false D^0's. In addition, for the $D^0\to K^{\mp}\pi^{\pm}$ samples we include a contribution from misidentified $D^0\to K^+K^-$ and $D^0\to\pi^+\pi^-$ decays, which also contribute measurable background.

$$\mathcal{L}_i=S(m_i,Q_i,t_i)+M(m_i,Q_i,t_i)+P(m_i,Q_i,t_i)+F(m_i,Q_i,t_i), \tag{7}$$

where m_i, Q_i, and t_i are the D mass, Q value and proper decay time of each candidate. A wrong-sign signal event is described by simple Gaussian terms in m_i and Q_i, multiplied by a sum of the three different decay time distributions that represent the DCS, mixing and interference contributions [see Eq. (1)]:

$$S(m_i,Q_i,t_i)=\frac{1}{N_{pred}}\frac{1}{\sqrt{2\pi}\sigma_D}e^{-(m_{D^0}-m_i)^2/2\sigma_D^2}\times\frac{1}{\sqrt{2\pi}\sigma_Q}e^{-(Q_{D^*}-Q_i)^2/2\sigma_Q^2}\{A_{dcs}B_{expt}(t_i)+A_{mix}B_{mix}(t_i)+A_{int}B_{int}(t_i)\}, \tag{8}$$

where

$$B_{expt}(t_i)=B^0_{expt}\epsilon(t_i)\int dt e^{-(t-t_i)^2/2\sigma_0^2}e^{-\Gamma t},$$
$$B_{mix}(t_i)=B^0_{mix}\epsilon(t_i)\int dt e^{-(t-t_i)^2/2\sigma_0^2}t^2e^{-\Gamma t}, \tag{9}$$
$$B_{int}(t_i)=B^0_{int}\epsilon(t_i)\int dt e^{-(t-t_i)^2/2\sigma_0^2}te^{-\Gamma t},$$

$\epsilon(t_i)$ is the reconstruction efficiency and σ_0 is the decay time resolution. Each $B(t_i)$ is normalized to unit integral so that A_{dcs}, A_{mix}, and A_{int} can be interpreted as the number of *observed* candidates of each type. The Gaussian smearing integrals are performed analytically with a smearing width $\sigma_0 = 0.05$ ps.

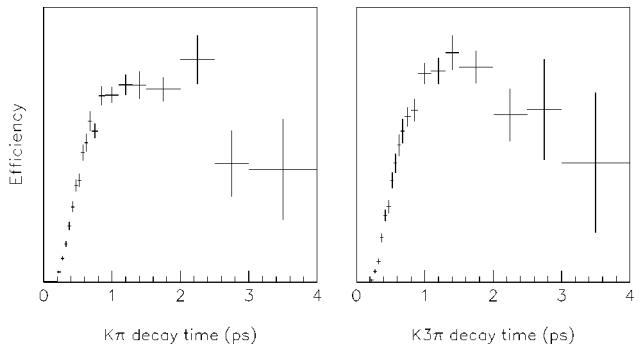

FIG. 2. Reconstruction efficiencies for $D^0 \to K\pi$ (left) and $D^0 \to K\pi\pi\pi$ (right), as measured from the right-sign data samples. The low efficiency at short decay times is typical of fixed-target experiments which identify charm decays by a secondary decay vertex. The drop in efficiency at long decay times is due to our selection criteria which remove decays occurring in downstream target foils. Efficiencies for charge conjugate final states (e.g., $\epsilon_{K^-\pi^+}$ and $\epsilon_{K^+\pi^-}$) are observed to be the same within errors, and have been combined in the above plots. The vertical scales are arbitrary.

The reconstruction efficiency $\epsilon(t)$ is the first of three functions that must be modeled for the fit. It is desirable to measure this function using real data rather than using Monte Carlo simulation. Fortunately, this can be accomplished with a sample of right-sign events. Since there is no mixing contribution to the right-sign decay rate, the true decay time distribution for right-sign decays is proportional to $e^{-\Gamma t}$ with $\Gamma = (0.415\ \text{ps})^{-1}$ [17]. The reconstructed distribution is proportional to $\epsilon(t)\int dt e^{-(t-t_i)^2/2\sigma_0^2} e^{-\Gamma t}$. Therefore, dividing the measured right-sign distribution (corrected for non-D^0 background using sideband subtraction) by the known smeared exponential gives a distribution proportional to the efficiency [18]. Figure 2 shows the results of that measurement for both the $K\pi$ and $K\pi\pi\pi$ final states, which we will use to represent $\epsilon(t)$ in the fit.

Despite the explicit mass cuts designed to reduce backgrounds from $D^0 \to KK$ and $D^0 \to \pi\pi$ decays described in Sec. IV, some contamination remains. The misidentified $D^0 \to KK$ and $D^0 \to \pi\pi$ events are described by

$$M(m_i, Q_i, t_i) = \frac{1}{N_{pred}} A_{KK,\pi\pi} U(m_i) V(Q_i) B_{expt}(t_i). \tag{10}$$

where the functions $U(m_i)$ and $V(Q_i)$ are obtained from Monte Carlo simulations of KK and $\pi\pi$ reflections remaining after all cuts, including explicit mass cuts designed to minimize these reflections. The parameters $A_{KK,\pi\pi}$ describe the number of events in the wrong-sign $K^+\pi^-$ and $K^-\pi^+$ samples. Similar backgrounds for the $K\pi\pi\pi$ mode are not significant (see Sec. VIII for further discussion).

The random pion background is described by

$$P(m_i, Q_i, t_i) = \frac{1}{N_{pred}} A_\pi \frac{1}{\sqrt{2\pi}\sigma_D} \times e^{-(m_{D^0} - m_i)^2/2\sigma_D^2} R(Q_i) B_{expt}(t_i). \tag{11}$$

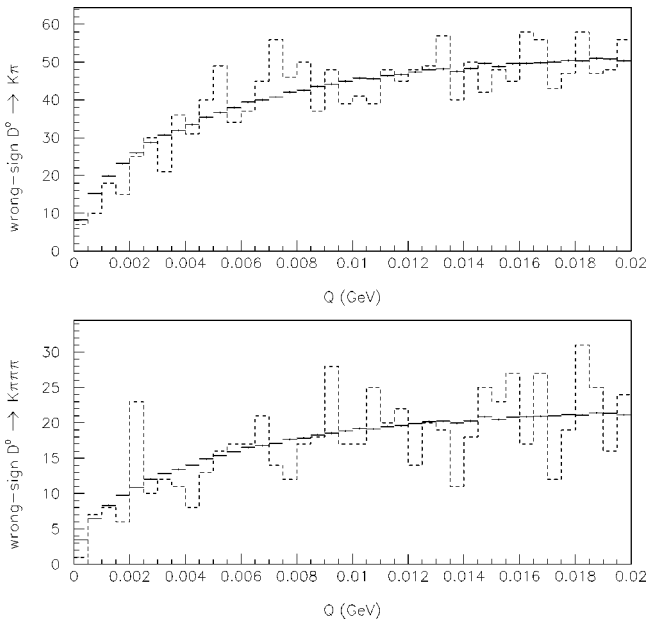

FIG. 3. Histograms show distributions of $Q \equiv m(K\pi\pi) - m(K\pi) - m(\pi)$ (top) and $Q \equiv m(K\pi\pi\pi\pi) - m(K\pi\pi\pi) - m(\pi)$ (bottom) for wrong-sign D candidates in the mass range 1.835–1.895 GeV/c^2. The points with error bars show the distributions from combining D^0 candidates and π's from separate events, normalized to the histograms. These distributions are used to represent $R(Q)$ in the likelihood fits.

