SLAC-PUB-9795
hep-ex/9703006

The Inclusive Decays B -> DX and B -> D(*)X

Submitted to Physical Review D

Stanford Linear Accelerator Center, Stanford University, Stanford, CA 94309

Work supported by Department of Energy contract DE–AC03–76SF00515.

CLNS 96/1454
CLEO 96-23

The Inclusive Decays $B \to DX$ and $B \to D^*X$

CLEO collaboration

Abstract

We report new measurements of the differential and total branching ratios for inclusive B decay to D^0, D^+ and D^{*+} and the first measurement of the same quantities for inclusive B decay to D^{*0}. Here B is the mixture of B_d and B_u from $\Upsilon(4S)$ decay. Furthermore, since more than one charm particle (or antiparticle) of the same kind can be produced in B decay, here "inclusive B branching ratio" is used to mean the average number of charm particles and their antiparticles of a certain species produced in B decay. We obtain the following results (the first error is statistical, the second systematic of this analysis, the third is propagated from other measurements): $\mathcal{B}(B \to D^0X) = (0.636 \pm 0.014 \pm 0.019 \pm 0.018), \mathcal{B}(B \to D^+X) = (0.235 \pm 0.009 \pm 0.009 \pm 0.024), \mathcal{B}(B \to D^{*0}X) = (0.247 \pm 0.012 \pm 0.018 \pm 0.018), \mathcal{B}(B \to D^{*+}X) = (0.239 \pm 0.011 \pm 0.014 \pm 0.009)$. The following ratio of branching ratios is not affected by most of the systematic errors: $\mathcal{B}(B \to D^{*0}X)/\mathcal{B}(B \to D^{*+}X) = (1.03 \pm 0.07 \pm 0.09 \pm 0.08)$. We also report the first measurement of the momentum-dependent D^{*0} polarization and a new measurement of the D^{*+} polarization in inclusive B decay. Using these measurements and other CLEO results and making some additional assumptions, we calculate the average number of c and $\bar{c}$ quarks produced in B decay to be $\langle n_c \rangle = 1.10 \pm 0.05$.

L. Gibbons,[1] S. D. Johnson,[1] Y. Kwon,[1] S. Roberts,[1] E. H. Thorndike,[1] C. P. Jessop,[2] K. Lingel,[2] H. Marsiske,[2] M. L. Perl,[2] S. F. Schaffner,[2] D. Ugolini,[2] R. Wang,[2] X. Zhou,[2] T. E. Coan,[3] V. Fadeyev,[3] I. Korolkov,[3] Y. Maravin,[3] I. Narsky,[3] V. Shelkov,[3] J. Staeck,[3] R. Stroynowski,[3] I. Volobouev,[3] J. Ye,[3] M. Artuso,[4] A. Efimov,[4] F. Frasconi,[4] M. Gao,[4] M. Goldberg,[4] D. He,[4] S. Kopp,[4] N. Horwitz,[4] G. C. Moneti,[4] R. Mountain,[4] Y. Mukhin,[4] S. Schuh,[4] T. Skwarnicki,[4] S. Stone,[4] M. Thulasidas,[4,a] G. Viehhauser,[4] X. Xing,[4] J. Bartelt,[5] S. E. Csorna,[5] V. Jain,[5] S. Marka,[5] A. Freyberger,[6] R. Godang,[6] K. Kinoshita,[6] I. C. Lai,[6] P. Pomianowski,[6] S. Schrenk,[6] G. Bonvicini,[7] D. Cinabro,[7] R. Greene,[7] L. P. Perera,[7] B. Barish,[8] M. Chadha,[8] S. Chan,[8] G. Eigen,[8] J. S. Miller,[8] C. O'Grady,[8] M. Schmidtler,[8] J. Urheim,[8] A. J. Weinstein,[8] F. Würthwein,[8] D. M. Asner,[9] D. W. Bliss,[9] W. S. Brower,[9] G. Masek,[9] H. P. Paar,[9] V. Sharma,[9] J. Gronberg,[10] R. Kutschke,[10] D. J. Lange,[10] S. Menary,[10] R. J. Morrison,[10] H. N. Nelson,[10] T. K. Nelson,[10] C. Qiao,[10] J. D. Richman,[10] D. Roberts,[10] A. Ryd,[10] M. S. Witherell,[10] R. Balest,[11] B. H. Behrens,[11] K. Cho,[11] W. T. Ford,[11] H. Park,[11] P. Rankin,[11] J. Roy,[11] J. G. Smith,[11] J. P. Alexander,[12] C. Bebek,[12] B. E. Berger,[12] K. Berkelman,[12] K. Bloom,[12] D. G. Cassel,[12] H. A. Cho,[12] D. M. Coffman,[12] D. S. Crowcroft,[12] M. Dickson,[12] P. S. Drell,[12] K. M. Ecklund,[12] R. Ehrlich,[12] R. Elia,[12] A. D. Foland,[12] P. Gaidarev,[12] B. Gittelman,[12] S. W. Gray,[12] D. L. Hartill,[12] B. K. Heltsley,[12] P. I. Hopman,[12] J. Kandaswamy,[12] N. Katayama,[12] P. C. Kim,[12] D. L. Kreinick,[12] T. Lee,[12] Y. Liu,[12] G. S. Ludwig,[12] J. Masui,[12] J. Mevissen,[12] N. B. Mistry,[12] C. R. Ng,[12] E. Nordberg,[12] M. Ogg,[12,b] J. R. Patterson,[12] D. Peterson,[12] D. Riley,[12] A. Soffer,[12] C. Ward,[12] M. Athanas,[13] P. Avery,[13] C. D. Jones,[13] M. Lohner,[13] C. Prescott,[13] S. Yang,[13] J. Yelton,[13] J. Zheng,[13] G. Brandenburg,[14] R. A. Briere,[14] Y.S. Gao,[14] D. Y.-J. Kim,[14] R. Wilson,[14] H. Yamamoto,[14] T. E. Browder,[15] F. Li,[15] Y. Li,[15] J. L. Rodriguez,[15] T. Bergfeld,[16] B. I. Eisenstein,[16] J. Ernst,[16] G. E. Gladding,[16] G. D. Gollin,[16] R. M. Hans,[16] E. Johnson,[16] I. Karliner,[16] M. A. Marsh,[16] M. Palmer,[16] M. Selen,[16] J. J. Thaler,[16] K. W. Edwards,[17] A. Bellerive,[18] R. Janicek,[18] D. B. MacFarlane,[18] K. W. McLean,[18] P. M. Patel,[18] A. J. Sadoff,[19] R. Ammar,[20] P. Baringer,[20] A. Bean,[20] D. Besson,[20] D. Coppage,[20] C. Darling,[20] R. Davis,[20] N. Hancock,[20] S. Kotov,[20] I. Kravchenko,[20] N. Kwak,[20] S. Anderson,[21] Y. Kubota,[21] M. Lattery,[21] J. J. O'Neill,[21] S. Patton,[21] R. Poling,[21] T. Riehle,[21] V. Savinov,[21] A. Smith,[21] M. S. Alam,[22] S. B. Athar,[22] Z. Ling,[22] A. H. Mahmood,[22] H. Severini,[22] S. Timm,[22] F. Wappler,[22] A. Anastassov,[23] S. Blinov,[23,c] J. E. Duboscq,[23] K. D. Fisher,[23] D. Fujino,[23,d] R. Fulton,[23] K. K. Gan,[23] T. Hart,[23] K. Honscheid,[23] H. Kagan,[23] R. Kass,[23] J. Lee,[23] M. B. Spencer,[23] M. Sung,[23] A. Undrus,[23,c] R. Wanke,[23] A. Wolf,[23] M. M. Zoeller,[23] B. Nemati,[24] S. J. Richichi,[24] W. R. Ross,[24] P. Skubic,[24] M. Wood,[24] M. Bishai,[25] J. Fast,[25] E. Gerndt,[25] J. W. Hinson,[25] N. Menon,[25] D. H. Miller,[25] E. I. Shibata,[25] I. P. J. Shipsey,[25] and M. Yurko[25]

[1]University of Rochester, Rochester, New York 14627
[2]Stanford Linear Accelerator Center, Stanford University, Stanford, California 94309
[3]Southern Methodist University, Dallas, Texas 75275
[4]Syracuse University, Syracuse, New York 13244
[5]Vanderbilt University, Nashville, Tennessee 37235
[6]Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
[7]Wayne State University, Detroit, Michigan 48202
[8]California Institute of Technology, Pasadena, California 91125
[9]University of California, San Diego, La Jolla, California 92093
[10]University of California, Santa Barbara, California 93106
[11]University of Colorado, Boulder, Colorado 80309-0390
[12]Cornell University, Ithaca, New York 14853
[13]University of Florida, Gainesville, Florida 32611
[14]Harvard University, Cambridge, Massachusetts 02138
[15]University of Hawaii at Manoa, Honolulu, Hawaii 96822
[16]University of Illinois, Champaign-Urbana, Illinois 61801
[17]Carleton University, Ottawa, Ontario, Canada K1S 5B6

[a]Permanent address: CPPM, Faculte' des Siences, Marseille, France .
[b]Permanent address: University of Texas, Austin TX 78712
[c]Permanent address: BINP, RU-630090 Novosibirsk, Russia.
[d]Permanent address: Lawrence Livermore National Laboratory, Livermore, CA 94551.

and the Institute of Particle Physics, Canada
[18]McGill University, Montréal, Québec, Canada H3A 2T8
and the Institute of Particle Physics, Canada
[19]Ithaca College, Ithaca, New York 14850
[20]University of Kansas, Lawrence, Kansas 66045
[21]University of Minnesota, Minneapolis, Minnesota 55455
[22]State University of New York at Albany, Albany, New York 12222
[23]Ohio State University, Columbus, Ohio 43210
[24]University of Oklahoma, Norman, Oklahoma 73019
[25]Purdue University, West Lafayette, Indiana 47907

I. INTRODUCTION

Measurements of weak decays of B mesons are essential to testing and understanding the standard model and determining the fundamental flavor-mixing parameters. These measurements also provide a unique opportunity to examine the short distance behaviour of weak decays [1]. Due to their large energy release, long distance corrections are expected to be less significant in hadronic B decay than in hadronic decays of charm and strange quarks. However, the formation of hadrons in the final state is still poorly understood and is an obstacle to predicting rates and spectra for hadronic decays [2,3].

For several years the small observed value of the B semileptonic decay branching fraction, $\mathcal{B}_{SL}(B)$, seemed to disagree with theoretical calculations [4,5,6,7,8,9]. Recently, higher order perturbative calculations, taking also into account the charm quark mass [10,11], come close to reconciling theory with experiment in this respect. However, the low value of $\mathcal{B}_{SL}(B)$ implies a larger than naively expected non-leptonic B decay rate [12]. Two mechanisms have been proposed and discussed in the literature [2,10,11,12,13,14]: an enhancement of $b \to c\bar{c}s$ or of $b \to c\bar{u}d$. The former would give a larger average number of c and $\bar{c}$ quarks, $\langle n_c \rangle$, per B decay. Since $B \to D^0$ and $B \to D^+$ transitions (where the D can also be the decay product of charm resonance) account for a large fraction of the charm quarks produced in B decay, it is important to measure accurately the branching fractions for these transitions. The shape of the momentum spectrum can also be compared to models of hadronic B decay and is sensitive to new production mechanisms [3,12]. In addition, the D and D^* inclusive spectra are of interest to future B experiments and high energy colliders as Monte Carlo simulations must be constrained to agree with the observed production at the $\Upsilon(4S)$.

In this paper, we describe high statistics measurements of the differential and total branching ratios for the inclusive B decay to D and D^* mesons, including the first measurement of $\mathcal{B}(B \to D^{*0}X)$. For previous measurements, see references [15,16].

We have also measured the D^{*+} and D^{*0} polarizations as a function of the D^* momentum. Since there is no complete reconstruction of the B final state, we do not distinguish between B^0 and B^+. We produce $B\bar{B}$ states from $\Upsilon(4S)$ decays, and our generic B is about an even [17] admixture of B_d and B_u mesons and their antiparticles . In the following, reference to the charge conjugate states is implicit unless explicitly stated.

It is possible for a B meson to decay to a final state containing two D mesons. By "inclusive B decay branching ratios" to a given D species measured in this analysis we mean the"average number of D and $\bar{D}$" per B decay.

After a brief description of the CLEO II apparatus and data sample used, we describe our analysis procedure in Sect.III. In Sect.IV-VI we give the B branching ratios and momentum spectra results and analysis details specific to the four individual channels. In Sect.VII, we show the results on D^* polarization and in Sect.VIII we discuss and summarize our results.

II. DETECTOR AND DATA SAMPLE

The data sample used in this inclusive B decay analysis was taken at CESR, the Cornell Electron Storage Ring, during 1990 – 1994. The sample consists of about 2,020 pb^{-1} of integrated luminosity of e^+e^- annihilation data taken near the peak of the $\Upsilon(4S)$ resonance and about 959 pb^{-1} just below the open bottom threshold (referred to in this paper as "continuum" data). The data used correspond to about 2.166 million $B\bar{B}$ events. We estimate the actual number of $B\bar{B}$ events by subtracting the number of events in the continuum data after scaling by the ratio of the luminosities and correcting for the center-of-mass energy dependence of the continuum annihilation cross-section. We also perform minor corrections due to minor differences in CESR running conditions at the two energies. This is equivalent to multiplying the integrated luminosity collected near the peak of the $\Upsilon(4S)$ resonance by the peak cross section $\sigma(e^+e^- \to \Upsilon(4S)) = (1.072 \pm 0.019)\ nb$ extracted from our own data. For the purpose of this analysis we assume that the $\Upsilon(4S)$ decays exclusively to $B\bar{B}$.

In the CLEO II detector [18] charged particles are tracked in a 1.5 Tesla magnetic field through three nested coaxial cylindrical drift chambers covering 94% of the solid angle. The innermost chamber is a six-layer straw-tube vertex detector of inner radius 4.5 cm with 50 μm position accuracy in the $r-\phi$ plane. It is followed by a ten-layer pressurized inner drift chamber with a position accuracy of 100 μm in $r-\phi$. The main cylindrical drift chamber [19] contains 51 anode layers, 11 of which are strung at angles to the z-axis progressing from about $\pm 4^\circ$ to about 7°. It has a position accuracy of 120 μm in $r-\phi$, it gives a transverse momentum resolution of $(\delta p_t/p_t)^2 = (0.0015 p_t)^2 + (0.005)^2$ (with p_t in GeV) and a dE/dx resolution (measured on Bhabha scattering events) of 6.5% for particle identification, giving good π/K separation up to 700 MeV/c. The outer radius of the main drift chamber is 1 m. Cathode layers are located at the inner and outer radii of the ten-layer inner drift chamber, and at the inner and outer radii of the main drift chamber, to improve information about the z coordinates (along the beam) of the tracks. Time of flight

counters with 154 *ps* resolution are located outside the drift chambers and provide additional information for particle identification (not used in this analysis).

Photon and π^0 detection as well as electron identification use the CsI electromagnetic shower detector [20]. It consists of 7800 CsI(Tl) crystals between the time of flight counters and the super-conducting magnet coil in the barrel region and, in endcaps, between the drift chamber plates and the magnet pole pieces, altogether covering 95% of the solid angle.

The material in the drift chamber endplates, electronics and cables degrades the performance of the calorimeter in the endcaps, especially at the two ends of the barrel region. Photon candidates are restricted to lie in the region of the calorimeter covering the angular region $|\cos\theta| < 0.707$. The energy calibration makes use of Bhabha scattering and radiative Bhabhas as well as $e^+e^- \to \gamma\gamma$ reactions and $\pi^0 \to \gamma\gamma$ decays [21]. For low multiplicity final states, the energy resolution in the barrel portion of the calorimeter is given by $\frac{\sigma_E}{E}(\%) = 0.35/E^{0.75} + 1.9 - 0.1E$, where E is the photon energy in GeV.

Muons can be identified by their penetration in the three 36 *cm* thick slabs of iron that surround the super-conducting coil in an octagonal geometry, and in the iron pole pieces of the magnet [22].

A. Monte Carlo simulation

To estimate detection efficiencies we generated Monte Carlo events using the Jetset 7.3 [23] package for continuum annihilation events and the CLEO model for $\Upsilon(4S) \to B\bar{B}$ decays [24]. Separate Monte Carlo datasets were generated for the analyses of $B \to D^0 X$ and of $B \to D^+ X$ decays. In the D^+ case, the D^+ is allowed to decay only into the final state we use, $K^-\pi^+\pi^+$, while the D^- is allowed to decay according to measured branching fractions of its decay modes. About 110,000 $B\bar{B}$ events and about 35,000 continuum events that contain a D^+ were generated. An analogous procedure is followed in generating $D^0 \to K^-\pi^+$, except that the D^0 is also allowed to decay into K^+K^- and $\pi^+\pi^-$ final states, according to their measured decay branching fractions relative to the $K^-\pi^+$ mode. About 170,000 $B\bar{B}$ events and about 150,000 continuum events that contain a D^0 were generated. The events are then processed through a GEANT-based [25] simulation of the CLEO II detector and reconstructed and analyzed as real data. We call these Monte Carlo datasets the "dedicated" Monte Carlo.

We also used a set of Monte Carlo events produced in a similiar manner, but with all D mesons decaying according to a model which incorporates the current knowledge of their decay modes. This statistically independent "generic" Monte Carlo (in which the "true" value of the quantities we aim to measure is known *a priori*) was used to check the analysis procedure. This type of consistency check, of course, does not exclude the possibility of a systematic flaw that is common to both the "dedicated" and "generic" Monte Carlo data samples.

The "generic" Monte Carlo was also used for a different purpose. The reconstructed tracks were associated with the simulated particles that generated them. We could thus generate high statistics, background free distributions of the signals and distributions of various backgrounds. We called these distributions Monte Carlo Tagged (MC-tag for short) distributions.

