CLNS 94/1285
CLEO 94-13

Measurement of the $\bar{B} \to D^* \ \ell \ \bar{\nu}$ Branching Fractions and $|V_{cb}|$

B. Barish,[1] M. Chadha,[1] S. Chan,[1] D.F. Cowen,[1] G. Eigen,[1] J.S. Miller,[1] C. O'Grady,[1] J. Urheim,[1] A.J. Weinstein,[1] D. Acosta,[2] M. Athanas,[2] G. Masek,[2] H.P. Paar,[2] J. Gronberg,[3] R. Kutschke,[3] S. Menary,[3] R.J. Morrison,[3] S. Nakanishi,[3] H.N. Nelson,[3] T.K. Nelson,[3] C. Qiao,[3] J.D. Richman,[3] A. Ryd,[3] H. Tajima,[3] D. Sperka,[3] M.S. Witherell,[3] M. Procario,[4] R. Balest,[5] K. Cho,[5] M. Daoudi,[5] W.T. Ford,[5] D.R. Johnson,[5] K. Lingel,[5] M. Lohner,[5] P. Rankin,[5] J.G. Smith,[5] J.P. Alexander,[6] C. Bebek,[6] K. Berkelman,[6] K. Bloom,[6] T.E. Browder,[6*] D.G. Cassel,[6] H.A. Cho,[6] D.M. Coffman,[6] D.S. Crowcroft,[6] P.S. Drell,[6] R. Ehrlich,[6] P. Gaidarev,[6] R.S. Galik,[6] M. Garcia-Sciveres,[6] B. Geiser,[6] B. Gittelman,[6] S.W. Gray,[6] D.L. Hartill,[6] B.K. Heltsley,[6] C.D. Jones,[6] S.L. Jones,[6] J. Kandaswamy,[6] N. Katayama,[6] P.C. Kim,[6] D.L. Kreinick,[6] G.S. Ludwig,[6] J. Masui,[6] J. Mevissen,[6] N.B. Mistry,[6] C.R. Ng,[6] E. Nordberg,[6] J.R. Patterson,[6] D. Peterson,[6] D. Riley,[6] S. Salman,[6] M. Sapper,[6] F. Würthwein,[6] P. Avery,[7] A. Freyberger,[7] J. Rodriguez,[7] S. Yang,[7] J. Yelton,[7] D. Cinabro,[8] S. Henderson,[8] T. Liu,[8] M. Saulnier,[8] R. Wilson,[8] H. Yamamoto,[8] T. Bergfeld,[9] B.I. Eisenstein,[9] G. Gollin,[9] B. Ong,[9] M. Palmer,[9] M. Selen,[9] J. J. Thaler,[9] K.W. Edwards,[10] M. Ogg,[10] A. Bellerive,[11] D.I. Britton,[11] E.R.F. Hyatt,[11] D.B. MacFarlane,[11] P.M. Patel,[11] B. Spaan,[11] A.J. Sadoff,[12] R. Ammar,[13] S. Ball,[13] P. Baringer,[13] A. Bean,[13] D. Besson,[13] D. Coppage,[13] N. Copty,[13] R. Davis,[13] N. Hancock,[13] M. Kelly,[13] S. Kotov,[13] I. Kravchenko,[13] N. Kwak,[13] H. Lam,[13] Y. Kubota,[14] M. Lattery,[14] M. Momayezi,[14] J.K. Nelson,[14] S. Patton,[14] D. Perticone,[14] R. Poling,[14] V. Savinov,[14] S. Schrenk,[14] R. Wang,[14] M.S. Alam,[15] I.J. Kim,[15] B. Nemati,[15] Z. Ling,[15] J.J. O'Neill,[15] H. Severini,[15] C.R. Sun,[15] F. Wappler,[15] G. Crawford,[16] C. M. Daubenmier,[16] R. Fulton,[16] D. Fujino,[16] K.K. Gan,[16] K. Honscheid,[16] H. Kagan,[16] R. Kass,[16] J. Lee,[16] R. Malchow,[16] Y. Skovpen,[16†] M. Sung,[16] C. White,[16] M.M. Zoeller,[16] F. Butler,[17] X. Fu,[17] G. Kalbfleisch,[17] W.R. Ross,[17] P. Skubic,[17] M. Wood,[17] J.Fast ,[18] R.L. McIlwain,[18] T. Miao,[18] D.H. Miller,[18] M. Modesitt,[18] D. Payne,[18] E.I. Shibata,[18] I.P.J. Shipsey,[18] P.N. Wang,[18] M. Battle,[19] J. Ernst,[19] L. Gibbons,[19] Y. Kwon,[19] S. Roberts,[19] E.H. Thorndike,[19] C.H. Wang,[19] J. Dominick,[20] M. Lambrecht,[20] S. Sanghera,[20] V. Shelkov,[20] T. Skwarnicki,[20] R. Stroynowski,[20] I. Volobouev,[20] G. Wei,[20] P. Zadorozhny,[20] M. Artuso,[21] M. Goldberg,[21] D. He,[21] N. Horwitz,[21] R. Kennett,[21] R. Mountain,[21] G.C. Moneti,[21] F. Muheim,[21] Y. Mukhin,[21] S. Playfer,[21] Y. Rozen,[21] S. Stone,[21] M. Thulasidas,[21] G. Vasseur,[21] X. Xing,[21] G. Zhu,[21] J. Bartelt,[22] S.E. Csorna,[22] Z. Egyed,[22] V. Jain,[22] D. Gibaut,[23] and K. Kinoshita[23]

(CLEO Collaboration)

[1]*California Institute of Technology, Pasadena, California 91125*
[2]*University of California, San Diego, La Jolla, California 92093*
[3]*University of California, Santa Barbara, California 93106*
[4]*Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213*
[5]*University of Colorado, Boulder, Colorado 80309-0390*
[6]*Cornell University, Ithaca, New York 14853*
[7]*University of Florida, Gainesville, Florida 32611*
[8]*Harvard University, Cambridge, Massachusetts 02138*
[9]*University of Illinois, Champaign-Urbana, Illinois, 61801*
[10]*Carleton University, Ottawa, Ontario K1S 5B6 and the Institute of Particle Physics, Canada*
[11]*McGill University, Montréal, Québec H3A 2T8 and the Institute of Particle Physics, Canada*
[12]*Ithaca College, Ithaca, New York 14850*
[13]*University of Kansas, Lawrence, Kansas 66045*
[14]*University of Minnesota, Minneapolis, Minnesota 55455*
[15]*State University of New York at Albany, Albany, New York 12222*
[16]*Ohio State University, Columbus, Ohio, 43210*
[17]*University of Oklahoma, Norman, Oklahoma 73019*
[18]*Purdue University, West Lafayette, Indiana 47907*
[19]*University of Rochester, Rochester, New York 14627*
[20]*Southern Methodist University, Dallas, Texas 75275*
[21]*Syracuse University, Syracuse, New York 13244*
[22]*Vanderbilt University, Nashville, Tennessee 37235*
[23]*Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 24061*

(October 5, 1994)

Abstract

We study the exclusive semileptonic B meson decays $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ and $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ using data collected with the CLEO II detector at the Cornell Electron-positron Storage Ring (CESR). We present measurements of the branching fractions, $\mathcal{B}(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}) = (0.5/f_{00})[4.49 \pm 0.32(\text{stat.}) \pm 0.39(\text{sys.})]\%$ and $\mathcal{B}(B^- \to D^{*0}\ \ell^-\ \bar{\nu}) = (0.5/f_{+-})[5.13 \pm 0.54(\text{stat.}) \pm 0.64(\text{sys.})]\%$, where f_{00} and f_{+-} are the neutral and charged B meson production fractions at the $\Upsilon(4S)$ resonance, respectively. Assuming isospin invariance and taking the ratio of charged to neutral B meson lifetimes measured at higher energy machines, we determine the ratio $f_{+-}/f_{00} = 1.04 \pm 0.13(\text{stat.}) \pm 0.12(\text{sys.}) \pm 0.10(\text{lifetime})$; further assuming $f_{+-} + f_{00} = 1$ we also determine the partial width $\Gamma(\bar{B} \to D^*\ \ell\ \bar{\nu}) =$

---

*Permanent address: University of Hawaii at Manoa

†Permanent address: INP, Novosibirsk, Russia

$[29.9 \pm 1.9(\mathrm{stat.}) \pm 2.7(\mathrm{sys.}) \pm 2.0(\mathrm{lifetime})]\ \mathrm{ns}^{-1}$ (independent of f_{+-}/f_{00}). From this partial width we calculate $\bar{B} \to D^* \ \ell \ \bar{\nu}$ branching fractions that do not depend on f_{+-}/f_{00} nor the individual B lifetimes, but only on the charged to neutral B lifetime ratio. The product of the CKM matrix element $|V_{cb}|$ times the normalization of the decay form factor at the point of no recoil of the D^* meson, $\mathcal{F}(y = 1)$, is determined from a linear fit to the combined differential decay rate of the exclusive $\bar{B} \to D^* \ \ell \ \bar{\nu}$ decays: $|V_{cb}|\mathcal{F}(1) = 0.0351 \pm 0.0019(\mathrm{stat.}) \pm 0.0018(\mathrm{sys.}) \pm 0.0008(\mathrm{lifetime})$. The value for $|V_{cb}|$ is extracted using theoretical calculations of the form factor normalization.

Typeset using REVTEX

I. INTRODUCTION

In the framework of the Standard Model of weak interactions the elements of the 3×3 Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix [1] must be determined experimentally. The element $|V_{cb}|$ is determined from studies of the semileptonic decays of B mesons. Measurements of $|V_{cb}|$ from the inclusive semileptonic rate [2–4] and from exclusive rates [5–10] are systematically limited by model dependence in the theoretical prediction of the decay rate. The recent development of Heavy Quark Effective Theory (HQET) [11] yields an expression for the $\bar{B} \to D^* \ \ell \ \bar{\nu}$ [12] decay rate in terms of a single unknown form factor [13] which, at the point of no recoil of the D^* meson, is absolutely normalized up to corrections of order $1/m_Q^2$ [14] (where m_Q is the b or c quark mass). It is currently believed that these corrections can be calculated with less than 5% uncertainty [15,16], which would permit a precise determination of $|V_{cb}|$ from the study of $\bar{B} \to D^* \ \ell \ \bar{\nu}$ as a function of the recoil of the D^* meson. The decay mode $\bar{B} \to D^* \ \ell \ \bar{\nu}$ is also preferred over other exclusive channels because D^* meson decays have a very clean experimental signature.

Throughout this paper the square of the four momentum transfer in $\bar{B} \to D^* \ \ell \ \bar{\nu}$ decays is denoted by $q^2 = M_{l\bar{\nu}}^2$ (where $M_{l\bar{\nu}}$ is the mass of the virtual W). The kinematic variable of HQET, which is a measure of the recoil of the D^* meson, is given by

$$y \equiv v \cdot v' = \frac{m_B^2 + m_{D^*}^2 - q^2}{2 m_B m_{D^*}}, \tag{1}$$

where v and v' are the four velocities of the B and D^* mesons and m_B and m_{D^*} are their respective masses.

We report on new measurements of the branching fractions and differential decay rate for the decays, $\bar{B}^0 \to D^{*+} \ \ell^- \ \bar{\nu}$ and $B^- \to D^{*0} \ \ell^- \ \bar{\nu}$ [17]. The product of the CKM matrix element $|V_{cb}|$ times the form factor normalization $\mathcal{F}(1)$ is determined from fits to the differential decay rate, and the model dependent estimates of $|V_{cb}|$ are extracted from the integrated rate. Using isospin invariance to equate the partial widths of $\bar{B}^0 \to D^{*+} \ \ell^- \ \bar{\nu}$ and $B^- \to D^{*0} \ \ell^- \ \bar{\nu}$, and recent B meson lifetime measurements, we also measure the $\Upsilon(4S)$ branching fractions,

$$f_{00} \equiv \mathcal{B}(\Upsilon(4S) \to B^0 \bar{B}^0) \tag{2}$$

$$f_{+-} \equiv \mathcal{B}(\Upsilon(4S) \to B^- B^+). \tag{3}$$

All measurements of exclusive B branching fractions at the $\Upsilon(4S)$ resonance currently assume equal production of charged and neutral B mesons, so the measurement of these production fractions affects both hadronic and semileptonic B branching fractions. A model dependent result for the inclusive branching fraction $\mathcal{B}(\bar{B} \to D^* \ X \ \ell \ \bar{\nu})$, where X is any possible hadronic state ($X \neq 0$), is also presented.

The paper is structured in the following way: in Section II the technique for obtaining yields of $\bar{B}^0 \to D^{*+} \ \ell^- \ \bar{\nu}$ and $B^- \to D^{*0} \ \ell^- \ \bar{\nu}$ events is discussed in general. Details of the candidate selection are then given in Section III. Backgrounds in the $\bar{B} \to D^* \ \ell \ \bar{\nu}$ samples and how their magnitudes are estimated are discussed in Section IV. Systematic studies of the reconstruction efficiencies are described in section V. Branching fraction results are presented in Section VI, followed by measurements of $|V_{cb}|$ in Section VII, and conclusions in Section VIII.

II. METHOD

The $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ and $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ efficiency-corrected yields, denoted by N_0 and N_- respectively, depend on the total number of $\Upsilon(4S)$ decays in the data sample, $N_{\Upsilon(4S)}$, and a product of branching fractions,

$$N_0 = 4N_{\Upsilon(4S)} f_{00} \mathcal{B}(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}) \mathcal{B}_{D^{*+}} \mathcal{B}_{D^0} \tag{4}$$

$$N_- = 4N_{\Upsilon(4S)} f_{+-} \mathcal{B}(B^- \to D^{*0}\ \ell^-\ \bar{\nu}) \mathcal{B}_{D^{*0}} \mathcal{B}_{D^0}, \tag{5}$$

where $\mathcal{B}_{D^{*+}} \equiv \mathcal{B}(D^{*+} \to D^0\pi^+)$, $\mathcal{B}_{D^{*0}} \equiv \mathcal{B}(D^{*0} \to D^0\pi^0)$ and $\mathcal{B}_{D^0} \equiv \mathcal{B}(D^0 \to K^-\pi^+)$. These three branching fractions have been measured by CLEO II [18,19] with D^* samples that are statistically independent of the sample considered here. The factors of 4 enter because each $\Upsilon(4S)$ decay produces a $B\bar{B}$ meson pair and because we are combining the e and μ lepton species.

We search for the decays $\bar{B} \to D^*\ \ell\ \bar{\nu}$ by combining reconstructed D^* mesons with "right sign" lepton candidates in the same event. By "right sign" we mean that a D^{*+} (D^{*0}) must be paired with a negative lepton, while a D^{*-} ($\bar{D}^{*0}$) requires a positive lepton. D^* mesons are reconstructed using the decay chains $D^{*+} \to D^0\pi^+$, $D^{*0} \to D^0\pi^0$ and $D^0 \to K^-\pi^+$ [20]. This technique uses our knowledge of the B meson momentum, $|\mathbf{p}_B|$. The energy of B mesons produced in symmetric e^+e^- annihilations, E_B, must equal the beam energy, which is precisely known from machine optics; hence $|\mathbf{p}_B|$ can be determined from E_B and the known B mass [21]. A kinematic constraint is obtained by writing the invariant mass of the emitted neutrino as

$$p_\nu^2 = (p_B - p_{D^*} - p_\ell)^2 \tag{6}$$

where p stands for the 4-vector of the particle in subscript. Expanding this equation results in

$$p_\nu^2 = (E_B - E_{D^*\ell})^2 - |\mathbf{p}_B|^{\ 2} - |\mathbf{p}_{D^*\ell}|^{\ 2} + 2|\mathbf{p}_B||\mathbf{p}_{D^*\ell}| cos\Theta \tag{7}$$

where $(E_B, \mathbf{p}_B)$ is the B meson 4-momentum, $(E_{D^*\ell}, \mathbf{p}_{D^*\ell})$ is the sum of the D^* and lepton 4-momenta, and Θ is the angle between the 3-momenta $\mathbf{p}_{D^*\ell}$ and $\mathbf{p}_B$. The first three terms on the right hand side constitute what is traditionally referred to as missing mass squared, symbolized MM^2. The factor multiplying $\cos\Theta$ will be denoted by C for *cosine multiplier*:

$$MM^2 \equiv (E_B - E_{D^*\ell})^2 - |\mathbf{p}_B|^{\ 2} - |\mathbf{p}_{D^*\ell}|^{\ 2}, \tag{8}$$

$$C \equiv 2|\mathbf{p}_B||\mathbf{p}_{D^*\ell}|. \tag{9}$$

$E_{D^*\ell}$ and $\mathbf{p}_{D^*\ell}$ are determined from the measured momenta of the lepton and D^* candidates. Since we know the magnitude of $\mathbf{p}_B$, but not its direction, $\cos\Theta$ is the only unknown.