The background shape in Q, represented by $R(Q)$, is independent of the candidate D mass. We model this shape by combining a D^0 candidate from one event with a π from another event. As long as the D^0 is not strongly correlated with other tracks in the event, and the selection cuts are not dependent on the spatial relation between the D^0 and other tracks in the event, this technique should provide an accurate model of background. Monte Carlo studies confirm the validity of this method. The resulting distribution is compared to the wrong-sign $D^0 \to K\pi$ and $D^0 \to K\pi\pi\pi$ data samples in Fig. 3.

The false D^0 background is adequately described by a linear function in m_i:

$$F(m_i, Q_i, t_i) = \frac{1}{N_{pred}} \frac{A_0[1 + A_1(m_i - m_0)]}{\Delta m} R(Q_i) B_{false}(t_i), \tag{12}$$

where m_0 is an arbitrary reference chosen to be 1.87 GeV/c^2 and Δm is the D^0 mass interval (0.2 GeV/c^2). The function $R(Q)$ is observed to be the same as in the case of the random pion and real D^0 background described by Eq. (11) above. The function $B_{false}(t_i)$ describes the time distribution of the false D^0 background. We model this distribution using candidates from the D mass sidebands of the right-sign event sample (Fig. 4). Since this background is very small, we do not need to model it with great precision, and the statistics of Fig. 4 are adequate.

The likelihood function for right-sign decays is constructed similarly. Since right-sign decays were used to model $\epsilon(t)$ and $B_{false}(t)$, we do not use the lifetime information for these events in the fit. Moreover, right-sign de-

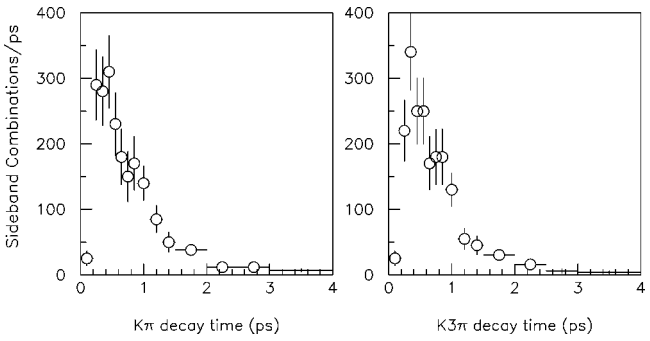

FIG. 4. The measured decay time distribution for $D^0 \to K\pi$ candidates (left) and $D^0 \to K\pi\pi\pi$ candidates (right), taken from the sidebands $1.77 < m_D < 1.81$ GeV/c^2 and $1.93 < m_D < 1.97$ GeV/c^2. These distributions are used to represent $B_{false}(t)$ in the likelihood fit.

cays are not subject to mixing or interference, and so the fit functions for these events are given by the simplified formulas

$$S(m_i, Q_i) = \frac{1}{N_{pred}} \frac{1}{\sqrt{2\pi}\sigma_D} \times e^{-(m_{D^0}-m_i)^2/2\sigma_D^2} \frac{1}{\sqrt{2\pi}\sigma_Q} e^{-(Q_{D^*}-Q_i)^2/2\sigma_Q^2} A_{rs},$$

$$P(m_i, Q_i) = \frac{1}{N_{pred}} A_\pi \frac{1}{\sqrt{2\pi}\sigma_D} e^{-(m_{D^0}-m_i)^2/2\sigma_D^2} R(Q_i), \tag{13}$$

$$F(m_i, Q_i) = \frac{1}{N_{pred}} \frac{A_0[1 + A_1(m_i - m_0)]}{\Delta m} R(Q_i).$$

We fit all the data, both right-sign and wrong-sign $D^0 \to K\pi$ and $D^0 \to K\pi\pi\pi$, simultaneously in one fit. Separate terms for charge conjugate final states are provided to allow for the most general possible form for CP violation. Under these conditions, we have four signal parameters (A_{DCS}, A_{mix}, A_{int} and $A_{KK,\pi\pi}$) and three background parameters (A_π, A_0 and A_1) for the two wrong-sign decay modes $D^0 \to K^+\pi^-$ and $\overline{D^0} \to K^-\pi^+$. We have three signal parameters (A_{DCS}, A_{mix} and A_{int}) and three background parameters (A_π, A_0 and A_1) for the two wrong-sign decay modes $D^0 \to K^+\pi\pi\pi$ and $\overline{D^0} \to K^-\pi\pi\pi$. For each right-sign mode ($\overline{D^0} \to K^+\pi^-$, $\overline{D^0} \to K^+\pi\pi\pi$, $D^0 \to K^-\pi^+$, and $D^0 \to K^-\pi\pi\pi$) we have one signal parameter (A_{rs}) and three background parameters (A_π, A_0 and A_1). Additionally, we have five mass parameters (m_{D^0}, $\sigma_{K\pi}$, $\sigma_{K\pi\pi\pi}$, Q_{D^*}, σ_Q) to describe the signal Gaussian functions. Separate D^0 mass resolutions are used for the $K\pi$ and $K\pi\pi\pi$ final states. The resolution in Q is dominated by the bachelor pion, and is therefore the same for the two final states.

With this list we have 47 parameters for a complete description of the data. However, we expect that the false D^0 backgrounds for right-sign and wrong-sign decays to the same D^0 final state should have the same slope parameter (A_1), although the level (A_0) may differ since they are combined with pions of different charge to form the D^* candidates. This observation reduces the number of parameters to 43.

TABLE I. Correction factors for the difference in integrated reconstruction efficiencies associated with the decay time distributions of the mixing and interference terms. See Eq. (15) in the text.

$c_{mix}(K\pi)$	0.388
$c_{int}(K\pi)$	0.499
$c_{mix}(K\pi\pi\pi)$	0.359
$c_{int}(K\pi\pi\pi)$	0.473

We also note that the values $r_{mix}(D^0 \to \overline{D^0})$ and $r_{mix}(\overline{D^0} \to D^0)$ should be independent of the D decay final state. Thus, without loss of generality, we can assume that $r_{mix}(D^0 \to K^+\pi^-) = r_{mix}(D^0 \to K^+\pi\pi\pi)$ and $r_{mix}(\overline{D^0} \to K^-\pi^+) = r_{mix}(\overline{D^0} \to K^-\pi\pi\pi)$. This eliminates two more parameters from our fit [19], leaving us with a total of 41 independent parameters to describe the full data set.

It is convenient to express the wrong-sign signal parameters A_{DCS}, A_{mix} and A_{int} in terms of the ratios of produced wrong-sign events to produced right-sign events, since these are the parameters of primary physics interest. For the wrong-sign $K^-\pi^+$ final state

$$r_{DCS}(K^-\pi^+) = \frac{A_{DCS}(\overline{D^0} \to K^-\pi^+)}{A_{rs}(\overline{D^0} \to K^+\pi^-)},$$

$$r_{mix}(K^-\pi^+) = \frac{A_{mix}(\overline{D^0} \to K^-\pi^+)}{A_{rs}(\overline{D^0} \to K^+\pi^-)} c_{mix}, \tag{14}$$

$$r_{int}(K^-\pi^+) = \frac{A_{int}(\overline{D^0} \to K^-\pi^+)}{A_{rs}(\overline{D^0} \to K^+\pi^-)} c_{int},$$

and similarly for $K^+\pi^-$, $K^-\pi\pi\pi$ and $K^+\pi\pi\pi$. The c's in the expressions for r_{mix} and r_{int} are given by

$$c_{mix} = \frac{\int dt_i \Gamma \epsilon(t_i) \int dt e^{-\Gamma t} e^{-(t-t_i)^2/2\sigma_0^2}}{\int dt_i \frac{1}{2}\Gamma^3 \epsilon(t_i) \int dt t^2 e^{-\Gamma t} e^{-(t-t_i)^2/2\sigma_0^2}}, \tag{15}$$

$$c_{int} = \frac{\int dt_i \Gamma \epsilon(t_i) \int dt e^{-\Gamma t} e^{-(t-t_i)^2/2\sigma_0^2}}{\int dt_i \Gamma^2 \epsilon(t_i) \int dt t e^{-\Gamma t} e^{-(t-t_i)^2/2\sigma_0^2}}.$$

These terms correct for the different integrated efficiencies for reconstructing wrong-sign DCS, mixing and interference events. Table I shows these correction factors for both the $K\pi$ and $K\pi\pi\pi$ final states.