III. ANALYSIS PROCEDURE

We reconstruct the D and D^* mesons using the following exclusive decay modes that have the largest signal to background ratios and for which the Monte Carlo can reliably estimate the reconstruction efficiency:

$$D^0 \to K^-\pi^+ \tag{1}$$
$$D^+ \to K^-\pi^+\pi^+ \tag{2}$$
$$D^{*0} \to D^0\pi^0 \to (K^-\pi^+)\pi^0 \tag{3}$$
$$D^{*+} \to D^0\pi^+ \to (K^-\pi^+)\pi^+ \tag{4}$$
$$D^{*+} \to D^+\pi^0 \to (K^-\pi^+\pi^+)\pi^0 \tag{5}$$

In this section, we describe the selection criteria and general procedures used in the analysis. In the sections which follow the details of the procedures for decay modes of the D^0, D^+, D^{*0} and D^{*+} will be given.

A. Selection Criteria

The continuum background is suppressed by excluding candidates from jet-like events. This is accomplished by using the Fox-Wolfram parameters [26]. We require $R_2 \equiv H_2/H_0 < 0.5$. This cut has an efficiency of $99.38 \pm 0.02\%$

while rejecting $29.0 \pm 0.2\%$ of continuum events in the B decay kinematical region. The efficiency for this requirement is determined from our Monte Carlo simulation and agrees with the estimate derived from the data.

Each charged track used to reconstruct a D or D^* is required to be consistent with originating from the primary vertex. If the momentum of a track is greater than 0.250 GeV, we require that the z coordinate (along the beam) of the point of closest approach of the track to the beam-line be within 3 cm of the z coordinate of the event vertex and that the track's impact parameter with respect to the beam line be less than 5 mm. For tracks with momentum less than 0.250 GeV these requirements are loosened to 5 cm and 10 mm respectively.

Particle identification requirements were imposed based on the specific ionization (dE/dx) measurements for the track, provided that more than 10 hits were recorded in the main drift chamber. The observed dE/dx had to be within three standard deviations of that expected for the particle species considered.

Each photon candidate shower is required to lie within the good barrel region ($\cos\theta < 0.707$) of the crystal calorimeter and to have a minimum energy of 30 MeV. Photon candidates are also required to be well separated from the extrapolated position of all charged tracks, and the lateral shape of the shower should be consistent with that expected from an electromagnetic shower.

Candidate π^0 mesons are reconstructed from pairs of photon candidates. If the effective mass of two photons is less than 2.58σ away from the expected π^0 mass, the combination is accepted as a π^0 candidate and then is kinematically fitted to the nominal π^0 mass.

B. Common procedure

The inclusive B decay spectra and branching fractions are obtained by the subtraction of the candidate mass distribution below $B\bar{B}$ threshold (scaled by the ratio of luminosities and of the e^+e^- annihilation cross section) from the candidate mass distribution on the $\Upsilon(4S)$ resonance. To illustrate the effect of continuum subtraction we show in Fig. 1 the $(K^-\pi^+)$ effective mass distribution of D^0 candidates in the sample taken at the $\Upsilon(4S)$ resonance, in the whole momentum interval allowed in B decay, and the corresponding distribution from the sample taken below $B\bar{B}$ threshold, scaled by a factor of 2.08, the combined ratio of luminosities and of the e^+e^- (non $b\bar{b}$) annihilation cross sections at the two energies.

In order to perform such subtraction it is convenient to use a scaling variable. We use the scaling variable x, defined as $x = p/p_{max}$ where $p_{max} = 4.950\ GeV/c$, the momentum for a D^0 produced in the reaction $e^+e^- \to D^0\overline{D}^0$ at a center of mass energy of 10.58 GeV ($p_{max} = 4.920\ GeV/c$ for the continuum sample, at a center of mass energy of 10.520 GeV).

The momenta of the charmed mesons are measured in the $\Upsilon(4S)$ rest frame rather than in the B rest frame. Since the $\Upsilon(4S)$ mass (10.58 GeV) is slightly above the threshold for B meson pair production (10.56 GeV), the B mesons are not at rest. The B momentum ranges from about 265 MeV to 355 MeV ($\pm$ one standard deviation). This motion smears the value of x relative to what it would be if the B were at rest. A Monte Carlo study shows that the smearing in the variable x varies from ± 0.013 to a maximum of ± 0.020. Taking this effect into account, the maximum value of x for B decay to D is 0.506, and 0.496 for decay to D^*.

1. Spectra and the B branching fractions

In order to measure the spectra, we divided our sample of charmed particle candidates into 10 x bins (20 for the D^0 where the statistics are high). For each x bin we generated the effective mass distributions of the D candidates (already selected as candidate D^* decay products in the case of D^*, see below) from "on resonance" and "continuum" data and performed bin-by-bin the scaled continuum subtraction, obtaining the mass distributions of the candidates from B decay. We then fitted this mass distribution, in each x bin, to the sum of the D signal and the various backgrounds. The shape of the signal, its parametrization and the different backgrounds will be described later.

We performed an identical analysis (except for the non-existent continuum subtraction) on the Monte Carlo simulated events, to find the detection efficiencies as a function of x. The inclusive B decay branching ratios are calculated from the integral of the appropriate D (or D^*) efficiency-corrected spectrum.

2. D^*-tagging and "background-free" D samples

We have studied the shape of the D signal and its momentum dependence, in data and in Monte Carlo simulation, using D^0s from $D^{*+} \to D^0\pi^+$ decays and D^+s from $D^{*+} \to D^+\pi^0$ decays. The D^0 sample selected this way is of known flavor because of the charge of the D^{*+}. It was also used to obtain the shape and momentum dependence of

the "switched-mass" D^0 background discussed later in Sect. IV A. We obtain a very low background sample of D^0s from D^{*+} by selecting events in a $\pm 2\sigma$ region around the peak of the mass difference $\delta m \equiv m(K^-\pi_1^+\pi_2^+) - m(K^-\pi_1^+)$ distribution ($\sigma \sim 0.8$ MeV). We obtain a low background sample of D^+s using a similar selection on the $\delta m \equiv m(K^-\pi^+\pi^+\pi^0) - m(K^-\pi^+\pi^+)$ distribution ($\sigma \sim 1.0$ MeV). The δm distributions for D candidates in the D mass signal region in the three channels considered in this analysis are shown in Figs. 2,3,4(data). Most of the residual background in the D^*-tagged D mass distribution is eliminated by subtraction of the δm sidebands.

Due to the fact that the signal in δm is very close to threshold, it is difficult to choose sidebands that are wide enough to balance the number of background events under the δm peak and far enough from the peak. Our choices for signal and sideband regions are shown in Table I. The signal region has a half width, and the sidebands a total width, of approximately twice the Gaussian σ of the δm signal. The position of the signal peak is obtained from the data.

Before subtraction, the number of events in the δm sidebands must be scaled to the estimated number of background events in the peak region. Two methods were used to estimate the scaling factor. The first method is to fit the δm distribution with a smooth threshold function background and a double-Gaussian signal and calculate the scale factor by integration of the background function. The second method is to obtain the scale factor by fitting the D invariant mass distribution in the δm peak and side band regions with a two-Gaussian D signal function and a polynomial background. The scale factor is then given by the ratio of the background levels under the D peaks. Fig. 5 was obtained by the second method. The results of the two methods agree and the scale factors are close to 1.1. The same procedure was also used for the analysis of the Monte Carlo sample. The candidate D mass distributions obtained from data by this D^*-tagging procedure are shown in Figs. 5 and 6.

C. Estimate of the detection efficiency

The detection efficiencies for the charm mesons studied were obtained analyzing the dedicated Monte Carlo simulation samples described in Sect. II A with the same procedures and selections used to analyze data. However, measurement errors are slightly underestimated in the simulation. In practice the only relevant parameter is the width of the reconstructed mass. Consequently, we analyzed the data using the parameters extracted from the data themselves, and analyzed the Monte Carlo samples using the parameters extracted from the Monte Carlo samples.

D. Calculation of the B decay branching fractions

The differential B decay branching fractions are calculated bin-by-bin from the following equations:

$$d\mathcal{B}(B \to DX) \times \mathcal{B}(D) = \frac{dN_D}{2N_{B\bar{B}} \times \epsilon} \tag{6}$$

$$d\mathcal{B}(B \to D^*X) \times \mathcal{B}(D^*) \times \mathcal{B}(D) = \frac{dN_{D^*}}{2N_{B\bar{B}} \times \epsilon} \tag{7}$$

where dN_D (dN_{D^*}) is the yield of D (D^*) in that bin, $N_{B\bar{B}}$ is the number of of $e^+e^- \to B\bar{B}$ events produced and ϵ is the x-dependent detection efficiency. We use CLEO results [27,28,30] for the D^* and D absolute branching fractions

$$\mathcal{B}(D^{*+} \to D^0\pi^+) = (68.1 \pm 1.0 \pm 1.3)\% \tag{8}$$
$$\mathcal{B}(D^{*+} \to D^+\pi^0) = (30.8 \pm 0.4 \pm 0.8)\% \tag{9}$$
$$\mathcal{B}(D^{*0} \to D^0\pi^0) = (63.6 \pm 2.3 \pm 3.3)\% \tag{10}$$
$$\mathcal{B}(D^0 \to K^-\pi^+) = (3.91 \pm 0.08 \pm 0.17)\% \tag{11}$$
$$\mathcal{B}(D^+ \to K^-\pi^+\pi^+) = (9.19 \pm 0.6 \pm 0.8)\% \tag{12}$$

The errors shown in our results for the differential spectra are relative, *i.e.* bin-to-bin errors. They do not include the error from smoothing the fitting function parameters and the errors of the detection efficiency. The overall statistical error on the B decay branching fraction is, however, derived from the integral of the spectrum when no smoothing was performed.

E. Systematic Errors

In this section we discuss the sytematic uncertainties of the $B \to D\ X$ and $B \to D^*\ X$ analyses. Uncertainties related to specific decay channels are deferred to the relevant sections. We estimate a 1% uncertainty in charged track detection efficiency (2% for $p(\pi)$ below about 200 MeV/c) and 5% in π^0 detection [32].

These errors are coherent for each kind of track and hence should be multiplied by the respective number of tracks in the decay under study. However, we obtain the B decay branching fractions by dividing the product of branching fractions (Eqs. (6) and (7)) by the D (and D^*) decay branching fractions measured in our own experiment. Hence, to the extent that the kinematical configurations and the data set used overlap, most tracking errors are eliminated. We estimate a residual systematic error of 0.5% per charged track (1.0% if below 200 MeV/c) and 2% for π^0 whenever there is compensation.

We have studied the effect on the branching fraction of the track quality, geometry cuts and event shape cuts by succesively removing them and then measuring the effect on the B branching fractions. The errors specific to each final state will be quoted in the appropriate sections.

In measuring $\mathcal{B}(B \to D^* X)$ the subtraction of the background due to association of a true D and a random π is performed by subtracting the D candidate mass distribution of the δm sidebands from that in the δm peak region. The result is slightly dependent on the choice of the width of the δm intervals (Table I). We varied the widths of these intervals by ± 0.5 MeV in each case and derived a relative uncertainty in the B branching fraction of 0.9% for $B \to D^{*+}$ and of 1.2% for $B \to D^{*0}$.

In the case of $B \to D^*$ decays, the branching fraction is also sensitive to the value of the scale factor used in the δm sideband subtraction This effect contributes a systematic error which is estimated to be the change in the branching fraction corresponding to a 1σ change in the scale factor.

The value of $N_{B\bar{B}}$ in Eqs. (6) and (7) is affected by a relative error of 1.8% (Sec. II). Finally, we take into account the statistical error on the efficiency, calculated from the Monte Carlo simulation. All the above errors are combined in quadrature to give the relative systematic error on our measurements.

The third error in the B branching fractions arises from the propagation of errors in the D and D^* decay branching fractions, in so far as they are not coherent with the errors in this analysis.

IV. INCLUSIVE $B \to D^0$ DECAY

The continuum-subtracted $m(K^-\pi^+)$ distribution in any given D^0 momentum interval is the sum of the D^0 signal plus various backgrounds. This is illustrated in Figs. 7 and 8 which show the $m(K^-\pi^+)$ distributions for one of the lowest momentum intervals ($0.050 < x < 0.075$), and for a higher momentum interval ($0.325 < x < 0.350$), respectively. We see a prominent signal and three backgrounds: (i) a combinatorial background; (ii) the "switched mass" background that contributes to the signal region and (iii) two "satellite bumps" on either side of the signal. Backgrounds (ii) and (iii) complicate the fit to the signal region and its immediate vicinity. It is not possible to establish the shape of the combinatorial background without a reliable knowledge of the shapes and amounts of backgrounds (ii) and (iii). A detailed discussion of the signal shape and of these backgrounds is presented below.

A. Switched mass background

$\overline{D^0} \to \pi^- K^+$ events may be misidentified as $D^0 \to K^-\pi^+$ when there is insufficient K/π discrimination. This background complicates the extraction of the D^0 yield, particularly at low D^0 momenta ($p < 700\ MeV$) where this "switched mass" background distribution is similiar in shape to the signal and peaks under it (see Fig. 7 and 8). For a D^0 at rest and without particle identification the area of this background within 2σ of the signal peak would equal the signal area. However, this background, within such limits, quickly decreases to about 10% of the area of the signal for D^0 momentum above 1 GeV and decreases further at higher momenta. For $x_D > 0.50$, the invariant mass distribution of these doubly misidentified events becomes so broad that it can be absorbed in the polynomial parameterization of the combinatorial background. The amount of this type of background is strongly affected by the use of particle identification.

We studied the shape and relative size of this background and its dependence on D^0 momentum in the data and in the Monte Carlo simulation using D^{*+}-tagged D^0s (Sec. III B 2), whose flavor is tagged by the charge of the π^+. We produced "switched mass" distributions, by interchanging the mass assignments of the K and π tracks for each event, provided that the inverted assignements were consistent with our particle identification criteria. We then subtracted the "switched mass" distribution in the δm sidebands from that in the δm peak region. Figs. 9 and 10 show the resulting $m(\pi^- K^+)$ distributions for a low D^0 momentum interval, and for a higher D^0 momentum interval. We also

calculated, as a function of momentum, the area of this background relative to the area of the correctly reconstructed D^0 signal also obtained as described in Sec. III B 2.

With our statistics it is difficult to smoothly parameterize the shape of this background as a function of momentum. Figs. 9 and 10 show, however, that the switched mass distributions obtained with the MC-tag analysis of the generic Monte Carlo simulation provide an excellent, parameter-free description of the D^{*+}-tagged data distributions (Sec. III B 2).

The preceding analysis is performed both on the data and on the dedicated Monte Carlo samples. The momentum dependence of the ratio of the switched mass distribution area relative to the area of the signal is shown in Fig. 11 as obtained from the MC-tag sample. Similar dependences are obtained using the D^{*+}-tagged samples of data and dedicated Monte Carlo. We have checked that changing from one dependence to another has minimal impact on the final result.

B. Background from misidentified $D^0 \to K^-K^+$, $D^0 \to \pi^-\pi^+$ and $D^0 \to K^-\pi^+\pi^0$

The Cabibbo-suppressed D^0 decay modes, $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$, produce asymmetric enhancements on opposite sides of the signal region when one of the tracks is misidentified as a π or K, respectively (see Fig. 7 and 8). The $K^-\pi^+$ effective mass distribution from the decay $D^0 \to K^-\pi^+\pi^0$ introduces an additional small background at the lower edge of our mass spectrum. Switching the kaon and pion masses in these decays produces an additional background at masses well below the D^0 mass but with tail up to the D^0 mass. If these contributions are not included in the shape of the background, it is possible to overestimate the level of combinatorial background and thereby underestimate the signal. Not taking into account the presence of these enhancements both in fitting the data and in fitting the dedicated Monte Carlo for obtaining the efficiency, results in a change of the order of 1% in the branching fraction. Also in this case we have found that the momentum dependence of the shape and of the area ratio obtained by the MC-tag procedure from the generic Monte Carlo is quite good and superior to that obtained using complicated analytical parametrizations. The size of these backgrounds relative to the signal was taken from the Monte Carlo simulation, but we have checked that letting their normalization float in the fit did not appreciably alter the result.

C. Combinatorial background

The combinatorial background, which is nearly flat in the $m(K^-\pi^+)$ region of interest, is parameterized by a second order polynomial.

D. Raw D^0 spectrum and the shape of the signal

It is very important to obtain an accurate representation of the signal shape and specifically of its tails, because of the presence of the mass-switched background and its correlation with the signal. We used two diferent signal parameterizations.

1. Analytical parameterization

A Gaussian function does not give a sufficiently accurate parameterization of the D^0 signal. Track measuring errors are different in different events because of the geometrical orientation of the D^0 decay products in the detector and their overlap with other tracks. In order to take into account this variation of the errors, one possibility is to use a double Gaussian with the two Gaussians constrained to have the same mean. The parameters of the signal shape are the mean, the width of the narrower Gaussian, σ_1, the ratio of the widths of the wider to the narrower Gaussian, σ_2/σ_1, and the ratio of the area of the wider Gaussian to the total area, A_2/A_{tot}. There are strong correlations among these three shape parameters. It is then difficult to obtain a smooth D^0 momentum dependence for them if they are allowed to float when fitting the signal in each momentum bin.

We obtained satisfactory results for these signal shape parameters by (i) combining "on resonance" and "continuum" D^0 candidates together to obtain higher statistics, especially close to the kinematic limit, or (ii) by using the D^*-tagged, background-free D^0 signal (Sec. III B 2). We chose the set that gave the best fit to the mass distributions and smaller errors for the parameters. In order to minimize the bin-to-bin statistical fluctuations, we re-binned the

D^0 sample in only 9 momentum intervals ($0.025 < x < 0.475$) and studied the position and shape of the D^0 signal peak as a function of momentum. We smoothed the momentum dependence of the signal shape parameters so obtained with polynomial functions of the momentum. We then obtained the raw D^0 spectrum by fitting the $m(K^-\pi^+)$ distribution in 20 x bins between 0.0 and 0.5 keeping σ_2/σ_1 and A_2/A_{tot} fixed at the smoothed values. σ_1 is allowed to float and varies from about 6 MeV at low momentum to about 8 MeV at the maximum momentum; σ_2/σ_1 and A_2/A_{tot} are approximately constant, the first at about 2.7 and the second at about 0.15.

