CLEO CONF 96-25
ICHEP96 PA05-069

Measurement of $\mathcal{B}\left(B^- \to D_J^0\,\pi^-\right)$ using Partial Reconstruction

CLEO Collaboration

(July 19, 1996)

Abstract

The unique kinematics of $B\overline{B}$ production at the $\Upsilon(4S)$ allow full reconstruction of the decay sequence $B \to D_J\,\pi$, $D_J \to D^\star\,\pi$, $D^\star \to D\,\pi$ without reconstruction of the D. This method yields an overall efficiency of over 20% for the reconstruction of the D_J. Using $2.39 fb^{-1}$ of data accumulated at the $\Upsilon(4S)$ by the CLEO-II detector, and a two-dimensional fit to the distributions of the $D^{\star +}\pi^-$ mass and the helicity angle, we measure:

$$\mathcal{B}\left(B^- \to D_1(2420)^0\,\pi^-\right) \times \mathcal{B}\left(D_1(2420)^0 \to D^{\star +}\,\pi^-\right) = (7.8\pm1.6\pm1.0\pm0.2) \times 10^{-4}$$

$$\mathcal{B}\left(B^- \to D_2^\star(2460)^0\,\pi^-\right) \times \mathcal{B}\left(D_2^\star(2460)^0 \to D^{\star +}\,\pi^-\right) = (4.2\pm1.6\pm0.6\pm0.1) \times 10^{-4}$$

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

[1]University of California, Santa Barbara, California 93106
[2]University of Colorado, Boulder, Colorado 80309-0390
[3]Cornell University, Ithaca, New York 14853

---

*Permanent address: University of Texas at Austin.

†Permanent address: BINP, RU-630090 Novosibirsk, Russia.

[4]University of Florida, Gainesville, Florida 32611
[5]Harvard University, Cambridge, Massachusetts 02138
[6]University of Hawaii at Manoa, Honolulu, Hawaii 96822
[7]University of Illinois, Champaign-Urbana, Illinois 61801
[8]Carleton University, Ottawa, Ontario, Canada K1S 5B6
and the Institute of Particle Physics, Canada
[9]McGill University, Montréal, Québec, Canada H3A 2T8
and the Institute of Particle Physics, Canada
[10]Ithaca College, Ithaca, New York 14850
[11]University of Kansas, Lawrence, Kansas 66045
[12]University of Minnesota, Minneapolis, Minnesota 55455
[13]State University of New York at Albany, Albany, New York 12222
[14]Ohio State University, Columbus, Ohio 43210
[15]University of Oklahoma, Norman, Oklahoma 73019
[16]Purdue University, West Lafayette, Indiana 47907
[17]University of Rochester, Rochester, New York 14627
[18]Stanford Linear Accelerator Center, Stanford University, Stanford, California 94309
[19]Southern Methodist University, Dallas, Texas 75275
[20]Syracuse University, Syracuse, New York 13244
[21]Vanderbilt University, Nashville, Tennessee 37235
[22]Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
[23]Wayne State University, Detroit, Michigan 48202
[24]California Institute of Technology, Pasadena, California 91125
[25]University of California, San Diego, La Jolla, California 92093

I. INTRODUCTION

Both CLEO and ARGUS have reported measurements of the decay $B^- \to D_1(2420)^0\,\pi^-$ [1,2] with large errors and upper limits for the decay $B^- \to D_2^\star(2460)^0\,\pi^-$. Detailed study of these two-body decays and the subsequent decays of the excited D_J mesons can yield much information about both the B-meson and the spectroscopy of the charm meson. These measurements probe the interference between competing B-decay amplitudes, the strong interaction between the charm quark and a light quark, and the approximate spin-flavor symmetry predicted by Heavy Quark Effective Theory (HQET).

In addition, one may use the relation between hadronic and semileptonic B decays provided by the factorization hypothesis to infer the abundance of excited charm, often referred to generically as $D^{\star\star}$, in B semileptonic decays. A major source of uncertainty in the $b \to c$ lepton spectrum can then be reduced, and commensurately better values of V_{cb} derived. When direct measurements of the semileptonic analogues are made, the factorization hypothesis may then be tested further.

The high statistics samples required to make progress in these pursuits are difficult to obtain due to the small branching fractions and large number of sequential decays involved. Full reconstruction of such samples is challenging even with the large dataset of B mesons collected by CLEO. Fortunately, because of the unique kinematics of $B\overline{B}$ production at the $\Upsilon(4S)$ resonance, it is possible to fully reconstruct the kinematics of the decay chain $B \to D^\star\,\pi\,\pi$, $D^\star \to D\,\pi$, by reconstructing only the three pions. This reconstruction results in zero additional constraints, and has one quadratic ambiguity. A similar method [2] has been previously used by ARGUS in their analysis of these decays. The need to reconstruct the D is eliminated by this method, increasing the signal efficiency by a large factor, while reducing the dependence of the final result on individual D branching ratios and reconstruction efficiencies. Naturally, moving from an over-constrained condition ('full' reconstruction) to a zero-constraint condition also has its costs. Those costs are higher backgrounds, greater complexity in reconstruction, and somewhat larger systematic errors from other sources. The statistical errors are still dominant given the current data sample.

In this paper, we describe the use of this method to reconstruct a high statistics sample of $B^- \to D^{\star +}\,\pi^-\,\pi^-$, focusing on the $B^- \to D_J^0\,\pi^-$ intermediate states. Using a two-dimensional, unbinned, maximum-likelihood fit; we measure the product branching ratio:

$$\mathcal{B}\left(B^- \to D_J^0\,\pi^-\right) \times \mathcal{B}\left(D_J^0 \to D^{\star +}\,\pi^-\right) \tag{1}$$

for the $D_1(2420)^0$ and the $D_2^\star(2460)^0$.

II. THEORETICAL CONSIDERATIONS

A. D_J Mesons and HQET

The D_J mesons consist of one light quark and one charm quark with total angular momentum J. The lowest-mass states of this kind have orbital angular momentum of 1. There are four $L = 1$ states with spin parity $J^P = 0^+$, 1^+, 1^+, and 2^+. These have been labeled by the Particle Data Group as $D_0^\star$, D_1 (for both 1^+ states) and $D_2^\star$ respectively.

TABLE I. The four neutral $L = 1$ charm mesons and the angular momenta allowed in their strong decays to $D^\star\pi$ in the HQET limit. L_f is the HQET-allowed orbital angular momentum of the the $D^\star\pi$ system in the final state. j is the total angular momentum carried by the light quark.

State	M(MeV)	Γ(MeV)	J^P	j	$L_f^{D_J \to D^\star\pi}$
$D_2^\star(2460)^0$	2458.9±2.0	23±5	2^+	3/2	2
$D_1(2420)^0$	2422.0±2.1	$18.9^{+4.6}_{-3.5}$	1^+	3/2	2
$D_1^0(j = 1/2)$	?	> 100?	1^+	1/2	0
$D_0^\star(j = 1/2)$	?	> 100?	0^+	1/2	—

Conservation of parity and angular momentum place restrictions on the strong decays of these states. The 0^+ state can only decay to $D\pi$ through S-wave decay. Either 1^+ can decay to $D^\star\pi$ through S-wave or D-wave decays. The 2^+ can decay to both $D\pi$ and $D^\star\pi$, but only through D-wave decay.

HQET predicts an approximate spin-flavor symmetry [3] for hadrons with only one heavy quark Q, where $M_Q \gg \Lambda_{QCD}$. In the limit $M_Q \to \infty$ the symmetry is perfect, and HQET makes definite predictions regarding the properties and decays of these states [4]. The quantum numbers of the eigenstates in this picture are S_Q, the spin of the heavy quark; $j = S_q + L$, the total angular momentum carried by the light quark; and $J = j + S_Q$, the total angular momentum of the meson. This arranges the four $L = 1$ states into two doublets, $j = 1/2$ and $j = 3/2$, as shown in Table I.