For each D^* and lepton combination in the same event the pair of variables C and MM^2 is calculated. For correctly reconstructed $\bar{B} \to D^*\ \ell\ \bar{\nu}$ decays (with perfect detector resolution), the values of C and MM^2 must lie within the kinematic boundary determined by Eq. (7) with $p_\nu^2 = 0$. This boundary is shown in Fig. 1 for the lepton momentum range $1.4 \leq |\mathbf{p}_\ell| \leq 2.4$ GeV. This lepton momentum range and kinematic boundary define the Signal Region [22]. To arrive at the number of true $\bar{B} \to D^*\ \ell\ \bar{\nu}$ decays in the data sample, we count the number of candidates observed in the Signal Region, subtract the expected number of background candidates which happen to fall inside the Signal Region, and divide by the Monte Carlo efficiency for signal events. Because the Signal Region spans a significant area of phase space, we cannot assume that backgrounds vary slowly in and near it, and a reliable estimate of the background inside the Signal Region cannot be obtained by interpolating from the number of candidates observed outside. Instead, we categorize all sources of background and estimate the total contribution from each source in the Signal Region by using data. The method is insensitive to the detailed C vs. MM^2 distribution of the signals and of all backgrounds except $\bar{B} \to D^* \; X \; \ell \; \bar{\nu}$.

$\bar{B} \to D^* \; X \; \ell \; \bar{\nu}$ decay is a significant source of background in this analysis, principally because this is the background physical process most similar to $\bar{B} \to D^* \; \ell \; \bar{\nu}$. It includes resonant $\bar{B} \to D^{**} \; \ell \; \bar{\nu}$ decays followed by $D^{**} \to D^* X$, as well as non-resonant decays. This background is estimated from data events in a different region of C vs. MM^2 with $0.8 < |\mathbf{p}_\ell| < 1.4$ GeV. This region in $|\mathbf{p}_\ell|$, C and MM^2 is called the Correlated Background Region. It is described in more detail in Section IVB, along with the method for subtracting $\bar{B} \to D^* \; X \; \ell \; \bar{\nu}$ background. The method is sensitive to the detailed C vs. MM^2 distribution of $\bar{B} \to D^* \; X \; \ell \; \bar{\nu}$ background, for which we use theoretical model predictions (and to which we assign conservative errors). Since this method must provide an estimate of the number of entries in the Signal Region due to $\bar{B} \to D^* \; X \; \ell \; \bar{\nu}$ decay, it also provides a model dependent measurement of $\mathcal{B}(\bar{B} \to D^* \; X \; \ell \; \bar{\nu})$.

III. EVENT RECONSTRUCTION

The data used in this analysis were produced in symmetric electron-positron collisions at the Cornell Electron-positron Storage Ring (CESR) and recorded with the CLEO II detector. The signal comes from an integrated luminosity of 1.55 fb^{-1} collected at the $\Upsilon(4S)$ center-of-mass energy [23]. An additional 0.69 fb^{-1} of data collected below the $B\bar{B}$ production threshold are used for continuum background determination.

The most crucial components of the CLEO II detector in this analysis are the tracking system, the CsI electromagnetic calorimeter, and the muon identification system. A detailed description of the CLEO II detector is given elsewhere [24]. The tracking system comprises a set of drift and straw tube chambers in a 1.5 Tesla magnetic field that measure the momenta of stable charged particles over approximately 92% of 4π with a transverse momentum resolution of $(\delta p_t/p_t)^2 = (0.0015 p_t)^2 + (0.005)^2$, where p_t is measured in GeV. Photons are detected in a CsI electromagnetic calorimeter with an angular acceptance of 95% of 4π. We restrict the fiducial volume for photons to the barrel portion of the calorimeter, $|\cos\theta_\gamma| < 0.8$, where θ_γ is the angle a photon makes with the beam line (polar angle). The calorimeter energy resolution is $\Delta E/E(\%) = 0.35/E^{3/4} + 1.9 - 0.1E$, where E is in GeV [24,25], which corresponds to 4% at 0.1 GeV. Both electron and muon candidates must lie within the polar angular region $|\cos\theta_\ell| < 0.71$. In this analysis, electrons with momenta above 0.8 GeV are identified by their electromagnetic interactions in the calorimeter, their energy loss in the drift chamber gas and their time of flight in the detector. The electron identification efficiency within the fiducial volume is over 94%, while a hadron in the momentum range 0.8 to 2.4 GeV has on average a $(0.3 \pm 0.1)\%$ probability of being misidentified as an electron. Muons are identified by their ability to penetrate at least 5 nuclear absorption lengths in iron, which puts a lower limit of 1.4 GeV on the muon momentum acceptance. Muons within the acceptance are identified with 93% efficiency, while hadrons have on average a $(1.4 \pm 0.2)\%$ probability of being misidentified as a muon.

For this analysis we select hadronic events [26] that have at least one track identified as a lepton with momentum $0.8 \leq |\mathbf{p}_\ell| \leq 2.4$ GeV. The ratio of the second to the zeroth Fox-Wolfram moments [27] of the event is required to be less than 0.4 to suppress background from continuum events. For each lepton in these events we search for D^0 candidates in the decay mode $D^0 \to K^-\pi^+$ using charge correlation with the lepton to make unambiguous mass assignments (the lepton and kaon charges must be the same).

We combine D^0 candidates with pion candidates to fully reconstruct D^* mesons in the modes $D^{*+} \to D^0\pi^+$ and $D^{*0} \to D^0\pi^0$. We call these pion candidates slow pions because their momentum is restricted to be less than 225 MeV in the laboratory frame. Charged slow pions are accepted if they lie in the polar angle region $|\cos\theta_{\pi^+}| < 0.71$ and have momentum above 65 MeV. Candidate π^0's are constructed from pairs of showers in the electromagnetic calorimeter which do not match the projection of any drift chamber track and have an invariant mass within 3σ of the measured π^0 mass ($\sigma = 5$ to 8 MeV, depending on shower energies and polar angles). Showers used in π^0 candidates must be in the polar angle region $|\cos\theta_\gamma| < 0.8$ and have an energy above 30 MeV. The π^0 momentum vector is reconstructed by constraining the shower position and energy measurements to produce the known π^0 mass. The momentum of D^* candidates must satisfy $|\mathbf{p}_{D^*}|/\sqrt{E_B^2 - m_{D^*}^2} < 0.5$ to be consistent with B decay.

The raw yield of events with a D^* meson and a lepton is obtained by fitting the D^0 mass peak after cutting on the $D^* - D^0$ candidate mass difference, $\delta_m \equiv M_{K\pi\pi_s} - M_{K\pi}$, and subtracting the scaled result of a D^0 mass fit to a δ_m sideband ($M_{K\pi}$ and $M_{K\pi\pi_s}$ denote the invariant masses of the D^0 and D^* candidates, respectively). This sideband subtraction is described in detail in the next section.

IV. BACKGROUNDS

The number of D^* and lepton pairs observed in a given region of the C vs. MM^2 plane is the sum of the signal and background sources in that region. The background sources fall into five categories, and we have evaluated the contribution from each of these sources using the data. In order of importance the five background categories are: combinatoric, correlated, uncorrelated, continuum, and fake lepton background.

A. Combinatoric Background

The dominant background in this analysis is the combinatoric background in D^{*+} and D^{*0} reconstruction. One class of background arises from combinations of random $K^-\pi^+$ pairs with any slow second pion, π_s. This class of background does not peak in $M_{K\pi}$, but a subclass of these events in which the π_s and either the K^- or π^+ are from a true $D^* \to D^0\pi$ decay does peak in δ_m. Combinations of $K^-\pi^+$ pairs from a correctly reconstructed D^0 meson with an unrelated pion candidate forms a second class of combinatoric background which peaks in $M_{K\pi}$, but not in δ_m. The $M_{K\pi}$ distributions in the δ_m signal and sideband regions are fit to a D^0 peak plus a Chebyshev polynomial, which removes the first class of background in both regions. To remove the second class and obtain a final signal yield, the D^0 yield in the δ_m sideband is scaled and subtracted from the yield in the δ_m signal region. The scale factor is determined from the background function in a fit to the δ_m distribution [28].

The distributions of δ_m for D^{*+} and D^{*0} are shown in Fig. 2 for events which lie in the C vs. MM^2 Signal Region and have $M_{K\pi}$ within 100 MeV of the measured D^0 mass. The δ_m signal region for D^{*+} is 8 MeV wide and is centered on the measured δ_m mean of 145.44 MeV [29], while the sideband region is $150 < \delta_m < 165$ MeV. The δ_m signal region for D^{*0} is 6 MeV wide and is centered on the measured δ_m mean of 142.12 MeV [30], while the sideband region is $147 < \delta_m < 162$ MeV. We have used a wider signal region for D^{*+} because the δ_m signal in this mode has non-Gaussian tails from systematic effects in the reconstruction of very low momentum charged tracks. These effects are understood and reproduced in the Monte Carlo simulation.

The $M_{K\pi}$ distributions for events in the δ_m signal and sideband regions are shown in Fig. 3. These are the four distributions that were fit to obtain the yields in the C vs. MM^2 Signal Region. The yields are listed separately for the δ_m signal and sideband regions in Tables I and II.

B. Correlated D^*–lepton Background

A $\bar{B}$ meson can decay to a final state with a D^* and an e^- or μ^- through channels other than $\bar{B} \to D^* \ell \bar{\nu}$, and these physical processes contribute a background that we call correlated. The principal source of correlated background is the decay $\bar{B} \to D^* X \ell \bar{\nu}$, where e.g., $X = \pi$ or other unreconstructed particle(s) ($X \neq 0$). To remove this background, we exploit two general features of $\bar{B} \to D^* X \ell \bar{\nu}$ decays. First, because we do not reconstruct X the MM^2 variable will tend to be shifted towards positive values. Second, because the D^*X invariant mass is larger than the D^* mass, the lepton spectrum will be softer than for $\bar{B} \to D^* \ell \bar{\nu}$ decays. Fig. 4 shows the C vs. MM^2 distribution in the two lepton momentum ranges $0.8 < |\mathbf{p}_\ell| < 1.4$ GeV and $1.4 < |\mathbf{p}_\ell| < 2.4$ GeV for Monte Carlo $\bar{B} \to D^{**} \ell \bar{\nu}$ decays generated according to the model of ISGW [31] (we take these events to be representative, at the level of precision required here, of generic $\bar{B} \to D^* X \ell \bar{\nu}$ decays). In the lower lepton momentum range there is a region of the C vs. MM^2 plane where the $\bar{B} \to D^{**} \ell \bar{\nu}$ efficiency is high, but there is almost no contribution from $\bar{B} \to D^* \ell \bar{\nu}$ decays. Therefore, the low lepton momentum range is used to measure the level of $\bar{B} \to D^* X \ell \bar{\nu}$, which is then scaled using Monte Carlo to determine the $\bar{B} \to D^* X \ell \bar{\nu}$ contribution to the high lepton momentum Signal Region [32]. We do not determine the level of $\bar{B} \to D^* X \ell \bar{\nu}$ from the events just outside the high momentum Signal Region, because the $\bar{B} \to D^{**} \ell \bar{\nu}$ efficiency is low in the high lepton momentum range and electron bremsstrahlung in the detector causes a small fraction of $\bar{B} \to D^* \ell \bar{\nu}$ decays to be reconstructed outside the signal boundary (Fig. 1).

The dashed lines in Fig. 4 define a region in the C vs. MM^2 plane for the lepton momentum range $0.8 < |\mathbf{p}_\ell| < 1.4$ GeV called the Correlated Background Region, because it has high efficiency for the $\bar{B} \to D^* X \ell \bar{\nu}$ background processes, but very low efficiency for our signal mode. The $M_{K\pi}$ distributions in the Correlated Background Region are shown in Fig. 5. Let N_S and N_C denote the total yields in the high lepton momentum Signal Region and the low lepton momentum Correlated Background Region, respectively, after subtraction of all other backgrounds as discussed in the following sections. They are related to $N_s(\bar{B} \to D^* \ \ell \ \bar{\nu})$, the number of $\bar{B} \to D^* \ \ell \ \bar{\nu}$ events in the high momentum Signal Region, and $N_c(\bar{B} \to D^* \ X \ \ell \ \bar{\nu})$, the number of $\bar{B} \to D^* \ X \ \ell \ \bar{\nu}$ events in the low momentum Correlated Background Region, via

$$\left.\begin{aligned} N_S &= \quad N_s(\bar{B} \to D^* \ \ell \ \bar{\nu}) \ + \ r_c N_c(\bar{B} \to D^* \ X \ \ell \ \bar{\nu}) \\ N_C &= r_s N_s(\bar{B} \to D^* \ \ell \ \bar{\nu}) \ + \quad N_c(\bar{B} \to D^* \ X \ \ell \ \bar{\nu}) \end{aligned}\right\} \tag{10}$$

where $r_s = 0.015 \pm 0.003$ is the ratio of the efficiency for $\bar{B} \to D^* \ \ell \ \bar{\nu}$ events in the Correlated Background Region to the efficiency in the Signal Region, and $r_c = 0.70 \pm 0.35$ is the ratio of the $\bar{B} \to D^* \ X \ \ell \ \bar{\nu}$ efficiency in the signal region to the efficiency in the Correlated Background Region. Even though the Correlated Background Region is outside the nominal kinematic boundary for signal, r_s is non-zero because of electron bremsstrahlung. The values for r_s and r_c were determined from Monte Carlo simulation, and the errors reflect uncertainties due to model dependence. We estimated r_c using resonant $\bar{B} \to D^{**} \ \ell \ \bar{\nu}$ decays. These were generated according to the ISGW model [31] for the ${}^1P_1$, ${}^3P_1$, and ${}^3P_2$ D^{**} states in their predicted relative abundances [33]. We expect that the $B \to D^{**}({}^3P_1)\ell\bar{\nu}$ decay is a reasonable approximation to non-resonant $B \to (D^*\pi)\ell\bar{\nu}$ decays because the ${}^3P_1$ is expected to be a very wide resonance. We take into account our rough modelling of $\bar{B} \to D^* \ X \ \ell \ \bar{\nu}$ decays by assigning a conservative error of 50% to r_c.

The observed numbers of events and the yields obtained by solving Eq. (10) with the quoted central values of r_s and r_c are summarized in Tables I and II. The uncertainties in r_s and r_c are included in the systematic errors for the final $\bar{B} \to D^* \ \ell \ \bar{\nu}$ and $\bar{B} \to D^* \ X \ \ell \ \bar{\nu}$ yields.

Sources other than $\bar{B} \to D^* \ X \ \ell \ \bar{\nu}$ contribute to the correlated background at much smaller levels. The decays $B \to D^*X$, where X fragments to light hadrons that decay semileptonically or are misidentified as leptons, are discussed in the section on lepton fakes. The processes $B \to D^*\tau^-\bar{\nu}$ followed by $\tau^- \to l^-\bar{\nu}\nu$, and $B \to D^*D_{(s)}^{(*)-}$ followed by $D_{(s)}^{(*)-} \to Xl^-\bar{\nu}$ lead to final states that include a D^* and a lepton with the same "right sign" charge correlation as signal. These sources have characteristics similar to $\bar{B} \to D^* \ X \ \ell \ \bar{\nu}$, but with an even softer lepton spectrum. Monte Carlo simulation predicts values for r_c of 0.01 ± 0.01 and 0.005 ± 0.005 for $B \to D^*\tau^-\bar{\nu}$ and $B \to D^*D_{(s)}^{(*)-}$ respectively. They can therefore be neglected in the Signal Region. We have estimated their contributions to the background region using available measurements or estimates of the various intermediate branching fractions [29,34] together with efficiencies determined in Monte Carlo simulations. The results are given in Tables I and II.

C. Uncorrelated D^* lepton background

"Uncorrelated" background encompasses events with a D^* from the decay of the $\bar{B}$ and a lepton from the B in a $B\bar{B}$ event. However, in order to contribute to the background in this analysis the D^* and the lepton must have the same charge correlation as signal decays. Therefore, the leptons in uncorrelated background are secondary B meson decay products from the decay chain $\bar{b} \to \bar{c} \to \ell$ (known as cascades) or $B^0 \to \bar{B}^0 \to \ell X$ (mixing). It is also possible to have leptons from the decay or misidentification of light hadrons (K or π).