Although the production characteristics of D^0 and $\overline{D^0}$ are different in our experiment, the ratios in Eq. (14) are designed to cancel this effect. In constructing these ratios we implicitly assume that the Cabibbo-favored amplitudes $\langle \overline{f}|H|D^0\rangle$ and $\langle f|H|\overline{D^0}\rangle$ are equal in magnitude, as men-

TABLE II. Signal and background parameters for the fit described in Sec. V, which assumes no CP violation in either the DCS or mixing terms. Thus, we have used the constraints $r_{DCS}(D^0\to K^+\pi^-)=r_{DCS}(\overline{D^0}\to K^-\pi^+)$, $r_{DCS}(D^0\to K^+\pi\pi\pi)=r_{DCS}(\overline{D^0}\to K^-\pi\pi\pi)$, and $r_{mix}(D^0\to\overline{D^0})=r_{mix}(\overline{D^0}\to D^0)$.

	$D^0\to K^-\pi^+$	$\overline{D^0}\to K^+\pi^-$	$D^0\to K^-\pi\pi\pi$	$\overline{D^0}\to K^+\pi\pi\pi$
A_{rs}	2269^{+49}_{-49}	2966^{+56}_{-56}	1314^{+38}_{-38}	1677^{+42}_{-42}
A_π	746^{+33}_{-33}	797^{+35}_{-35}	311^{+21}_{-21}	368^{+23}_{-23}
A_0	338^{+24}_{-24}	423^{+27}_{-27}	278^{+19}_{-19}	356^{+21}_{-21}
	$\overline{D^0}\to K^-\pi^+$	$D^0\to K^+\pi^-$	$\overline{D^0}\to K^-\pi\pi\pi$	$D^0\to K^+\pi\pi\pi$
r_{DCS}(%)	$0.90^{+1.20}_{-1.09}\pm 0.44$		$-0.20^{+1.17}_{-1.06}\pm 0.35$	
r_{mix} (%)	$0.39^{+0.36}_{-0.32}\pm 0.16$			
r_{int} (%)	$-0.46^{+0.89}_{-0.97}\pm 0.41$	$-0.84^{+0.92}_{-1.00}\pm 0.43$	$-0.29^{+0.89}_{-0.96}\pm 0.37$	$-0.25^{+0.87}_{-0.94}\pm 0.36$
A_π	737^{+32}_{-32}	749^{+30}_{-30}	323^{+19}_{-19}	315^{+19}_{-19}
A_0	243^{+24}_{-24}	333^{+22}_{-22}	214^{+16}_{-16}	238^{+17}_{-17}
$A_{KK,\pi\pi}$	60^{+15}_{-14}	49^{+15}_{-14}		
	$K^-\pi^+$	$K^+\pi^-$	$K^-\pi\pi\pi$	$K^+\pi\pi\pi$
A_1 (c^2/GeV)	$-3.29^{+0.81}_{-0.81}$	$-3.73^{+0.68}_{-0.68}$	$-2.86^{+0.76}_{-0.76}$	$-2.65^{+0.70}_{-0.70}$

tioned previously [9]. With this assumption, r_{mix} of Eq. (14) can be interpreted according to convention as

$$r_{mix}=\frac{1}{2\Gamma^2}\left|\frac{q}{p}\right|^2\left((\Delta M)^2+\frac{(\Delta\Gamma)^2}{4}\right). \tag{16}$$

At this point, the fit is completely general, with no physics assumptions applied. However, as discussed in Sec. II, it is unlikely that CP is violated in the DCS and pure mixing terms of the wrong-sign rates, even in most extensions to the standard model. Under these circumstances, there are three additional constraints, namely $r_{DCS}(D^0\to K^+\pi^-)=r_{DCS}(\overline{D^0}\to K^-\pi^+)$, $r_{DCS}(D^0\to K^+\pi\pi\pi)=r_{DCS}(\overline{D^0}\to K^-\pi\pi\pi)$, and $r_{mix}(D^0\to\overline{D^0})=r_{mix}(\overline{D^0}\to D^0)$. These constraints remove three more parameters from the fit, leaving a total of 38. We will use this fit to give us our primary result, summarized in Table II.

VI. RESULTS

We fit the data over the range 1.77–1.97 GeV/c^2 in m_D, 0.0–0.020 GeV/c^2 in Q and 0.0–4.0 ps in t. Tables II and III show the resulting 38 parameters from our primary fit, described in the previous section. The wrong-sign ratios are all small or consistent with zero, indicative of small DCS to Cabibbo-favored ratios and very little mixing. Using the criterion $\Delta\ln\mathcal{L}=0.82$ (neglecting systematic errors), we calculate the one-sided, 90% C.L. upper limit for mixing to be $r_{mix}<0.85\%$. There is also no evidence for CP violation. Figures 5 and 6 show the fit results overlaid on the data distributions for m_D, t, and Q. Good agreement is evident in every distribution.

The lego plots of Fig. 1 demonstrate that the largest background comes from real D decays combining with random pions to produce false D^* candidates. This phenomenon is also apparent in Figs. 5 and 6 where we see many wrong-sign candidates accumulated at the D^0 mass (left column) but very few of these candidates show the correct Q value (right column) to have come from D^* decays. A true wrong-sign signal from mixing or DCS decays would be manifest as a simultaneous peak in both the $M_{K\pi}$ (or $M_{K\pi\pi\pi}$) and Q distributions.

It is important to note that the excess of candidates at $Q\approx 0.006$ GeV in the wrong-sign decays $D^0\to K^+\pi^-$ and $\overline{D^0}\to K^-\pi^+$ (lower right plots of Fig. 5) is due primarily to $D^*\to D^0\pi$ with $D^0\to K^+K^-$ or $\pi^+\pi^-$, which is misidentified as $D^0\to K\pi$. These candidates are reconstructed at the right Q value for D^* decays, but appear outside the D^0 mass region in $M_{K\pi}$. Although it is hard to see these candidates in the lego plots of Fig. 1, they show up as the enhancements in the projected Q distributions in Fig. 5.

Figure 5 also shows the misreconstructed K^+K^- and $\pi^+\pi^-$ mass, time and Q distributions from our Monte Carlo studies as the crosshatched histograms in the bottom six plots. The normalization is determined by our fitted values for $A_{KK,\pi\pi}$ from Eq. (10). Although the reflections are barely

TABLE III. Mass parameters for the fit described in Sec. V, which assumes no CP violation in either the DCS or mixing terms. Thus, we have used the constraints $r_{DCS}(D^0\to K^+\pi^-)=r_{DCS}(\overline{D^0}\to K^-\pi^+)$, $r_{DCS}(D^0\to K^+\pi\pi\pi)=r_{DCS}(\overline{D^0}\to K^-\pi\pi\pi)$, and $r_{mix}(D^0\to\overline{D^0})=r_{mix}(\overline{D^0}\to D^0)$. All values are in MeV/$c^2$. There are systematic uncertainties on these parameters that are bigger than the statistical errors shown here, but they have inconsequential effects on the parameters of Table II.

m_{D^0}	$1865.8^{+0.1}_{-0.1}$
$\sigma_{K\pi}$	$15.16^{+0.16}_{-0.16}$
$\sigma_{K\pi\pi\pi}$	$10.76^{+0.15}_{-0.15}$
Q_{D^*}	$5.92^{+0.01}_{-0.01}$
σ_Q	$0.76^{+0.01}_{-0.01}$