In the HQET limit, the $j = 3/2$ mesons may decay only via D-wave, and the $j = 1/2$ only via S-wave. Thus, the (1^+, $j = 3/2$) state is predicted to be relatively narrow (Γ of order 10 MeV) while the $j = 1/2$ states are predicted to be relatively broad (Γ of hundreds of MeV). Due to the approximate nature of this symmetry for the charm quark, some mixture of S-wave and D-wave components is expected in the decays of the (1^+, $j = 3/2$) and (1^+, $j = 1/2$) states. Note that this is not the case for the (2^+, $j = 3/2$) state, where the D-wave requirement arises from conservation of parity and angular momentum.

The helicity angular distribution of the $D^\star$ in the decay chain $D_J \to D^\star\pi_2$, $D^\star \to D\pi_3$ can be used to analyze the spin of the D_J and the partial wave components of its decay regardless of the initial polarization of the D_J. This angle, denoted θ_3, is the polar angle between the $D^\star$ boost direction in the D_J frame and the D direction measured in the $D^\star$ frame.

This angle, denoted θ_3, is the polar angle between the $D^\star$ direction and the D direction measured in the $D^\star$ frame (Fig. 1). The predicted distributions for θ_3 are:

$$\frac{dN}{d\cos\theta_3} = \begin{cases} \sin^2\theta_3 & (\text{ } 2^+ \text{ state }) \\ 1 & (\text{ S-wave } 1^+ \text{ state }) \\ 1 + 3\cos^2\theta_3 & (\text{ D-wave } 1^+ \text{ state }) \end{cases} \tag{2}$$

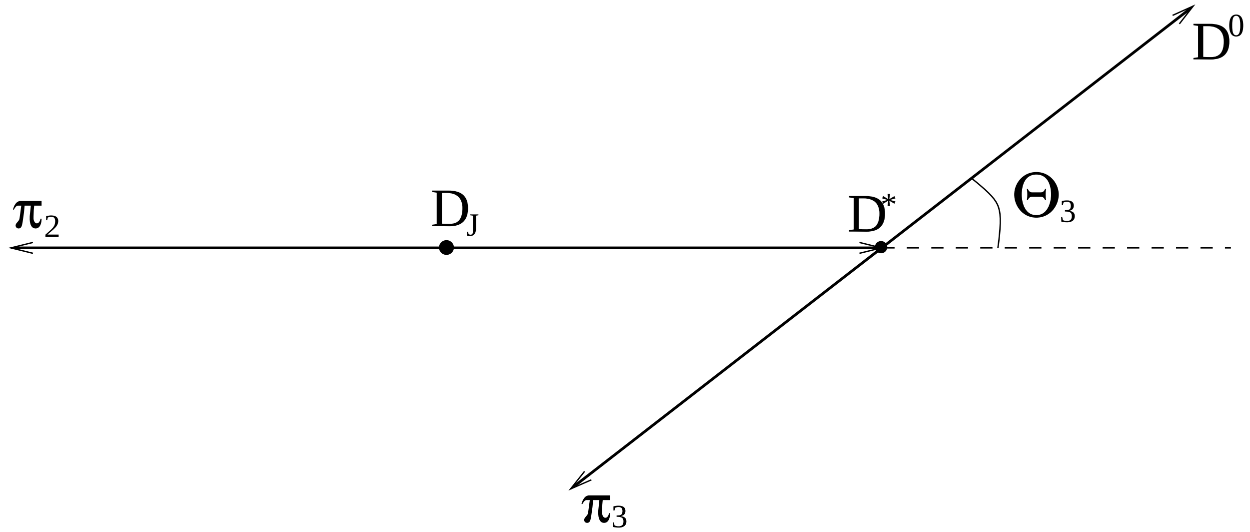

FIG. 1. The angle θ_3 . The figure is drawn in the rest frame of the $D^{\star}$.

TABLE II. Results of helicity analysis of D_J^0 by CLEO. The data were fitted to the distribution $(1 + B\cos^2\theta_3)$ to determine the experimental value of B. These results clearly support the D-wave prediction for $D_1(2420)^0 \to D^{\star +}\pi^-$ decays made by HQET.

State	Predicted B	Measured B
$D_1(2420)^0$	3	$2.74^{+1.04}_{-0.93}$
$D_2^{\star}(2460)^0$	-1	$-0.74^{+0.49}_{-0.38}$

Previous helicity analysis using high statistics samples of D_J^0 at CLEO [6] have revealed two narrow states which appear to correspond to the $j = 3/2$ states. In that study the data were fitted to the angular distribution $(1+B\cos^2\theta_3)$ where B is a fit parameter. The results, summarized in Table II, only loosely constrain S-wave decays of the $D_1(2420)^0$ but tend to support the HQET prediction that D-wave decays dominate. The result for the $D_2^{\star}(2460)^0$ agrees well with the D-wave requirement provided by conservation of parity and angular momentum.

III. DATA SAMPLE AND EVENT SELECTION

The data used in this analysis were selected from hadronic events produced in e^+e^- annihilations at the Cornell Electron Storage Ring (CESR). The data sample consists of $2.39 fb^{-1}$ collected by the CLEO-II detector at the $\Upsilon(4S)$ resonance (referred to as 'on-resonance') and $1.13 fb^{-1}$ at a center-of-mass energy just below the threshold for production of $B\overline{B}$ pairs (referred to as 'off-resonance' or 'continuum'). The on-resonance data corresponds to $2,560,000\pm46,000$ $B\overline{B}$ pairs. The off-resonance data is used to model the continuum background. Hadronic events are selected by requiring a minimum of three charged tracks, a total visible energy greater than 15% of the center-of-mass energy, and a primary vertex within ± 5 cm in the z direction and ± 5 mm in the $r-\phi$ plane from the nominal collision point.

The CLEO II detector [7] measures both charged and neutral particles with excellent resolution and efficiency. Charged particles are measured in a 1.5 Tesla magnetic field through three nested cylindrical drift chambers covering 94% of the solid angle. These chambers are followed by time-of-flight (TOF) counters. Particle identification is accomplished by combining specific ionization (dE/dx) information from the drift chambers with TOF information. Information from an electromagnetic calorimeter and muon chambers can be used to help reject electrons and muons as π candidates.

Since the B mesons are produced almost at rest, their decay products are, on average, uniformly distributed throughout the volume of the detector. This produces events that tend to be 'spherical' in shape. Continuum events are more jetlike in shape. In order to select events that are consistent with the hypothesis of $B\overline{B}$ production, we choose events that are spherical using the shape variable R_2. This variable is the ratio of the second Fox-Wolfram moment to the zeroth moment and has been found to be useful in discriminating between $B\overline{B}$ events and continuum events. Only events with $R_2 < 0.3$ are used in this analysis. We also require that the events we use have no charged tracks with momentum greater than 2.45 GeV. The efficiency of these cuts for the signal we seek is $\approx 87\%$.

The decay sequence $B^- \to D^{\star +}\,\pi^-\,\pi^-$, $D^{\star +} \to D^0\,\pi^+$ is reconstructed using only the three pions. For this general case the pions from the B decay are indistinguishable. However, for the resonant components $B^- \to D^0_J\,\pi_1{}^-$, $D^0_J \to D^{\star +}\,\pi_2{}^-$, $D^{\star +} \to D^0\,\pi_3{}^+$ it is important to remember that the π_1 and π_2 play very different roles: For the $D_1(2420)^0$ and the $D_2^\star(2460)^0$, their momentum spectra do not overlap as shown in Fig. 2.

All candidate tracks are required to be consistent with being produced at the primary event vertex. In addition, we use only tracks with $|\cos\theta| < 0.91$ with respect to the beamline, requiring that they pass through the central drift chambers described above. The detector acceptance for charged tracks below 50 MeV is negligible. We restrict the candidate π_3 to have momentum greater than 50 MeV, discarding a small fraction of π_3 from signal events.

We require consistency with the pion hypothesis for all π candidates by requiring that the (dE/dX) and TOF information are within three standard deviations of the expected values when the information is available. When not available, no cut is made.

We use information from the muon counters and the calorimeter to reduce the number of muons and electrons which can fake our π_1 and π_2. These vetos are only applied when the information from these detector elements is available and are not applied to π_3 candidates.