The same procedure was followed to analyze the events from the Monte Carlo. The comparison of the data and Monte Carlo shape parameters shows that the Monte Carlo simulation underestimates the track measuring errors (reflected essentially in σ_1) by about 10%.

2. Use of Monte Carlo simulation with track tagging

An alternative way of parameterizing the D^0 signal shape and a check of the double-Gaussian parameterization, is to use the histograms provided by our generic Monte Carlo simulation and tagging the reconstructed tracks (MC-tag, Sect. II A). In constructing these histograms we have used the same selection criteria used in the analysis of data and of dedicated Monte Carlo already described.

We have just seen that our Monte Carlo simulation underestimates by about 10% the overall width σ of the D^0 signal over the whole D^0 spectrum. It is however likely that the Monte Carlo simulation reproduces more accurately the dependence of σ on track position, orientation and overlap with other tracks, all factors that cause the non-Gaussian spread in measuring errors seen in the data. We then expect that the *shape* of the D^0 signal obtained by tagging the Monte Carlo tracks will be a good representation of the data if corrected for the overall width. We have expressed this correction by one parameter (then fitted on the data) that changes the width of the signal mass distributions without otherwise altering its shape.

We used the histograms obtained by this procedure as fitting functions of the "background-free" signal distributions obtained by $D^{\star+}$-tagging (Sect. III B 2). We obtained excellent fits, altough not superior to the double-Gaussian fits. We obtained D^0 spectra using these distributions as fitting functions. The D^0 yields differed from those obtained with the double-Gaussian by less than a tenth of the statistical error.

E. Results

Particle identification (Sect. III A) reduces the combinatorial background by about a factor of 2 and improves the statistical error while slightly increasing the systematic error due to the small uncertainty on the accuracy of the simulation of the ionization measurements. We performed our analysis both with and without particle identification requirements and obtained results that were in good agreement. The results presented here were obtained using particle identification and the double-Gaussian parameterization of the signal. In Figs. 12 and 13 we show the fitted mass distributions in the 20 x bins. Summing the spectrum over the 20 bins for $0.0 < x_D < 0.5$ gave a raw yield of $62,648 \pm 1,394$ D^0 from B decay.

The dedicated Monte Carlo simulated events were analyzed as real events to produce the histogram of detection efficiency versus momentum that was smoothed and used to correct the raw spectra (Fig. 14). The detection efficiency is nearly independent of momentum in the region of interest, except at the lowest momenta where there is an increase of angular acceptance due to the near alignment of the two D^0 decay products.

The continuum-subtracted, efficiency-corrected, inclusive D^0 momentum spectrum in $B \to D^0 X$ decay is shown in Fig. 15. Summing this spectrum over the interval $0 < x < 0.5$, we obtain the efficiency corrected yield of $108,507 \pm 2,407$ D^0s (resulting in a spectrum averaged detection efficiency = 0.578) out of 4.3314 million B decays. This gives the product of branching fractions,

$$\mathcal{B}_B \mathcal{B}_D = (2.51 \pm 0.06)\%, \tag{13}$$

where $\mathcal{B}_B \equiv \mathcal{B}(B \to D^0 X)$ and $\mathcal{B}_D \equiv \mathcal{B}(D^0 \to K^-\pi^+) + \mathcal{B}(D^0 \to K^+\pi^-)$ [1] and the error is only statistical.

Using the CLEO results $\mathcal{B}(D^0 \to K^-\pi^+) = (3.91 \pm 0.08 \pm 0.17)\%$ [28] and $\Gamma(D^0 \to K^+\pi^-)/\Gamma(D^0 \to K^-\pi^+) = (0.0077 \pm 0.0025 \pm 0.0025)$ [29], we obtain from Eq. (13) the branching fraction

[1] Our procedure finds both the Cabibbo-allowed, $D^0 \to K^-\pi^+$, and the doubly Cabibbo-suppressed, $D^0 \to K^+\pi^-$, decay modes, so we must divide by the sum of the two decay branching fractions.

$$\mathcal{B}(B \to D^0 X) = 0.636 \pm 0.014 \pm 0.019 \pm 0.018. \quad (14)$$

The first error is statistical, while the second is the systematic error. The contributions to the second (systematic) error, as percentages of the branching fraction are given in table II. The third error is due to the error on the absolute branching fraction $\mathcal{B}(D^0 \to K^-\pi^+)$.

As part of a different project [33], an analysis of the $B \to D^0 X$ decay was carried out over a somewhat increased data sample and with more stringent particle identification requirements, using also time of flight information. If both the $D^0 \to K^-\pi^+$ and $\pi^- K^+$ hypotheses were acceptable, but one gave a particle ID $\chi^2_{worse} > 4 + \chi^2_{better}$, the hypothesis with the worse χ^2 was rejected. This selection in fact reduced the switched mass background to zero. The result of this analysis for $\mathcal{B}(B \to D^0 X)$ is fully consistent with our result.

V. INCLUSIVE $B \to D^+$ DECAY

The complications from satellite peaks or switched mass backgrounds are not present in the measurement of the D^+ spectrum and other kinematic reflections are quite small (they will be discussed in the next subsection). It is then possible to estimate the shape of the combinatorial background by fitting the $K^-\pi^+\pi^+$ mass distribution excluding the signal region. However, it is difficult to determine the momentum dependence of the signal shape parameters given the large combinatorial background and low signal to background ratio. The $K^-\pi^+\pi^+$ mass distributions summed over all momenta of interest for the data at the $\Upsilon(4S)$ and those below $B\bar{B}$ threshold (scaled according to luminosity and square of center-of-mass energy ratios) are shown in Fig. 16. If we parameterize the signal as a double Gaussian, a modest change in the curvature of the combinatorial background in the $m(K^-\pi^+\pi^+)$ distribution strongly correlates with a wider and larger second component of the double Gaussian signal.

We analysed the $m(K^-\pi^+\pi^+)$ distributions using the double Gaussian parameterization of the signal. As in the analysis of the D^0 spectra, we used D^+ tagged by the $D^{*+} \to D^+\pi^0$ decay to determine the momentum dependence of the signal parameters. We also performed the analysis using the simple Gaussian signal parameterization. This is quite adequate to fit the data, given our statistical errors. We found that the results from the two different signal parameterizations are the same within a small fraction of the statistical error. Here we shall present the results obtained with the simple Gaussian parameterization.

A. Background from D_s^+ Decays

We have small backgrounds from D_s^+ decays where a K^+ is misidentified as a π^+. We took this into account using the following procedure. We generated $B\bar{B}$ Monte Carlo events that contain at least one D_s^+, which decays into $K^-\pi^+K^+$ directly or through the $\overline{K^*}^0 K^+$ or $\phi\pi^+$ resonance channel. We processed them through the simulation of our detector and track reconstruction. We selected events that passed the selection criteria for the decay $D^+ \to K^-\pi^+\pi^+$. We plotted these (fake) $m(K^-\pi^+\pi^+)$ distributions for each momentum bin and normalized them using our knowledge of the differential $\mathcal{B}(B \to D_s^+ X)$ [34] and of the D_s decay branching ratios. We added these histograms to the combinatorial background and signal functions when fitting the data. Taking into account this feed-down we changed our result for $\mathcal{B}(B \to D^+ X)$ by $\Delta\mathcal{B} = -0.0021 \pm 0.0002$. We have examined other possible feed-downs and verified that they do not affect our measurement.

B. D^+ spectrum

We fitted the continuum subtracted sample of $D^+ \to K^-\pi^+\pi^+$ candidates. We divided the sample in 20 momentum bins between $0.00 < x < 1.00$ and performed the bin-by-bin fit of the $m(K^-\pi^+\pi^+)$ distributions using the signal and background functions described above. In Fig. 17 we show the fitted mass distributions in the 10 x bins. The sum over the 10 bins for $0.0 < x_D < 0.5$ gave a raw yield of $35,804 \pm 1,297$ D^+ from B decay. The detection efficiency as a function of x is shown in Fig. 18. We have smoothed the x dependence in different ways. The resulting average efficiencies differ from one another by 0.18% at most. The error introduced by the choice of the smoothing function has been taken into account in the systematic error. The smoothing actually used is shown in Fig. 18.

C. Result

The efficiency corrected D^+ spectrum is shown in Fig. 19. Summing this spectrum over the interval $0.0 < x_D < 0.5$, we obtained an efficiency corrected yield of $93,493 \pm 3,386$ D^+ (resulting in a spectrum averaged detection efficiency $= 0.383$) that gives the product of branching fractions,

$$\mathcal{B}_B \mathcal{B}_D = (2.16 \pm 0.08)\%, \tag{15}$$

where $\mathcal{B}_B \equiv \mathcal{B}(B \to D^+ X)$, $\mathcal{B}_D \equiv \mathcal{B}(D^+ \to K^- \pi^+ \pi^+)$ and the error is statistical only.

Using CLEO's result for the absolute branching fraction $\mathcal{B}(D^+ \to K^- \pi^+ \pi^+) = (9.19 \pm 0.6 \pm 0.8)\%$ [30] gives:

$$\mathcal{B}(B \to D^+ X) = 0.235 \pm 0.009 \pm 0.009 \pm 0.024. \tag{16}$$

The first error is statistical, the second is systematic, and the third error is due to the statistical error on $\mathcal{B}(D^+ \to K^- \pi^+ \pi^+)$ and to that part of its systematic error that propagates to our measurement. Table III gives the components of the second error as percentages of the branching fraction.

Also this analysis, as in the $B \to D^0\ X$ case, has been carried out with different procedures and selection criteria that produced results within a fraction of the systematic error of the one reported here.

We will show later that the $B \to D^{*+} X$ branching fractions measured by reconstructing D^{*+} through the decay chains $D^{*+} \to D^0 \pi^+ \to (K^- \pi^+) \pi^+$ and $D^{*+} \to D^+ \pi^0 \to (K^- \pi^+ \pi^+) \pi^0$ are consistent with each other, their ratio being 1.01 ± 0.09. This is a consistency check of the results given above.

VI. INCLUSIVE $B \to D^*$ DECAY

We selected D^* candidates by combining a π of the appropriate charge with a D candidate and then imposing requirements on δm, the difference between the mass of the D^* candidate and that of the D candidate. However, this sample also contains background composed of pions randomly associated with a correctly reconstructed D meson. We eliminated this "fake" D^* background by subtracting properly scaled δm "sidebands" as described in Sec. III B 2. The mass distribution of the D candidates so obtained was fitted to find the D yield as in the previous sections.

A. Inclusive $B \to D^{*0}$ Decay

We selected D^0 candidates in the decay mode $D^0 \to K^- \pi^+$ (as described in Sect. IV) and combined them with a π^0 to form D^{*0} candidates. The D^{*0} yield in each momentum bin was found by fitting the $m(K^- \pi^+)$ distribution obtained with the δm selection and sideband subtraction procedure. This decay channel is affected by the same backgrounds as the D^0 channel. To fit the $m(K^- \pi^+)$ mass distributions, we used the parameterization and procedures from the analysis of $B \to D^0\ +\ X$ described in Sect. IV. However, the large number of free parameters involved with the multi-Gaussian parameterization was not suitable with the lower statistics available in this channel. We have thus used the histograms of the satellite peaks and of the switched $K - \pi$ distributions from Monte Carlo tagging while for the signal we used either a single Gaussian or, the histogram from Monte Carlo tagging. Both procedures gave $\mathcal{B}(B \to D^{*0}\ +\ X)$ values differing by a small fraction of the statistical error. We report here the results from the single Gaussian parameterization of the signal. In Fig. 22 we show the fitted mass distributions in the 10 x bins. The sum over the 10 bins gave a raw yield of $3,539 \pm 175$ D^{*0} from B decay. After correcting bin-by-bin for detection efficiency (Fig. 20), we obtained the inclusive D^{*0} spectrum shown in Fig. 23. Fig. 21 shows the continuum subtracted $K^- \pi^+$ effective mass distribution of D^{*0} candidates from the decay chain $B \to D^{*0} X \to (D^0 \pi^0) X$ in the momentum interval $(0.0 < x_{D^*} < 0.50)$.

As in the previous cases, by summing the spectrum in the interval $0.0 < x < 0.5$ we found a corrected yield of $26,840 \pm 1,331$ D^{*0} from B decay (resulting in a spectrum averaged detection efficiency $= 0.132$) and the product of branching fractions,

$$\mathcal{B}_B \mathcal{B}^*_D \mathcal{B}_D = (0.620 \pm 0.031)\% \tag{17}$$

where $\mathcal{B}_B \equiv \mathcal{B}(B \to D^{*0} X)$, $\mathcal{B}^*_D \equiv \mathcal{B}(D^{*0} \to D^0 \pi^0)$, $\mathcal{B}_D \equiv \mathcal{B}(D^0 \to K^- \pi^+)$ and only the statistical error is shown.

Using the CLEO results $\mathcal{B}(D^{*0} \to D^0 \pi^0) = (63.6 \pm 2.3 \pm 3.3)\%$ [27] and $\mathcal{B}(D^0 \to K^- \pi^+)\ = (3.91 \pm 0.08 \pm 0.17)\%$ [28], and $\Gamma(D^0 \to K^+ \pi^-)/\Gamma(D^0 \to K^- \pi^+) = (0.0077 \pm 0.0025 \pm 0.0025)$ [29], from Eq. (17) we obtain the branching fraction:

$$\mathcal{B}(B \to D^{*0} X) = 0.247 \pm 0.012 \pm 0.018 \pm 0.018. \tag{18}$$

The first error is statistical. The second error is systematic and includes the components listed in Table IV. These are given as percentages of the branching fraction itself. The third error is determined from the statistical error on $\mathcal{B}(D^0 \to K^-\pi^+)$.

B. Inclusive $B \to D^{*+}$ Decay

We have analysed the inclusive decay $B \to D^{*+}$ in two D^{*+} decay modes. The $D^{*+} \to D^0\pi^+$ channel has significantly more events because of the larger D^* branching fraction and higher detection efficiency. However, the detection efficiency is a steep function of the D^{*+} momentum (Fig. 24) because of the short range of the low momentum π^+ and absorption in the beampipe. The detection efficiency for the charged pion is nearly zero for $x_{D^*} < 0.15$. Summing over the spectrum gives $\mathcal{B}(B \to D^{*+}X)$ in the charged pion mode for $x_{D^*} > 0.15$. The $D^{*+} \to D^+\pi^0$ channel, which has fewer events and much larger backgrounds (Fig. 3) has an efficiency which is nearly constant with momentum (Fig. 28) and is the only source of information for the low momentum region. We will separately describe the analyses using the two D^{*+} decay modes and then discuss how to combine them to obtain $\mathcal{B}(B \to D^{*+}X)$.

1. Using the $D^{*+} \to D^0\pi^+$ decay channel

Candidate D^0 mesons were reconstructed in the $D^0 \to K^-\pi^+$ decay mode (as described in Sec. IV) and combined with a π^+ (referred to as the "slow π") to form D^{*+} candidates. In order to maximize the detection efficiency of the slow π^+ we did not require dE/dx information to be available for this track. The D^{*+} yield in each of the 7 momentum bins ($0.15 < x_{D^*} < 0.50$) was found by fitting the $m(K^-\pi^+)$ distribution obtained through the δm selection and sideband subtraction procedure described in Sec. III B 2. In fitting the D^0 peak, we used the same parameterization that was used in the analysis of $B \to D^0 + X$, but no background due to double misidentification of $D^0 \to K^-\pi^+$ is present in this mode. In Fig. 25 we show the fitted mass distributions in the 7 x bins. Summing the spectrum in the interval $0.15 < x_{D^*} < 0.5$, gave a raw yield of $4,214 \pm 120$ D^{*+}.

Correcting bin-by-bin for detection efficiency (Fig. 24), we obtained the inclusive D^{*+} spectrum from B decay shown in Fig. 26. The efficiency corrected yield for $x_{D^*} > 0.15$ is $24,691 \pm 902$ D^{*+} from B decay (resulting in a spectrum averaged detection efficiency = 0.171). This gives the product of branching fractions,

$$\mathcal{B}_B\mathcal{B}^*_D\mathcal{B}_D(x_{D^*} > 0.15) = (0.570 \pm 0.021)\% \tag{19}$$

where $\mathcal{B}_B \equiv \mathcal{B}(B \to D^{*+}X)$, $\mathcal{B}^*_D \equiv \mathcal{B}(D^{*+} \to D^0\pi^+)$, $\mathcal{B}_D \equiv (\mathcal{B}(D^0 \to K^-\pi^+) + \mathcal{B}(D^0 \to K^+\pi^-))$ and the error quoted is only statistical.

Fig. 27 shows the continuum subtracted $K^-\pi^+$ effective mass distribution of D^{*+} candidates from the decay chain $B \to D^{*+}\ X \to (D^0\pi^+)X$ in the momentum interval interval ($0.15 < x_{D^*} < 0.50$), after δm sideband subtraction.