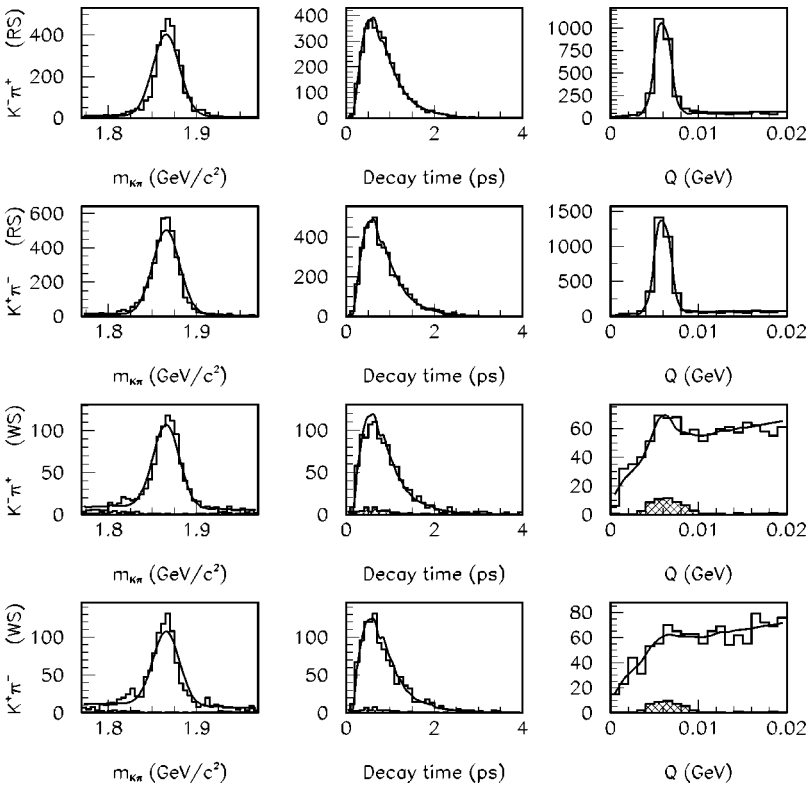

FIG. 5. Projections of the four $D\to K\pi$ data samples onto each of three distributions $m_{K\pi}$, t and Q. Data are from the D^0 mass range $[1.770 < m(K\pi) < 1.970]$ GeV/c^2. Solid curves show the projections of our primary fit, summarized in Table II. The broad component of the peaks in the wrong-sign Q plots is described well by reflections of K^+K^- and $\pi^+\pi^-$ signals. MC simulations of reflected K^+K^- and $\pi^+\pi^-$ signals normalized to the fit values of $A_{KK,\pi\pi}$ are shown as crosshatched histograms in the wrong-sign plots.

visible in the $m_{K\pi}$ and time distributions, they contribute a broad enhancement at 0.006 GeV in the Q distributions. Figure 7 shows the fitted contribution of $D^0\to K^+K^-$ and $\pi^+\pi^-$ misidentified decays scaled up by a factor of 20 and superimposed on the wrong-sign mass plot. The reflected signal is depleted in the D^0 mass signal region (indicated by arrows). More relevant to the mixing rate determination is Fig. 3a which shows the combined Q distributions from $D^0\to K^+\pi^-$ and $\overline{D^0}\to K^-\pi^+$, but with tighter cuts around the D^0 mass. Clearly, very little of the excess remains in the central D^0 mass region. The fit attributes only about 34 candidates to the total wrong-sign $K\pi$ signal.

We have also investigated the effects of other charm backgrounds which might feed into our wrong-sign samples using Monte Carlo studies and replotting correctly identified states as if they were misidentified. The largest such source

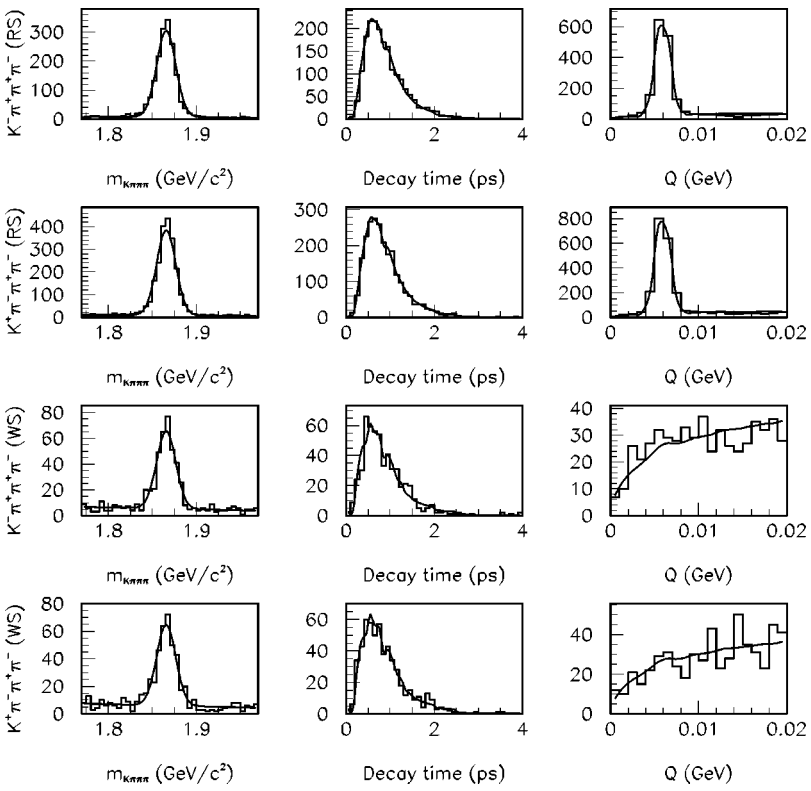

FIG. 6. Projections of the four $D\to K\pi\pi\pi$ data samples onto each of three distributions $m_{K\pi\pi\pi}$, t, and Q. Data are from the D^0 mass range $[1.770 < m(K\pi\pi\pi) < 1.970]$ GeV/c^2. Solid curves show the projections of our primary fit, summarized in Table II.

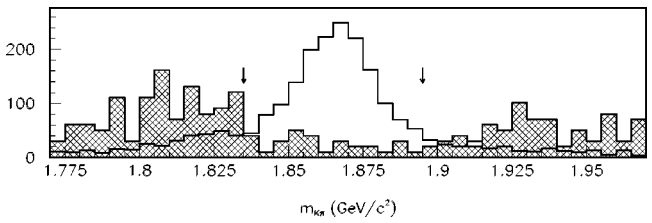

FIG. 7. The sum of the wrong-sign mass distributions in the lower two rows of Fig. 5. The crosshatched histogram is the reflected mass distribution of $D^0 \to K^+K^-$ and $D^0 \to \pi^+\pi^-$ decays from Monte Carlo simulations normalized to 20 times the amount favored by our primary fit, summarized in Table II. Notice the depletion of reflected signal in the D^0 mass signal region. Figure 3a shows the Q distribution for events within the restricted mass region indicated by the arrows.