The efficiency of these tracking and pion consistency cuts for the three pions from $B^- \to D_1(2420)^0\,\pi_1{}^-$, $D_1(2420)^0 \to D^{\star -}\,\pi_2{}^-$, $D^{\star +} \to D^0\,\pi_3{}^+$ are 0.75, 0.69, and 0.51 for the π_1, π_2, and π_3 respectively.

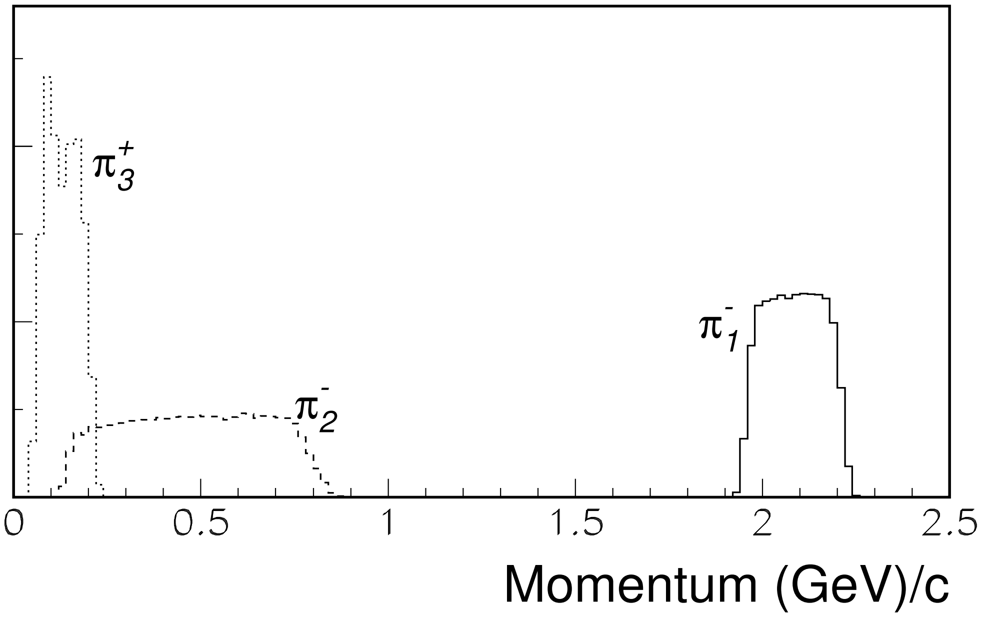

FIG. 2. The spectra of the pions produced by the decay sequence $B^- \to D_1(2420)^0\, \pi_1{}^-$, $D_1(2420)^0 \to D^{\star -}\, \pi_2{}^-$, $D^{\star +} \to D^0\, \pi_3{}^+$.

IV. RECONSTRUCTION METHOD

A. Variables and Constraints

The reconstruction method employed by this analysis allows full reconstruction of the decay sequence without any reconstruction of the D. A similar method has been used previously by ARGUS [2]. For the purpose of describing the reconstruction of $B \to D^\star \pi\, \pi$, $D^\star \to D\, \pi$ it will be convenient to create a pseudo-particle $X = (\pi\pi)$ and consider the more general decay sequence:

$$B \to D^\star X\ ,\ D^\star \to D\pi_{\rm slow}. \tag{3}$$

There are five particles in the sequence described by Eq. 3 (B,$D^\star$,X,D,π). Four variables (a 4-momentum) are required to describe each particle. A total of $20 = 5{\times}4$ variables are then needed to completely describe the sequence in Eq. 3.

The constraints we exploit are:

- Energy-momentum conservation in the two decays ($4{\times}2 = 8$ constraints).
- The B mass (1 constraint).
- The B energy, given by the beam energy of CESR (1 constraint).
- The $D^\star$ mass (1 constraint).

- The four reconstructed variables that describe the X (4 constraints).
- The four reconstructed variables that describe the π_{slow} (4 constraints).
- The mass of the D (1 constraint).

This gives 20 constraints, just enough to cover the variables and produce a complete solution for the kinematics, a '0C' fit. The solution we get is quadratic, however, meaning that:

1. In general, the solutions are complex valued. When a non-zero imaginary part is required there is no physically acceptable solution.
2. For most cases with a physically acceptable solution, there are two acceptable solutions.

For the analysis of $B \to D^\star \pi \pi$, we may then search for resonances between the $D^\star$ and each π in X, namely the D_J. Since we have a complete solution for the kinematics of the decay sequence, we may also employ the decay angular distributions to help distinguish the various D_J from each other and from background events.

For the $B \to D^\star X$ decay we have:

- m_B (assumed), E_B (measured by CESR).
- $m_{D^\star}$ (assumed).
- $\vec{p}_X$, m_X (reconstructed assuming $X = (\pi\pi)$).

By requiring conservation of energy-momentum in this decay we can:

1. Compute the magnitude of $\vec{p}_{D^\star}$ in the lab.
2. Constrain the direction of $\vec{p}_{D^\star}$ to lie on a cone centered about the $-\vec{p}_X$ direction in the lab (see Fig. 3).

The opening angle of this cone, denoted $\theta_{XD^\star}$, is given in terms of lab quantities by:

$$\cos\theta_{XD^\star} = \frac{{m_B}^2 - {m_{D^\star}}^2 - {m_X}^2}{2|\vec{p}_{D^\star}||\vec{p}_X|} - \frac{1}{\beta_{D^\star}\beta_X} \tag{4}$$

When reconstructing random tracks as X, these constraints usually give $|\cos\theta_{XD^\star}| > 1$, providing strong suppression of background events.

For the $D^\star \to D\pi$ decay we have:

- $m_{D^\star}$ (assumed), $E_{D^\star}$ (calculated from constraints above).
- m_D (assumed).
- $\vec{p}_\pi$ (reconstructed), m_π (assumed).

Note that we have the analogous information to that in the decay $B \to D^\star X$. Again, by requiring four-momentum conservation we constrain the $D^\star$ to lie on a cone, this time centered about the $\pi_{\rm slow}$ (see Fig. 3).

The opening angle of this cone, denoted $\theta_{\pi D^\star}$, is given in terms of lab quantities by:

$$\cos\theta_{\pi D^\star} = -\frac{{m_{D^\star}}^2 + {m_\pi}^2 - {m_{D^0}}^2}{2|\vec{p}_{D^\star}||\vec{p}_\pi|} + \frac{1}{\beta_{D^\star}\beta_\pi} \tag{5}$$

When reconstructing a random track as the $\pi_{\rm slow}$, these constraints often give $|\cos\theta_{\pi D^\star}| > 1$, again providing supression of background events. Due the large number of $\pi_{\rm slow}$ candidates, this requirement rejects background combinations less strongly than the similar requirement for $B \to D^\star X$.

B. Solutions for $\vec{p}_{D^\star}$ and Identification of Signal

We now know the magnitude of the $D^\star$ momentum and have constrained it to lie simultaneously on two distinct cones, one centered about $\vec{p}_X$ and one about $\vec{p}_\pi$ (see Fig. 3). There are two possibilities:

1. These cones do not cross and there is no acceptable solution for $\vec{p}_{D^\star}$.
2. These cones do cross. There are then, in general, two acceptable solutions for $\vec{p}_{D^\star}$.

From this information, the directions of the two solutions for $\vec{p}_{D^\star}$ are easily calculated. With solutions for the $D^\star$ four-momentum, it is trivial to reconstruct the solutions for p_B and p_{D^0} as well.

One of these solutions typically corresponds to the 'right' solution for $\vec{p}_{D^\star}$. The other solution is an artifact of the reconstruction method. Due to the small sizes of the cones ($\vec{p}_B$,$p_{\pi_{\rm slow}} \ll \vec{p}_{D^\star}$) the smearing of $\vec{p}_{D^\star}$ for the artifact solution is typically small.