The cross section for producing uncorrelated background events is measured using the data. Because the B and $\bar{B}$ mesons in an event decay independently, the uncorrelated background cross section is a product of the inclusive D^* cross section from $\Upsilon(4S)$ decay, $\sigma'(e^+e^- \to \Upsilon(4S) \to D^*)$, and the $B \to \ell^-$ inclusive branching fraction, $\mathcal{B}'(B \to \ell^-)$. The primes indicate that the quantities of interest are the *detected* ("raw") cross section and branching fraction, which include detector acceptance and reconstruction efficiencies as well as the underlying cross section and branching fraction. We measure these two quantities with the same event selection and cuts used to obtain $\bar{B} \to D^*\, \ell\, \bar{\nu}$ yields.

We find that the cross sections for producing and detecting a D^* from $\Upsilon(4S)$ decay (integrated over momentum) are 3.1 ± 0.15 pb and 3.2 ± 0.33 pb for D^{*+} and D^{*0} respectively.

Measuring $\mathcal{B}'(B \to \ell^-)$ is not as straightforward as counting leptons because of the need to preserve the correct D^* and lepton charge correlation, *i.e.*, we need specifically $\mathcal{B}'(B \to \ell^-)$, rather than $\mathcal{B}'(\Upsilon(4S) \to \ell^{\pm})$. We measure $\mathcal{B}'(B \to \ell^-)$ using like charge dilepton events, in which one of the leptons tags one daughter of the $\Upsilon(4S)$ as either a B or a $\bar{B}$ [35]. We find that $\mathcal{B}'(B \to \ell^-)$ is $(0.7 \pm 0.05)\%$ and $(1.2 \pm 0.05)\%$ in the momentum range $1.4 < |\mathbf{p}_\ell| < 2.4$ and $0.8 < |\mathbf{p}_\ell| < 1.4$ GeV, respectively. These detected branching fractions include many sources, such as misidentified hadrons from the "other" B, photon conversions, etc., but these all contribute to uncorrelated background and therefore contribute to the quantity we are trying to measure.

Multiplying $\sigma'(e^+e^- \to \Upsilon(4S) \to D^*)$ by $\mathcal{B}'(B \to \ell^-)$ yields the total cross section for uncorrelated background. However, we wish to estimate the number of uncorrelated background events in a particular region of the C vs. MM^2 plane; therefore the distribution of uncorrelated background in these variables must be known. These variables depend only on the inclusive D^* and lepton momentum distributions (which we have already measured) and the angle between the D^* and the lepton, α. The distribution of $\cos\alpha$ is flat because the two B mesons are produced nearly at rest. We use these three known distributions to simulate the C vs. MM^2 distribution of uncorrelated background. From this simulation and the measured uncorrelated background cross section we estimate the numbers of uncorrelated background events in the Signal Region and Correlated Background Region (given in Tables I and II).

D. Continuum and Lepton Fakes

The level of continuum background is estimated using the 0.69 fb^{-1} of data recorded at energies slightly below the $\Upsilon(4S)$ resonance. These data are analyzed in the same manner as resonance data in order to estimate the continuum backgrounds listed in Tables I and II.

There is a small background of $D^*\ell$ candidates where the lepton is actually a hadron that has been misidentified as a lepton. We call this the fake lepton background. Fake $D^*\ell$ pairs are predominantly uncorrelated, simply because there is more energy available to produce light hadrons in the decay of the B meson that did not produce the D^*. These uncorrelated fake leptons are already included in the cross section for uncorrelated background.

The correlated lepton fake contribution to the raw number of $D^*\ell$ pairs is estimated by performing a similar analysis of the data where D^* candidates are paired with light hadron candidates instead of leptons, and scaling the result by the known electron and muon fake rates. For this study a light hadron is defined as any detected charged particle that fails very loose electron and muon identification cuts. We use Monte Carlo simulation to determine the efficiency for correlated D^* and hadron pairs to fall within the signal and background regions of the C vs. MM^2 plane. The correlated lepton fake contributions turn out to be so small that we can easily afford a large uncertainty due to the use of Monte Carlo. The results are given in Tables I and II.

V. DETECTION EFFICIENCIES

The efficiency for $\bar{B} \to D^* \ell \bar{\nu}$ events to pass our event selection criteria is estimated from a Monte Carlo simulation. The generator produces events obeying all kinematic constraints due to angular momentum conservation and the $V - A$ nature of the pseudoscalar to vector decay [36]. These events are passed through a full detector simulation [37] and then analyzed as described above.

An uncertainty in the efficiency for $\bar{B} \to D^* \ell \bar{\nu}$ exists because the dynamics of the decay depend on unknown form factors [38]. Several phenomenological models of these form factors are available. For the results presented here, we have used the predictions of Neubert [39,40] to determine the central values of our efficiencies, and have compared these results to those obtained with the ISGW [31], BSW [41], and KS [42] models to estimate the model uncertainty (See Tables VI and VII). After comparing these models as well as varying the form factors within the Neubert model, it is estimated that the model dependence contributes a 3% systematic uncertainty in the detection efficiency.

The efficiencies for $\bar{B} \to D^* \ell \bar{\nu}$ events to be counted as signal are

$$\epsilon_s(\bar{B}^0 \to D^{*+} \ell^- \bar{\nu}) = [9.54 \pm 0.23 \pm 0.71]\% \tag{11}$$

$$\epsilon_s(B^- \to D^{*0} \ell^- \bar{\nu}) = [7.18 \pm 0.52 \pm 0.53]\% \tag{12}$$

where the first (systematic) error is uncorrelated between the two efficiencies and the second is correlated. These efficiencies do not include the D^* and D^0 branching fractions. The errors include: Monte Carlo statistics, model dependence, an uncertainty reflecting changes in efficiency corrected yields due to variation of cuts, the uncertainty in the normalization of the δ_m sidebands, the uncertainty in the lepton reconstruction and identification, a 2% per track uncertainty in the reconstruction of K and π tracks with momenta above 250 MeV, and the uncertainty in modelling the efficiency of the slow pions. The methods used to estimate the latter uncertainty are described below. The various contributions to the systematic uncertainty in the efficiency are listed in Table III.

In the laboratory frame, the slow pion from the $\bar{B} \to D^* \ell \bar{\nu}$, $D^* \to D\pi$ decay chain has a momentum of less than 225 MeV with the peak of the distribution at 100 MeV. The efficiency to reconstruct π^0's is fairly flat in this momentum range but difficult to model because of the large background of low energy showers in the calorimeter. The efficiency for reconstructing π^+'s decreases sharply below 100 MeV, falling to 0 at 50 MeV. It is crucial to this analysis that these efficiencies be correctly reproduced by the Monte Carlo and we have performed several detailed studies to evaluate the reliability of our simulation [43].

A. π^+ Efficiency

The shape of the efficiency curve for charged pions of momentum less than 225 MeV can be measured using inclusive D^{*+} decays in the data. In the strong decay $D^{*+} \to D^0\pi^+$, $dN/d\cos\varphi_{D^*\pi}$ must be symmetric about $\cos\varphi_{D^*\pi} = 0$, where $\varphi_{D^*\pi}$ is the angle between the slow π^+ momentum in the D^{*+} rest frame and the D^{*+} direction in the lab frame. However, the observed $dN/d\cos\varphi_{D^*\pi}$ distribution in the data may be asymmetric because the slow π^+ efficiency varies with pion momentum, which is highly correlated with $\cos\varphi_{D^*\pi}$.

The charged pion efficiency between 0 and 225 MeV was measured in data by simultaneously fitting the observed $dN/d\cos\varphi_{D^*\pi}$ distributions in 8 bins of D^{*+} momentum between 0 and 5 GeV. It was found that the Monte Carlo reproduces the data efficiency shape over the pion momentum range 0 to 225 MeV (Fig. 6). This study of the $\cos\varphi_{D^*\pi}$ distribution tests the simulation of the shape of the efficiency curve, but not its absolute normalization. We test the normalization by comparing the yield of fully reconstructed $D^0 \to K^-\pi^+\pi^0$ to the yield of partially reconstructed $D^0 \to K^-\pi^0(\pi^+)$, where the π^+ is not used. The momentum range for these pions is 0.2 to 1.0 GeV. The Monte Carlo efficiency agrees well with the absolute π^+ efficiency measured in the data, within the statistical precision of 2%. Together, these studies check at the 5% level the simulation of the slow π^+ efficiency.

B. π^0/π^+ Efficiency

The accuracy of the simulation of the slow π^0 efficiency is difficult to test. The method outlined above cannot be used for the decay $D^{*0} \to D^0\pi^0$ because of the large combinatoric background in the π^0 reconstruction. Instead, we compare the ratio of efficiencies, $\epsilon(\pi^0)/\epsilon(\pi^+)$, in the pion momentum range 0 to 225 MeV between data and Monte Carlo. We then combine the results of this study with the uncertainty in the charged pion absolute efficiency described above in order to obtain the uncertainty in the π^0 absolute efficiency.

We determine the ratio of efficiencies below 225 MeV, $c_1 = \epsilon(\pi^0)/\epsilon(\pi^+)$, using a combination of two methods. The first part uses the decay $\eta \to \pi^+\pi^-\pi^0$ as a source of both neutral and charged soft pions. Candidate η's are selected with either a π^0 in the low (0 to 225 MeV) momentum range and both charged pions above 250 MeV or a $\pi^\pm$ in the low momentum range and the other charged and neutral pions with momentum above 250 MeV. The ratio of the number of events in these two bins depends on the efficiency to find the pions times a kinematic factor that depends on the η Dalitz plot and the η momentum spectrum. The kinematic factor can be interpreted as the average probability for the three pions from the η decay to populate the desired momentum ranges. The η Dalitz plot has been precisely studied in previous experiments [44], and the η momentum spectrum in the Monte Carlo was tuned to the data using the $\eta \to \gamma\gamma$ decay mode. The ratio of yields is proportional to the ratio c_1/c_2 where c_1 is defined above and c_2 is the ratio of neutral to charged pion efficiencies above 250 MeV. The ratio c_1/c_2 is determined in data and compared with the same ratio calculated from Monte Carlo.

A second study using K^0_S's then determines the product $c_1 c_2$. The K^0_S study takes advantage of the fact that the ratio of branching fractions $\mathcal{B}(K^0_S \to \pi^0\pi^0)/\mathcal{B}(K^0_S \to \pi^+\pi^-)$ is well known. Therefore, the ratio of neutral to charged final state yields is the product of a known ratio of branching fractions, an efficiency ratio and an acceptance factor which is determined from Monte Carlo. For kaons of momentum 250 to 500 MeV, one daughter pion is in the range 60 to 225 MeV and the other daughter is above 225 MeV so that the ratio of yields $N(K^0_S \to \pi^0\pi^0)/N(K^0_S \to \pi^+\pi^-)$ in the K^0_S momentum range 250 to 500 MeV is proportional to the product $c_1 c_2$.

When the results of the η and K^0_S studies are combined we find that c_1 (the ratio of neutral to charged pion efficiency for momentum 60 to 225 MeV), as determined from the data, agrees with the predictions of the Monte Carlo to within the statistical precision of 7%. We therefore assign a systematic error on the efficiency ratio of neutral to charged slow pions of 7%. Combining this with our study of the absolute slow charged pion efficiency, we assign an error of 8.6% to the absolute determination of the neutral pion efficiency in the momentum range below 225 MeV. The Monte Carlo prediction of the efficiency to reconstruct neutral pions is shown in Fig. 7.

VI. BRANCHING FRACTION RESULTS

In this analysis we measure two independent quantities: N_0 and N_- of Eqs. (4) and (5). These are in turn given by

$$N_0 = \frac{N_s(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu})}{\epsilon_s(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu})} \tag{13}$$

$$N_- = \frac{N_s(B^- \to D^{*0}\ \ell^-\ \bar{\nu})}{\epsilon_s(B^- \to D^{*0}\ \ell^-\ \bar{\nu})} \tag{14}$$

where $N_s(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu})$ and $N_s(B^- \to D^{*0}\ \ell^-\ \bar{\nu})$ are the background subtracted yields for the given signal modes (Tables I and II), and $\epsilon_s(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu})$ and $\epsilon_s(B^- \to D^{*0}\ \ell^-\ \bar{\nu})$ are the efficiencies for the given signal modes (Section V). Dividing N_0 and N_- by four times the number of $\Upsilon(4S)$ decays in the data sample yields the product branching fractions given in Table IV. The number of $\Upsilon(4S)$ events was determined by studying the hadronic cross section at energies both on the $\Upsilon(4S)$ resonance and slightly below the resonance in the continuum. For our data sample, $N_{\Upsilon(4S)} = (1.65 \pm 0.01) \times 10^6$ events.

The top two product branching fractions of Table IV depend on the unknowns to be determined, $\mathcal{B}(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu})$ and $\mathcal{B}(B^- \to D^{*0}\ \ell^-\ \bar{\nu})$, as well as the two unmeasured quantities f_{+-} and f_{00}. In all, there are four unknowns and only two measurements. The assumption that the $\Upsilon(4S)$ decays only to $B\bar{B}$ ($f_{+-} + f_{00} = 1$) reduces the number of unknowns to three. This assumption is supported by the observation that the $\Upsilon(4S)$ width is three orders of magnitude larger than lower Υ states which are not massive enough to decay to a pair of B mesons [45]. There are two ways to further reduce the number of unknowns. The first method is to make the traditional assumption that $f_{+-} = f_{00}$, which is uncertain at the $5 - 10\%$ level [46]. Using this assumption we present the two independent measurements: $\mathcal{B}(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu})$ and $\mathcal{B}(B^- \to D^{*0}\ \ell^-\ \bar{\nu})$. Alternatively, we can assume that $\mathcal{B}(B^- \to D^{*0}\ \ell^-\ \bar{\nu})/\mathcal{B}(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}) = \tau_{B^-}/\tau_{\bar{B}^0}$. This is equivalent to assuming that the partial widths for the exclusive semi-leptonic decays of the charged and neutral B's are equal: $\Gamma(B^- \to D^{*0}\ \ell^-\ \bar{\nu}) = \Gamma(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu})$. With this assumption and a measurement of $\tau_{B^-}/\tau_{\bar{B}^0}$ from other experiments [47], our measurement of N_0 and N_- can be used to determine $\Gamma(\bar{B} \to D^*\ \ell\ \bar{\nu})$ and f_{+-}/f_{00}. We will first discuss the $\bar{B} \to$

$D^* \, \ell \, \bar{\nu}$ branching fractions obtained using the two different assumptions, and then the $\Upsilon(4S)$ branching fractions obtained with the second assumption.

A. $\bar{B} \to D^* \, \ell \, \bar{\nu}$ Branching Fractions

To extract the B meson semileptonic branching fractions from Table IV we require values for the rest of the branching fractions in each product. For the D^* branching fractions we use the CLEO II results [18],

$$\begin{aligned} \mathcal{B}(D^{*+} \to D^0\pi^+) &= [68.1 \pm 1.0 \pm 1.3]\% \\ \mathcal{B}(D^{*0} \to D^0\pi^0) &= [63.6 \pm 2.3 \pm 3.3]\% \end{aligned} \tag{15}$$

For the D^0 branching fraction we use the CLEO II result [19]:

$$\mathcal{B}(D^0 \to K^-\pi^+) = [3.91 \pm 0.08 \pm 0.07 \pm 0.16]\% \tag{16}$$

which is the value given in Ref. [19] without final state radiation. The systematic error on K and π track reconstruction (third error) has been separated from other systematics (second error) because it is correlated with row 5 of Table III. This correlation is an important advantage of using this CLEO II measurement of $\mathcal{B}(D^0 \to K^-\pi^+)$, as the error due to $K\pi$ track reconstruction will cancel in the determination of the $\bar{B} \to D^* \, \ell \, \bar{\nu}$ branching fractions. Thus the fractional systematic error on our results for $\mathcal{B}(\bar{B} \to D^* \, \ell \, \bar{\nu})$ will be reduced relative to the error on the product branching fractions listed in Table IV.