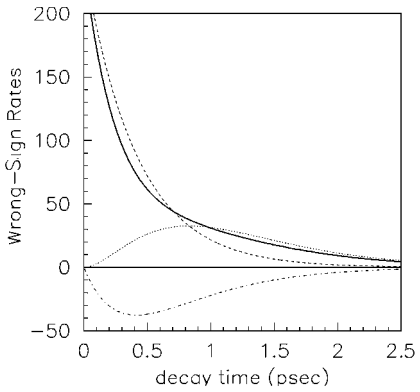

FIG. 8. A hypothetical plot of the time dependence of wrong-sign decays taken from [7]. The dashed line represents the DCS contribution. The dotted line shows the contribution due to mixing. The dash-dotted line shows the contribution from destructive interference of DCS and mixing amplitudes when the interference is 30% of its maximum. The solid line is the sum of all three contributions. The vertical scale is arbitrary.

of background comes from doubly misidentified decays of $D^0 \to K\pi$ or $K\pi\pi\pi$, in which the K and a π of opposite charge are both misidentified (as π and K) by the Čerenkov detector. Although we explicitly cut against these misidentified decays in our data selection, a small fraction is expected to pass our cuts. Our Čerenkov measurements allow D^0 decays to be doubly misidentified around 1.3% of the time. However, only about 15% of these candidates have an invariant mass within 20 MeV/c^2 of the D^0 mass. The selection cut on the reflected mass of each candidate further reduces this background by about a factor of 20. Furthermore, since the background is very broadly distributed in mass, rather than peaked in the signal region, we expect the fit to respond only weakly by changing the wrong-sign ratios, probably at the level of a few times 10^{-4} or lower. Since this background has an exponential decay time distribution, it will be interpreted as a signal for r_{DCS}, and will not affect the measurement of r_{mix} or r_{int} at all.

The remaining potential sources of charm background come from $D^* \to \pi D$ decays with the D decaying to a mode other than $K\pi$ or $K\pi\pi\pi$. For the $K\pi$ mode, the $D^0 \to K^+K^-$ and $\pi^+\pi^-$ were the most significant, and were handled as described previously. In addition, we have examined decays $D^0 \to K^-\pi^+\pi^0$, $K^-\mu^+\nu$ (doubly misidentified), and $D^0 \to \pi^+\pi^-\pi^0$ (singly misidentified), which might contribute as background to $D^0 \to K\pi$. As a general rule, the misidentification rates for these modes are similar to what was observed for the double misidentification above (all misidentification is dominated by the Čerenkov selection criteria for the kaon candidate), while the misidentified masses (in some cases after losing a neutral particle) are well outside the signal region. None of the other decays were seen to contribute a significant background to our data samples.

In performing the fit, we discovered that the r_{DCS} and r_{mix} terms are strongly anticorrelated with the r_{int} terms, and strongly correlated with each other. Figure 8 demonstrates how these correlations come about in a hypothetical case where the interference contribution approximately cancels the contribution from pure mixing. This plot demonstrates that even when the full time evolution deviates only slightly from the pure exponential form of DCS decays, a large contribution from mixing can be present if it is offset by a destructive interference contribution. This implies that the fitted values for the interference contribution and the mixing contribution are strongly anticorrelated.

Figure 9 illustrates the correlations in our particular fit by showing the likelihood contour plots for representative pairs of parameters for the $D^0 \to K^+\pi^-$ mode. These strong correlations account for much of the uncertainty in the wrong-sign ratios. Table IV gives the correlation coefficients for the different wrong-sign ratios. The correlations of these ratios with all other parameters of the fit are negligible. We note that the correlations would be slightly reduced in an experiment with better efficiency at short decay times where there is good discrimination between r_{DCS} and the other terms.

TABLE IV. Correlation coefficients for the wrong-sign ratios from the fit of Table II. Only the lower halves of the symmetric matrices are shown. Correlations with the fit parameters that are not shown are negligible.

	$r_{DCS}(K\pi)$	$r_{DCS}(K\pi\pi\pi)$	$r_{int}(K^-\pi^+)$	$r_{int}(K^+\pi^-)$	$r_{int}(K^-\pi\pi\pi)$	$r_{int}(K^+\pi\pi\pi)$	r_{mix}
$r_{DCS}(K\pi)$	1.00						
$r_{DCS}(K\pi\pi\pi)$	0.58	1.00					
$r_{int}(K^-\pi^+)$	-0.92	-0.68	1.00				
$r_{int}(K^+\pi^-)$	-0.90	-0.68	0.95	1.00			
$r_{int}(K^-\pi\pi\pi)$	-0.71	-0.90	0.84	0.85	1.00		
$r_{int}(K^+\pi\pi\pi)$	-0.70	-0.90	0.82	0.83	0.95	1.00	
r_{mix}	0.78	0.74	-0.92	-0.93	-0.92	-0.90	1.00

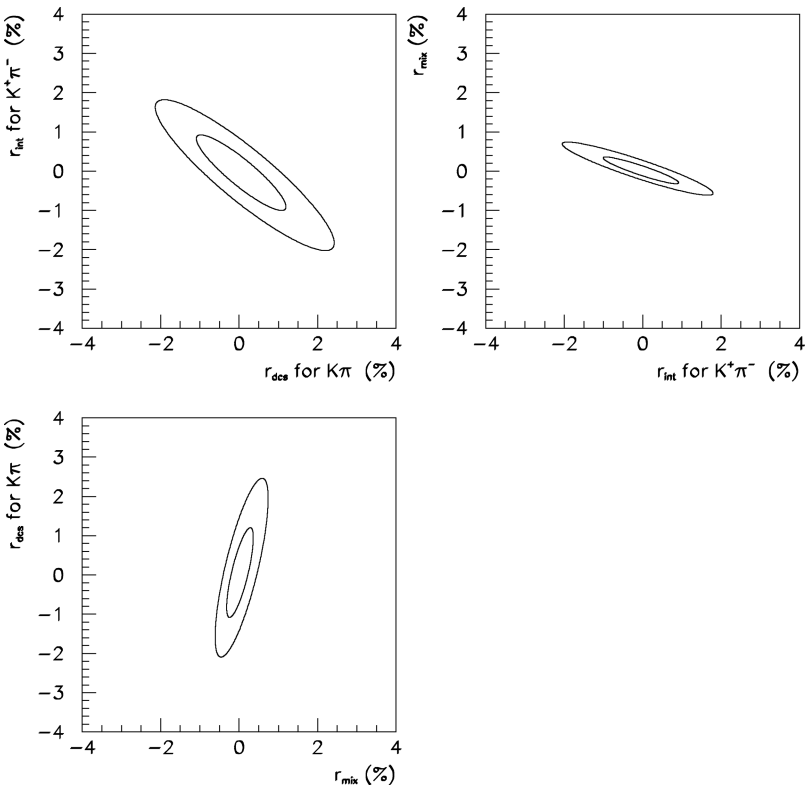

FIG. 9. Likelihood contours corresponding to $\Delta\ln\mathcal{L}=0.5$ and 2.0 for the fit of Table II, illustrating the correlations among the three parameters $r_{DCS}(K\pi)$, r_{mix} and $r_{int}(K^+\pi^-)$. Strong correlations among these parameters are apparent. The correlations among other wrong-sign ratios are similar.

VII. OTHER FITS

Table II shows our primary results in the search for mixing. These results assume that CP violation can only occur in the interference terms of the fit, an assumption supported by most extensions to the standard model (see the discussion in Sec. II). However, to answer any concerns about this assumption, we have also performed a fit in which the CP constraints are relaxed. Table V shows the results for the wrong-sign ratios of that 41 parameter fit. As expected, the central values for r_{DCS} and r_{mix} bracket the corresponding combined terms in Table II, and all the fit errors have increased. Using the criterion $\Delta\ln\mathcal{L}=0.82$ (neglecting systematic errors), we calculate the one-sided, 90% C.L. upper limits to be $r_{mix}(\overline{D^0}\to D^0)<0.74\%$ and $r_{mix}(D^0\to\overline{D^0})<1.45\%$.