Once the decay sequence is reconstructed, the problem of how to identify signal events must still be solved. We may not make the usual plot of beam-constrained mass, having already used the beam energy constraint to reconstruct the decay chain. However, the following variables are available:

- $M(D^{\star+}\pi^-)$, where we combine the $D^\star$ and one of the π^- in X to search for the D_J^0.
- The helicity angular distribution, θ_3.
- Two other decay angular distributions are available. We do not use the information in these distributions for the current result and will not discuss them further.

While the most powerful of these for the $D_1(2420)^0$ and $D_2^\star(2460)^0$ is $M(D^{\star+}\pi^-)$, a two-dimensional fit using both $M(D^{\star+}\pi^-)$ and θ_3 will give greater signal separation and will allow us to probe the properties of the D_J^0.

The efficiency for this method *including* the event and pion selection cuts already discussed is 21% and 22% for the decays $B^- \to D_1(2420)^0\,\pi^-$ and $B^- \to D_2^\star(2460)^0\,\pi^-$ respectively. This does not include the branching fraction for $D^{\star+} \to D^0\,\pi^+$.

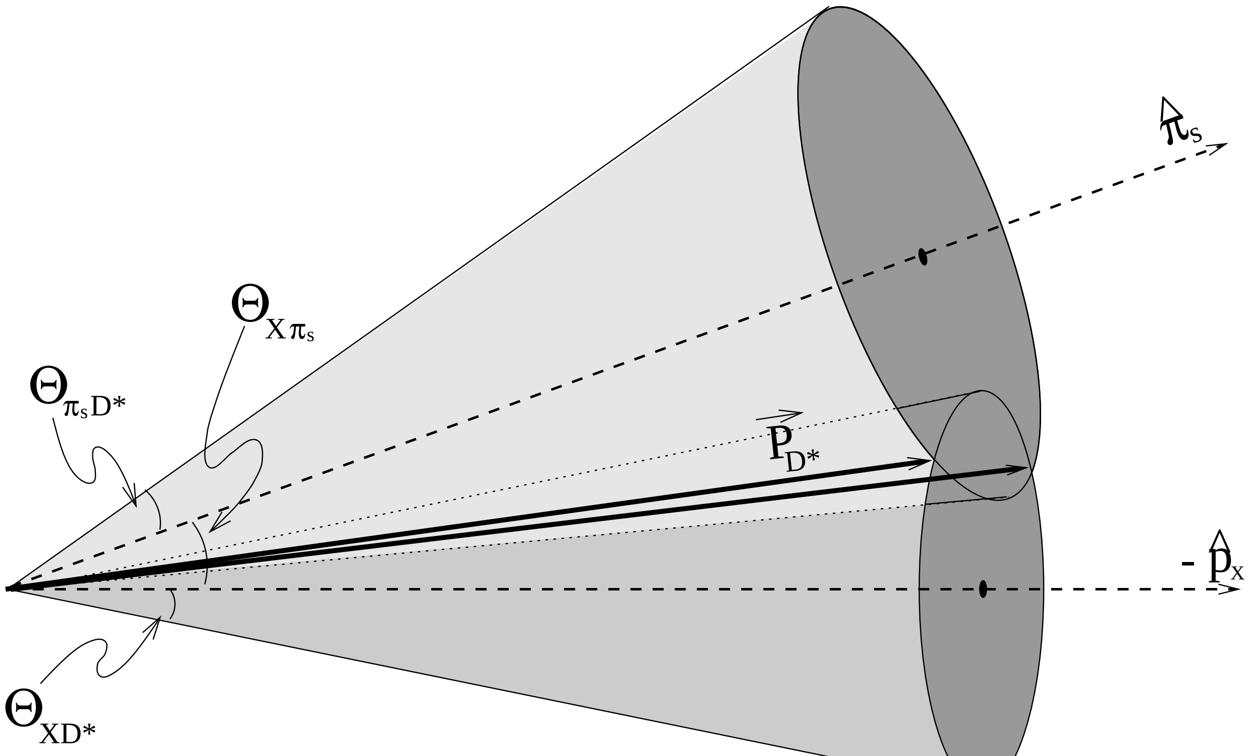

FIG. 3. The two cones which constrain the $D^{\star}$ direction as described in the text. The solution for $\vec{p}_{D^{\star}}$ must lie on both cones. There are then, in general, either two solutions as shown, or no solutions, in the case that the cones do not cross. The latter corresponds to the case $\theta_{X\pi} > \theta_{XD^{\star}} + \theta_{\pi D^{\star}}$.

C. Multiple Entries

This method of reconstruction can produce several solutions per event for the above variables. This multiplicity is from three sources:

1. The indistinguishability of π_1 and π_2 means that either may be combined with the $D^{\star +}$ to form $M(D^{\star +}\pi^-)$.

2. Random tracks often succesfully substitute for one of the three pion candidates in a $B^- \to D_J^0\, \pi^-$ signal event.

3. The quadratic ambiguity produces two solutions per three-π combination.

To simplify the statistical treatment of the data, we eliminate these respective sources of multiple entries in the following ways:

1. Since the modes we are most interested in $(B^- \to D_J^0\, \pi^-)$ produce π_1 and π_2 in very distinct momentum regions, we only consider the lower mass $D^{\star +}\pi^-$ combination. The higher mass combination always has $M(D^{\star +}\pi^-) > 3\, GeV/c^2$. This is well outside the signal region for the narrow D_J so that no signal is lost.

2. Entries where a fake track successfully substitutes for a real π in a signal event are completely dominated by fake π_2 from the 'other B'. The momentum spectrum of these random fakes is roughly the same as the charged particle spectrum at the $\Upsilon(4S)$: They are mostly low momentum tracks (65 – 300 MeV/c). For this reason, we plot only the three–π combination in each event with the highest p_{π_2}. This enhances the ratio of right/wrong combinations in the signal by $\approx 50\%$ over selecting a combination randomly for each event.

3. The two quadratic solutions for $\vec{p}_B$ will generally be at different angles with respect to the beamline. Because B production at the $\Upsilon(4S)$ has a distribution of $(dN/d\cos\theta_B = \sin^2\theta_B)$ with respect to the beamline, we select the quadratic solution with smaller $|\cos\theta_B|$. This enhances the ratio of right/wrong solutions in the signal by $\approx 60\%$ over selecting one of the two solutions randomly, yielding a somewhat narrower distribution in $M(D^{\star +}\pi^-)$ for the narrow D_J.

V. BACKGROUNDS

Rejection of the large combinatoric backgrounds relies entirely upon the physicality constraints imposed by the reconstruction method. These constraints enforce a certain kinematic relationship among the candidate tracks. Specifically, it must be possible for the three tracks to have come from the decay chain we seek. This fact is key in understanding the specific sources of these backgrounds, leading us to consider backgrounds of two types. This first type we call *random* background.

Random background is the combinatoric background produced by random tracks in the detector. The tracks that are successfully reconstructed to create the candidate decay chain have no special kinematic relationship to one another. They only happen to satisfy the physicality requirement by 'random chance'. Background of this type is distributed across the the three-body phase space for $B^- \to D^{\star+}\pi^-\pi^-$, and is smooth in all of our fitting variables compared to $B^- \to D_1(2420)^0\pi^-$ and $B^- \to D_2^\star(2460)^0\pi^-$. The background from continuum is obviously of this type since the constraint for $B \to D^\star X$ has no meaning for these events. $B\overline{B}$ events also contain such 'random' combinations, along with backgrounds of a second kind which we refer to as *semi-correlated*.

In $B\overline{B}$ events there are decays which mimic the kinematics of the signal mode closely enough to enter into the signal plots with relatively large efficiency. Monte Carlo studies of ≈ 20 of these potential feedthrough modes reveal a pattern:

- Feedthrough modes that are kinematically 'similar' to the signal modes have higher feedthrough rates.
- Feedthrough modes that are kinematically 'similar' to the signal modes tend to peak more strongly in $M(D^{\star+}\pi^-)$ and the decay angular distributions.