2. Using the $D^{*+} \to D^+\pi^0$ decay channel

Candidate D^+s are reconstructed in the decay mode $D^+ \to K^-\pi^+\pi^+$ (as described in Sec. V) and combined with a π^0 to form D^{*+} candidates. The D^{*+} yield in each of the 20 momentum bins ($0.0 < x_{D^*} < 1.0$) was found by fitting the $m(K^-\pi^+\pi^+)$ distribution obtained from the δm selection and sideband subtraction procedure described in Sec. III B 2. The D^+ peak was fitted using the same parameterization and procedure used in the analysis of $B \to D^+X$. Summing the spectrum over the interval $0.0 < x_{D^*} < 0.5$ gave a raw yeld of $2,925 \pm 250$ D^{*+}. After correcting for detection efficiency (Fig. 28) and summing over x bins, we obtained the inclusive D^{*+} spectrum from B decay shown in Fig. 29. The corrected yield is $27,683 \pm 2,339$ D^{*+} from B decay (resulting in a spectrum averaged detection efficiency = 0.106) and the product of branching fractions is,

$$\mathcal{B}_B\mathcal{B}^*_D\mathcal{B}_D = (0.639 \pm 0.054)\% \tag{20}$$

where $\mathcal{B}_B \equiv \mathcal{B}(B \to D^{*+}X)$, $\mathcal{B}^*_D \equiv \mathcal{B}(D^{*+} \to D^+\pi^0)$, $\mathcal{B}_D \equiv \mathcal{B}(D^+ \to K^-\pi^+\pi^+)$ and the error quoted is statistical only.

Fig. 30 shows the continuum subtracted $K^-\pi^+\pi^+$ invariant mass distribution of D^{*+} candidates from the decay chain $B \to D^{*+} + X \to (D^+\pi^0)X$ in the whole momentum interval interval ($0.0 < x_{D^*} < 0.50$).

C. Combined results for the inclusive $B \to D^{*+}$ Decay

Using the corrected differential branching fractions $d\mathcal{B}(B \to D^{*+}X)/dx$ obtained in two independent D^* decay modes, we combined them as follows. In the momentum region $0.0 < x_{D^*} < 0.15$ we use the only measurement available, that from the $D^{*+} \to D^+\pi^0$ decay mode, $\mathcal{B}(B \to D^{*+}X)(0.0 < x_{D^*} < 0.15) = 0.031 \pm 0.009 \pm 0.0025 \pm 0.0027$. In the momentum region $0.15 < x_{D^*} < 0.50$ we calculated the weighted average of the two measurements. The resulting spectrum is shown in Fig. 31.

The sum over all x bins gives the branching fraction,

$$\mathcal{B}(B \to D^{*+}X) = 0.239 \pm 0.011 \pm 0.014 \pm 0.009. \tag{21}$$

The first error is statistical and is dominated by the error on the D^* branching fraction for $0.0 < x_{D^*} < 0.15$. The second error is systematic and includes the components listed in Table V. These are quoted as percentages of the branching fraction itself. The third error is propagated from the statistical error on the D^0, D^+ and D^{*+} decay branching fractions.

VII. POLARIZATION

The polarization of D^* mesons has been predicted for semileptonic B decays and for two body hadronic decays [3,36]. A model dependent estimate of momentum dependence of the polarization for directly produced D^* mesons is available for inclusive decays [3]. CLEO has previously measured the D^{*+} polarization [15]. Here we present a new measurement of the D^{*+} polarization with higher statistics. However, the present detector is operated with a higher magnetic field that makes impossible to extend the measurement to low D^{*+} momenta for the $D^{*+} \to D^0\pi^+$ decay mode. We also present the first measurement of D^{*0} polarization in inclusive B decays. In this case, the measurement can be extended to the lowest momenta without difficulty. These polarization measurements served also as a check of the accuracy of our Monte Carlo simulation of B decay.

The polarization as a function of x is determined from the distribution of the D^* decay angle, θ. This is the angle between the direction of flight of the D^* in the laboratory frame and the direction of the daughter D^0 in the D^* rest frame. The distribution of this decay angle can be parameterized as

$$\frac{dW}{d\cos\theta} = \frac{3}{6+2\alpha}[1 + \alpha\cos^2\theta] \tag{22}$$

This is equivalent to the expression in terms of the spin-density matrix element ρ_{00}:

$$W(\cos\theta) = \frac{3}{4}[(1-\rho_{00}) + (3\rho_{00}-1)\cos^2\theta] \tag{23}$$

Here, the spin-density matrix is determined in a coordinate system with the quantization axis along the direction of motion of D^* in the laboratory frame. The element ρ_{00} is the probability for the D^* to be in the $J_z = 0$ state.

The polarization parameter α is related to the longitudinal and transverse decay rates of D^* as

$$\alpha = \frac{\Gamma_L}{\Gamma_T} - 1$$

When α is close to its lower bound of -1, the distribution is $\sin^2\theta$ implying transverse polarization while large values of α imply longitudinal polarization.

Only the $D^* \to D^0\pi, D^0 \to K^-\pi^+$ decay mode is used. We applied the same selection criteria that were used in the branching fraction measurements to obtain the D^* samples. The sample was divided into 5 intervals in $0.0 < x < 0.5$ and in 5 intervals in $\cos\theta$. Each of the $K\pi$ distributions was fitted to a Gaussian shape plus polynomial background. We repeat the same analysis procedure on the Monte Carlo simulated data to find the dependence of efficiencies on x and $\cos\theta$.

Figures 32 and 33 show the efficiency corrected and background subtracted $\cos\theta$ distributions for D^{*+} and D^{*0} respectively. We have compared these distributions with those predicted by our Monte Carlo simulation. The simulation appears to correctly model the data. This checks our results on spectra and B decay branching ratios because an incorrect simulation of the polarization may result in an incorrect determination of the detection efficiency.

Figure 34 shows α as a function of the scaled momentum variable x for D^{*+} and D^{*0} .

VIII. DISCUSSION AND SUMMARY

In Table VI we compare our results with those from previous measurements. In order to make the comparison independent of D^* and D decay branching fractions used in the different experiments, we give the product of the branching fractions.

Only the statistical error is reported here for our current results (third column) because the systematic error on these products of branching ratios cannot take advantage of some cancellations when we divide them by D and D^* branching ratios measured in our own experimant. Our results are consistent with previous measurements, except for the $B \to D^0 X$ case where our branching fraction is significantly higher. The branching fraction $\mathcal{B}(B \to D^{*0} X)$ is measured here for the first time. The mode $\mathcal{B}(B \to D^{*+} X)$ is reconstructed in the $D^{*+} \to D^+ \pi^0$ decay mode for the first time. The result for $\mathcal{B}(B \to D^{*+} X)$ in the $D^{*+} \to D^0 \pi^+$ decay mode refers only to $x_D > 0.15$ and the previous one from CLEO only to $x_D > 0.10$. Table VII summarizes the results on the inclusive B decay branching fractions.

From the measurement of $\mathcal{B}(B \to D^{*0} X)$ we can determine the ratio

$$\frac{\mathcal{B}(B \to D^{*0} X)}{\mathcal{B}(B \to D^{*+} X)} = 1.03 \pm 0.07 \pm 0.09 \pm 0.08$$

which is consistent with the naive expectation of 1.00.

Using these new measurements, and previous ones shown in Table VIII , we can now calculate the average number of charm quarks produced in B decay,

$$\langle n_c \rangle = 1.10 \pm 0.05,$$

This value is consistent with the naive expectation of 1.15, but it is considerably lower than the value (~ 1.30) required to account for the low value of the B semileptonic branching ratio in models where the channel $b \to c\bar{c}s$ is enhanced [2,10,11,12,13,14].

ACKNOWLEDGEMENTS

We gratefully acknowledge the effort of the CESR staff in providing us with excellent luminosity and running conditions. J.P.A., J.R.P., and I.P.J.S. thank the NYI program of the NSF, M.S. thanks the PFF program of the NSF, G.E. thanks the Heisenberg Foundation, K.K.G., M.S., H.N.N., T.S., and H.Y. thank the OJI program of DOE, J.R.P., K.H., M.S. and V.S. thank the A.P. Sloan Foundation, and A.W. and R.W. thank the Alexander von Humboldt Stiftung for support. M.S. is supported as a Scholar of the Cottrell Research Corporation. This work was supported by the National Science Foundation, the U.S. Department of Energy, and the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

[1] For a review of the experimental and theoretical status of B meson decays, see"B Decays", S. Stone ed., 2nd Ed., World Scientific, Singapore, 1994; "B Mesons", T.E. Browder and K. Honscheid, to appear in *Progress in Nuclear and Particle Physics*, Vol. 35, 81-220 (1995).

[2] A.F. Falk, M.B. Wise, I.Dunietz, *Phys. Rev.* **D 51**, 1183 (1985).

[3] M. Wirbel and Y.-L. Wu *Phys. Lett.* **B 228**, 430 (1989).

[4] I.I. Bigi, M. Shifman, N.G. Uraltsev and A.I. Vainshtein, *Phys. Rev. Lett.* **71**, 496 (1993); B. Blok, L. Koyrakh, M. Shifman and A.I. Vainshtein, *Phys. Rev.* **D49**, 3356 (1994).

[5] A. Manohar and M.B. Wise, *Phys. Rev.* **D49**, 1310 (1994).

[6] T. Mannel, *Nucl. Phys.* **B413**, 396 (1994).

[7] A.F. Falk, Z. Ligeti, M. Neubert and Y. Nir, *Phys. Lett.* **B326**, 145 (1994); L. Koyrakh, *Phys. Rev.* **D49**, 3379 (1994); S. Balk, J.G. Körner, D. Pirjol and K. Schilcher, *Z.Phys.* **C64**, 37 (1994).

[8] I.I. Bigi, N.G. Uraltsev and A.I. Vainshtein, *Phys. Lett.* **B293**, 430 (1992); Erratum, *Phys. Lett.* **B297**, 477 (1993); I.I. Bigi, B. Blok, M. Shifman, N.G. Uraltsev and A.I. Vainshtein, 7th Meeting of the Division of Particles Fields of the APS (DPF 92), Batavia, IL, 10-14 Nov 1992, published in DPF Conf.1992:610-613.

[9] B. Blok and M. Shifman, *Nucl. Phys.* **B399**, 441 (1993); B. Blok and M. Shifman, *Nucl. Phys.* **B399**, 459 (1993); N. Bilic, B. Guberina and J. Trampetic, *Nucl. Phys.* **B248**, 261 (1984); M. Voloshin and M. Shifman, *Sov. J. Nucl. Phys.* **41**, 120 (1985); M. Voloshin and M. Shifman, *Sov. Phys.–JETP* **64**, 698 (1986); V.A. Khoze, M. Shifman, N.G. Uraltsev and M. Voloshin, *Sov. J. Nucl. Phys.* **46**, 112 (1987).

[10] E. Bagan *et al.*, *Nucl. Phys.* **B 432**, 3 (1995); *Phys. Lett.* **B 342**, 362 (1995); *Phys. Lett.* **B 351**, 546 (1995), [E: *Phys. Lett.* **B 374**, 363 (1996)].
[11] M. Neubert and C.T. Sachrajda, "Spectator Effects in Inclusive Decays of Beauty Hadrons", preprint CERN-TH/96-19 (hep-ph/9603202), submitted to *Nucl. Phys.* **B**.
[12] G. Buchalla, I. Dunietz and H. Yamamoto, *Phys. Lett.* **B364**, 188 (1995).
[13] B. Blok and T. Mannel, *Mod. Phys. Lett.* **A11**, 1263, (1996).
[14] I.L. Grach, I.M. Narodetskii, S. Simula and K.A. Ter-Martirosyan, "Exclusive and Inclusive Weak Decays of the B-Meson", Istituto Superiore di Sanita', preprint no. INFN-ISS 95/17 (hep-ph/9603239) submitted to *Nucl. Phys.* **B**.
[15] D. Bortoletto *et al.* (CLEO), *Phys. Rev.* **D 45**, 21 (1992).
[16] H. Albrecht *et al.* (ARGUS), *Z. Phys.* **C 52**, 353 (1991).
[17] B. Barish *et al.* (CLEO) *Phys. Rev.* **D 51**, 1014 (1995).
[18] Y. Kubota *et al.* (CLEO), *Nucl. Inst. and Meth.* **A 320**, 66 (1992).
[19] D.G. Cassel *et al.*, *Nucl. Inst. and Meth.* **A 252**, 325 (1986).
[20] E. Blucher *et al.*, *Nucl. Inst. and Meth.* **A 249**, 201 (1989).
[21] R. Morrison *et al.* (CLEO), *Phys. Rev. Lett.* **67**, 1696 (1991).
[22] D. Bortoletto *et al.*, *Nucl. Inst. and Meth.* **A 320**, 114 (1992).
[23] T. Sjöstrand, Comp. Phys. Comm. **39**, 347 (1986), T. Sjöstrand, M. Bengston, Comp. Phys. Comm. **43**, 367 (1987).
[24] Our Monte Carlo simulation does not include the possibility of $B \to D\bar{D}KX$ decay. Since our detection efficiency has been obtained bin-by-bin in the D momentum, this omission does not affect our result.
[25] R. Brun *et al.*, GEANT 3.14, CERN DD/EE/84-1.
[26] G.C. Fox and S. Wolfram, *Phys. Rev. Lett.* **41**, 1581 (1978).
[27] F. Butler *et al.* (CLEO), *Phys. Rev. Lett.* **69**, 2041 (1992).
[28] D. Akerib *et al.* (CLEO), *Phys. Rev. Lett.* **71**, 3070 (1993). We use the value 3.91% before radiative corrections because our D^0 yield is not corrected for radiative corrections.
[29] D. Cinabro *et al.* (CLEO), *Phys. Rev. Lett.* **72**, 1406 (1994)
[30] R. Balest *et al.* (CLEO), *Phys. Rev. Lett.* **72**, 2328 (1994).
[31] Review of Particle Properties, *Phys. Rev.* **D 54**, 1 (1996).
[32] For a discussion of charged track and π^0 detection efficiency and their errors see ref. [17]
[33] CLEO work in preparation.
[34] D. Gibaut *et al.* (CLEO), *Phys. Rev.* **D 53**, 4734 (1996).
[35] CLEO internal note CBX 96-43
[36] J.L. Rosner *Phys. Rev.* **D 42**, 3732 (1990).
[37] M. Zoeller, Ph.D. dissertation, S.U.N.Y. at Albany, (1994)
[38] D. Cinabro *et al.* (CLEO), paper contributed to the ICHEP conference, Glasgow, 1994, preprint CLEO CONF 94-08, and J.P. Alexander *et al.* (CLEO) *Phys. Rev. Lett.* **74**, 3113 (1995), E: *Phys. Rev. Lett.* **75**, 4155 (1995).
[39] M.B. Voloshin, *Phys.Lett.* **B385**, 369 (1996).
[40] R. Balest *et al.* (CLEO) *Phys. Rev.* **D52**, 2661 (1995).

TABLES

TABLE I. δm cut and width of the sidebands (in MeV).

channel	width	signal region	lower-band	upper-band
$D^{*+} \to D^0\pi^+$	$\pm$1.5	144.1-147.1	141.1-142.6	148.6-150.1
$D^{*+} \to D^+\pi^0$	$\pm$1.5	139.1-142.1	136.1-137.6	143.6-145.1
$D^{*0} \to D^0\pi^0$	$\pm$2.0	140.6-144.6	137.6-139.6	145.6-147.6

TABLE II. Relative Systematic Errors on $\mathcal{B}(B \to D^0 X)$

residual track finding efficiency uncertainty	1.0%
variation of track quality and geometrical cuts	0.5%
choice of signal shape parameterization	2.0%
error in particle identification efficiency	0.8%
error in number of B and $\bar{B}$	1.8%
statistical error on efficiency	0.4%
Total	3.0%

TABLE III. Relative Systematic Errors on $\mathcal{B}(B \to D^+ X)$

residual track finding efficiency uncertainty	1.5%
variation of track quality and geometrical cuts	1.2%
choice of signal shape parameterization	2.0%
error in number of B and $\bar{B}$	1.8%
statistical error on efficiency	0.7%
uncertainty in background shape	1.7%
smoothing of the efficiency vs x	0.14%
estimate of kinematical reflections	0.10%
Total	3.8%

TABLE IV. Relative Systematic Errors on $\mathcal{B}(B \to D^{*0} X)$

residual particle finding efficiency uncertainty	5.1%
variation of track quality and geometrical cuts	0.5%
choice of signal shape parameterization	2.0%
error in number of B and $\bar{B}$	1.8%
statistical error on efficiency	2.1%
choice of the δm region widths	1.2%
scale factor in the δm sideband subtraction	3.3%
Total	7.1%

TABLE V. Relative Systematic Errors on $\mathcal{B}(B \to D^{*+} X)$

residual track finding efficiency uncertainty	5.2%
variation of track quality and geometrical cuts	0.5%
choice of signal shape parameterization	2.0%
error in number of B and $\bar{B}$	1.8%
choice of the δm region widths	1.2%
statistical error on efficiency (combined)	1.0%
Total	6.1%

TABLE VI. Comparison of our results on "product branching ratios" (in %) with those of previous measurements. The error in this work's result is only the statistical one (see text).

Channel	Measured Br. Ratio	This work	CLEO [15]	ARGUS [16]
$B \to D^0 X$	$\mathcal{B}_B \mathcal{B}_D$	2.51 ± 0.06	$2.33 \pm 0.12 \pm 0.14$	$1.94 \pm 0.15 \pm 0.25$
$B \to D^+ X$	$\mathcal{B}_B \mathcal{B}_D$	2.16 ± 0.08	$2.26 \pm 0.30 \pm 0.18$	$2.09 \pm 0.27 \pm 0.40$
$B \to D^{*0} X$	$\mathcal{B}_B \mathcal{B}_{D^*} \mathcal{B}_D$	0.620 ± 0.031	–	–
$B \to D^{*+} X$	$\mathcal{B}_B \mathcal{B}_{D^*} \mathcal{B}_{D^0}$	0.570 ± 0.021	$0.56 \pm 0.03 \pm 0.05$	$0.71 \pm 0.06 \pm 0.12$
$B \to D^{*+} X$	$\mathcal{B}_B \mathcal{B}_{D^*} \mathcal{B}_{D^+}$	0.639 ± 0.054	—	—

TABLE VII. Inclusive B decay branching fractions to D and D^*.