We note that the earlier measurement by the E691 collaboration [3] assumed that the interference terms r_{int} were negligible. Recently, there has been lively discussion concerning the validity of this assumption [7,8,20]. Although some authors suggest that the phase between DCS and mixing amplitudes may be small, and therefore that the interference terms r_{int} should also be small, we prefer to quote our results without this constraint. Nonetheless, to compare our measurements with the previous results from E691, we have performed a fit in which we set the interference terms to zero. Our results for mixing are $r_{mix}=0.21^{+0.09}_{-0.09}\pm0.02\%$, which is to be compared with the E691 result $r_{mix}=(0.05\pm0.20)\%$. The reduction of the fit errors from Table II is indicative of the strong correlations with the r_{int} parameters, which are now fixed at zero. E691 also touched on this point by considering several different fixed values of the interference term (only one interference term was allowed in their model). Their results showed behavior similar to what we see in our data: strong correlation between wrong-sign ratios, and reduced fit errors when fixing the interference term to zero.

Finally, we explore the possibility that mixing is completely negligible, as one would expect from purely standard model contributions. In this case, we fit only for the DCS terms, obtaining $r_{DCS}(K\pi)=(0.68^{+0.34}_{-0.33}\pm0.07)\%$ and $r_{DCS}(K\pi\pi\pi)=(0.25^{+0.36}_{-0.34}\pm0.03)\%$. The result for $D^0\to K\pi$ demonstrates a two-sigma excess in the signal region which we believe is the result of real DCS decays. Figure 10 shows this excess after background subtraction. We note that our value for the DCS rate is consistent with the CLEO measurement [6] for the total wrong-sign rate: $r_{WS}(K\pi)=(0.77\pm0.25\pm0.25)\%$.

VIII. SYSTEMATIC UNCERTAINTIES

Systematic uncertainties in the fit arise primarily from our modeling of the three functions $\epsilon(t)$, $R(Q)$, and $B_{false}(t)$.

TABLE V. Fit results for the wrong-sign ratios of the most general fit, with no assumptions about CP.

	$\overline{D^0}\to K^-\pi^+$	$\overline{D^0}\to K^-\pi\pi\pi$	$D^0\to K^+\pi^-$	$D^0\to K^+\pi\pi\pi$
$r_{DCS}(\%)$	$0.80^{+1.46}_{-1.37}\pm0.47$	$-0.67^{+1.44}_{-1.35}\pm0.41$	$1.26^{+1.94}_{-1.79}\pm0.49$	$0.33^{+1.91}_{-1.70}\pm0.32$
r_{mix} (%)	$0.18^{+0.43}_{-0.39}\pm0.17$		$0.70^{+0.58}_{-0.53}\pm0.18$	
r_{int} (%)	$-0.11^{+1.08}_{-1.16}\pm0.43$	$0.22^{+1.12}_{-1.18}\pm0.41$	$-1.46^{+1.49}_{-1.59}\pm0.47$	$-0.89^{+1.35}_{-1.46}\pm0.35$

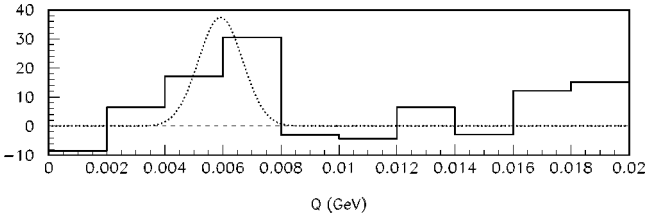

FIG. 10. Distribution in Q for wrong-sign $D \to K\pi$ candidates in the mass range 1.845–1.885 GeV/c^2. The plot has been background subtracted using the fit with no mixing (standard model case). The Gaussian overlay shows the size of the signal attributed to DCS decays by the fit.

By using the right-sign data samples to estimate these functions, we have minimized our dependence on Monte Carlo models, but some uncertainty remains. The results of our studies of systematic uncertainties are summarized in Tables VI and VII. Below we describe the entries in Table VI which are the systematic errors obtained from studies on the fit of Table II. The entries in Table VII are obtained similarly. We find that the systematic uncertainties in the analysis are small compared with the statistical errors from the fit.

The uncertainties in the first row arise from our estimates of the size of the reflection in $K\pi$ from misidentified $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ events. Most of this uncertainty arises from our knowledge of the branching ratios for these modes. We have estimated this error as 20% of the difference between the results of fits with and without terms describing these reflections.

The uncertainties listed in the third row (''statistics of effy and bkgd distrs'') result from the uncertainties in our estimates of $\epsilon(t)$, $R(Q)$, and $B_{false}(t)$ due to the finite size of our right-sign data sample. The error bars on the corresponding histograms in Figs. 2, 3, and 4 show the level of uncertainty involved. Statistical uncertainties in the model of the function $R(Q)$ (Fig. 3) have been greatly reduced by combining each D^0 candidate with pions from many different events, so that the uncertainties in this histogram are negligible in comparison with the uncertainties of Figs. 2 and 4. In order to propagate these statistical uncertainties to uncertainties in the fitted parameters, we perform many fits to the data, modifying each bin in each of the histograms of $\epsilon(t)$, $R(Q)$ and $B_{false}(t)$ by a Gaussian fluctuation with the resolution given by the error bars. The rms spreads of the fitted parameters from 25 such fits are given in Table VI.

The uncertainties listed in the fourth row of Table VI (''binning of effy and bkgd distrs'') are due to the fact that we have represented the $\epsilon(t)$, $R(Q)$, and $B_{false}(t)$ functions by binned histograms rather than smooth functions. We have, of course, tried to choose bin sizes small enough so that binning effects are not significant. In order to verify this claim, we replaced the histograms in the fit with smoothed functions which were derived from the histogram data, and repeated the fit. The differences between the parameter values with the smoothed functions and the parameter values with the histograms are quoted in Table VI as the uncertainties due to binning. We expect this method to give an overestimate of the binning effect, since it also includes the effect of some statistical fluctuations in the measured histograms which are adjusted by the smoothing function. As anticipated, the binning effects are small.

The uncertainties listed in the fifth row of the table (''time resolution'') are due to the resolution on the measured decay time. Since the smearing is small, a good assumption is that $\epsilon(t)$ is almost the same with and without smearing. For this analysis, the most likely resolution on the decay time is about 0.03 ps, with some measurements having a resolution as large as 0.08 ps. In order to quantify the error due to smearing, we replace the functions of Eq. (9) by functions convoluted with a fixed Gaussian resolution:

$$B_{expt}(t_i) = B^0_{expt} \int dt e^{-(t-t_i)^2/2\sigma_0^2} e^{-\Gamma t} \epsilon(t),$$

$$B_{mix}(t_i) = B^0_{mix} \int dt e^{-(t-t_i)^2/2\sigma_0^2} t^2 e^{-\Gamma t} \epsilon(t), \qquad (17)$$

$$B_{int}(t_i) = B^0_{int} \int dt e^{-(t-t_i)^2/2\sigma_0^2} t e^{-\Gamma t} \epsilon(t),$$

TABLE VII. Systematic uncertainties in the no-mixing fit, corresponding to the standard model case. Entries are explained in the text.

	$r_{DCS}(K\pi)(\%)$	r_{DCS} $(K\pi\pi\pi)(\%)$
Fit value	$0.68^{+0.34}_{-0.33}$	$0.25^{+0.36}_{-0.34}$
K^+K^-, $\pi^+\pi^-$ reflections	± 0.04	± 0.00
Statistics of effy and bkgd distrs	± 0.02	± 0.02
Binning of effy and bkgd distrs	± 0.01	± 0.00
Time resolution	± 0.01	± 0.01
Mass resolution	± 0.04	± 0.02
Total systematic uncertainty	± 0.07	± 0.03

TABLE VI. Systematic uncertainties for the key parameters in the fit of Table II. This fit describes the case where we have allowed for CP violation only in the interference term. Entries are explained in the text.