These feedthroughs are dominated by events of the type $B \to D^\star X\pi$, $D^\star \to D\pi_{\rm slow}$ where X consists of either zero or one light particle. In these events, the B decay provides a pion in the correct momentum region for a π_1 (or π_2) candidate, while the $D^\star$ decay provides a slow pion traveling in the proper direction for a $\pi_{\rm slow}$ candidate. The π_2 is usually provided by the decay of the other B in the event. Only the π_1 and π_3 are 'correlated', while the π_2 is 'random'. This is closely related to the occurrence of fake π_2 combinations in signal events that tend to create multiple entries.

Similarly, these random π_2 have the momentum distribution of charged tracks at the $\Upsilon(4S)$, and are dominated by low momentum tracks. This creates the largest feedthrough efficiencies for modes where the reconstruction requires a low momentum π_2 candidate. These modes are characterized by having a relatively fast π_1 (> 2 GeV/c).

The specific modes which are found to dominate these feedthroughs are:

1. $\overline{B^0} \to D_J^+\pi^-$, $D_J^+ \to D^{\star+}\pi^0$, $D^{\star+} \to D^0\pi^+$
2. $\overline{B^0} \to D^{\star+}\pi^-$, $D^{\star+} \to D^0\pi^+$
3. $\overline{B^0} \to D^{\star+}\rho^-$, $D^{\star+} \to D^0\pi^+$

Only the first two have a shape significantly different than that of 'random' background in our reconstructed quantities. However, all three have relatively high feedthrough rates (see Table III), and special care must be taken to account for their effects.

There is also some contribution from modes of the type $B \to DX\pi$, where the $\pi_{\rm slow}$ and π_2 are both 'random' but the feedthrough efficiencies are down by a factor of 10 from their cousins with a $D^\star$.

In addition to the above backgrounds, there are two sources of *real* $B^- \to D^{\star+}\pi^-\pi^-$ decays which we shall consider as backgrounds for the decays $B^- \to D_1(2420)^0\pi^-$ and $B^- \to D_2^\star(2460)^0\pi^-$. The first of these is the allowed decay $B^- \to D_1^0(j=1/2)\pi^-$, $D_1^0(j=1/2) \to D^{\star+}\pi^-$. The second is the non-resonant decay $B^- \to D^{\star+}\pi^-\pi^-$. In the future we hope to fit these signals by using all of the decay angular distributions to best constrain their contributions.

TABLE III. The feedthrough efficiencies and number of events expected in our data sample after all cuts. Also shown is the correlation coefficient between each background and each signal shape from the fit described in the next section. The values for $B\overline{B}$ background as a whole are shown for comparison. The yield for $\overline{B^0} \to D_J^+ \pi^-$ is taken from the fit.

			Correlation w/ $B^- \to D_J^0 \pi^-$	
Background	ϵ	$N_{expected}$	$D_1(2420)^0$	$D_2^\star(2460)^0$
$B\overline{B}$ generic	2.4×10^{-3}	—	0.043	-0.033
$\overline{B^0} \to D_J^+ \pi^-$	0.157	370±113	-0.276	-0.182
$\overline{B^0} \to D^{\star +} \pi^-$	5.9×10^{-2}	307±26	0.142	0.157
$\overline{B^0} \to D^{\star +} \rho^-$	4.2×10^{-2}	532±109	-0.002	0.078

VI. RESULTS

A. Two-dimensional Fit — $M(D^{\star +}\pi^-) vs. \cos\theta_3$

To extract the quantities of interest we perform unbinned maximum likelihood fits using Monte Carlo distributions [8]. We use this method to perform a two-dimensional fit in the plane of the variables $M(D^{\star +}\pi^-)$ and $\cos\theta_3$. The background components are modeled in the following ways:

- *Continuum.* To model the shape of the continuum background, we use the off-resonance data. Since variations in beam energy within each 'dataset' are very small, we model the beam energy for the off-resonance data by using the average on-resonance beam energy for each dataset. Carefully modeling the beam energy allows us to fix the normalization of this contribution using the relative luminosity of the off-resonance data.

- *$B\overline{B}$.* To model background from other B decay modes, we use a large sample of generic $B\overline{B}$ Monte Carlo events to model this shape and allow the normalization to vary in the fit.

- *Semi-correlated feedthroughs.* In order to properly account for the semi-correlated feedthroughs, separate Monte Carlo samples of these are generated. We add these shapes to our fit and allow their normalizations to float. In cases where measurements of these modes exist, we check that the yields we extract for these fractions agree with previous measurements.

- $B^- \to D_1^0(j = 1/2)\,\pi^-$. To model the effect of $B^- \to D_1^0(j = 1/2)\,\pi^-$ decays, we generate a large Monte Carlo sample of $B^- \to D^{\star +}\pi^-\pi^-$. This sample is generated flat in phase space as well as in all angular distributions. These events are then re-weighted by the fitter, allowing the mass, width and normalization of this component to vary assuming an S-wave distribution for the $D_1^0(j = 1/2) \to D^{\star +}\pi^-$ decays.

TABLE IV. Signal yields and values for B for the two fits described in the text. The $D_1(2420)^0$ is modeled in the fit using the distribution $(1 + B\cos^2\theta_3)$, allowing variation of the S-wave to D-wave ratio in $D_1(2420)^0$ decays.

	Signal Yields		
Fit	$B^- \to D_1(2420)^0\,\pi^-$	$B^- \to D_2^\star(2460)^0\,\pi^-$	B
B floating	268±67	170±62	$4.17^{+9.40}_{-2.60}$
B fixed	281±56	165±61	2.74

- *Non-resonant* $B^- \to D^{\star +}\,\pi^-\,\pi^-$. To model the effect of non-resonant contributions to the signal, we generate a Monte Carlo sample flat in phase space and decay angles and include it in our fit with floating normalization.

The signal components are modeled in the following ways:

- $B^- \to D_1(2420)^0\,\pi^-$. To model the shape of the narrow 1^+ state, we generate $B^- \to D_1(2420)^0\,\pi^-$ Monte Carlo with a flat distribution in $\cos\theta_3$. We then re-weight this distribution by $(1 + B\cos^2\theta_3)$ and allow B to vary.

- $B^- \to D_2^\star(2460)^0\,\pi^-$. To model the shape of the narrow 2^+ state, we generate $B^- \to D_2^\star(2460)^0\,\pi^-$ Monte Carlo with a pure $\sin^2\theta_3$ distribution. This constraint comes from conservation of parity and angular momentum.

We generate large Monte Carlo samples for each component of the fit that are flat in the parameters we wish to vary. The contribution from each fit component at a given point in the data is computed by counting the number of events in a neighborhood of the two-dimensional space about that point [8]. The *reconstructed* Monte Carlo events in those neighborhoods are then re-weighted with each pass of the fitter using the *generated* quantities for each Monte Carlo event. This allows us to build up a Monte Carlo distribution corresponding to any set of generated parameters with each pass of the fitter, enabling us to fit for those parameters. This method also properly accounts for the effects of smearing and acceptance in the data.

When using this fitting method, one typically extracts only the parameters that define the shape of the distribution, and the total yield of the fit shape is fixed to the data yield. However, since our yield is a quantity of interest it is desirable to find a method which accounts for this properly. For this reason, we use the method of extended maximum likelihood [9] to define our likelihood function. This method allows the total yield in the fit shape to vary.

A fit was performed, allowing the parameter B to vary. This fit gives $B = 4.17^{+9.40}_{-2.60}$. Since these errors are larger than those from the previous CLEO measurement of this quantity, we re-fit using the best measured value (see Table II) . The results of both fits are summarized in Table IV. We shall use the fit with 'B' fixed to determine the branching fractions, and all further discussion will pertain to this fit.

Fig. 4 shows a projection of the fit onto the $M(D^{\star+}\pi^-)$ axis. Fig. 5 shows projections of the fit onto the $M(D^{\star+}\pi^-)$ axis in two regions of $\cos\theta_3$, showing the separation between the two D_J states in this variable.

Fig. 6 shows projections of the fit onto the $\cos\theta_3$ axis for sidebands above and below the signal region. This shows the agreement between the background shape and the data in these regions adjacent to the signal. Fig. 7 shows projections of the fit onto the $\cos\theta_3$ axis for two bands within the signal region. This shows the agreement between the angular distribution of the fit and the data in the regions populated by the $D_1(2420)^0$ and the $D_2^\star(2460)^0$.