If we assume equal production of charged and neutral meson pairs at the $\Upsilon(4S)$, $f_{+-} = f_{00} = 0.5$, we obtain,

$$\mathcal{B}_{0.5}(\bar{B}^0 \to D^{*+} \, \ell^- \, \bar{\nu}) = [4.49 \pm 0.32 \pm 0.39]\%, \tag{17}$$

$$\mathcal{B}_{0.5}(B^- \to D^{*0} \, \ell^- \, \bar{\nu}) = [5.13 \pm 0.54 \pm 0.64]\%, \tag{18}$$

where the first error is statistical, and the second (systematic) error includes the errors on the D^0 and D^* branching fractions. On the other hand, by assuming $\Gamma(\bar{B}^0 \to D^{*+} \, \ell^- \, \bar{\nu}) = \Gamma(B^- \to D^{*0} \, \ell^- \, \bar{\nu})$, which follows from isospin invariance, and $f_{+-} + f_{00} = 1$, Eqs. (4) and (5) can be combined to give

$$\Gamma(\bar{B} \to D^* \, \ell \, \bar{\nu}) = \frac{1}{4N_{\Upsilon(4S)}\mathcal{B}_{D^0}} \left[\frac{N_0}{\tau_{\bar{B}^0}\mathcal{B}_{D^{*+}}} + \frac{N_-}{\tau_{B^-}\mathcal{B}_{D^{*0}}} \right]. \tag{19}$$

Using the average of LEP and CDF measurements for the lifetimes τ_{B^-} and τ_{B^0} [47] yields,

$$\Gamma(\bar{B} \to D^* \, \ell \, \bar{\nu}) = [29.9 \pm 1.9 \pm 2.7 \pm 2.0] \ \mathrm{ns}^{-1}, \tag{20}$$

independent of f_{+-}/f_{00}. The second (systematic) error takes into account correlations between the errors of Eqs. (17) and (18). The third error is due to the B lifetime measurements and is determined with the conservative assumption that the errors on the charged and neutral lifetimes are fully correlated.

The partial width can be converted to either the charged or the neutral $\bar{B} \to D^* \, \ell \, \bar{\nu}$ branching fraction by multiplying Eq. (19) times the appropriate lifetime, but note that

$\mathcal{B}(B^- \to D^{*0}\ \ell^-\ \bar\nu)$ and $\mathcal{B}(\bar B^0 \to D^{*+}\ \ell^-\ \bar\nu)$ are not measured independently this way. The result is

$$\mathcal{B}(B^- \to D^{*0}\ \ell^-\ \bar\nu) = \frac{1}{2}\left[\frac{\tau_{B^-}}{\tau_{\bar B^0}}\mathcal{B}_{0.5}(\bar B^0 \to D^{*+}\ \ell^-\ \bar\nu) + \mathcal{B}_{0.5}(B^- \to D^{*0}\ \ell^-\ \bar\nu)\right], \tag{21}$$

independent of f_{+-}/f_{00}. Note that this does not depend on the individual B lifetimes, but only on their ratio. Taking $\tau_{B^-}/\tau_{\bar B^0} = 1.1 \pm 0.11$ from other experiments [47] leads to

$$\mathcal{B}(B^- \to D^{*0}\ \ell^-\ \bar\nu) = 1.1 \times \mathcal{B}(\bar B^0 \to D^{*+}\ \ell^-\ \bar\nu) = [5.03 \pm 0.32 \pm 0.45 \pm 0.24]\%, \tag{22}$$

where the uncertainty in our knowledge of the lifetime ratio leads to the third systematic error. Use of a different central value for the lifetime ratio will change the right hand side of Eq. (22) according to Eq. (21). The factor 1.1 carries no error to indicate that we have assumed that the ratio of branching fractions exactly equals the lifetime ratio, which means that we can not independently determine the charged and neutral branching fractions.

One can compare the above branching fractions to previous results from ARGUS and CLEO, which assume $f_{+-} = f_{00} = 0.5$, after correcting all results for the new CLEO II D^* and D^0 branching fractions [48]. The results of this comparison are shown in Table V. Our results are in good agreement with previous measurements.

B. $\Upsilon(4S)$ Branching Fractions

By taking the ratio of Eq. (5) over (4) and assuming only that $\Gamma(\bar B^0 \to D^{*+}\ \ell^-\ \bar\nu) = \Gamma(B^- \to D^{*0}\ \ell^-\ \bar\nu)$, we have

$$\frac{f_{+-}}{f_{00}}\frac{\tau_{B^-}}{\tau_{\bar B^0}} = \frac{N_-}{N_0}\frac{\mathcal{B}_{D^{*+}}}{\mathcal{B}_{D^{*0}}} = 1.14 \pm 0.14 \pm 0.13. \tag{23}$$

where the first error is statistical and the second (systematic) error is dominated by the uncertainty on the slow pion efficiency ratio, $\epsilon_{\pi^0}/\epsilon_{\pi^+}$. Substituting the value for $\tau_{B^-}/\tau_{\bar B^0}$ from Ref. [47],

$$\frac{f_{+-}}{f_{00}} = 1.04 \pm 0.13 \pm 0.12 \pm 0.10, \tag{24}$$

where the third error is due to the input lifetime ratio. Finally, the assumption $f_{+-} + f_{00} = 1$ can be used to extract a value for f_{+-} or f_{00},

$$f_{+-} = 1 - f_{00} = 0.510 \pm 0.052, \tag{25}$$

where the error is combined statistical and systematic, including the error in the B lifetime ratio.

C. $\mathcal{B}(\bar{B} \to D^* \; X \; \ell \; \bar{\nu})$ from Correlated Background Region Yields

The product branching fractions for $\mathcal{B}(\bar{B} \to D^* \; X \; \ell \; \bar{\nu})$ in Table IV have been computed from the correlated background yields in the Correlated Background Region from Tables I and II divided by an efficiency estimated with Monte Carlo $\bar{B} \to D^{**} \; \ell \; \bar{\nu}$ events, as explained below.

All possible resonant and non-resonant sources of correlated background contribute to the data yields:

$$\begin{aligned} N_c(\bar{B} \to D^* \; X \; \ell \; \bar{\nu}) = \mathcal{N} \sum_i \mathcal{B}(\bar{B} \to D_i^{**} \; \ell \; \bar{\nu}) \mathcal{B}(D_i^{**} \to D^* X) \epsilon_i \\ + \mathcal{N} \sum_j \mathcal{B}(\bar{B} \to D^* X_j \; \ell \; \bar{\nu}) \epsilon_j, \end{aligned} \tag{26}$$

where i ranges over all possible resonant states and j over all non-resonant states, and $\mathcal{N}$ contains the total number of B decays in the data times the D^* and D^0 branching fractions. The efficiencies for the individual channels, ϵ_i and ϵ_j, are potentially all different. Therefore, an accurate estimate of the efficiency for the $\bar{B} \to D^* \; X \; \ell \; \bar{\nu}$ inclusive process, $\epsilon_c(\bar{B} \to D^* \; X \; \ell \; \bar{\nu})$, requires knowledge of the relative abundances of all the resonant and non-resonant exclusive modes. However, if the different exclusive modes all have similar efficiencies, one only needs to simulate one or a few of them in order to estimate $\epsilon_c(\bar{B} \to D^* \; X \; \ell \; \bar{\nu})$. We compared the Monte Carlo efficiencies for $\bar{B} \to D^{**} \; \ell \; \bar{\nu}$ where the D^{**} was a 1^1P_1, 1^3P_1 or 1^3P_2 state generated according to the ISGW model [31] in the $D^*\pi$ decay channel. Since the 1^3P_1 resonance is very wide, we expect it to resemble non-resonant decays where X_j is a pion. For these three exclusive channels we find the following efficiencies in the Correlated Background Region: $(3.4 \pm 0.2)\%(1^1P_1)$, $(2.3 \pm 0.2)\%(1^3P_1)$, and $(2.9 \pm 0.2)\%(1^3P_2)$ for a D^{*+} in the final state, and $(2.5 \pm 0.2)\%(1^1P_1)$, $(1.7 \pm 0.2)\%(1^3P_1)$, and $(2.3 \pm 0.2)\%(1^3P_2)$ for a D^{*0} in the final state. We take the average of these numbers as a rough estimate of the Correlated Background Region efficiency for any channel which leads to a D^* in the final state,

$$\epsilon_c(\bar{B} \to D^{*+} \; X \; \ell^- \; \bar{\nu}) = [2.9 \pm 0.6]\% \tag{27}$$

$$\epsilon_c(\bar{B} \to D^{*0} \; X \; \ell^- \; \bar{\nu}) = [2.2 \pm 0.4]\%, \tag{28}$$

where the errors were estimated from the spread in the efficiencies of the three individual modes considered, together with the uncertainties from Table III.

Dividing rows 3 and 4 of Table IV by the D^0 and D^* branching fractions, we obtain,

$$\mathcal{B}(\bar{B} \to D^{*+} \; X \; \ell^- \; \bar{\nu}) = [0.6 \pm 0.3 \pm 0.1]\% \tag{29}$$

$$\mathcal{B}(\bar{B} \to D^{*0} \; X \; \ell^- \; \bar{\nu}) = [0.6 \pm 0.6 \pm 0.1]\% \tag{30}$$

These inclusive branching fractions do not depend on the unknown relative abundances of the different resonant and non-resonant D^*X states in semileptonic B decay, as long as they all have similar efficiencies in the Correlated Background Region. To obtain an upper limit for $\mathcal{B}(\bar{B} \to D^{**}\ell\bar{\nu})$, we assume that non-resonant channels are all zero, and use the relative abundances of the first radial excitation D^{**} states in the ISGW model to estimate $\mathcal{B}(D^{**} \to D^*\pi) = 77\%$ [33]. Together with results (29) and (30) this leads to

$\sum_i \mathcal{B}(\bar{B} \to D_i^{**}\ell\bar{\nu}) = (1.5 \pm 0.8 \pm 0.2)\%$, which does not include an error on the ISGW estimate of the relative D^{**} abundances. Converting this to an upper limit yields,

$$\sum_i \mathcal{B}(\bar{B} \to D_i^{**}\ell\bar{\nu}) < 2.8\% \tag{31}$$

at the 95% confidence level. This result is consistent with previous model-dependent determinations of the $\bar{B} \to D^{**}\ \ell\ \bar{\nu}$ branching fraction [9].

VII. DETERMINATION OF $|V_{cb}|$

The decay amplitude of $\bar{B} \to D^*\ \ell\ \bar{\nu}$ for massless leptons [49] is commonly expressed in terms of the three meson form factors $A_1(q^2)$, $A_2(q^2)$ and $V(q^2)$ [41]. These form factors characterize the transition between the two strongly bound states B and D^*, and are therefore not calculable. In order to calculate the decay rate for $\bar{B} \to D^*\ \ell\ \bar{\nu}$ and hence determine $|V_{cb}|$ from the measured rate it is necessary to know the normalization as well as the q^2 dependence of these form factors. Several phenomenological models have been devised in order to estimate the form factors, but these typically assume some q^2 dependence and use ad hoc wave function representations for the B and D^* bound states. Hence it is difficult to estimate the accuracy of model predictions. In order to compare different models we define the form factor ratios, A_2/A_1 and V/A_1. Table VI lists the expressions for A_1, A_2/A_1, and V/A_1 for the different models examined.

Recent work on Heavy Quark Effective Theory and its application to the decay $\bar{B} \to D^*\ \ell\ \bar{\nu}$ has led to constraints on the form factors that permit a less model dependent determination of $|V_{cb}|$. In the limit of infinite b and c quark masses the meson form factors A_1, A_2 and V are well defined functions of a single form factor $\xi(y)$, known as the Isgur-Wise function [11], where y is the kinematic variable of HQET, and is related to q^2 by Eq. (1). The Isgur-Wise function $\xi(y)$ contains the non-perturbative QCD dynamics of the light degrees of freedom in the mesons, and it is therefore not calculable. However, at $y = 1$ (point of maximum q^2) it is normalized to unity.

For finite b and c quark masses the $\bar{B} \to D^*\ \ell\ \bar{\nu}$ differential decay rate can be expressed in terms of a single unknown form factor $\mathcal{F}(y)$ which incorporates the three form factors A_1, A_2, and V. The relationships between these and the Isgur-Wise function are given by HQET only to leading order, and the corrections necessary to account for the deviation from the heavy quark symmetry limit are absorbed into $\mathcal{F}(y)$. Following Neubert [39,50] we write,

$$\begin{aligned}\frac{d\Gamma}{dy} = {} & \frac{G_F^2}{48\pi^3} m_{D^*}^3 (m_B - m_D^*)^2 |V_{cb}|^2 \mathcal{F}^2(y) \times \\ & \sqrt{y^2-1}\Big[4y(y+1)\frac{1-2yr+r^2}{(1-r)^2} + (y+1)^2\Big],\end{aligned} \tag{32}$$

where $r = m_{D^*}/m_B$. The function $\mathcal{F}(y)$ can be related to the Isgur-Wise function and correction terms that vanish in the infinite b and c mass limit, $\mathcal{F}(1) = \eta_A \xi(1) + \mathcal{O}\left(\frac{\Lambda_{QCD}}{m_Q}\right)^2$, where η_A is a perturbatively calculable QCD radiative correction and Λ_{QCD} is the QCD scale parameter. A next to leading order calculation gives $\eta_A = 0.986 \pm 0.006$ [39]. For $\bar{B} \to D^*\ \ell\ \bar{\nu}$ decay, it has been shown that corrections of order $1/m_c$ and $1/m_b$ are identically zero [14], and it is currently estimated that the second order deviations from the symmetry limit can be calculated with approximately 4% uncertainty [15,16,39]. Therefore, a precise measurement of $|V_{cb}|\mathcal{F}(1)$ will result in an accurate determination of $|V_{cb}|$ with no need for predictions of the shape of the decay form factors. Measurements of the differential decay rate $d\Gamma/dy$ will also determine the shape of $\mathcal{F}(y)$, which provides information about the non-perturbative QCD dynamics of the decay.

In the next section we use model predictions for the A_1, A_2 and V form factors to extract model dependent values of $|V_{cb}|$ from our measurement of $\mathcal{B}(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu})$ and $\mathcal{B}(B^- \to D^{*0}\ \ell^-\ \bar{\nu})$. These model dependent extractions have unknown theoretical errors. Following the model dependent extractions we describe the measurement of $|V_{cb}|\mathcal{F}(1)$ from the differential decay distribution $d\Gamma(\bar{B} \to D^*\ \ell\ \bar{\nu})/dy$. Using a model prediction for the corrections to the heavy quark symmetry limit at $y = 1$ we then present a value for $|V_{cb}|$. We also compare the results for the shape of $\mathcal{F}(y)$ with model predictions.

A. Model dependent determinations

There are several quark models that predict the normalization and q^2 dependence of the form factors describing the decay $\bar{B} \to D^*\ \ell\ \bar{\nu}$. When extracting model dependent values for $|V_{cb}|$ based on predictions of the form factors and their q^2 dependence it is important to determine the efficiency for the model under investigation because the acceptance can vary with different values of A_1, A_2/A_1 and V/A_1. The form factor ratios and their q^2 dependence for the different models considered here are in Table VI. Using the partial width from Eq. (20), $|V_{cb}|$ can be calculated from model predictions of the rate for $\bar{B} \to D^*\ \ell\ \bar{\nu}$. The calculation includes a correction for the difference in the efficiency predicted with each model; the values for $|V_{cb}|$ obtained with four different models are given in Table VII.

B. $|V_{cb}|$ determined from the $d\Gamma(\bar{B} \to D^*\ \ell\ \bar{\nu})/dy$ distribution

In this section the extraction of $|V_{cb}|\mathcal{F}(1)$ from the $d\Gamma(\bar{B} \to D^*\ \ell\ \bar{\nu})/dy$ distribution is described. Three fits are performed. Using the assumption that $f_{+-} = f_{00} = 0.5$, we first extract $|V_{cb}|\mathcal{F}(1)$ separately from the differential decay distributions for $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ and for $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ events. We then simultaneously fit the $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ and the $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ events, which results in a determination of $|V_{cb}|\mathcal{F}(1)$ that is independent of f_{00}/f_{+-}.

We use the unbinned maximum likelihood method developed for fitting the multidimensional decay distribution of $D \to K^* l \bar{\nu}$ events [51]. Application to the extraction of $|V_{cb}|\mathcal{F}(y)$ is straightforward since only one dimension is used in the fit. It is convenient to make the likelihood a function of V_{cb} by normalizing our probability distribution function to the number of observed events. This requires the inclusion of a Poisson probability factor [52].

D^* and lepton pairs are selected in the Signal Region of the C vs. MM^2 plane as already described. However, since we are now performing an unbinned likelihood fit, we apply a $M_{K\pi}$ cut as well as a δ_m cut. The $M_{K\pi}$ cuts are $|M_{K\pi} - m_{D^0}| < 25$ MeV and $|M_{K\pi} - m_{D^0}| < 20$ MeV for $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ and $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ respectively, where m_{D^0}

is the nominal D^0 mass. In order to have low background levels in the unbinned likelihood fit, the tighter $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ cut was chosen to optimize signal to background as much as possible, without introducing a systematic bias from reproduction of signal shapes by the Monte Carlo simulation.