	$r_{DCS}(K\pi)(\%)$	$r_{DCS}(K\pi\pi\pi)(\%)$	$r_{mix}(\%)$
Fit value	$0.90^{+1.20}_{-1.09}$	$-0.20^{+1.17}_{-1.06}$	$0.39^{+0.36}_{-0.32}$
K^+K^-, $\pi^+\pi^-$ reflections	± 0.02	± 0.02	± 0.01
Statistics of effy and bkgd distrs	± 0.27	± 0.17	± 0.09
Binning of effy and bkgd distrs	± 0.25	± 0.21	± 0.11
Time resolution	± 0.17	± 0.19	± 0.06
Mass resolution	± 0.18	± 0.11	± 0.05
Total systematic uncertainty	± 0.44	± 0.35	± 0.16

where $\epsilon(t)$ is obtained as in Sec. V and the integrals are obtained numerically. We then perform the fit with three different values for σ_0: 0.02 ps, 0.05 ps and 0.08 ps. We quote the average of the central value differences (from the fits for 0.02 ps and 0.05 ps and from the fits for 0.08 ps and 0.05 ps) as variations in Table VI, line 5. When the exponential lifetime is modified by our detector acceptance which is poor at low lifetimes, there is a ''peak'' at $\approx$0.5 ps. We observe that the time smearing affects the likelihood most near this ''peak,'' while mixed events are most likely around higher values of decay time. Consequently, the DCS ratios exhibit the largest variation, while the mixing ratio is relatively stable.

The uncertainties in the sixth line of the table (''mass resolution'') come from the assumption of a constant Gaussian resolution in $m(K\pi)$ and $m(K\pi\pi\pi)$. In truth, the mass resolution should depend on the D^0 momentum and on the kinematics of the decay. We have studied the dependence on momentum and verified a noticable correlation between resolution and momentum. For the $K\pi$ decay mode, about 90% of the events have a mass resolution between 12 and 16 MeV/c^2, with a tail reaching out to about 25 MeV/c^2 at high momentum. For the $K\pi\pi\pi$ mode, the variation is much smaller, with all events exhibiting a mass resolution in the range 8.5–11 MeV/c^2. To quantify this effect, we have varied the mass resolutions ± 2 MeV/c^2 in the fit and recorded the maximum variations in wrong-sign ratios in Table VI. The fit results are quite insensitive to variation of the resolution on $m(K\pi\pi\pi)$, but change slightly with the resolution on $m(K\pi)$. A change in $m(K\pi)$ resolution primarily affects $r_{DCS}(K\pi)$, but because of the correlations in the fit, it will also alter r_{mix} and $r_{DCS}(K\pi\pi\pi)$, as shown.

Some other assumptions and biases in our fit model bear further comment. First of all, we have assumed that the efficiency function $\epsilon(t)$ for reconstructing a D^0 from a D^* decay [signal terms from Eq. (8)] is the same as the efficiency function for reconstructing a primary D^0 [random pion term from Eq. (11)]. Since our reconstruction and selection criteria are only weakly dependent on the D^0 production kinematics, we expect to find very little difference in efficiency for these two sources of D mesons. Studies of reconstructed D^0 decays which are *not* associated with a bachelor pion appear to confirm that the difference is negligible. Second, the $K\pi\pi\pi$ final state may result from different resonant substructures in the Cabibbo-favored and DCS amplitudes. This can, in principle, lead to different efficiencies for the DCS and interference terms in the fit. (The two-body $K\pi$ mode is, of course, immune to this problem.) Once again, however, the fact that our reconstruction depends only slightly on the decay kinematics leads to effects at the level of only 1%. The very similar time dependence of the efficiency functions for the $K\pi$ and $K\pi\pi\pi$ final states (Fig. 2) demonstrates how little $\epsilon(t)$ depends on the D decay.

We are also aware that training the D^* neural net on a sample of right-sign D^* decays can, in principle, produce a small bias in that sample (but not the wrong-sign samples which were not used for training the net). Careful selection of input variables for the neural net that are not correlated with Q (our variables depend only on parameters that describe the bachelor pion) should prevent any significant bias. We have investigated this effect by subdividing the training sample into 10 subsamples, training a neural net on one subsample, and applying the resulting net to the remaining samples as a test of the bias. We then repeat the process on each subsample to get a better statistical average. By comparing the sensitivity (measured as $S/\sqrt{B}$) of the training samples with the sensitivity of the test samples, we determine the level of bias. We find that the number of right-sign signal events could be biased upwards by about 1%. This is a negligible effect compared to our statistical error.

The last line in Table VI shows the contribution from all the systematic errors added in quadrature. These totals are less than half the size of the statistical errors in Table II.

IX. DISCUSSION

At the current level of sensitivity, mixing searches begin to constrain some models [21]. There are also other search methods that are promising. Using the same D^* decay chain to identify the produced D meson, but looking at semileptonic decays of the D, is one possibility. Although semileptonic decays are harder to reconstruct due to missing neutrinos, they are not subject to contributions from DCS amplitudes, and therefore do not suffer from the main limitations discussed in this paper. In a separate publication we describe such a search [5] with the result $r_{mix}<0.50\%$ at the 90% C.L. The possibility exists for even higher statistics searches in future experiments. Alternatively, it may be possible to detect mixing via the lifetime difference between the two physical eigenstates by comparing the measured lifetimes for different CP final states [22]. Of course, this approach will only detect mixing if it is associated with a substantial lifetime difference as opposed to mixing that only results from a mass difference. We are investigating this method as well. Finally, the cleanest signal for mixing might be found at a τ-charm factory which produces D^0-$\overline{D^0}$ pairs on resonance. As has been discussed previously [23], certain hadronic final states from these D^0-$\overline{D^0}$ pairs can only be produced by mixing and not by DCS amplitudes. We remain hopeful that one of these techniques may be used to detect D^0-$\overline{D^0}$ mixing, and thus provide information about the existence of new physics.

X. SUMMARY

We have searched for evidence of D^0-$\overline{D^0}$ mixing and DCS decays by looking for wrong-sign decays in the decay chain $D^*\rightarrow\pi D$ with $D\rightarrow K\pi$ or $D\rightarrow K\pi\pi\pi$. Our results are summarized in Table VIII.

We have seen no evidence for mixing in either D^0 decay mode. The results of a maximum likelihood fit to the data are given in Table II. The possibility of additional sources of wrong-sign decays from DCS amplitudes limits our sensitivity for detecting mixing alone. Using the criterion $\Delta\ln\mathcal{L}=0.82$, we calculate the one-sided, 90% C.L. upper limit to be $r_{mix}<0.85\%$. If, in order to account for the most general case possible, we relax the assumption that CP is conserved in the mixing and DCS terms of the fit (as in Table V), we calculate the upper limits for mixing to be $r_{mix}(\overline{D^0}\rightarrow D^0)<0.74\%$ and $r_{mix}(D^0\rightarrow\overline{D^0})<1.45\%$.

Our quoted sensitivity to mixing is similar to that of Fer-

TABLE VIII. A summary of values from our four fits. The top line describes the most likely case for extensions to the standard model which produce large mixing. The second line describes our most general fit, providing the most conservative results. The third line matches the assumptions of previous experiments, which we do not feel are justifiable. The bottom line describes the standard model case. Results from other experiments are also listed for comparison.