The fitted values of the semi-correlated feedthroughs have large errors and agree well with previous measurements.

B. Systematic Errors

Several sources of systematic error must be considered. Estimates of these errors are summarized in Table V.

The error in event selection acceptance is dominated by the cut on event shape. We estimate this error by varying the R_2 cut and re-fitting. From this we estimate this error at 2.9% for the $D_1(2420)^0$ and 4.6% for the $D_2^\star(2460)^0$.

The errors in efficiency for the the three pions are correlated and must be added linearly. We estimate a 1% uncertainty in efficiency for the π_1. For the slower π_2 and low-momentum π_3 we estimate 2% and 5% errors respectively. The pion consistency cuts also introduce errors in efficiency which add linearly. We estimate the error due to these cuts at 1% for the π_2 and π_3 and 2% for the faster π_1 due to uncertainties in the muon veto efficiency for this track.

Mis-estimation of the beam energy changes the physicality constraints imposed by the reconstruction method. This produces an error in the signal efficiency and the relative normalization of the off-resonance data. We estimate the effect on signal efficiency by variation of the beam energy within known errors. We estimate the effect on continuum normalization by varying the beam energy for continuum data within the extremes of the on-resonance energy for each dataset. The combined errors from these sources are 3.1% and 3.5% for the $D_1(2420)^0$ and $D_2^\star(2460)^0$ respectively.

Uncertainty in the overall background shape is dominated by the errors in the normalization of each background fraction. For the backgrounds that we allow to float in our fit, the systematic errors due to these errors in normalization are accounted for in the fit errors (see Table IV). This includes the $B^-\to D_1^0(j=1/2)\,\pi^-$ and non-resonant $B^-\to D^{\star+}\,\pi^-\,\pi^-$ as well as the semi-correlated feedthroughs; $\overline{B^0}\to D^{\star+}\,\pi^-$ and $\overline{B^0}\to D^{\star+}\,\rho^-$; and the random background from $B\overline{B}$ events. The error on the fixed continuum normalization in the fit is dominated by modeling of the on-resonance beam energy for the off-resonance data as previously discussed.

Systematic error due to uncertainty in the signal shapes are estimated by variation of these parameters within the errors from previous measurements. For variation of masses and width, the world average [10] of these values is used. For the parameter 'B', the CLEO value (see Table II) was used. We estimate the error from interference between $D_1(2420)^0$ decays and $D_1^0(j=1/2)$ decays by assuming complete constructive or destructive interfer-

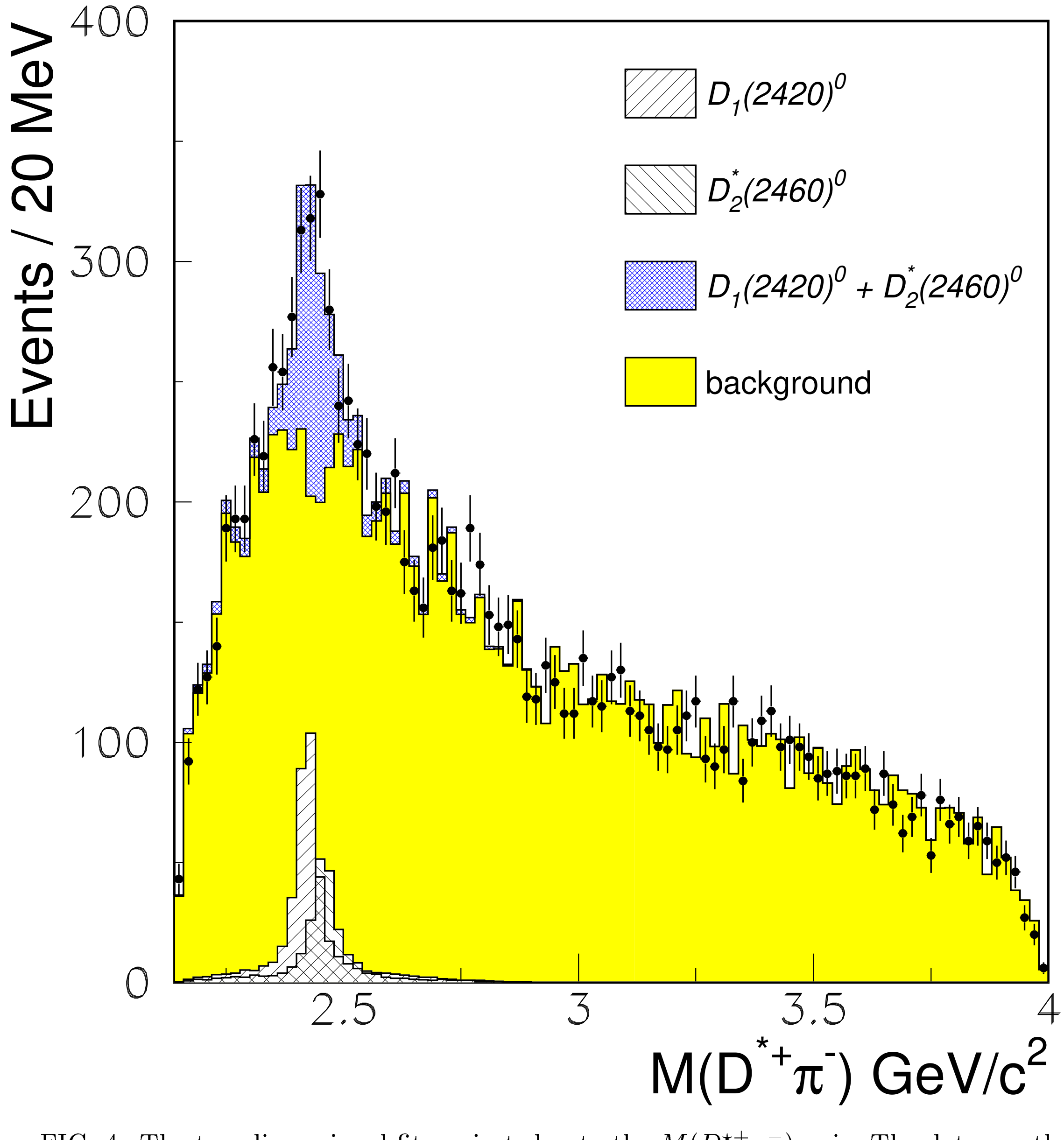

FIG. 4. The two-dimensional fit projected onto the $M(D^{\star+}\pi^-)$ axis. The data are the plotted points with errors. The long, flat tails in the signal shapes result from combinations where a random track has faked the real π_2.