$\mathcal{B}(B \to D^0 X)$	$0.636 \pm 0.014 \pm 0.019 \pm 0.018$
$\mathcal{B}(B \to D^+ X)$	$0.235 \pm 0.009 \pm 0.009 \pm 0.024$
$\mathcal{B}(B \to D^{*0} X)$	$0.247 \pm 0.012 \pm 0.018 \pm 0.018$
$\mathcal{B}(B \to D^{*+} X)$	$0.239 \pm 0.015 \pm 0.014 \pm 0.009$

TABLE VIII. Charm counting

$\mathcal{B}(\bar{B} \to D^0 X) + \mathcal{B}(\bar{B} \to D^+ X)$	0.871 ± 0.035
$\mathcal{B}(\bar{B} \to D_s X)$ [2]	0.118 ± 0.031
$\mathcal{B}(\bar{B} \to \Lambda_c X)$ [3]	0.039 ± 0.020
$\mathcal{B}(\bar{B} \to \Xi_c^+ X)$ [4]	0.008 ± 0.005
$\mathcal{B}(\bar{B} \to \Xi_c^0 X)$	0.012 ± 0.009
$2 \times \mathcal{B}(\bar{B} \to$ Charmonia $X)$ [5]	0.054 ± 0.007

[2] We took the product of branching ratios measured by CLEO [34]: $\mathcal{B}(\bar{B} \to D_s X) \cdot \mathcal{B}(D_s^+ \to \phi\pi^+) = (0.424 \pm 0.014 \pm 0.031)\%$ and divided it by the world average $\mathcal{B}(D_s^+ \to \phi\pi^+) = (0.36 \pm 0.09)\%$ [31].

[3] We have used the product branching ratio $\mathcal{B}(B \to \Lambda_c\ X)\mathcal{B}(\Lambda_c \to pK^-\pi^+) = (0.181 \pm 0.022 \pm 0.024)\%$ from [37] and divided it by the world average $\mathcal{B}(\Lambda_c \to pK^-\pi^+) = (4.4 \pm 0.6)\%$ [31].

[4] This branching fraction and the following one are derived from the product branching ratios reported in [38]. The $\mathcal{B}(B \to \Xi_c\ X)$ branching fractions in these papers were calculated under the assumptions that $\Gamma(\Xi_c \to X\ell\bar{\nu}) = \Gamma(D \to X\ell\bar{\nu})$. M.B. Voloshin [39] estimated that, because of Pauli interference of the strange quark, $\Gamma(\Xi_c \to X\ell\bar{\nu})$ is in fact larger. His estimate (M.B. Voloshin, private communication) is $\Gamma(\Xi_c \to X\ell\bar{\nu}) = (2.5 \pm 1.0)\Gamma(D \to X\ell\bar{\nu})$. Furthermore, here we assume $\Gamma(\Xi_c \to \Xi\ell\bar{\nu})/\Gamma(\Xi_c \to X\ell\bar{\nu}) = 0.8 \pm 0.2$ rather that 1.0 as in ref. [38]. Altogether, the values of $\mathcal{B}(\Xi_c^+ \to \Xi^-\pi^+\pi^+)$ and $\mathcal{B}(\Xi_c^0 \to \Xi^-\pi^+)$ are decreased by a factor of 2.0 ± 0.9 with respect to those in ref. [38].

[5] We took the inclusive, direct B decay branching fractions to ψ, $\psi\prime$, χ_{c1} and χ_{c2} from the CLEO result [40] and added 0.009 ± 0.003 (the upper limit on $\mathcal{B}(B \to \eta_c\ X)$) to take into account all other charmonium production from B decay.

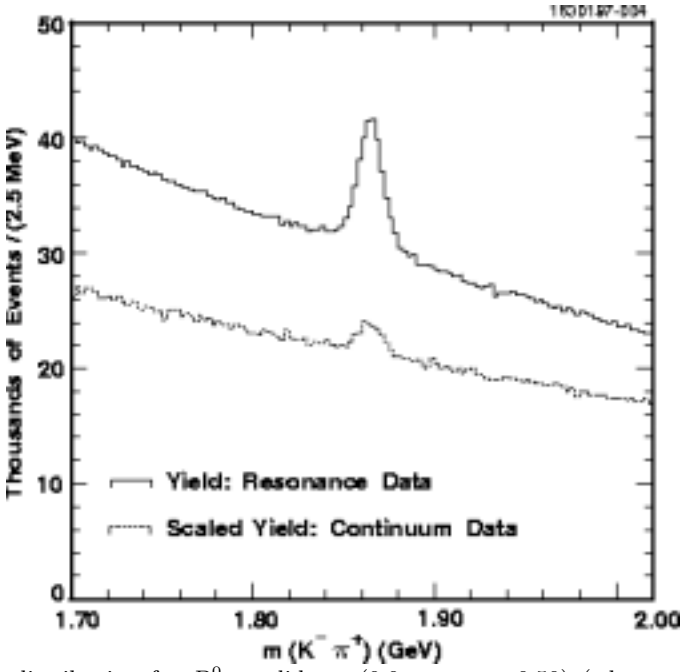

FIG. 1. The $m(K^-\pi^+)$ distribution for D^0 candidates ($0.0 < x_D < 0.50$) (where $x_D = p_D/p_{max}$) from "on resonance" data ($p_{max} = 4.950\ GeV/c$) and from "continuum" data ($p_{max} = 4.920\ GeV/c$). The "continuum" distribution is scaled by the luminosity and cross section factor 2.080.

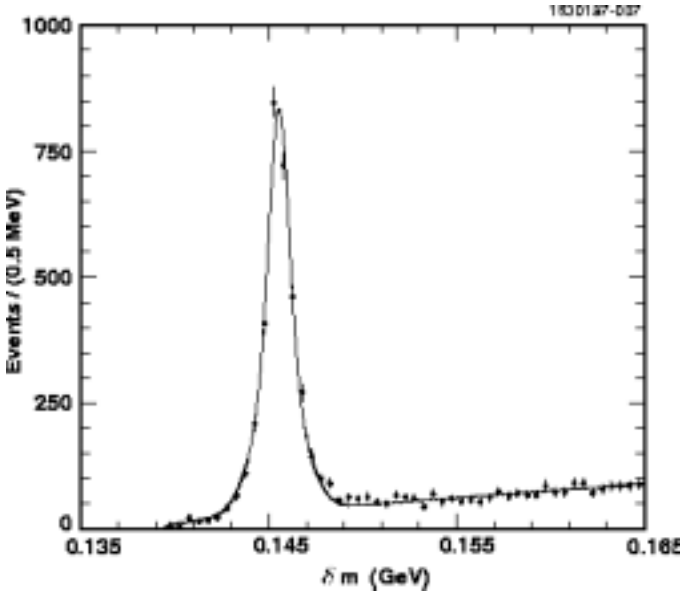

FIG. 2. The δm distribution for $D^{*+} \to D^0\pi^+$ candidates (data).

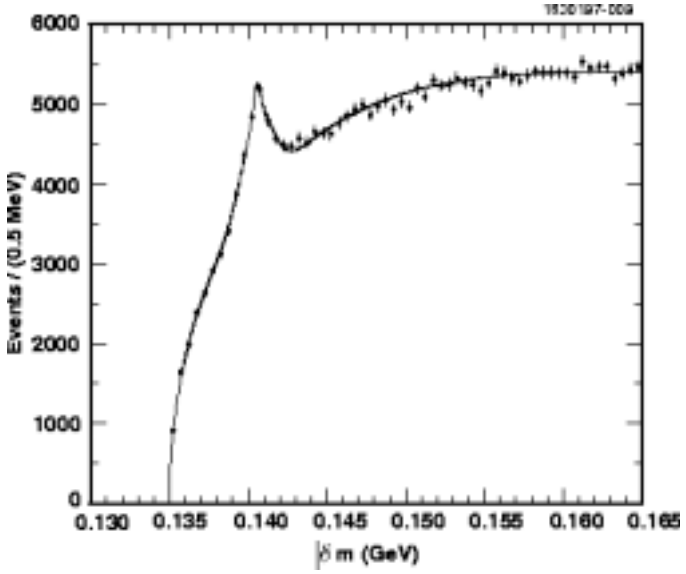

FIG. 3. The δm distribution for $D^{*+} \to D^+\pi^0$ candidates (data).

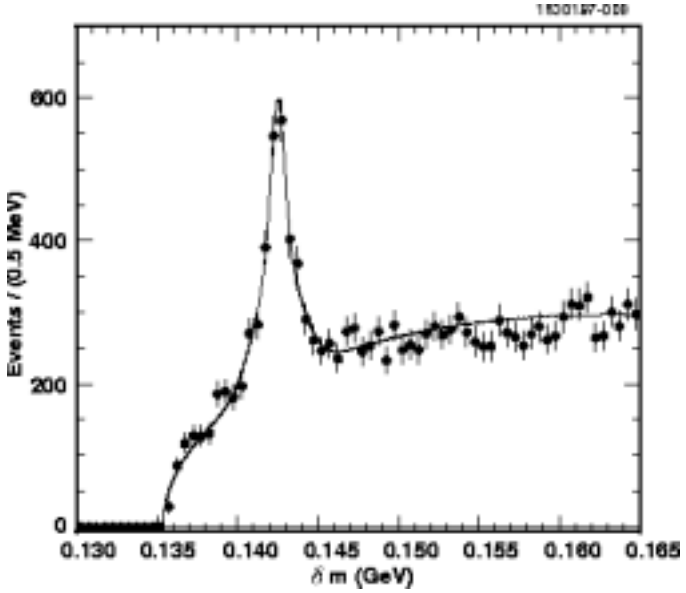

FIG. 4. The δm distribution for $D^{*0} \to D^0\pi^0$ candidates (data).

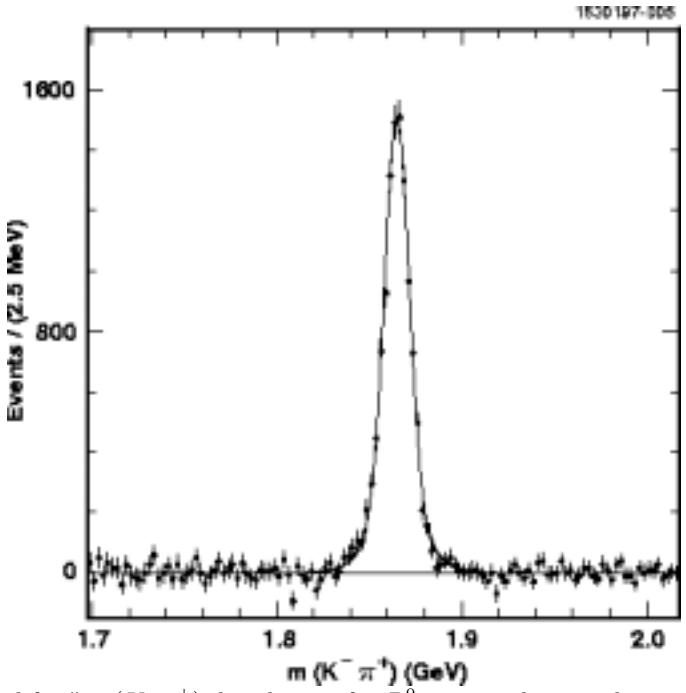

FIG. 5. The "background-free" $m(K^-\pi^+)$ distribution for D^0 mesons that are decay products of $D^{*+} \to D^0\pi^+$. $0.0 < x_D < 0.50$. The two-Gaussian fit of the distribution is also shown.

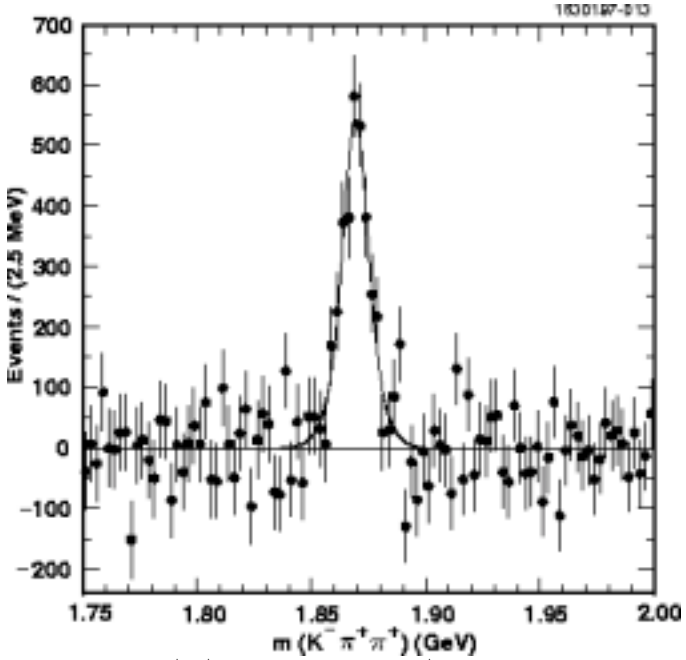

FIG. 6. The "background-free" $m(K^-\pi^+\pi^+)$ distribution for D^+ mesons that are decay products of $D^{*+} \to D^+\pi^0$. $0.0 < x_D < 0.50$. The two-Gaussian fit of the distribution is also shown.

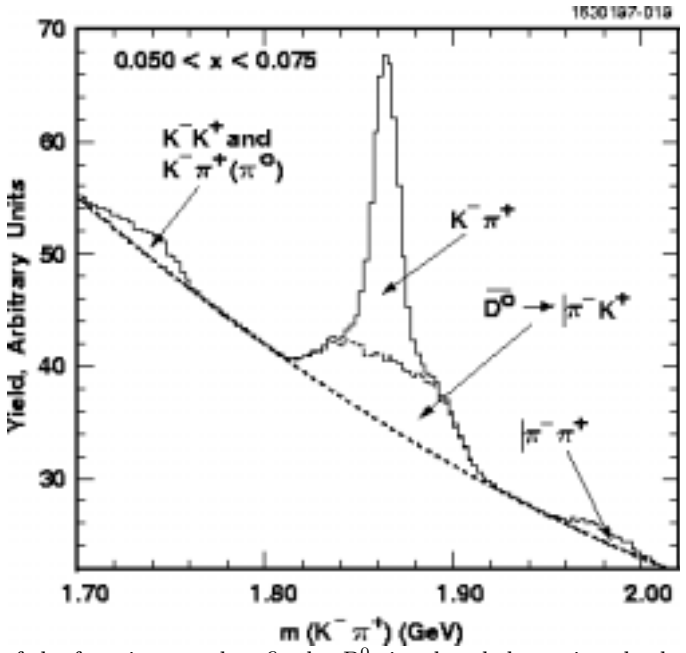

FIG. 7. Representation of the functions used to fit the D^0 signal and the various backgrounds for the momentum bin $0.050 < x < 0.075$. Note the offset of the vertical scale.

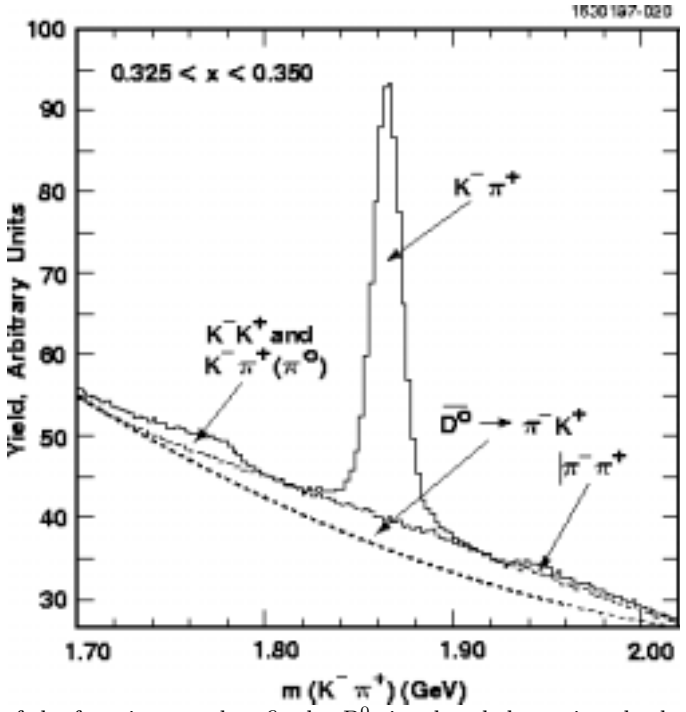

FIG. 8. Representation of the functions used to fit the D^0 signal and the various backgrounds for the momentum bin $0.325 < x < 0.350$. Note the offset of the vertical scale.

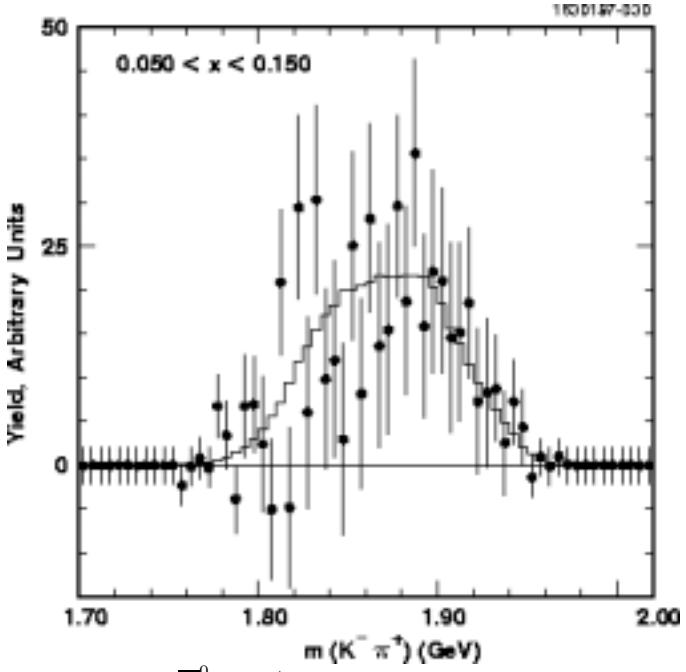

FIG. 9. The $m(K^-\pi^+)$ distribution for $\overline{D}^0 \to K^+\pi^-$ after switching particle identity. Points with error bars are from data, the histogram is obtained from Monte Carlo using track tagging (MC-tag). Momentum bin $0.05 < x < 0.15$.