Eq. (32) can be rewritten as

$$\frac{d\Gamma}{dy} = \mathcal{G}(y)|V_{cb}|^2\mathcal{F}^2(y), \tag{33}$$

where all the known quantities have been folded into a single function $\mathcal{G}(y)$. $|V_{cb}|$ and $\mathcal{F}(y)$ are to be determined by fitting the data. Since we want to determine the product $|V_{cb}|\mathcal{F}(y)$ at $y = 1$, we approximate the unknown function $\mathcal{F}(y)$ with an expansion about $y = 1$,

$$\mathcal{F}(y) = \mathcal{F}(1)\left[1 - a^2(y-1) + b(y-1)^2\right], \tag{34}$$

thus defining the fit parameters to be $|V_{cb}|\mathcal{F}(1)$, a^2 and b. We use a^2 rather than a because the B and D^* meson wave functions have maximum overlap at $y = 1$, so the first derivative of $\mathcal{F}(y)$ must be negative at that point. Our results for a^2 and b can be used to compare the shape of $\mathcal{F}(y)$ determined from data with theoretical model predictions [53].

The variable y is the D^* energy to mass ratio in the rest frame of the decaying B. However, in data we measure E_{D^*}/m_{D^*} in the laboratory frame, symbolized by $\tilde{y}$. This quantity is a function of y and the B meson momentum $\mathbf{p}_B$, but since the direction of $\mathbf{p}_B$ is unknown, the dN/dy distribution of the data is not directly observed. We therefore fit $dN/d\tilde{y}$, and naturally incorporate the smearing due to the B motion in the same way as the smearing due to detector resolution (which is much smaller). This is accomplished by a function $\mathcal{R}(\tilde{y}, y)$, which is convoluted with $\mathcal{G}(y)\mathcal{F}^2(y)$ in order to arrive at a probability distribution function that can be compared with the data,

$$G(\tilde{y}, a, b) = \int_1^{y_0} \mathcal{G}(y)\mathcal{F}^2(y)\mathcal{R}(\tilde{y}, y)dy, \tag{35}$$

where y_0 is the upper kinematic limit of y, given by $q^2 = 0$. The resolution function $\mathcal{R}(\tilde{y}, y)$ is determined from Monte Carlo simulation, and allows $G(\tilde{y}, a, b)$ to correctly include changes in detection efficiency with variations in the parameters of $\mathcal{F}$, thus reducing model dependence.

The full likelihood function is

$$\mathcal{L}(|V_{cb}|, a, b) = e^{-N(|V_{cb}|,a,b)-N_b}\prod_{i=1}^{n}\left(N(|V_{cb}|, a, b)G(\tilde{y}_i, a, b) + \sum_b n_b P_b(\tilde{y}_i)\right), \tag{36}$$

with

$$N(|V_{cb}|, a, b) = 4N_{\Upsilon(4S)}f\mathcal{B}_{D^*}\mathcal{B}_{D^0}\tau_B|V_{cb}|^2\int_1^{y_0}\epsilon(y)\mathcal{G}(y)\mathcal{F}^2(y)dy \tag{37}$$

$$N_b = n_{\rm comb} + n_{D^{**}l\bar{\nu}} + n_{\rm uncorr} \tag{38}$$

$$\sum_b n_b P_b(\tilde{y}_i) = n_{\rm comb}P_{\rm comb}(\tilde{y}_i) + n_{D^{**}l\bar{\nu}}P_{D^{**}l\bar{\nu}}(\tilde{y}_i) + n_{\rm uncorr}P_{\rm uncorr}(\tilde{y}_i). \tag{39}$$

$N(|V_{cb}|, a, b)$ is the expected number of reconstructed signal events, where $\epsilon(y)$ is the efficiency for detecting $\bar{B} \to D^*\ \ell\ \bar{\nu}$ events determined from the Monte Carlo, τ_B is the B

meson lifetime and f (meaning either f_{+-} or f_{00}), $N_{\Upsilon(4S)}$, $\mathcal{B}_{D^*}$ and $\mathcal{B}_{D^0}$ have been defined in Eqs. (4) and (5). The total number of background events passing our cuts is N_b. The exponential factor in Eq. (36) comes from the Poisson probability mentioned earlier. The backgrounds are represented by Eqs. (38) and (39). They have been divided into three types, combinatorial, $\bar{B} \to D^* X \ell \bar{\nu}$, and uncorrelated background, and their normalization and $\tilde{y}$ dependence is represented by Eq. (39). The combinatoric background yield and $\tilde{y}$ distribution, $(n_{\rm comb} P_{\rm comb}(\tilde{y}))$, are obtained from sidebands in the δm distribution. The correlated background yield, $n_{D^{**}l\bar{\nu}}$, is obtained from Table I or II [54] and its $\tilde{y}$ distribution, $P_{D^{**}l\bar{\nu}}(\tilde{y}_i)$, is taken from Monte Carlo generated as in Section VI C according to the ISGW model. The uncorrelated background yield and $\tilde{y}$ distribution, $n_{\rm uncorr} P_{\rm uncorr}(\tilde{y})$ is obtained from the measured inclusive $B \to D^*$ and $B \to l^-$ spectra, assuming an isotropic angular distribution between the uncorrelated leptons and D^*'s. The three $\tilde{y}$ distributions of the backgrounds to $\bar{B}^0 \to D^{*+} \ell^- \bar{\nu}$ and $B^- \to D^{*0} \ell^- \bar{\nu}$ are shown in Fig. 9.

The data yields as a function of $\tilde{y}$ for $\bar{B}^0 \to D^{*+} \ell^- \bar{\nu}$ and $B^- \to D^{*0} \ell^- \bar{\nu}$ events are shown in Fig. 10. The values for $|V_{cb}|\mathcal{F}(1)$ and a^2 from linear fits to these data are given in Table VIII. We have fixed b to zero in these fits because the statistics of the individual modes are not sufficient to constrain the curvature of $\mathcal{F}(y)$. The first error listed in Table VIII is statistical and the second error is an estimate of the systematic error, including the uncertainty in the B lifetimes. The statistical uncertainty in the backgrounds has been taken into account by varying the background levels in the fit by ± 1 standard deviation of their measured central values, and is included in the statistical error.

The contributions to the systematic error on $|V_{cb}|\mathcal{F}(1)$ and a^2 are listed in Table IX. The errors on the absolute efficiencies are the same as for the branching fraction measurements. The errors that are detailed in Table IX are specific to the $d\Gamma/dy$ fits and account for the uncertainty in the efficiency as a function of $\tilde{y}$ and uncertainties in the fitting procedure. As a check of the fitting technique we have performed a binned fit to the efficiency corrected $d\Gamma/d\tilde{y}$ distribution, using an analytically smeared form [43],

$$G'(\tilde{y}, a, b) = \frac{1}{2\gamma\beta} \int_{y_{min}(\tilde{y})}^{y_{max}(\tilde{y})} \frac{\mathcal{G}(y)\mathcal{F}^2(y)}{\sqrt{y^2-1}} dy, \tag{40}$$

where $\gamma\beta$ is the known momentum to mass ratio of the decaying B meson and (y_{min}, y_{max}) is the range of possible values of y given a value of $\tilde{y}$ (see Fig. 8). The results agree well with the results obtained from the unbinned likelihood fit. The systematic uncertainties due to the fitting procedure were estimated by comparing the results of the unbinned and binned methods in 36 Monte Carlo samples, each with the same statistics as our data sample.

We have also fit the differential decay distribution of the two decay modes $\bar{B}^0 \to D^{*+} \ell^- \bar{\nu}$ and $B^- \to D^{*0} \ell^- \bar{\nu}$ simultaneously by calculating the likelihood of the two modes separately and maximizing their product. The result of the simultaneous fit is shown in Fig. 11, along with the combined data and background level for the $\bar{B}^0 \to D^{*+} \ell^- \bar{\nu}$ and $B^- \to D^{*0} \ell^- \bar{\nu}$ modes. From a linear fit to these data we obtain the most precise measurement to date of $|V_{cb}|\mathcal{F}(1)$ and a^2 [55],

$$|V_{cb}|\mathcal{F}(1) = 0.0351 \pm 0.0019 \pm 0.0018 \pm 0.0008 \tag{41}$$
$$a^2 = 0.84 \pm 0.12 \pm 0.08, \tag{42}$$

independent of f_{+-}/f_{00}; where the first error is statistical, the second is systematic, and the third is due to the input lifetimes [56]. With the statistics of both modes combined we can also perform a three parameter fit, where the quadratic term in the expansion for $\mathcal{F}(y)$ is retained. The results of this fit are shown in the last row of Table VIII. For comparison, we provide in Table X the a^2 (and b) values predicted by the different decay models.

To extract a value of $|V_{cb}|$ one requires a prediction for $\mathcal{F}(1)$. Table XI gives the values of $|V_{cb}|$ extracted from our result using the most updated calculations of $\mathcal{F}(1)$. These results are consistent with previously published values of $|V_{cb}|$ (after scaling for differences in D and D^* branching fractions and B lifetimes) from both inclusive and exclusive analyses, although the unknown model dependence of the previous $|V_{cb}|$ extractions makes detailed comparisons difficult. Model dependent values of $|V_{cb}|$ in Table VII are also consistent with these results.

The product $|V_{cb}|\mathcal{F}(y)$ as determined from our fits to the combined $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ and $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ data is shown in Fig. 12. Fig. 12(a) shows the result obtained by keeping only the linear term in the expansion of $\mathcal{F}(y)$, while (b) shows the result for a second order polynomial. The dotted lines show the contours for 1σ variations of the fit parameters. All the lines are functions of the variable y, which is equal to the D^* energy to mass ratio, E_{D^*}/m_{D^*}, evaluated in the B rest frame. The points are the data. As discussed before, we do not measure y but rather $\tilde{y}$, which is E_{D^*}/m_{D^*} in the laboratory frame. We have therefore binned the data in the average of the maximum and minimum possible values of y for each measured $\tilde{y}$. This average is denoted by $y_A(\tilde{y})$. Unfortunately $y_A(\tilde{y})$ is not an unbiased estimator or y, and although $y_A(\tilde{y})$ has the same kinematic limits as y, the smearing between these two variables is significant compared to the bin size. Therefore, the data shown this way do not exactly correspond to the product $|V_{cb}|\mathcal{F}(y)$. Note, however, that the lines shown are not the result of a fit to these data points, but were obtained from our unbinned fit which correctly accounts for the boost between the B frame and the laboratory frame.

The shape of $\mathcal{F}(y)$ determined from the fit to the data can be used to obtain the functional form of $d\Gamma(\bar{B} \to D^*\ \ell\ \bar{\nu})/dq^2$; (where q^2 is the true q^2 of the decay, not the smeared variable measurable experimentally) which is useful to test hypotheses of factorization in B meson decay [21]. Fig. 13 shows $d\Gamma/dq^2$ as extracted from our data (upper solid line) along with the $d\Gamma/dq^2$ distributions derived from several popular models for comparison. The data points are also shown for illustrations. As discussed above we do not measure q^2 experimentally and the lines shown are not a result of a fit to the data shown in the figure. Values of $d\Gamma(\bar{B} \to D^*\ \ell\ \bar{\nu})/dq^2$ as determined from the data are given in Table XII for commonly used q^2 points.

VIII. CONCLUSIONS

Using the CLEO II data sample we have reconstructed both $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ and $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ decay modes in the D^* decay chains: $D^{*+} \to D^0\pi^+$ and $D^{*0} \to D^0\pi^0$ with $D^0 \to K^-\pi^+$. We have measured the exclusive branching fractions,

$$\mathcal{B}(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}) = (0.5/f_{00})[4.49 \pm 0.32(\text{stat.}) \pm 0.39(\text{sys.})]\% \tag{43}$$

$$\mathcal{B}(B^- \to D^{*0}\ \ell^-\ \bar{\nu}) = (0.5/f_{+-})[5.13 \pm 0.54(\text{stat.}) \pm 0.64(\text{sys.})]\%, \tag{44}$$

which depend on the relative production of charged and neutral B mesons at the $\Upsilon(4S)$ resonance. We have also presented these results in a way that requires no knowledge of

f_{+-}/f_{00}, but does depend on B lifetime measurements from other experiments. With the assumption of equal partial widths, $\Gamma(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}) = \Gamma(B^- \to D^{*0}\ \ell^-\ \bar{\nu})$, we have determined the product

$$\frac{f_{+-}}{f_{00}}\frac{\tau_{B^+}}{\tau_{B^0}} = 1.14 \pm 0.14(\text{stat.}) \pm 0.13(\text{sys.}), \tag{45}$$

and by using existing B lifetime ratio measurements [47] we have solved for $f_{+-}/f_{00} = 1.04 \pm 0.13(\text{stat.}) \pm 0.12(\text{sys.}) \pm 0.10(\text{lifetime})$. This measurement verifies at the 10% level the assumption that $f_{+-} = f_{00}$, widely used to calculate B meson branching fractions with data collected at the $\Upsilon(4S)$ resonance. Also with the assumptions of equal partial widths and $f_{+-} + f_{00} = 1$, and using external B lifetime measurements, we have determined the partial width

$$\Gamma(\bar{B} \to D^*\ \ell\ \bar{\nu}) = [29.9 \pm 1.9(\text{stat.}) \pm 2.7(\text{sys.}) \pm 2.0(\text{lifetime})]\text{ns}^{-1}, \tag{46}$$

independent of f_{+-}/f_{00}. From this width we have determined the $\bar{B} \to D^*\ \ell\ \bar{\nu}$ branching fractions with dependence only on the charged to neutral B lifetime ratio (but not the individual lifetimes),

$$\begin{aligned}\mathcal{B}(B^- \to D^{*0}\ \ell^-\ \bar{\nu}) &= 1.1 \times \mathcal{B}(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}) \\ &= [5.03 \pm 0.32(\text{stat.}) \pm 0.45(\text{sys.}) \pm 0.24(\text{lifetime})]\%. \end{aligned}\tag{47}$$

where the factor 1.1 carries no error because we are assuming that the ratio of branching fractions exactly equals the value used for the lifetime ratio (the uncertainty in our knowledge of the lifetime ratio results in the third error).

Taking advantage of theoretical constraints on the normalization and q^2 dependences of the form factors for the $\bar{B} \to D^*\ \ell\ \bar{\nu}$ decay provided by HQET, a combined fit to the differential decay distributions for $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ and $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ results in a determination of $|V_{cb}|\mathcal{F}(1)$,

$$|V_{cb}|\mathcal{F}(1) = 0.0351 \pm 0.0019(\text{stat.}) \pm 0.0018(\text{sys.}) \pm 0.0008(\text{lifetime}), \tag{48}$$

which is also independent of f_{+-}/f_{00}. The explicit dependence on the B^+ and B^0 lifetimes is given in Ref. [56]. The form factor normalization $\mathcal{F}(1)$ is believed to be predictable with small theoretical uncertainty, and this permits a precise determination of the CKM matrix element. Using the predictions for the normalization of $\mathcal{F}(1)$ [57,58], we obtain new values for $|V_{cb}|$.

IX. ACKNOWLEDGEMENTS

We gratefully acknowledge the effort of the CESR staff in providing us with excellent luminosity and running conditions. J.P.A. and P.S.D. thank the PYI program of the NSF, I.P.J.S. thanks the YI program of the NSF, G.E. thanks the Heisenberg Foundation, K.K.G., I.P.J.S., and T.S. thank the TNRLC, K.K.G., H.N.N., J.D.R., T.S. and H.Y. thank the OJI program of DOE and P.R. thanks the A.P. Sloan Foundation for support. This work was supported by the National Science Foundation and the U.S. Dept. of Energy.

TABLES

TABLE I. A summary of the events in the $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ Signal Region and Correlated Background Region. In the top 2 rows the first error is statistical and the second is the systematic uncertainty due to the choice of functions used to fit the data. The bottom two rows have been computed using Eq. (10) and the uncertainties in r_s and r_c have been propagated into the systematic error. The numbers in parenthesis are not independent measurements, but merely functions of the other numbers in the bottom two rows.

$D^{*+}l^-$	Signal Region	Correlated Background Region
δ_m Signal	$457 \pm 23 \pm 9$	$70 \pm 11 \pm 1$
Scaled δ_m Sideband	$47 \pm 6 \pm 1$	9 ± 5
Uncorrelated Background	5.0 ± 0.5	19 ± 2
Continuum	5 ± 7	0 ± 6
Correlated Lepton Fakes	2 ± 1	1 ± 1
$B \to D^*\tau\bar{\nu}$ and $B \to D^* D_s$	–	4 ± 3
Net Yield (N_S and N_C)	$398 \pm 25 \pm 9$	$37 \pm 14 \pm 1$
$N_s(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu})$	$376 \pm 27 \pm 16$	(6 ± 0.5)
$N_c(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu})$	$(22 \pm 9 \pm 11)$	$31 \pm 14 \pm 1$

TABLE II. A summary of the events in the $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ Signal Region and Correlated Background Region. In the top 2 rows the first error is statistical and the second is the systematic uncertainty due to the choice of functions used to fit the data. The bottom two rows have been computed using Eq. (10) and the uncertainties in r_s and r_c have been propagated into the systematic error. The numbers in parenthesis are not independent measurements, but merely functions of the other numbers in the bottom two rows.