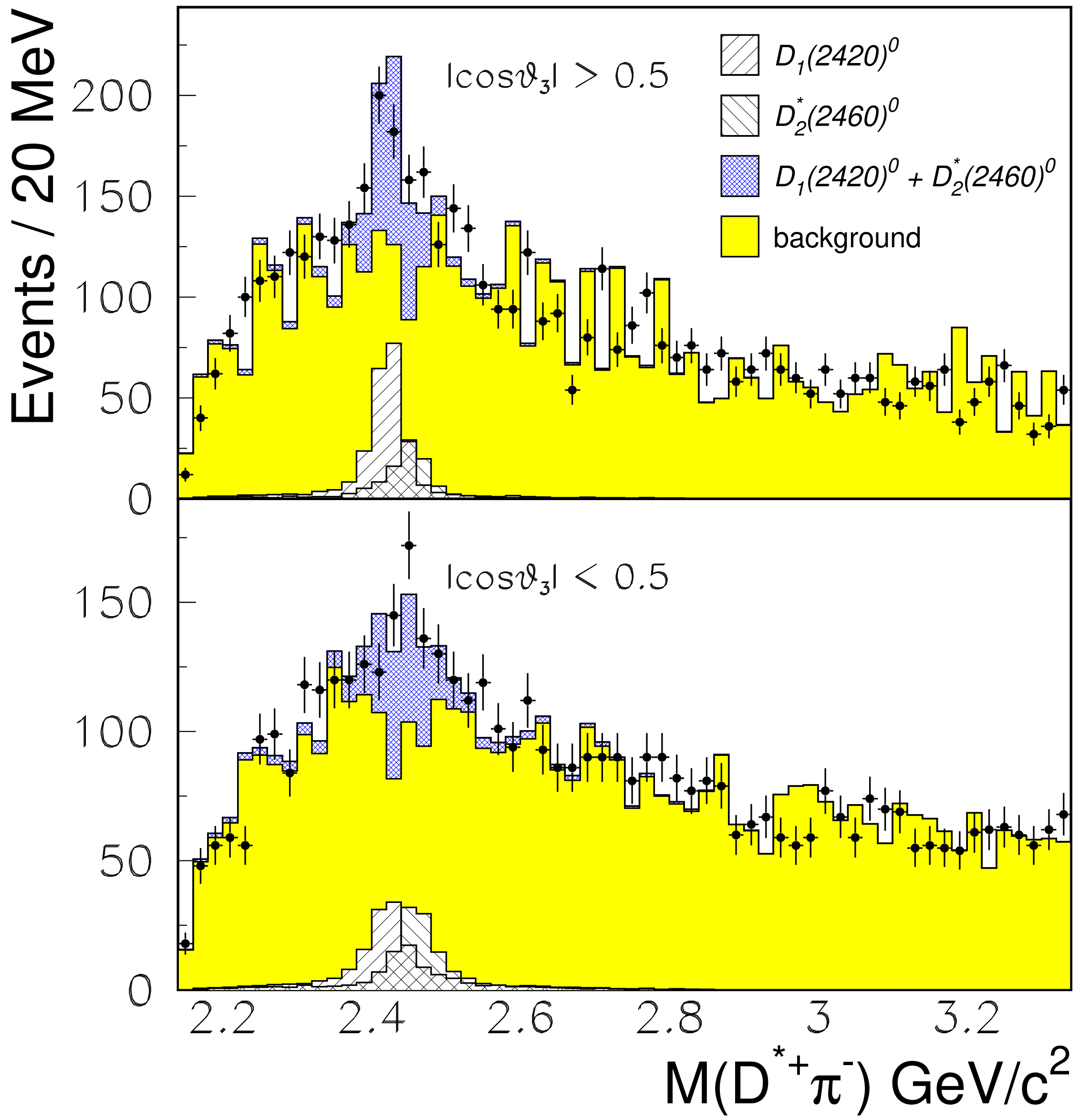

FIG. 5. The two-dimensional fit projected onto the $M(D^{\star +}\pi^-)$ axis for $|\cos\theta_3| > 0.5$ and $|\cos\theta_3| < 0.5$. The data are the plotted points with errors. The long, flat tails in the signal shapes result from combinations where a random track has faked the real π_2.

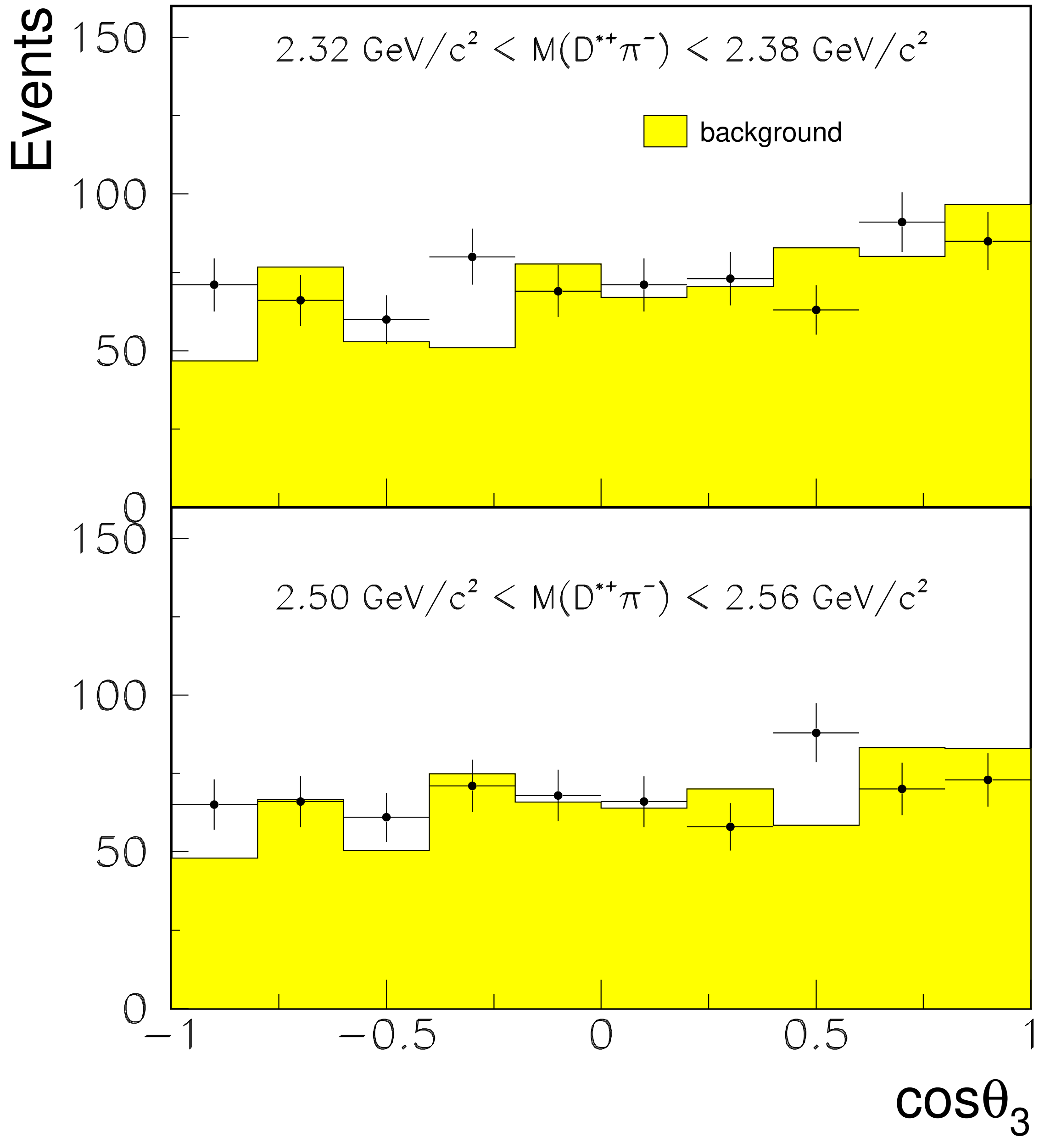

FIG. 6. The two-dimensional fit projected onto the $\cos\theta_3$ axis for two sidebands around the signal region. The data are the plotted points with errors.