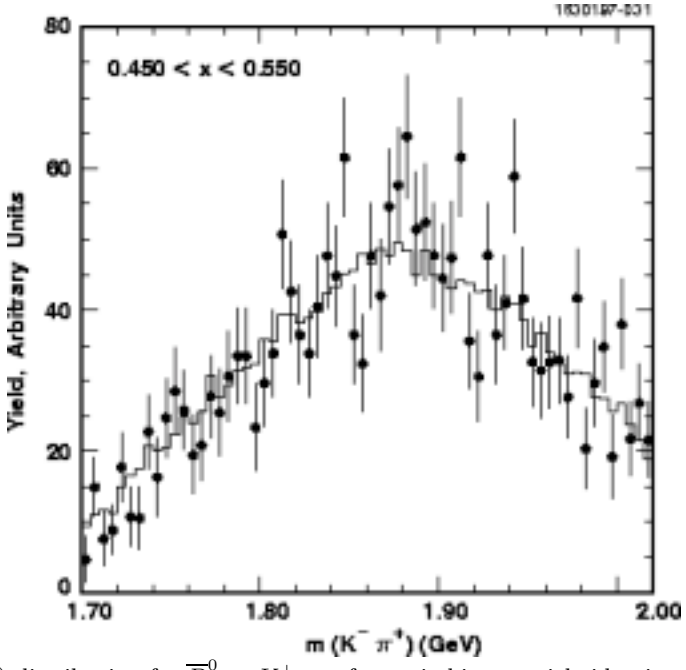

FIG. 10. The $m(K^-\pi^+)$ distribution for $\overline{D}^0 \to K^+\pi^-$ after switching particle identity. Points with error bars are from data, the histogram is obtained from Monte Carlo using track tagging (MC-tag). Momentum bin $0.45 < x < 0.55$.

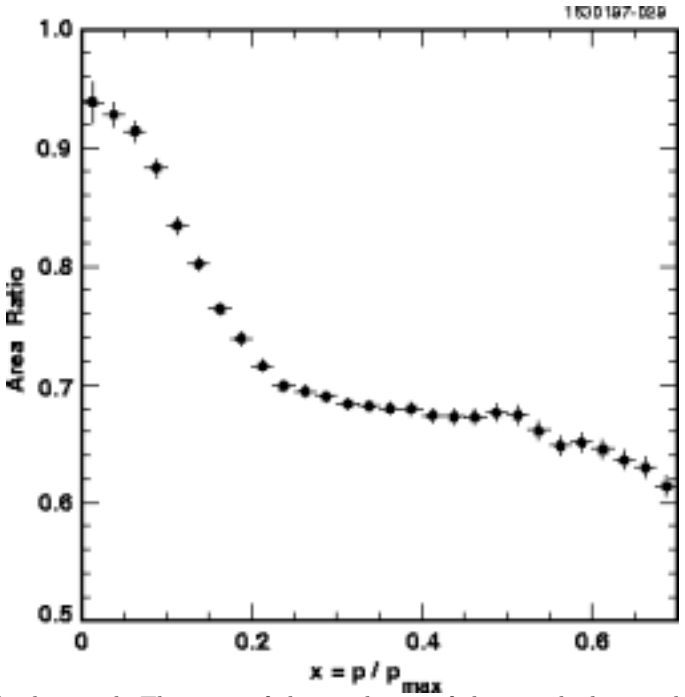

FIG. 11. Switched $K\pi$ background. The ratio of the total area of the switched mass distributions (satisfying our particle identification selections) to the signal area, vs scaled momentum $x = p/p_{max}$.

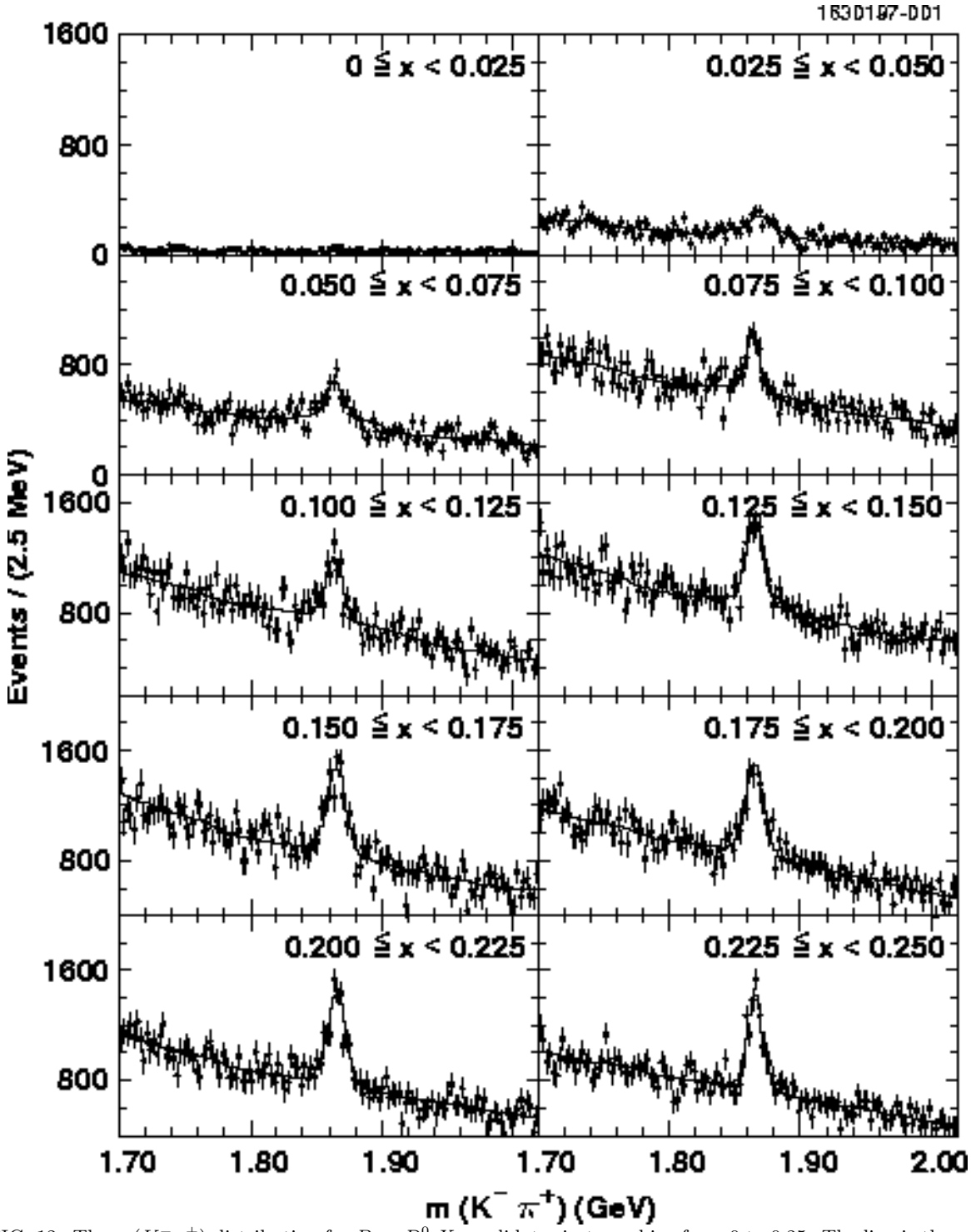

FIG. 12. The $m(K^-\pi^+)$ distribution for $B \to D^0\ X$ candidates in ten x bins from 0 to 0.25. The line is the result of the fit.

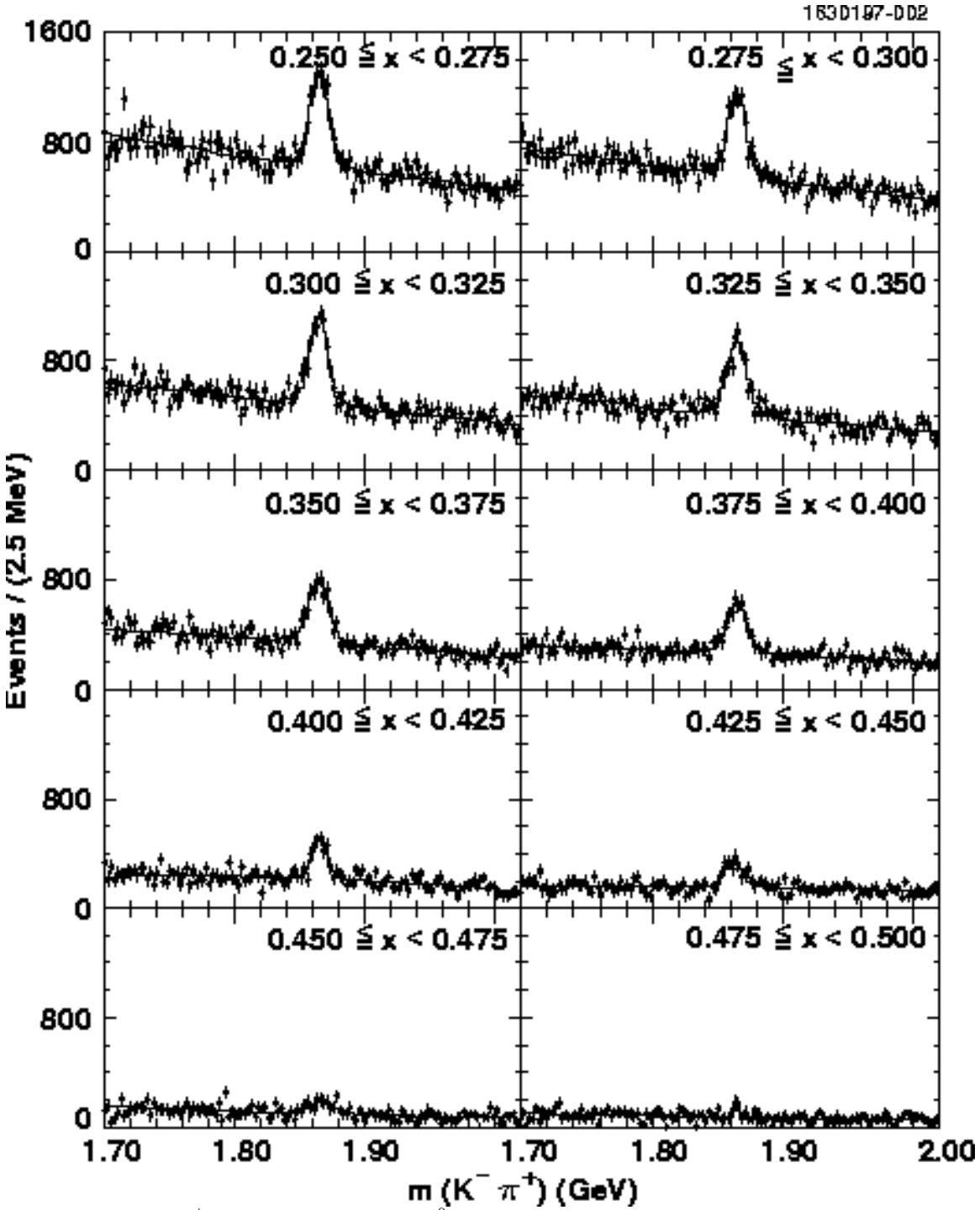

FIG. 13. The $m(K^-\pi^+)$ distribution for $B \to D^0\, X$ candidates in ten x bins from 0.25 to 0.50. The line is the result of the fit.

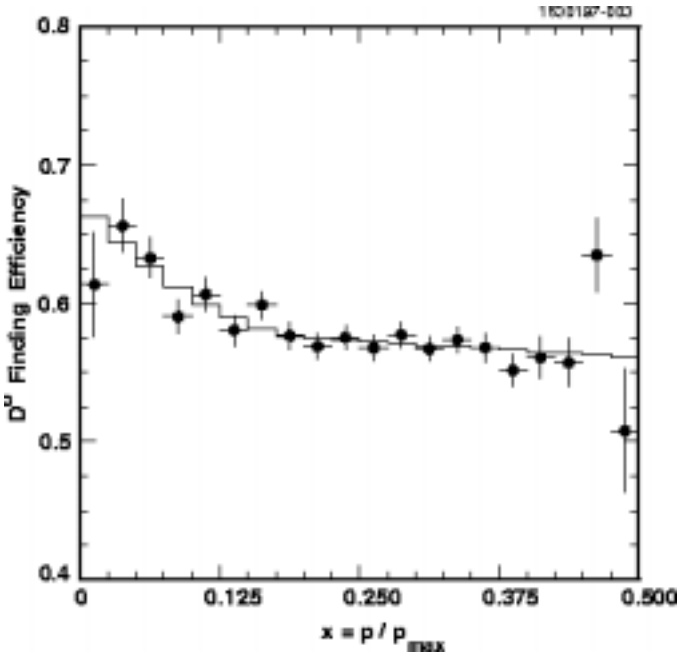

FIG. 14. D^0 finding efficiency as a function of scaled momentum x. The histogram is the result of the smoothing fit, binned.

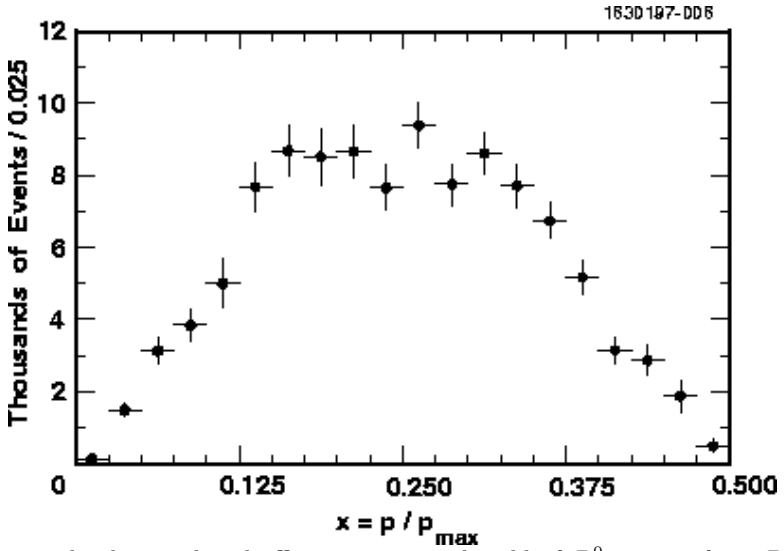

FIG. 15. The background subtracted and efficiency corrected yield of D^0 mesons from $B \to D^0 X$ decay, as a function of scaled momentum ($p_{max} = 4.950\ GeV/c$).

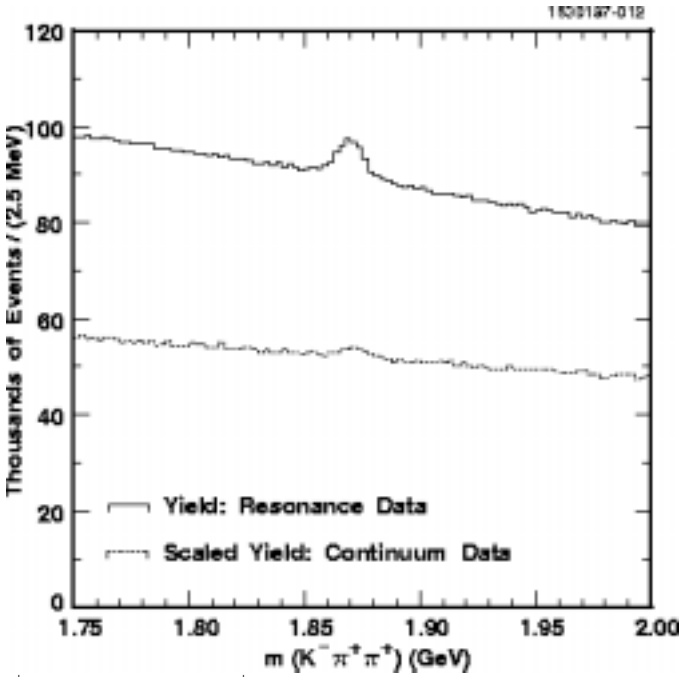

FIG. 16. The $m(K^-\pi^+\pi^+)$ distribution for D^+ candidates from data at the $\Upsilon(4S)$ resonance and from below $B\bar{B}$ threshold (scaled), $0.0 < x_D < 0.5$. The line is the result of the fit.