Fit type	This result	Other comparable results
CP violation only in interference term	$r_{mix}=(0.39^{+0.36}_{-0.32}\pm 0.16)\%$	$r_{mix}=(0.11^{+0.30}_{-0.27})\%$ (E791 [5]) (semileptonic decays)
Most general, no CP assumptions	$r_{mix}(\overline{D^0}\to D^0)=(0.18^{+0.43}_{-0.39}\pm 0.17)\%$ $r_{mix}(D^0\to\overline{D^0})=(0.70^{+0.58}_{-0.53}\pm 0.18)\%$	
No CP violation, no interference	$r_{mix}=(0.21^{+0.09}_{-0.09}\pm 0.02)\%$	$r_{mix}=(0.05\pm 0.20)\%$ (E691 [3])
No mixing	$r_{DCS}(K\pi)=(0.68^{+0.34}_{-0.33}\pm 0.07)\%$ $r_{DCS}(K\pi\pi\pi)=(0.25^{+0.36}_{-0.34}\pm 0.03)\%$	$r_{ws}(K\pi)=(0.77\pm 0.25\pm 0.25)\%$ (CLEO [6])

milab E691, but our analysis is notably more general in its assumptions concerning DCS-mixing interference and CP violation. Assuming no DCS-mixing interference constrains the mixing and DCS contributions much more severely, but we do not feel this assumption is justifiable. Nevertheless, for comparison we include the mixing results for this case in Table VIII. All our results for the $K\pi$ final state are also consistent with the CLEO measurement of $r_{WS}(K\pi)=(0.77\pm 0.25\pm 0.25)\%$ [6] for wrong-sign decays. In particular, if the mixing amplitude is set to zero, we find a two-sigma enhancement in the $K\pi$ mode and no significant enhancement in the $K\pi\pi\pi$ mode: $r_{DCS}(K\pi)=(0.68^{+0.34}_{-0.33}\pm 0.07)\%$ and $r_{DCS}(K\pi\pi\pi)=(0.25^{+0.36}_{-0.34}\pm 0.03)\%$.

ACKNOWLEDGMENTS

We gratefully acknowledge the assistance of the staff of Fermilab and of all the participating institutions. This research was supported by the Brazilian Conselho Nacional de Desenvolvimento Científico e Technológico, CONACyT (Mexico), the Israeli Academy of Sciences and Humanities, the U.S. Department of Energy, the U.S.-Israel Binational Science Foundation, and the U.S. National Science Foundation. Fermilab is operated by the Universities Research Association, Inc., under contract with the United States Department of Energy.

[1] See, for example, A. Datta and D. Kumbhakar, Z. Phys. C **27**, 515 (1985); J. F. Donoghue *et al.*, Phys. Rev. D **33**, 179 (1986); T. Ohl, G. Ricciardi, and E. H. Simmons, Nucl. Phys. **B403**, 605 (1993).

[2] For a recent review of many new physics contributions to D^0-$\overline{D^0}$ mixing see J. L. Hewett, in *Searching for New Physics with Charm*, LISHEP '95 cbt Proceedings, Rio de Janeiro, Brazil, 1995, edited by F. Caruso *et al.* (Editions Frontieres, Gif-sur-Yvette, France, 1995); Another resource is Y. Nir, in *VI International Symposium on Heavy Flavor Physics*, Proceedings, Pisa, Italy, 1995, edited by F. Constantini and M. Girogi [Nuovo Cimento **109A** 991 (1996)], hep-ph/9507290, and references therein. See also M. Leurer, Phys. Rev. Lett. **71**, 1324 (1993); M. Leurer, Phys. Rev. D **49**, 333 (1994).

[3] J. C. Anjos *et al.*, Phys. Rev. Lett. **60**, 1239 (1988).

[4] W. C. Louis *et al.*, Phys. Rev. Lett. **56**, 1027 (1986). The mixing limit in this paper is based on the assumption that the charm production cross section rises linearly with atomic mass of the target. The limit would be less strict if the cross section rises more slowly.

[5] E. M. Aitala *et al.*, Phys. Rev. Lett. **77**, 2384 (1996).

[6] D. Cinabro *et al.*, Phys. Rev. Lett. **72**, 1406 (1994).

[7] G. Blaylock, A. Seiden, and Y. Nir, Phys. Lett. B **355**, 555 (1995).

[8] L. Wolfenstein, Phys. Rev. Lett. **75**, 2460 (1995).

[9] We have assumed in this expression that the Cabibbo-favored amplitudes, which are described by standard model processes, are dominated by a single weak phase. Thus, $\langle f|H|\overline{D^0}\rangle_{CF}=\langle \overline{f}|H|D^0\rangle_{CF}$ to good approximation. This will also be of relevance in our interpretation of r_{mix}, discussed later in this paper.

[10] J. A. Appel, Annu. Rev. Nucl. Part. Sci. **42**, 367 (1992), and references therein; D. J. Summers *et al.*, in Proceedings of the XXVII Recontre de Moriond, Electroweak Interactions and Unified Theories, Les Arcs, France, 1992, p. 417; S. Amato *et al.*, Nucl. Instrum. Methods Phys. Res. A **324**, 535 (1993); S. Bracker *et al.*, IEEE Trans. Nucl. Sci. **NS-43**, 2457 (1996).

[11] D. Bartlett *et al.*, Nucl. Instrum. Methods Phys. Res. A **260**, 55 (1987).

[12] See, e.g., R. Beale and T. Jackson, *Neural Computing: An Introduction* (Hilger, Bristol, 1990).

[13] See, e.g., S. S. Haykin, *Neural Networks: A Comprehensive Foundation* (Maxwell Macmillan International, New York, 1994).

[14] G. Strang, *Linear Algebra and its Applications*, 3rd ed. (Harcourt, Brace, Jovanovich, San Diego, 1988). See also G. H. Golub and C. F. Van Loan, *Matrix Computations* (Johns Hop-

kins University Press, Baltimore, MD, 1989).

[15] P. P. Kanjilal and D. N. Banerjee, IEEE Trans. Neural Netw. **6**, 1061 (1995).

[16] For a general reference on binary decision trees see L. Brieman, J.H. Friedman, R.A. Olshen and C.J. Stone, *Classification and Regression Trees* (Chapman and Hall, New York, 1984).

[17] Particle Data Group , R. M. Barnett *et al.*, Phys. Rev. D **54**, 1 (1996), p. 33.

[18] We have made the simplifying assumption that the efficiency can be expressed as a function of t alone. In truth, it can depend on a number of parameters involving the location of the decay vertex, the D momentum, and the decay kinematics. Consequently, our formalism is rigorously correct only for the DCS term, which has the same time distribution as the right-sign sample used to measure $\epsilon(t)$. However, in this analysis, we believe the efficiency is mostly dependent on the decay time of the D^0 and is therefore the same for all three time distributions of Eq. (9). Of course, the efficiency may more correctly be described as a function of the true decay time instead and should therefore be included inside the integrals of Eq. (9). Since the smearing is small, we do not do so and discuss the associated systematic error in Sec. VII below.

[19] Instead of constraining r_{mix} to be the same for both $K\pi$ and $K\pi\pi\pi$ modes, we could of course fit the modes separately and combine the results to obtain an estimate of the common value of r_{mix}. However, because of the strong correlations present, it is easier to let the fit perform the constraint than to propagate the full correlation matrices in a weighted mean of the separate modes. The results are essentially the same in either case.

[20] For a discussion of the interference between DCS and mixing amplitudes in studies of D^0-$\overline{D^0}$ mixing, see the following as well as Refs. [7,8]: T. E. Browder and S. Pakvasa, Phys. Lett. B **383**, 475 (1996).

[21] Examples of models which lead to large mixing are described in K. S. Babu *et al.*, Phys. Lett. B **205**, 540 (1988); L. Hall and S. Weinberg, Phys. Rev. D **48**, 979 (1993).

[22] The utility of searching for D^0-$\overline{D^0}$ mixing by examining the lifetime difference of different CP final states of neutral D decay was first brought to our attention by the following reference: T. Liu, ''Proceedings of the Workshop on the Future of High Sensitivity Charm Experiments: CHARM2000,'' Batavia, IL, 1994, Harvard Report No. HUTP-94-E021, 1994, and hep-ph/9408330.

[23] I. Bigi and A. I. Sanda, Phys. Lett. B **171**, 320 (1986).