$D^{*0}l^-$	Signal Region	Correlated Background Region
δ_m Signal	$476 \pm 25 \pm 10$	$94 \pm 15 \pm 1$
Scaled δ_m Sideband	$144 \pm 10 \pm 3$	34 ± 7
Uncorrelated Background	8 ± 1	23 ± 3
Continuum	6 ± 7	9 ± 11
Correlated Lepton Fakes	2 ± 2	1 ± 1
$B \to D^*\tau\bar{\nu}$ and $B \to D^* D_s$	-	2 ± 2
Net Yield (N_S and N_C)	$316 \pm 28 \pm 10$	$25 \pm 20 \pm 1$
$N_s(B^- \to D^{*0}\ \ell^-\ \bar{\nu})$	$302 \pm 32 \pm 13$	(5 ± 0.5)
$N_c(B^- \to D^{*0}\ \ell^-\ \bar{\nu})$	$(14 \pm 14 \pm 7)$	$20 \pm 20 \pm 1$

TABLE III. Contributions to the fractional errors of the detection efficiency for the $\bar{B}^0 \to D^{*+}\,\ell^-\,\bar{\nu}$ and $B^- \to D^{*0}\,\ell^-\,\bar{\nu}$ decays. Errors common to both modes are entered only once in the column labeled "Both".

Source	$\Delta\epsilon_s/\epsilon_s$ (%)		
	$\bar{B}^0 \to D^{*+}\,\ell^-\,\bar{\nu}$	Both	$B^- \to D^{*0}\,\ell^-\,\bar{\nu}$
Form Factors		3.0	
δ_m Sideband Normalization	0.8		1.7
Variation of Cuts	2.3		1.0
Lepton Efficiency		2.2	
$D^0 \to K^-\pi^+$ Efficiency		4.0 *	
Slow π^+ Efficiency		5.0	
Slow π^0/π^+ Efficiency Ratio	–		7.0
Total Uncorrelated	2.4		7.3
Total Correlated		7.4	

$*$ This error is correlated with the systematic error of the CLEO II $D^0 \to K^-\pi^+$ branching fraction.

TABLE IV. Results for product branching fractions.

Product Branching Fraction	Result $\times 10^4$
$\mathcal{B}(\Upsilon(4S) \to B^0\bar{B}^0)\mathcal{B}(\bar{B}^0 \to D^{*+}\,\ell^-\,\bar{\nu})\mathcal{B}(D^{*+} \to D^0\pi^+)\mathcal{B}(D^0 \to K^-\pi^+)$	$6.0 \pm 0.43 \pm 0.55$
$\mathcal{B}(\Upsilon(4S) \to B^+B^-)\mathcal{B}(B^- \to D^{*0}\,\ell^-\,\bar{\nu})\mathcal{B}(D^{*0} \to D^0\pi^0)\mathcal{B}(D^0 \to K^-\pi^+)$	$6.4 \pm 0.68 \pm 0.73$
$\mathcal{B}(\bar{B} \to D^{*+}\,X\,\ell^-\,\bar{\nu})\mathcal{B}(D^{*+} \to D^0\pi^+)\mathcal{B}(D^0 \to K^-\pi^+)$	$1.6 \pm 0.7 \pm 0.3$
$\mathcal{B}(\bar{B} \to D^{*0}\,X\,\ell^-\,\bar{\nu})\mathcal{B}(D^{*0} \to D^0\pi^0)\mathcal{B}(D^0 \to K^-\pi^+)$	$1.4 \pm 1.5 \pm 0.3$

TABLE V. Comparison with previously published results for $\mathcal{B}(\bar{B}^0 \to D^{*+}\,\ell^-\,\bar{\nu})$ and $\mathcal{B}(B^- \to D^{*0}\,\ell^-\,\bar{\nu})$. All previous results have been rescaled to use the CLEO II D^0 and D^* branching ratios (except for the CLEO1.5 B^- number which depends nonlinearly on these branching fractions), and use $f_{+-} = f_{00} = 0.5$.

Experiment	$\mathcal{B}(\bar{B}^0 \to D^{*+}\,\ell^-\,\bar{\nu})$ (%)	$\mathcal{B}(B^- \to D^{*0}\,\ell^-\,\bar{\nu})$ (%)
CLEO II ($\tau_{B^+}/\tau_{B^0} = 1.10 \pm 0.11$)	$4.57 \pm 0.29 \pm 0.41 \pm 0.22$ *	$5.03 \pm 0.32 \pm 0.45 \pm 0.24$ *
CLEO II ($f_{+-} = f_{00} = 0.5$)	$4.49 \pm 0.32 \pm 0.39$	$5.13 \pm 0.54 \pm 0.64$
ARGUS [9,8]	$4.9 \pm 0.5 \pm 0.6$	$6.4 \pm 1.5 \pm 1.4$
ARGUS [10]	$4.5 \pm 0.3 \pm 0.4$	
CLEO 1.5 [6,7]	$4.1 \pm 0.5 \pm 0.6$	$4.1 \pm 0.8 \pm 0.9$

$*$ When $\tau_{B^+}/\tau_{B^0} = 1.10$ is used to extract the CLEO II branching fractions, all errors for the two different exclusive modes are fully correlated.

TABLE VI. Predictions for the form factor ratios and their q^2 or y dependence, and the predictions and q^2 dependence for the common form factor A_1. The pole forms have been abbreviated as: $P_1 = 1 - q^2/6.34^2$ and $P_2 = 1 - q^2/6.73^2$ where q is in GeV. $\mathcal{E}(q^2)$ stands for $\exp[-0.03(q^2_{max} - q^2)]$. For the form factor ratios of the Neubert model [40] $N(y) = 2.5/(y+1)$.

	ISGW [31]	BSW [41]	KS [42]	Neubert [39]
$V(q^2)/A_1(q^2)$	1.27	$1.09P_2/P_1$	$1.00/P_1$	$N(y)[1.35 - 0.22(y-1) + 0.09(y-1)^2]$
$A_2(q^2)/A_1(q^2)$	1.14	1.06	$1.00/P_1$	$N(y)[0.79 + 0.15(y-1) - 0.04(y-1)^2]$
$A_1(q^2)$	$0.94\mathcal{E}(q^2)$	$0.65/P_2$	$0.70/P_1$	$0.86[2/(y+1)]^{0.6}$

TABLE VII. Model dependent predictions of the $\bar{B} \to D^* \ell \bar{\nu}$ partial width, detection efficiencies for each model, and $|V_{cb}|$ values derived from the measured partial width. An additional 3.5% systematic error due to B lifetime measurements is common to all the $|V_{cb}|$ values given.

	ISGW [31]	BSW [41]	KS [42]	Neubert [39]
$\Gamma(\bar{B} \to D^* \ell \bar{\nu})\|V_{cb}\|^{-2}$	24.6 ps^{-1}	21.9 ps^{-1}	25.8 ps^{-1}	29.0 ps^{-1}
$\epsilon_{\rm model}/\epsilon_{\rm Neubert}$	1.00	0.97	0.98	1
$\|V_{cb}\| \times 10^3$	$34.8 \pm 1.1 \pm 1.6$	$37.5 \pm 1.2 \pm 1.7$	$34.4 \pm 1.1 \pm 1.5$	$32.2 \pm 1.0 \pm 1.4$

TABLE VIII. Values for $|V_{cb}|\mathcal{F}(1)$, a^2 and b determined by the fits to the $d\Gamma/d\tilde{y}$ distributions. The first error is statistical and the second systematic (including B lifetimes). The values for the individual modes, $\bar{B}^0 \to D^{*+} \ell^- \bar{\nu}$ and $B^- \to D^{*0} \ell^- \bar{\nu}$, assume $f_{00} = f_{+-} = 0.5$. The results of the combined fit (bottom two rows) are independent of this assumption.

Mode	$\|V_{cb}\|\mathcal{F}(1) \times 10^3$	a^2	b
$\bar{B}^0 \to D^{*+} \ell^- \bar{\nu}$	$34.7 \pm 2.5 \pm 1.8$	$0.80 \pm 0.17 \pm 0.08$	0.0
$B^- \to D^{*0} \ell^- \bar{\nu}$	$35.7 \pm 2.8 \pm 2.4$	$0.94 \pm 0.20 \pm 0.08$	0.0
$\bar{B} \to D^* \ell \bar{\nu}$	$35.1 \pm 1.9 \pm 1.9$	$0.84 \pm 0.13 \pm 0.08$	0.0
	$35.3 \pm 3.2 \pm 3.0$	$0.92 \pm 0.64 \pm 0.40$	$0.15 \pm 1.24 \pm 0.90$

TABLE IX. Estimates of the systematic error on V_{cb} and a^2 determined from the fits to the $d\Gamma/dy$ distribution. The fractional error on V_{cb} is the same as the fractional error on the product $|V_{cb}|\mathcal{F}(1)$ (the uncertainty in theoretical estimates of $\mathcal{F}(1)$ is not included in this table). The error on the absolute efficiency is taken from Table III, excluding the 4% contribution from $K\pi$ reconstruction, which is also excluded from the error for $\mathcal{B}_{D^0}$.

source	$\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$		$B^- \to D^{*0}\ \ell^-\ \bar{\nu}$		$\bar{B} \to D^*\ \ell\ \bar{\nu}$	
	$\Delta V_{cb}/V_{cb}$	$\Delta a^2/a^2$	$\Delta V_{cb}/V_{cb}$	$\Delta a^2/a^2$	$\Delta V_{cb}/V_{cb}$	$\Delta a^2/a^2$
B lifetime	2.9%		3.6%		2.3%	
Absolute Efficiency	3.3%		4.8%		3.3%	
$\mathcal{B}_{D^*} \times \mathcal{B}_{D^0}$	1.9%		3.4%		2.1%	
Slow π efficiency shape	1.2%	3.5%			0.8%	2.4%
Fitting systematics	3%	9%	3%	9%	3%	9%
Total	5.8%	9.7%	7.5%	9.0%	5.5%	9.3%

TABLE X. Model predictions for the form factor parameters a^2 and b. These values were determined by fitting the $d\Gamma/dy$ distributions predicted by each model. Results are given with and without b fixed to zero.

Parameter	This Exp.	ISGW [31]	BSW [41]	KS [42]	Neubert [39]
a^2 $(b=0)$	$0.84 \pm 0.13 \pm 0.08$	0.91	0.77	0.83	0.48
a^2	$0.92 \pm 0.64 \pm 0.40$	0.88	0.84	1.10	0.59
b	$0.15 \pm 1.24 \pm 0.90$	-0.1	0.1	0.6	0.28

TABLE XI. The most recent calculations of the normalization $\mathcal{F}(1)$ by theoretical models and values of $|V_{cb}|$ extracted from our result with a linear fit to the data.

Model	$\mathcal{F}(1)$	$\lvert V_{cb}\rvert \times 10^3$
Neubert [57]	0.93 ± 0.03	$37.7 \pm 2.0 \pm 2.1 \pm 1.2$
Shifman et al. [58]	0.89 ± 0.03	$39.4 \pm 2.1 \pm 2.2 \pm 1.3$

TABLE XII. Values for the differential decay rate at selected q^2 points. These were calculated from Eqs. (32) and (1) with the results of the maximum likelihood fit given in Eqs. (41) and (42). The first error is statistical, the second is the systematic uncertainty due to $\mathcal{B}_{D^*} \times \mathcal{B}_{D^0}$, the third is the systematic uncertainty due to B lifetimes and the fourth is the rest of the systematic uncertainties as listed in Table IX.

q^2	$d\Gamma/dq^2$ ($\mathrm{ns}^{-1}\mathrm{GeV}^{-2}$)
$m_{\pi^+}^2$	$1.65 \pm 0.23 \pm 0.07 \pm 0.07 \pm 0.26$
m_{ρ}^2	$1.90 \pm 0.23 \pm 0.08 \pm 0.09 \pm 0.28$
$m_{a_1}^2$	$2.33 \pm 0.22 \pm 0.10 \pm 0.10 \pm 0.31$
$m_{D_s}^2$	$3.17 \pm 0.16 \pm 0.13 \pm 0.14 \pm 0.34$
$m_{D_s^*}^2$	$3.33 \pm 0.16 \pm 0.14 \pm 0.15 \pm 0.34$

FIGURES

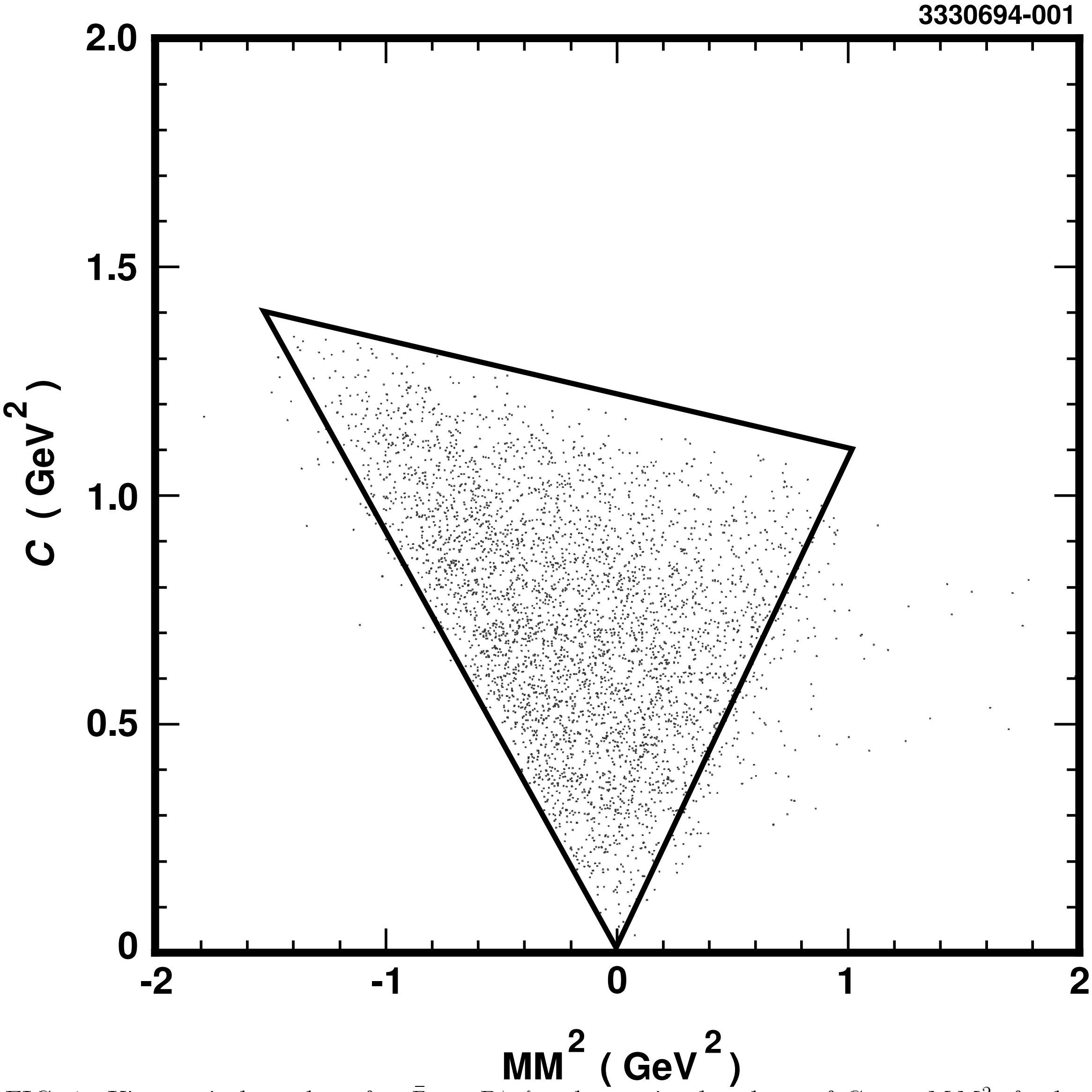

FIG. 1. Kinematic boundary for $\bar{B} \to D^* \ \ell \ \bar{\nu}$ decays in the plane of C vs. MM^2, for lepton momentum in the range $1.4 < |\mathbf{p}_\ell| < 2.4$ GeV (solid line). This is the Signal Region for this analysis. The dots, including those outside the kinematic boundary, are Monte Carlo signal events. Final state radiation and bremsstrahlung occasionally force reconstructed events outside of the kinematic boundary.