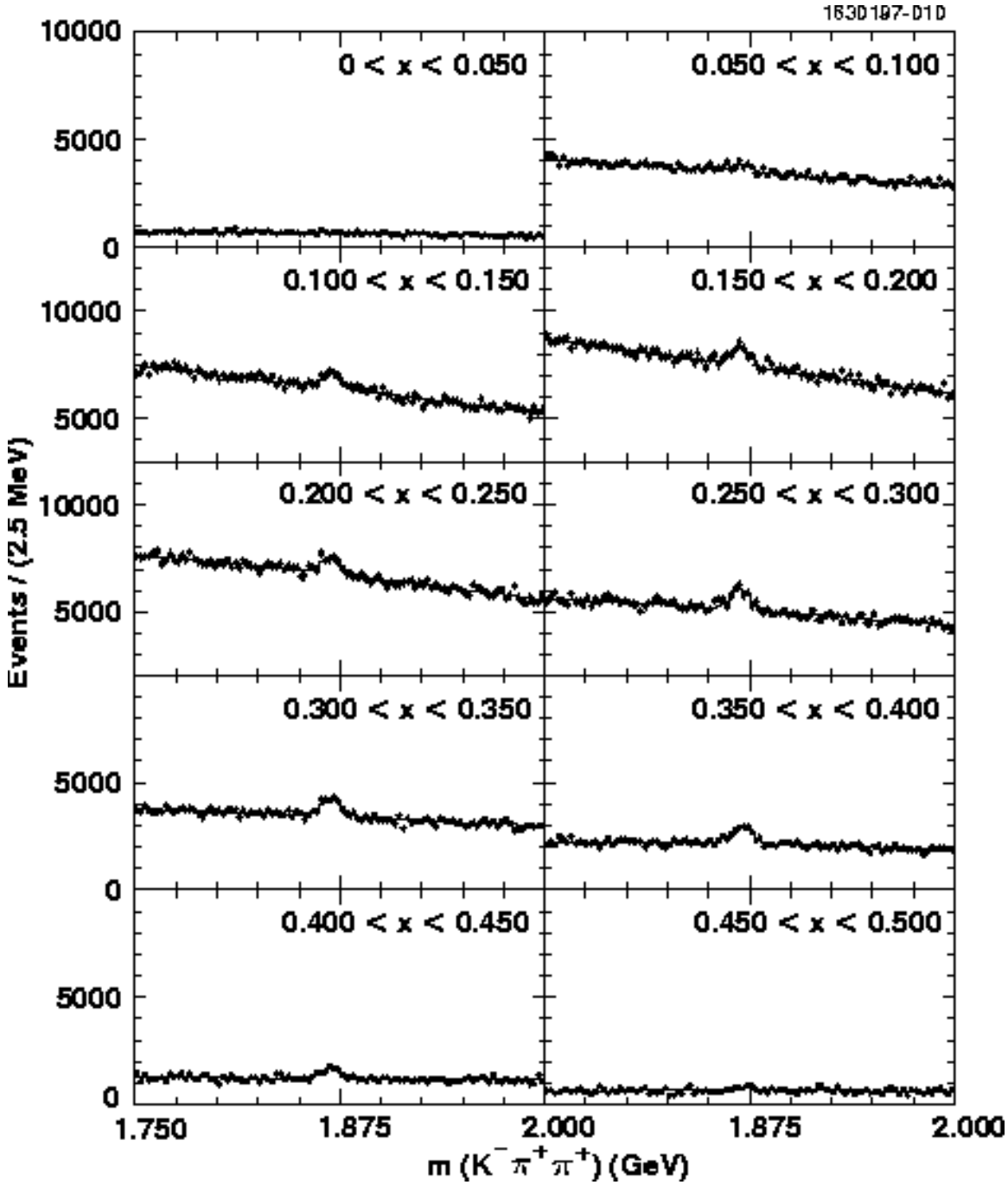

FIG. 17. The $m(K^-\pi^+\pi^+)$ distribution for $B \to D^+ X$ candidates in ten x bins from 0.0 to 0.50.

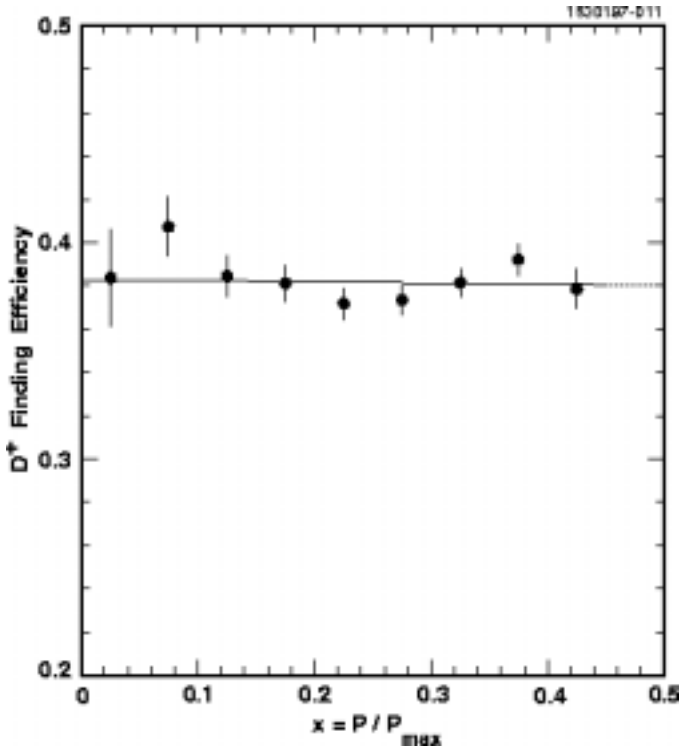

FIG. 18. D^+ finding efficiency as a function of scaled momentum x. The histogram is the result of the smoothing fit, binned.

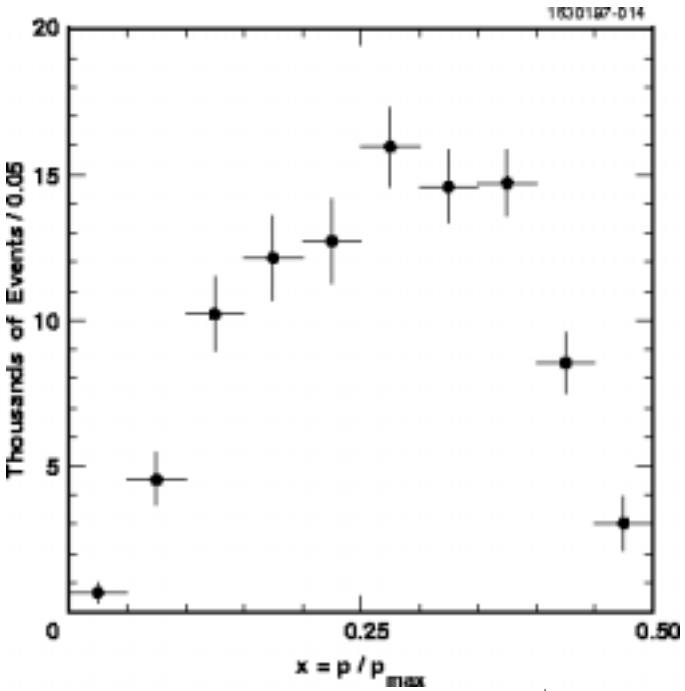

FIG. 19. The background subtracted and efficiency corrected yield of D^+ mesons from $B \to D^+X$ decay, as a function of scaled momentum ($p_{max} = 4.950\ GeV/c$).

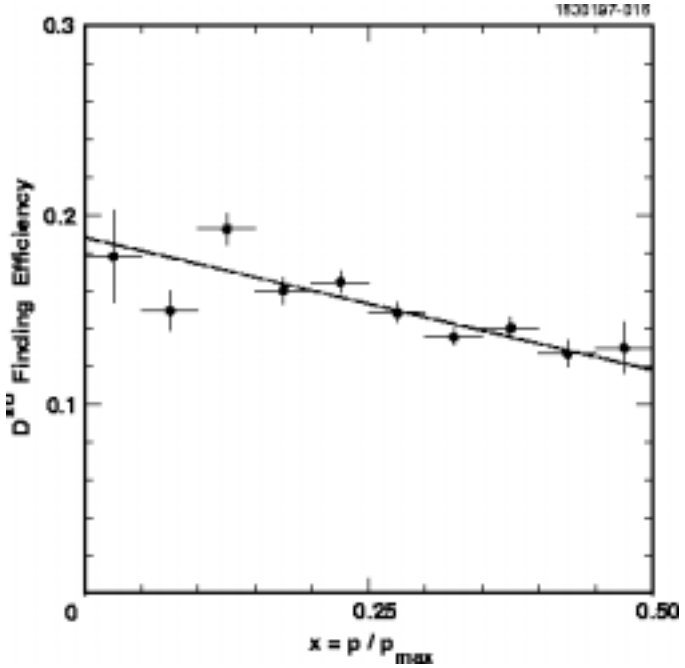

FIG. 20. D^{*0} finding efficiency as a function of scaled momentum x. The histogram is the result of the smoothing fit, binned.

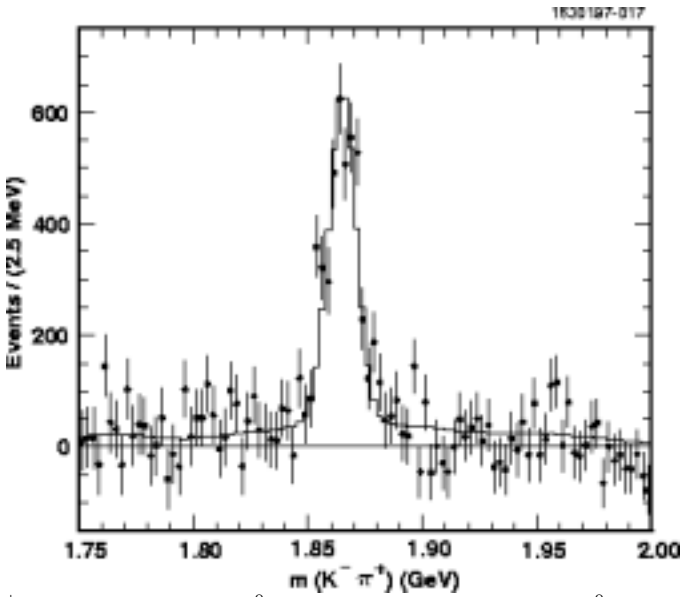

FIG. 21. The $m(K^-\pi^+)$ distribution for D^0 candidates from $B \to D^{*0} X$ in the momentum interval $(0.0 < x_{D^*} < 0.50)$.

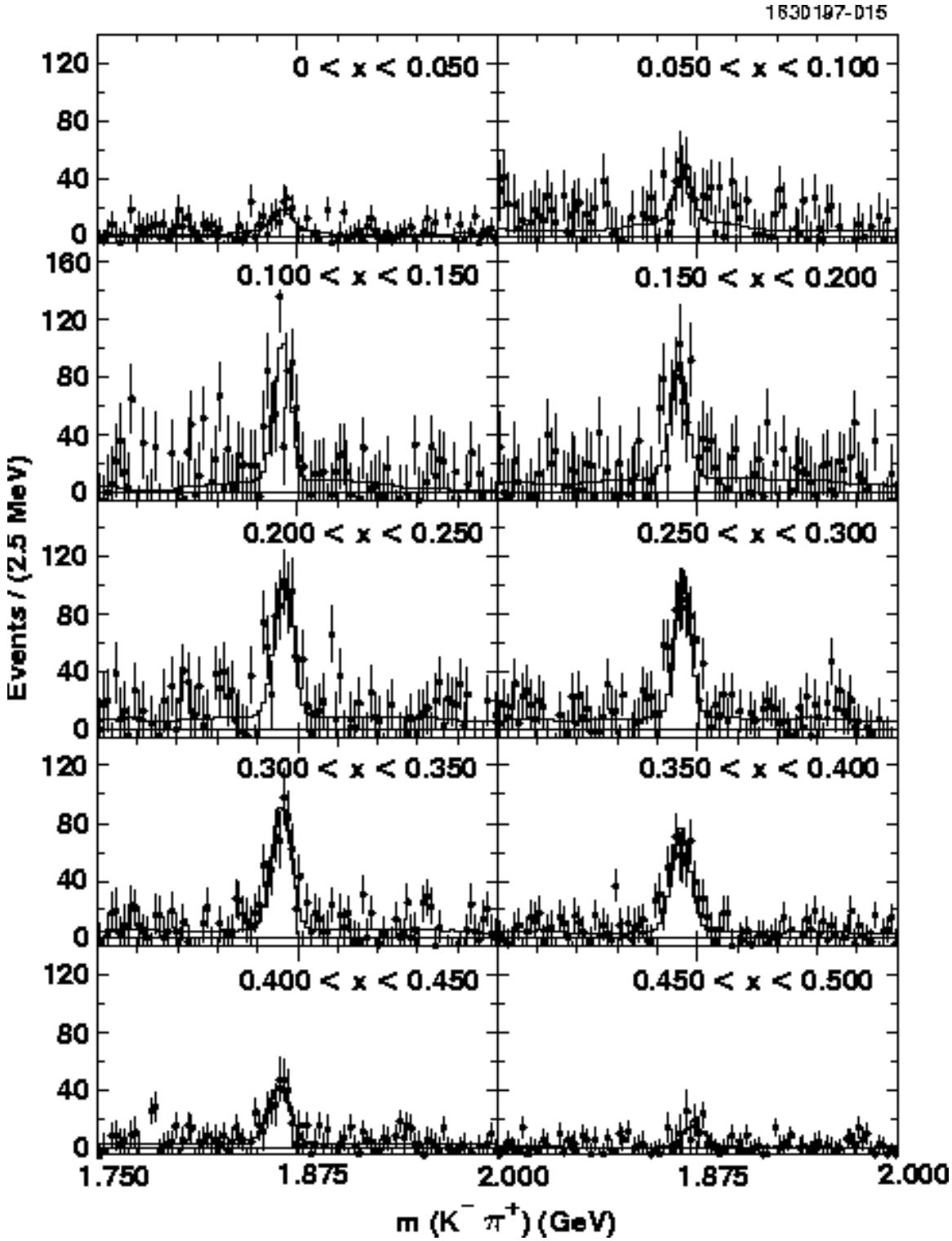

FIG. 22. The $m(K^-\pi^+)$ distribution for D^{*0} in ten x bins from 0 to 0.50. The line is the result of the fit.

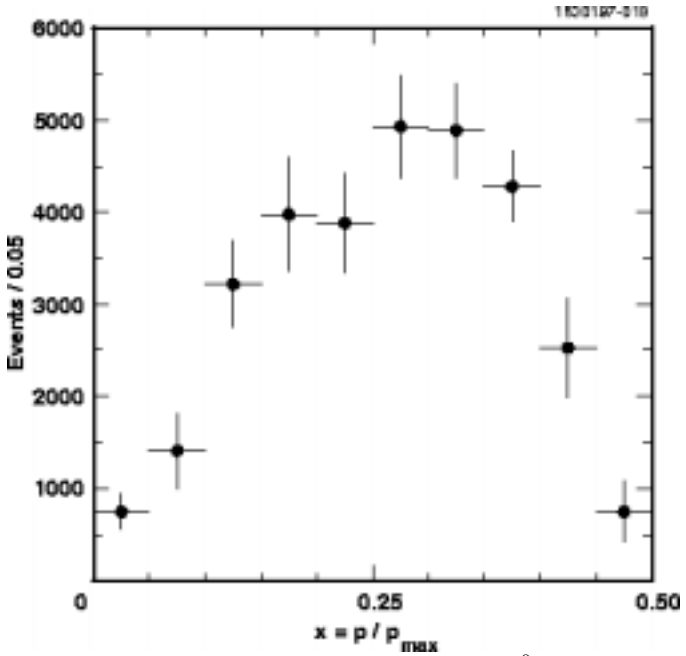

FIG. 23. The background subtracted and efficiency corrected yield of D^{*0} mesons from $B \to D^{*0}X \to (D^0\pi^0) + X$ decay, as a function of scaled momentum ($p_{max} = 4.950\ GeV/c$).

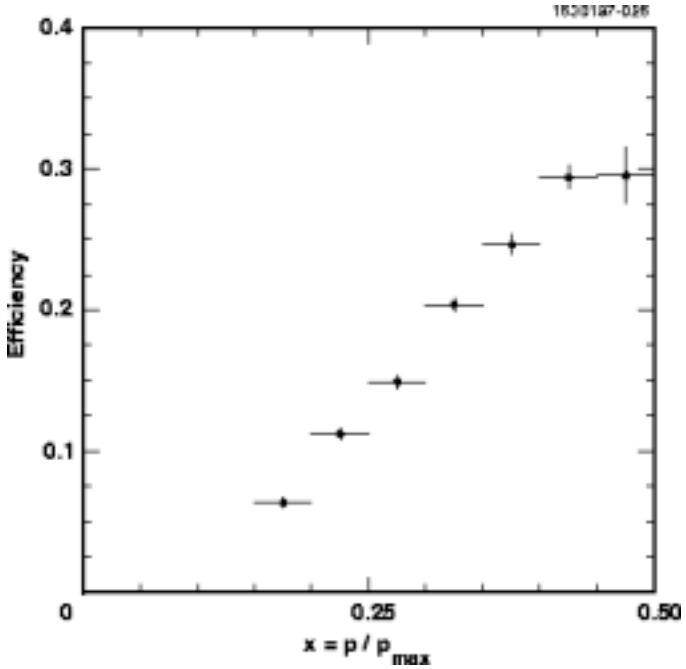

FIG. 24. The $D^{*+} \to D^0\pi^+$ finding efficiency as a function of scaled momentum x.