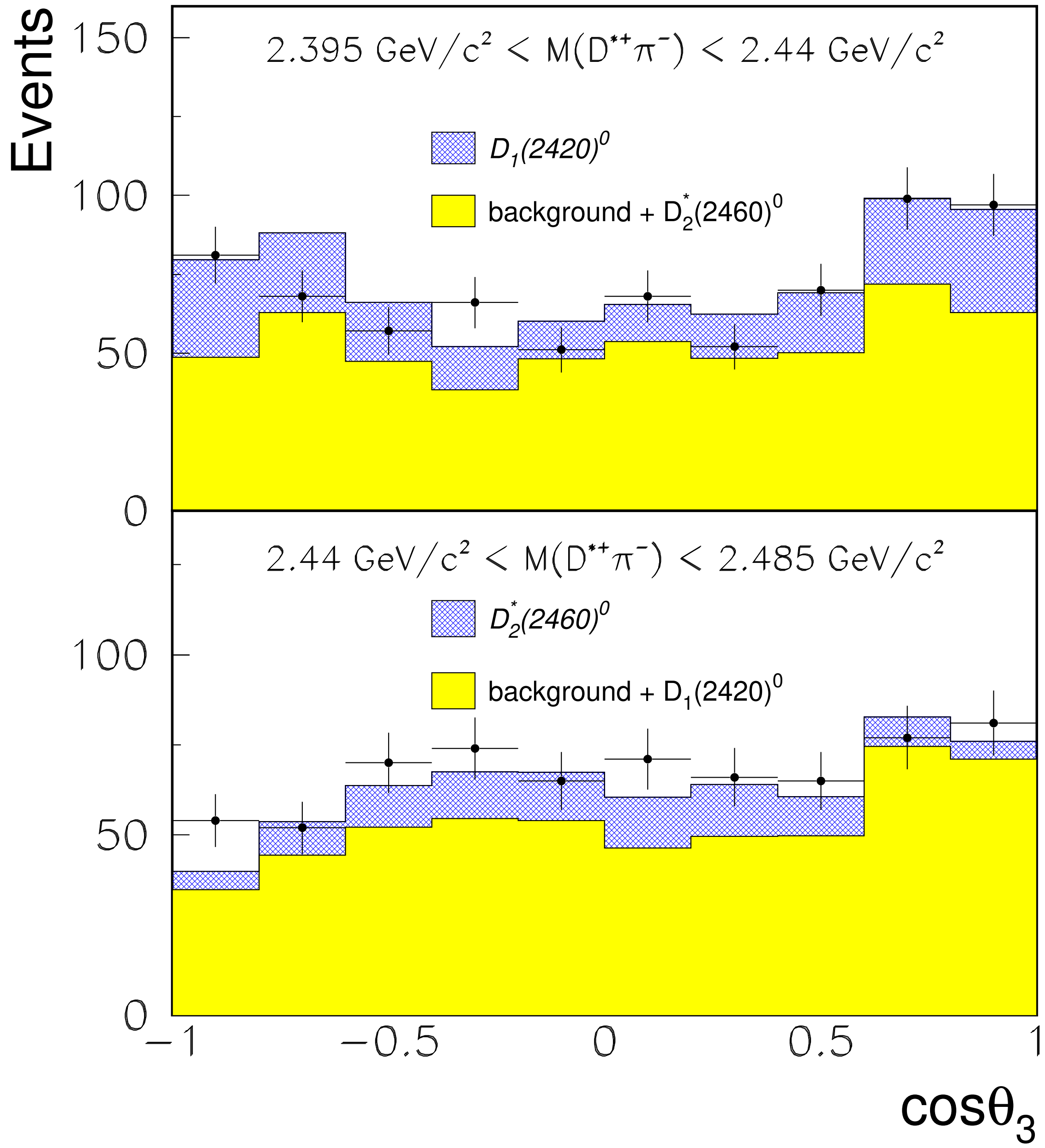

FIG. 7. The two-dimensional fit projected onto the $\cos\theta_3$ axis for two bands inside the signal region. The lower-mass band contains mostly $D_1(2420)^0$ while the upper-mass band contains a larger contribution from $D_2^{\star}(2460)^0$. The data are the plotted points with errors.

TABLE V. Summary of systematic errors (%).

Source	$B^- \to D_1(2420)^0\,\pi^-$	$B^- \to D_2^\star(2460)^0\,\pi^-$
event selection	2.9	4.6
track reconstruction & selection	8	8
π consistency	4	4
partial reconstruction efficiency	3.1	3.5
signal shape	9.4	8.0
Fit	5.7	6.2
$N_{B\overline{B}}$	1.8	1.8
Total	14.9	14.8

ence between the two 1^+ states reduced by a factor from our relatively flat acceptance in all decay angular distrubutions. We estimate the error due to the S-wave assumption for $D_1^0(j=1/2) \to D^{\star +}\,\pi^-$ decays and the effects of interference on the $D_1(2420)^0$ *shape* as the difference between the 'B-fixed' and 'B-floating' fits. These combined errors are 9.4% for the $D_1(2420)^0$ and 8.0% for the $D_2^\star(2460)^0$.

To estimate the error from the fitting method, we study the variation of the most likely fit parameters as a function of the size of the neigborhoods used when calculating the density function for each fit component. We vary the neighborhoods by a factor of four in area and take the errors as the resulting variation in the signal yields. This contributes a 5.7% error for the $D_1(2420)^0$ and a 6.2% error for the $D_2^\star(2460)^0$.

VII. CONCLUSION

When these yields and errors are converted into branching ratios, we get the results:

$$\mathcal{B}\left(B^- \to D_1(2420)^0\,\pi^-\right) \times \mathcal{B}\left(D_1(2420)^0 \to D^{\star +}\,\pi^-\right) \times \mathcal{B}\left(D^{\star +} \to D^0\,\pi^+\right) = (5.3\pm1.1\pm0.8) \times 10^{-4}$$

$$\mathcal{B}\left(B^- \to D_2^\star(2460)^0\,\pi^-\right) \times \mathcal{B}\left(D_2^\star(2460)^0 \to D^{\star +}\,\pi^-\right) \times \mathcal{B}\left(D^{\star +} \to D^0\,\pi^+\right) = (2.9\pm1.1\pm0.4) \times 10^{-4}$$

We assume here that the $\Upsilon(4S)$ decays equally to B^+B^- and $B^0\overline{B^0}$. Using the previously measured CLEO value of the $D^{\star +} \to D^0\,\pi^+$ [11] branching fraction gives:

$$\mathcal{B}\left(B^- \to D_1(2420)^0\,\pi^-\right) \times \mathcal{B}\left(D_1(2420)^0 \to D^{\star +}\,\pi^-\right) = (7.8\pm1.6\pm1.0\pm0.2) \times 10^{-4}$$
$$\mathcal{B}\left(B^- \to D_2^\star(2460)^0\,\pi^-\right) \times \mathcal{B}\left(D_2^\star(2460)^0 \to D^{\star +}\,\pi^-\right) = (4.2\pm1.6\pm0.6\pm0.1) \times 10^{-4}$$

The last error is from the $D^{\star +} \to D^0\,\pi^+$ branching fraction.

Using isospin conservation and neglecting decay modes to other than $D^\star$ and a single π, one arrives at $\mathcal{B}\,(D_1(2420)^0 \to D^{\star +}\,\pi^-) = 2/3$. Measurements of $D_2^\star(2460)^0 \to D^+\pi^-$ [10] coupled with isospin conservation and neglect of decay modes with multiple pions gives $\mathcal{B}\,(D_2^\star(2460)^0 \to D^{\star +}\,\pi^-) = 0.20$. *Assuming* these values we calculate:

$$\mathcal{B}\left(B^- \to D_1(2420)^0\,\pi^-\right) = (1.17 \pm 0.24 \pm 0.16 \pm 0.03) \times 10^{-3}$$
$$\mathcal{B}\left(B^- \to D_2^\star(2460)^0\,\pi^-\right) = (2.1 \pm 0.8 \pm 0.3 \pm 0.05) \times 10^{-3}$$

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., and M.S. thank the A.P. Sloan Foundation, and A.W., and R.W. thank the Alexander von Humboldt Stiftung for support. 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] CLEO Collab., M.S. Alam *et al.*, Phys. Rev. D **50** (1994) 43.
[2] ARGUS Collab., H. Albrecht *et al.*, Phys. Lett. B **335** (1994) 526.
[3] N. Isgur and M. Wise, Phys. Lett. B **232** (1989) 113.
[4] A.F. Falk and M. Luke, Phys. Lett. B **292** (1992) 119.
[5] A.F. Falk and M.E. Peskin, Phys. Rev. D **49** (1994) 3320.
[6] CLEO Collab., P. Avery *et al.*, Phys. Lett. B **331** (1994) 236.
[7] CLEO Collab., Y. Kubota *et al.*, Nucl. Instr. and Meth. **A320** (1992) 66.
[8] D. Schmidt, R. Morrison and M. Witherell, Nucl. Instr. and Meth. **A328** (1993) 547.
[9] R. Barlow, Nucl. Instr. and Meth. **A297** (1990) 496.
[10] L. Montanet *et al.*, Phys. Rev. D **50** (1994) 1173 and 1995 off year partial update for the 1996 edition (URL: http://pdg.lbl.gov/).
[11] CLEO Collab., Butler *et al.*, Phys. Rev. Lett. **69** (1992) 2041.