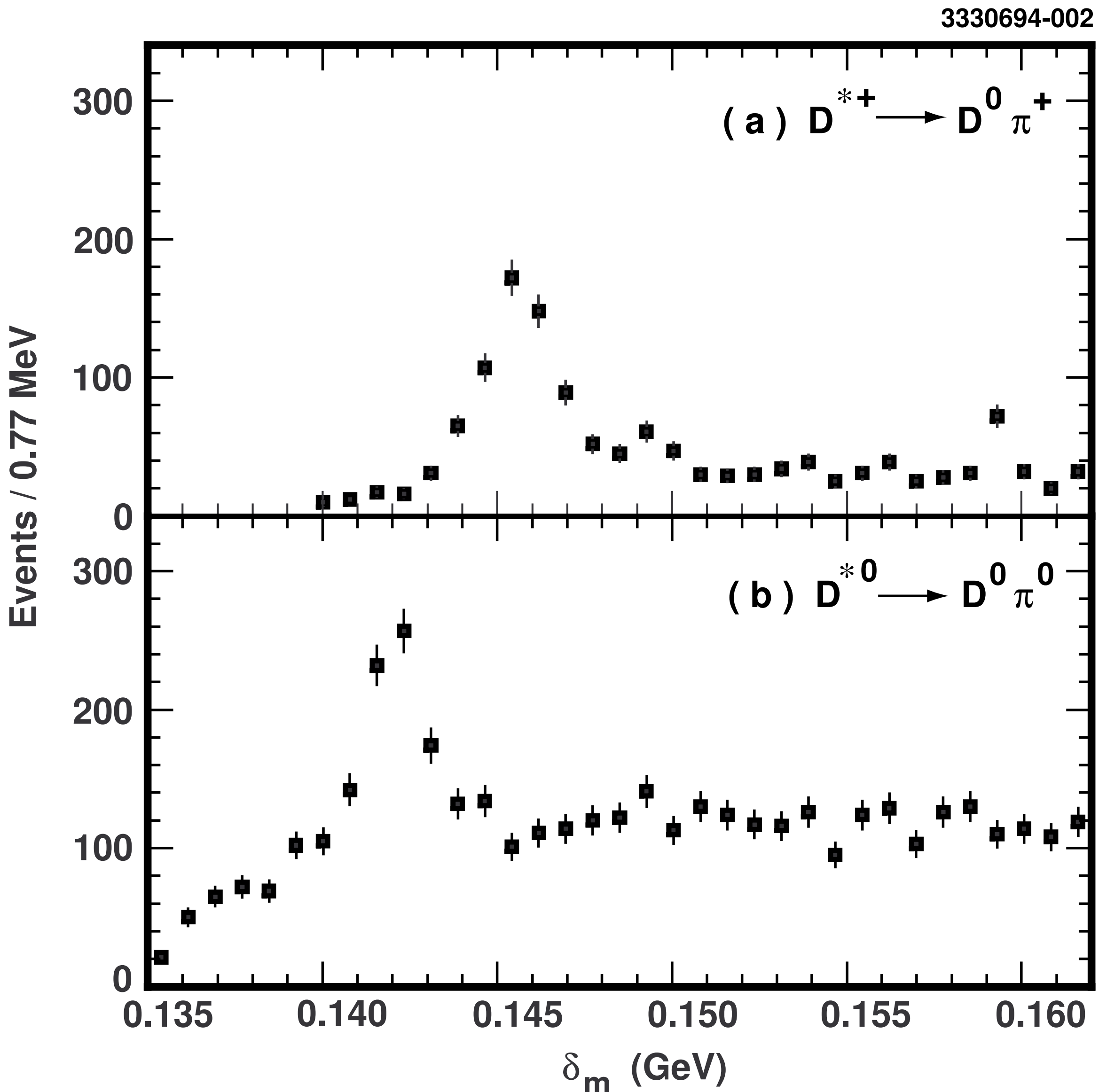

FIG. 2. Mass difference (δ_m) distributions of events falling in the Signal Region of the decay modes (a) $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ and (b) $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$. All candidates are required to have $M_{K\pi}$ within 100 MeV of the measured D^0 mass.

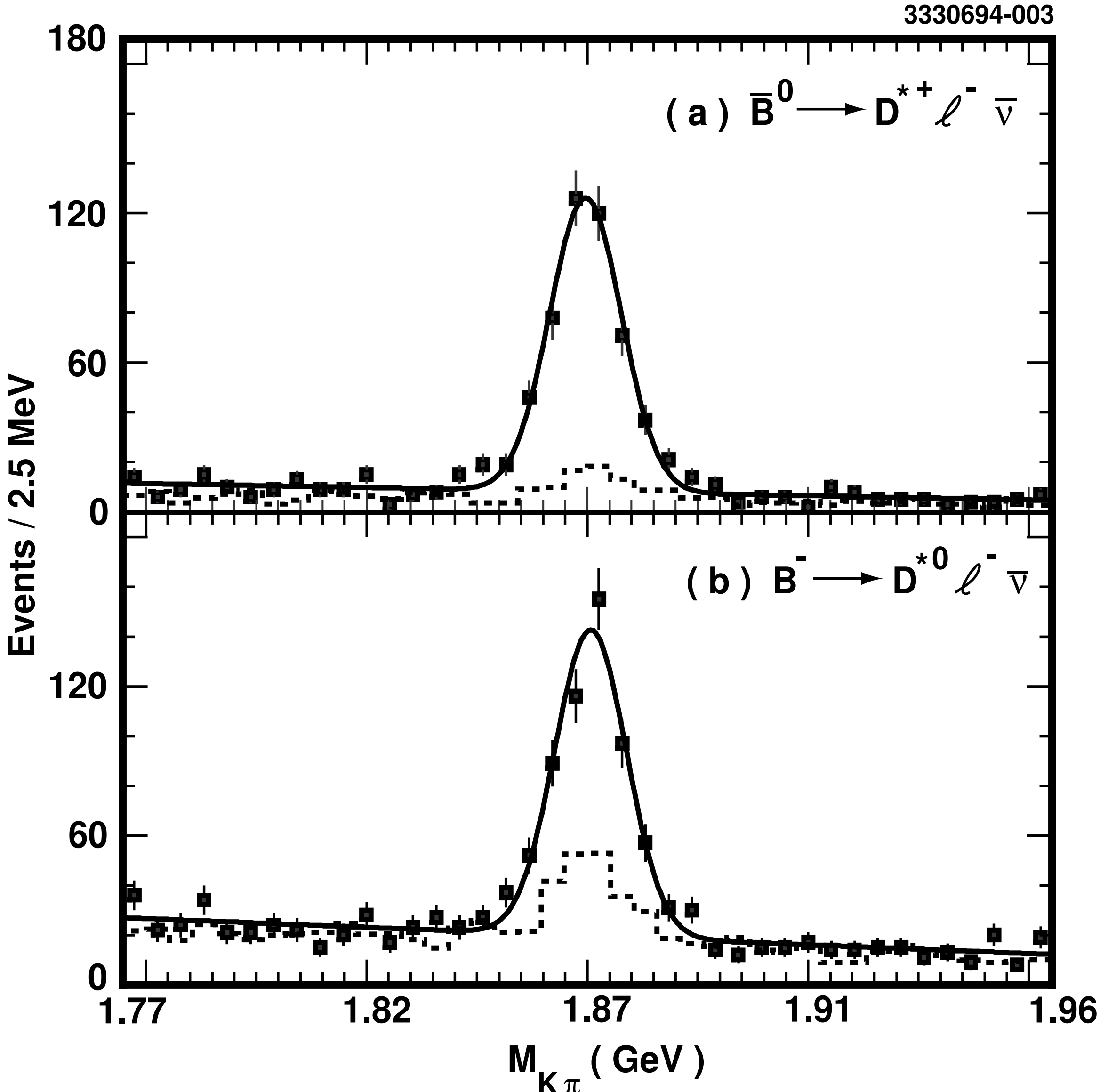

FIG. 3. Invariant mass of $K^-\pi^+$ combinations passing a δ_m cut in the Signal Region of the decay modes (a) $\bar{B}^0 \to D^{*+}\, \ell^-\, \bar{\nu}$ and (b) $B^- \to D^{*0}\, \ell^-\, \bar{\nu}$. The dashed histograms show the scaled $M_{K\pi}$ distributions of events in the Signal Region, but in the δ_m sideband. The solid lines show the fits used to determine the yields.

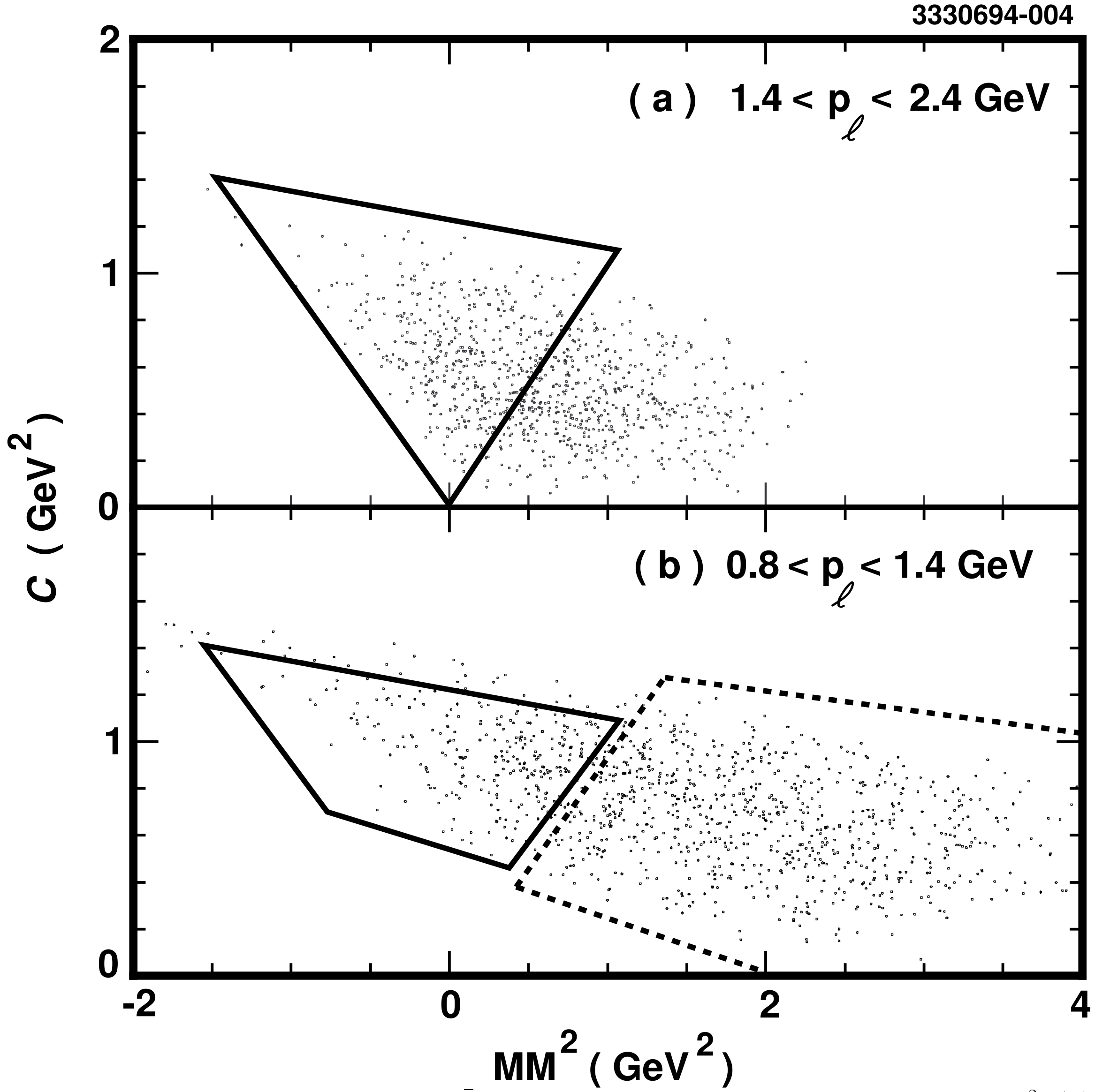

FIG. 4. Distribution of Monte Carlo $\bar{B} \to D^{**}\ \ell\ \bar{\nu}$ events in the plane of C vs. MM^2, (a) for lepton momentum in the range $1.4 < |\mathbf{p}_\ell| < 2.4$ GeV and (b) $0.8 < |\mathbf{p}_\ell| < 1.4$ GeV. The kinematic boundary for $\bar{B} \to D^{*}\ \ell\ \bar{\nu}$ decays is also shown in each plot (solid line). The dashed line in the right plot indicates the boundary of the Correlated Background Region.

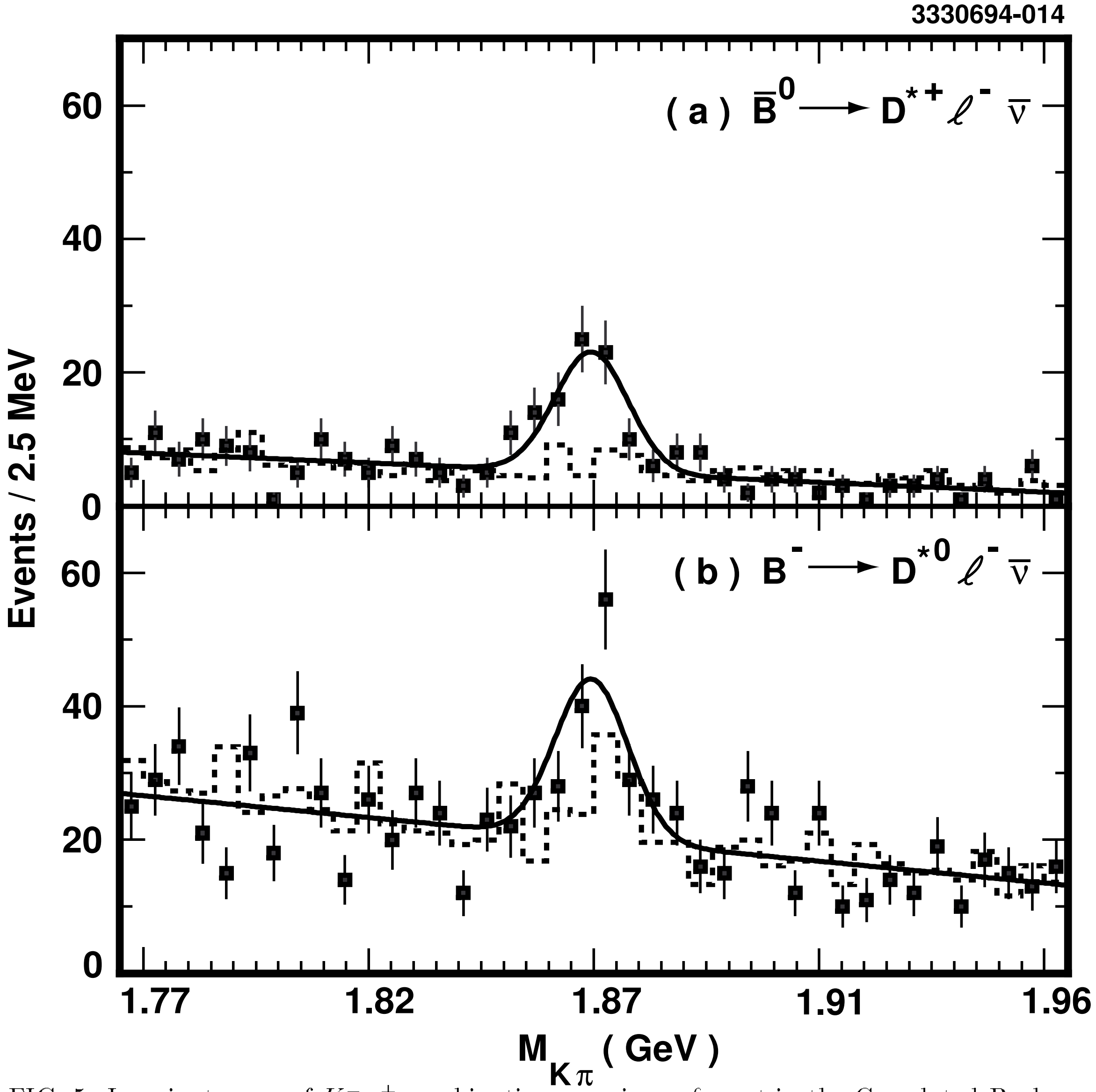

FIG. 5. Invariant mass of $K^-\pi^+$ combinations passing a δ_m cut in the Correlated Background Region of the decay modes (a) $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ and (b) $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$. The dashed histograms show the scaled $M_{K\pi}$ distributions of events in the δ_m sideband. The solid lines show the fits used to determine the yields.