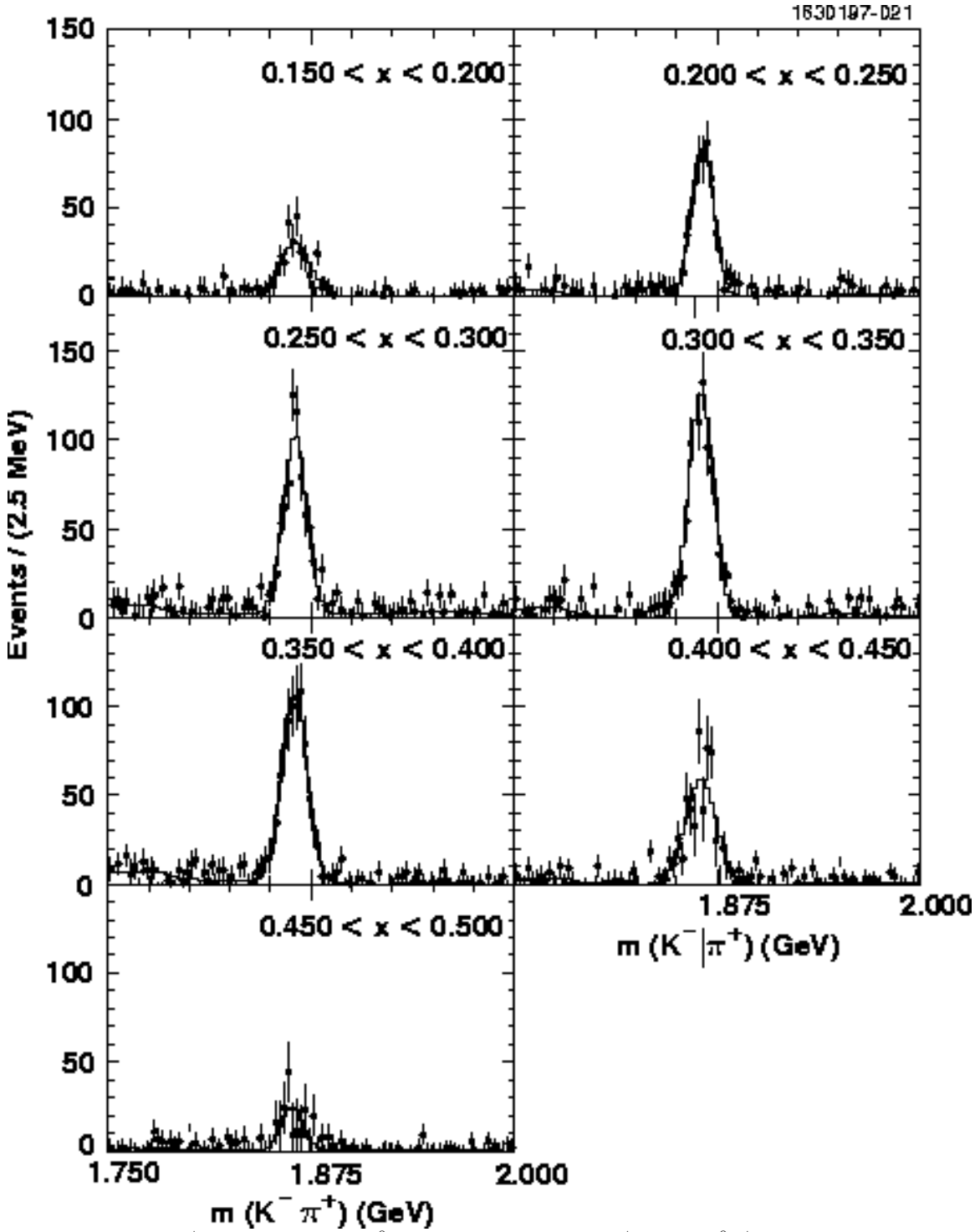

FIG. 25. The $m(K^-\pi^+)$ distribution for D^0 candidates from $B \to D^{*+}\ X \to (D^0\pi^+)\ X$ in 7 x bins from 0.15 to 0.50. The δm sidebads have been subtracted. The line is the result of the fit.

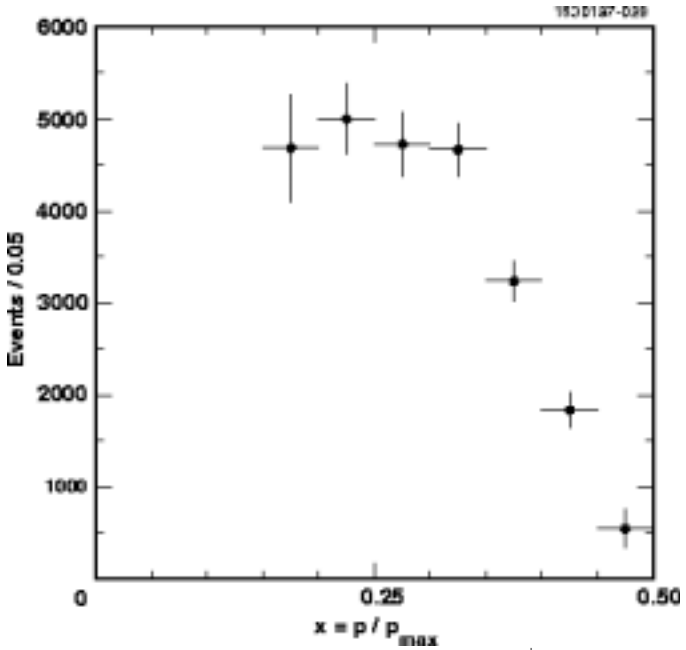

FIG. 26. The background subtracted and efficiency corrected yield of D^{*+} mesons from $B \to D^{*+}X \to (D^0\pi^+)\ X$ decay, as a function of scaled momentum ($p_{max} = 4.950\ GeV/c$).

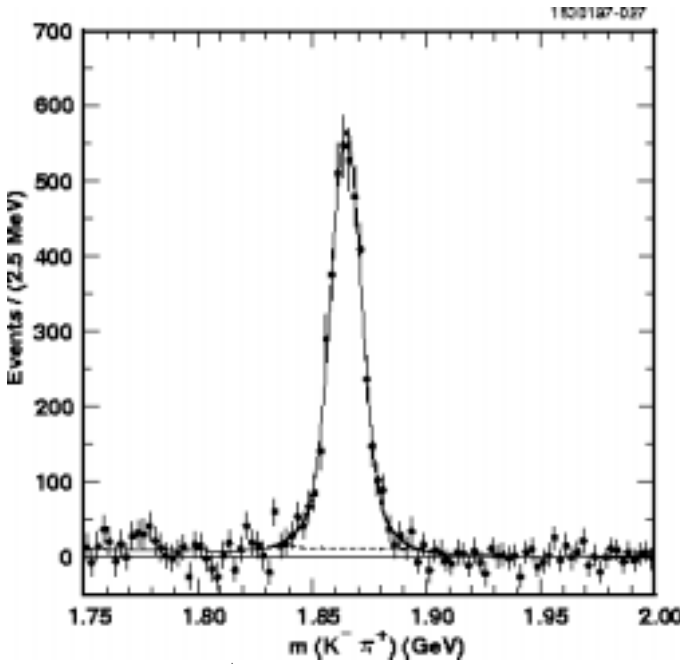

FIG. 27. The $m(K^-\pi^+)$ distribution for D^{*+} candidates from B decay ($0.15 < x_{D^*} < 0.50$), after δm sideband subtraction. The dashed line shows the background under the signal.

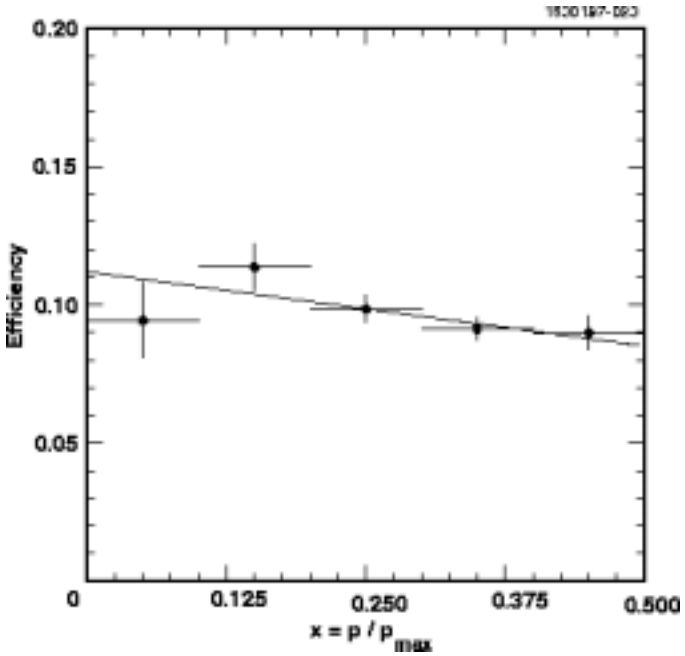

FIG. 28. $D^{*+} \to D^+\pi^0$ finding efficiency as a function of the scaled momentum x. The histogram is the result of the smoothing fit, binned.

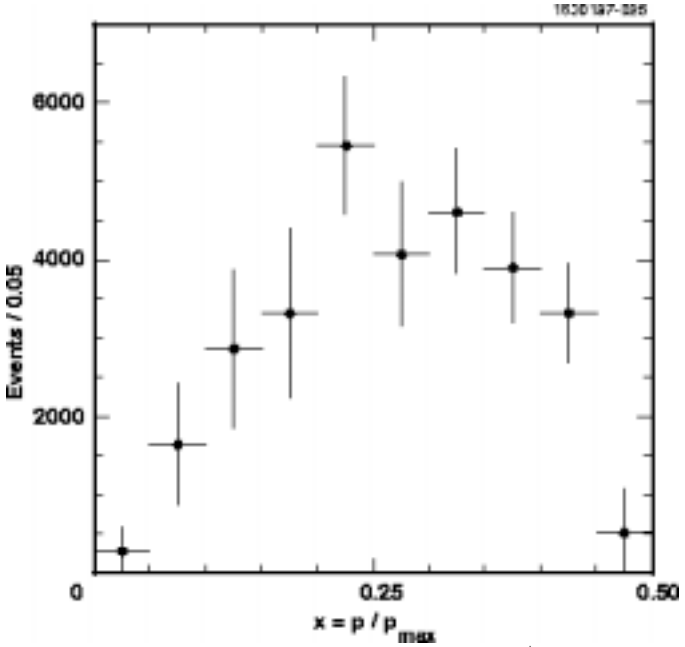

FIG. 29. The background subtracted and efficiency corrected yield of D^{*+} mesons from $B \to D^{*+}X \to (D^+\pi^0)\ X$ decay, as a function of scaled momentum ($p_{max} = 4.950\ GeV/c$).

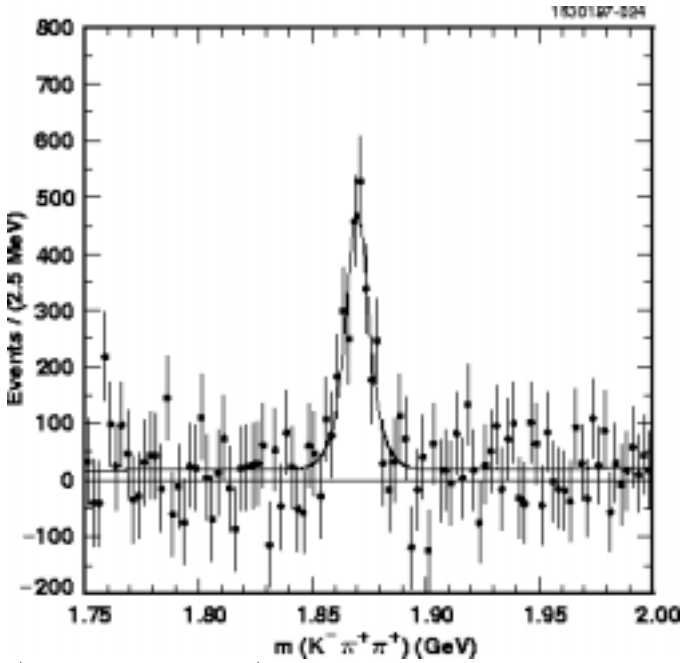

FIG. 30. The $m(K^-\pi^+\pi^+)$ distribution for D^{*+} candidates from B decay ($0.0 < x_{D^*} < 0.50$), after δm sideband subtraction.

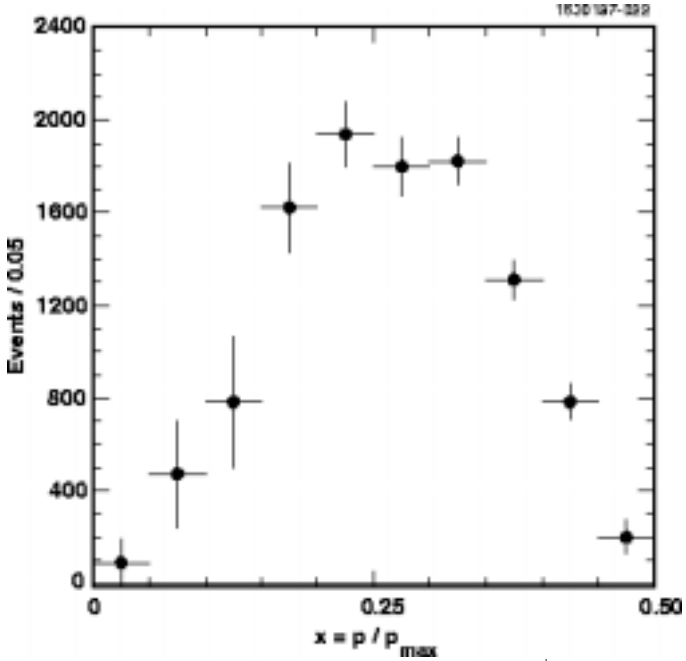

FIG. 31. The background subtracted and efficiency corrected yield of D^{*+} mesons from $B \to D^{*+}X$ decay, as a function of scaled momentum ($p_{max} = 4.950\ GeV/c$). For $0 < x < 0.15$, only the $D^{*+} \to D^+\pi^0$ measurement is used. For $0.15 < x < 0.5$, measurements from both $D^{*+} \to D^0\pi^+$ and $D^{*+} \to D^+\pi^0$ are combined.

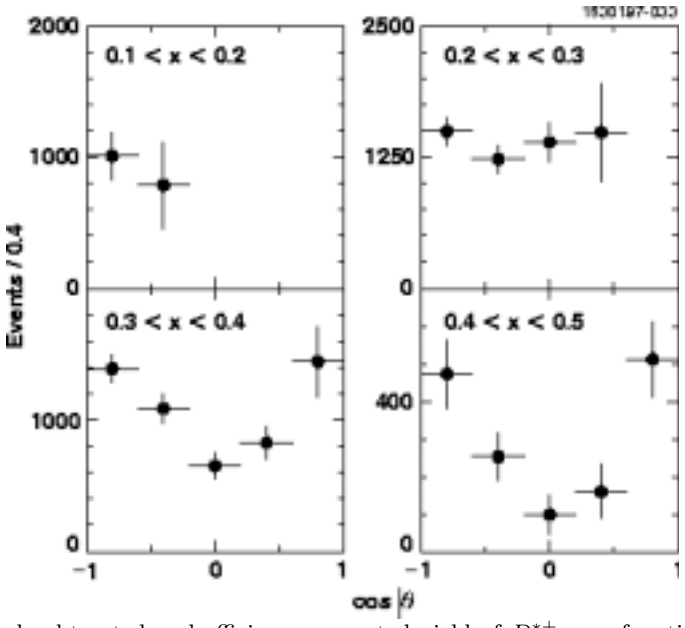

FIG. 32. The background subtracted and efficiency corrected yield of D^{*+} as a function of $\cos\theta$ for four x bins from 0.1 to 0.5.

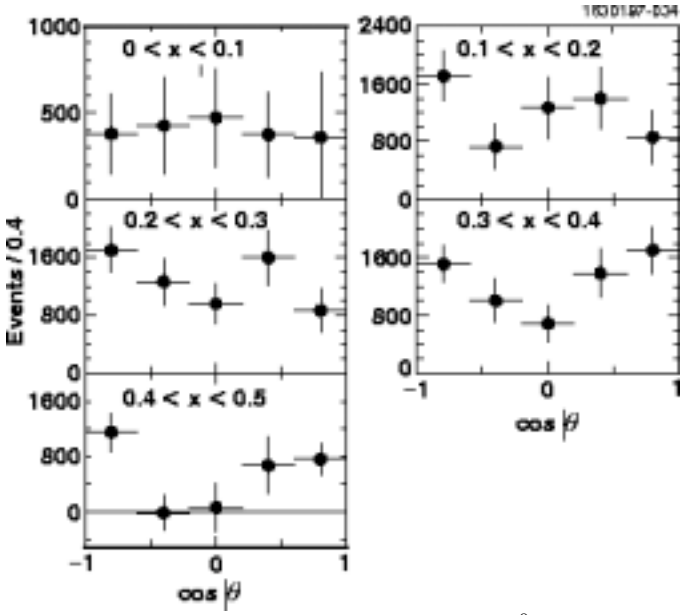

FIG. 33. The background subtracted and efficiency corrected yield of D^{*0} as a function of $\cos\theta$ for five x bins from 0.1 to 0.5.

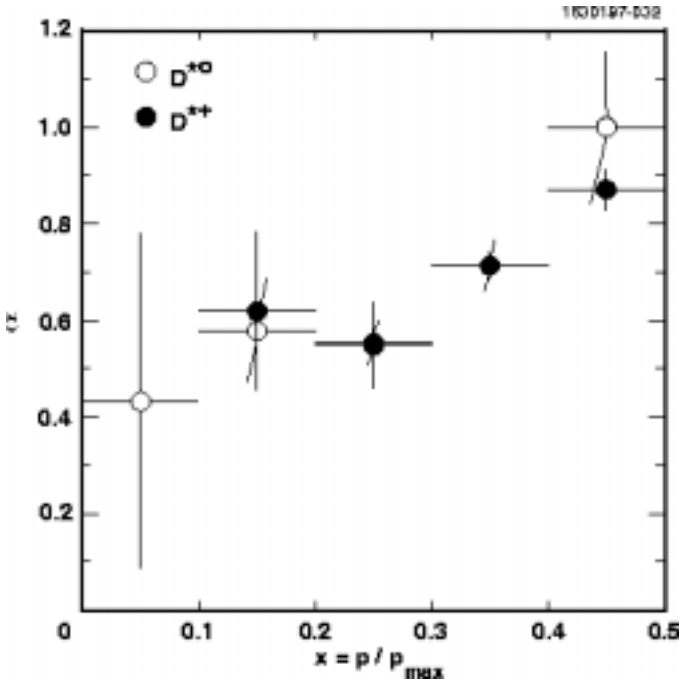

FIG. 34. $\alpha(D^*)$ vs $x = p/p_{max}$ ($p_{max} = 4.950\ GeV/c$).