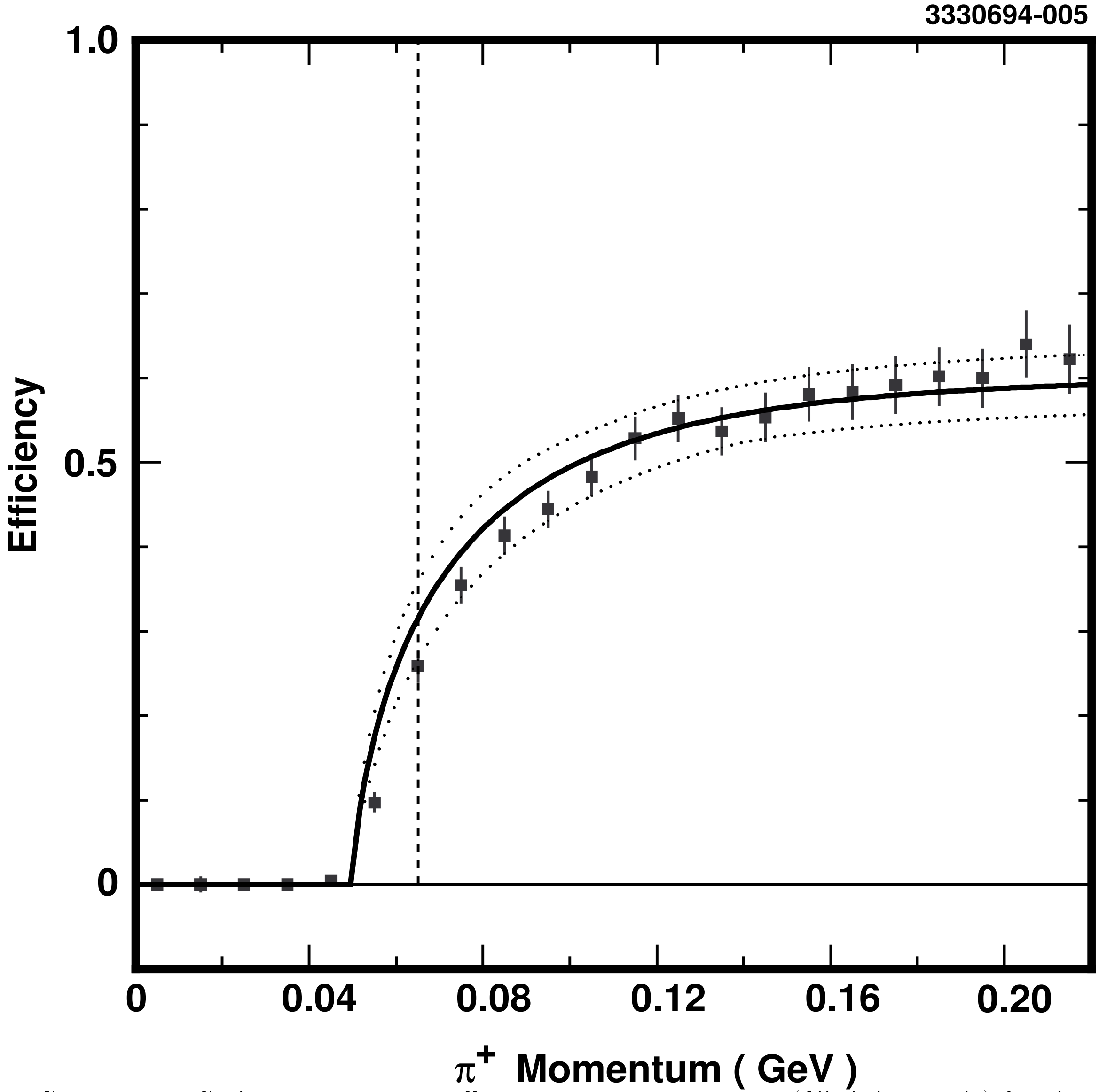

FIG. 6. Monte Carlo reconstruction efficiency versus momentum (filled diamonds) for charged slow pion candidates. Overlaid (solid curve) is the efficiency as determined from the inclusive D^{*+} decay angle distribution in data. The 1σ variations in the parameters of this curve are shown by the dotted lines. The dashed line indicates the minimum momentum cut used in this analysis. The efficiency shown here includes the geometric acceptance of 71%.

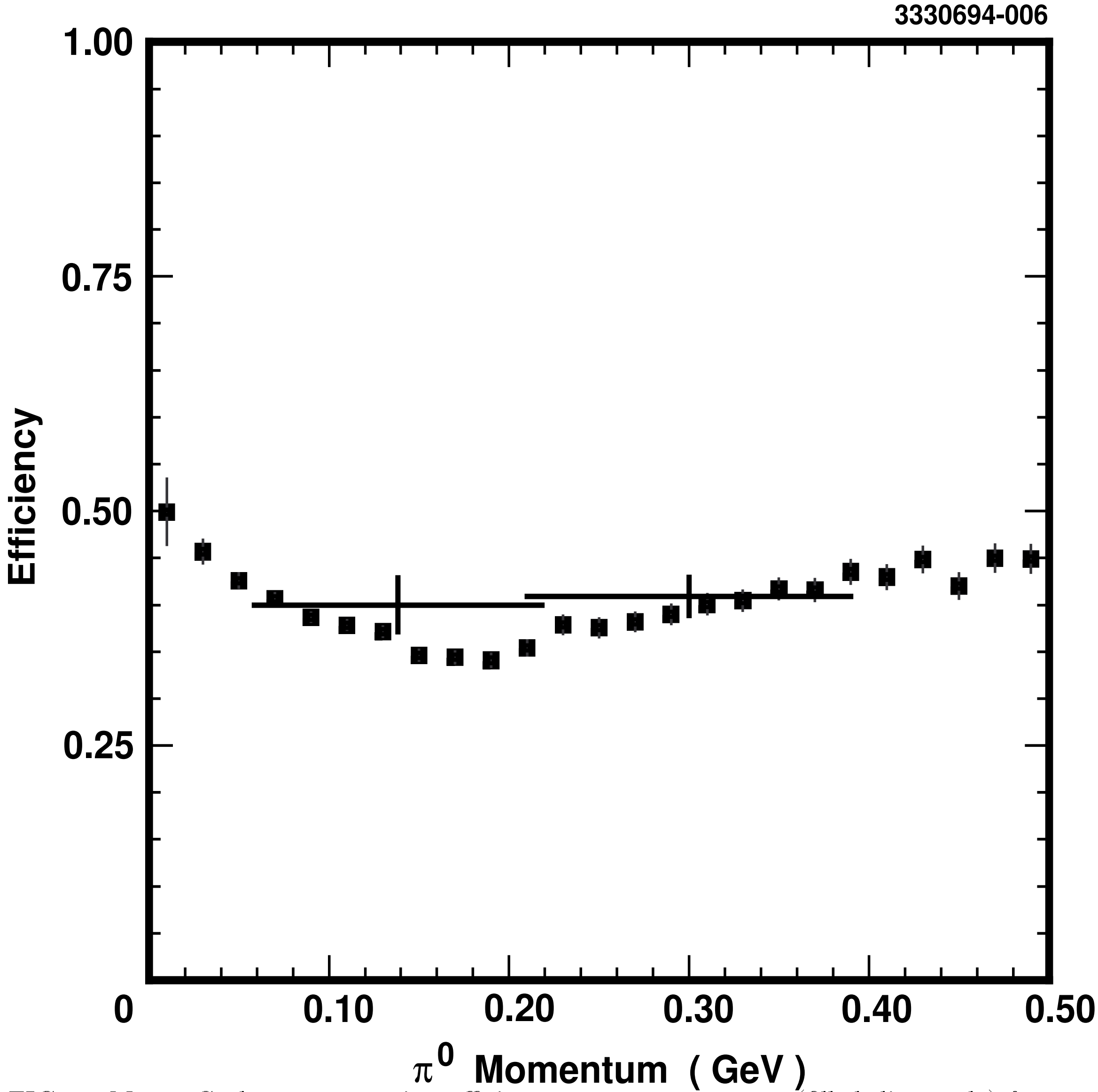

FIG. 7. Monte Carlo reconstruction efficiency versus momentum (filled diamonds) for neutral slow pion candidates. Overlaid (crosses) are the results of the η and K_S^0 studies which compare the ratio of neutral to charged pion efficiency in data and Monte Carlo. These two points were calculated assuming that the Monte Carlo exactly simulates the charged pion efficiency. The efficiency shown here includes the geometric acceptance (not constant with π^0 momentum), which accounts for much of the inefficiency.

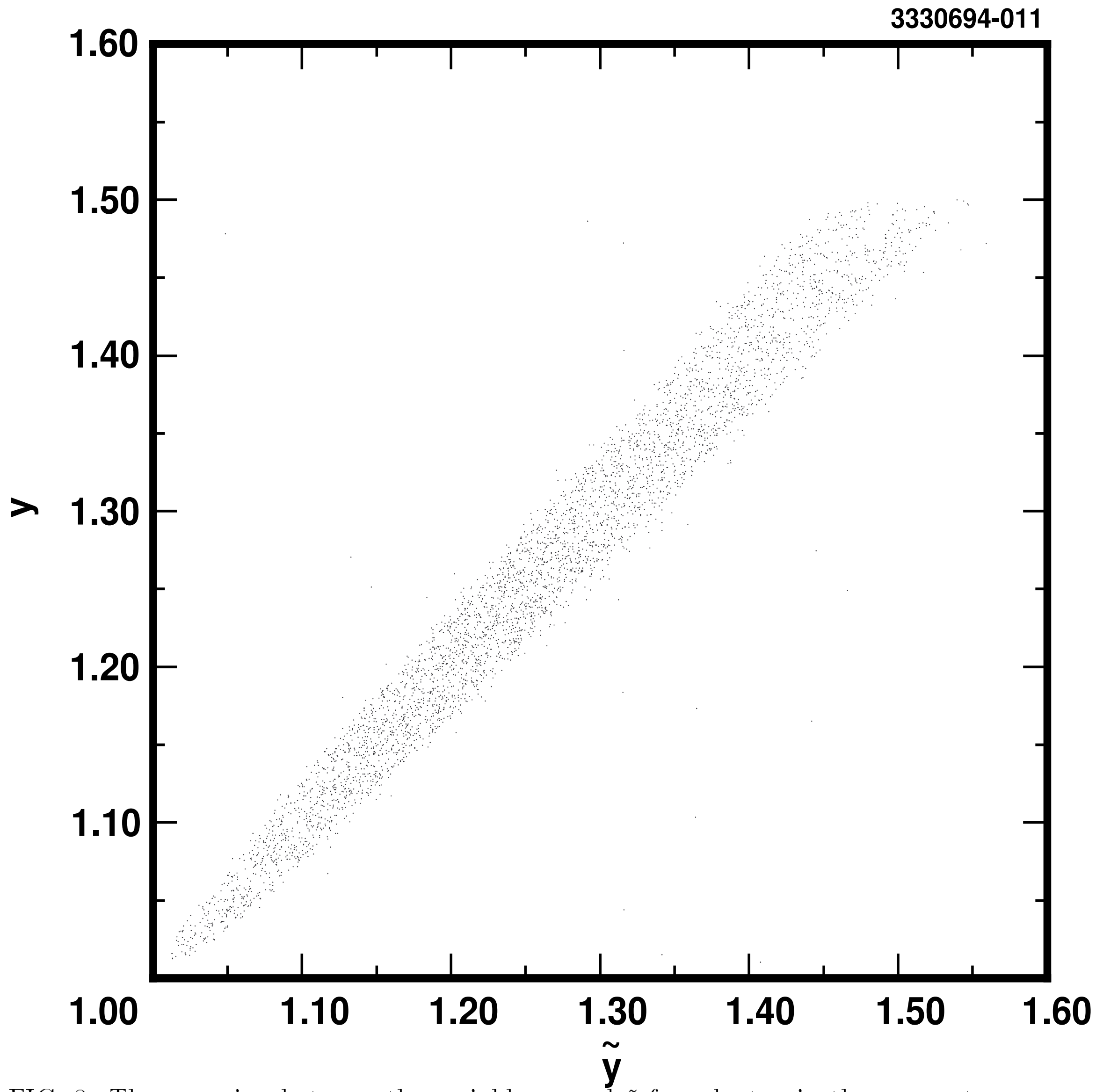

FIG. 8. The smearing between the variables y and $\tilde{y}$ for a lepton in the momentum range 1.4 to 2.4 GeV. The points are from Monte Carlo events generated according to the model of Ref. [39].

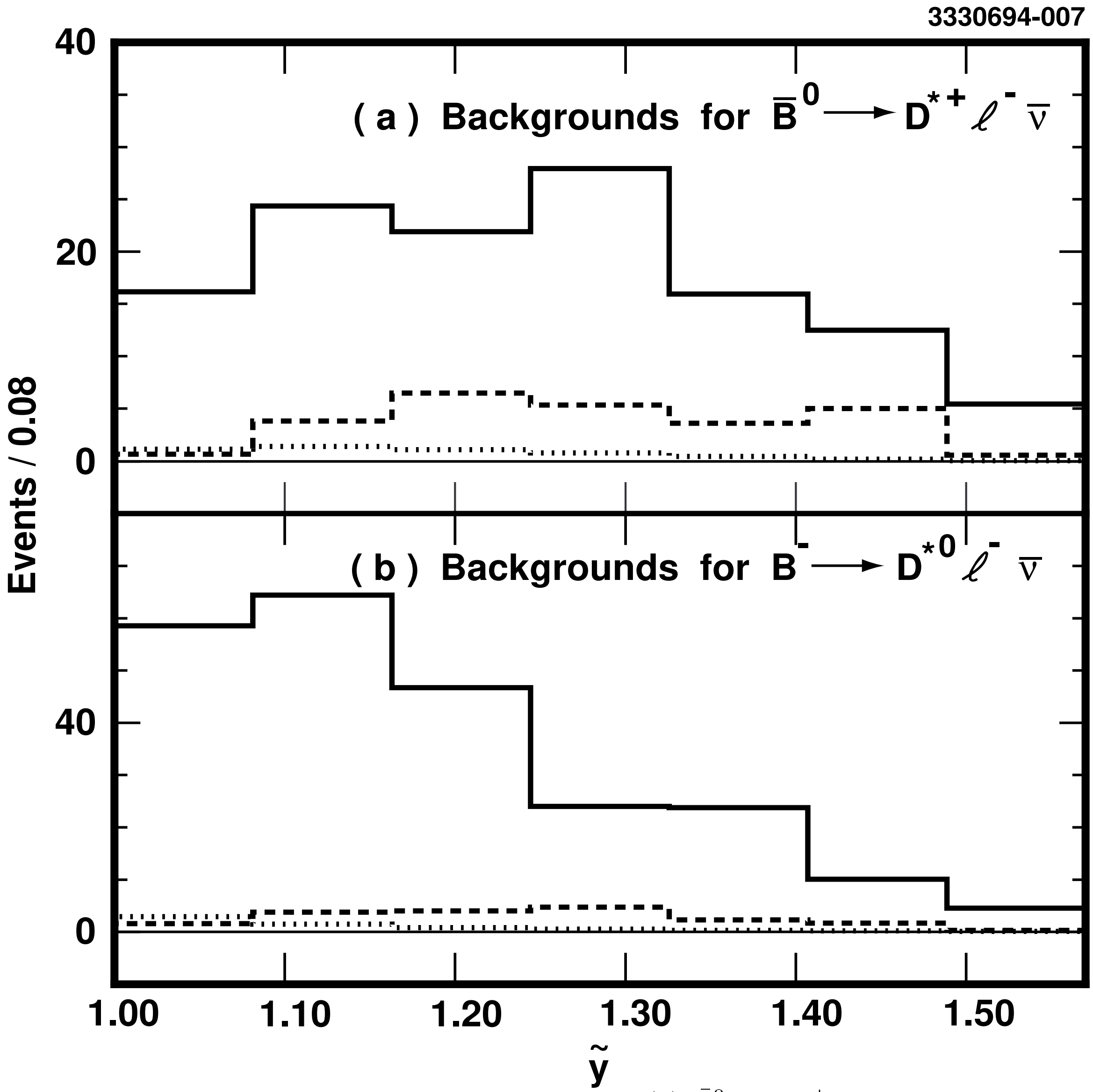

FIG. 9. The $\tilde{y}$ distributions of the backgrounds for (a) $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ events and for (b) $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ events. The solid line is the combinatoric background, the dashed line is the correlated background, and the dotted line is the uncorrelated background. The area of the curves are normalized to represent the background levels in the data.

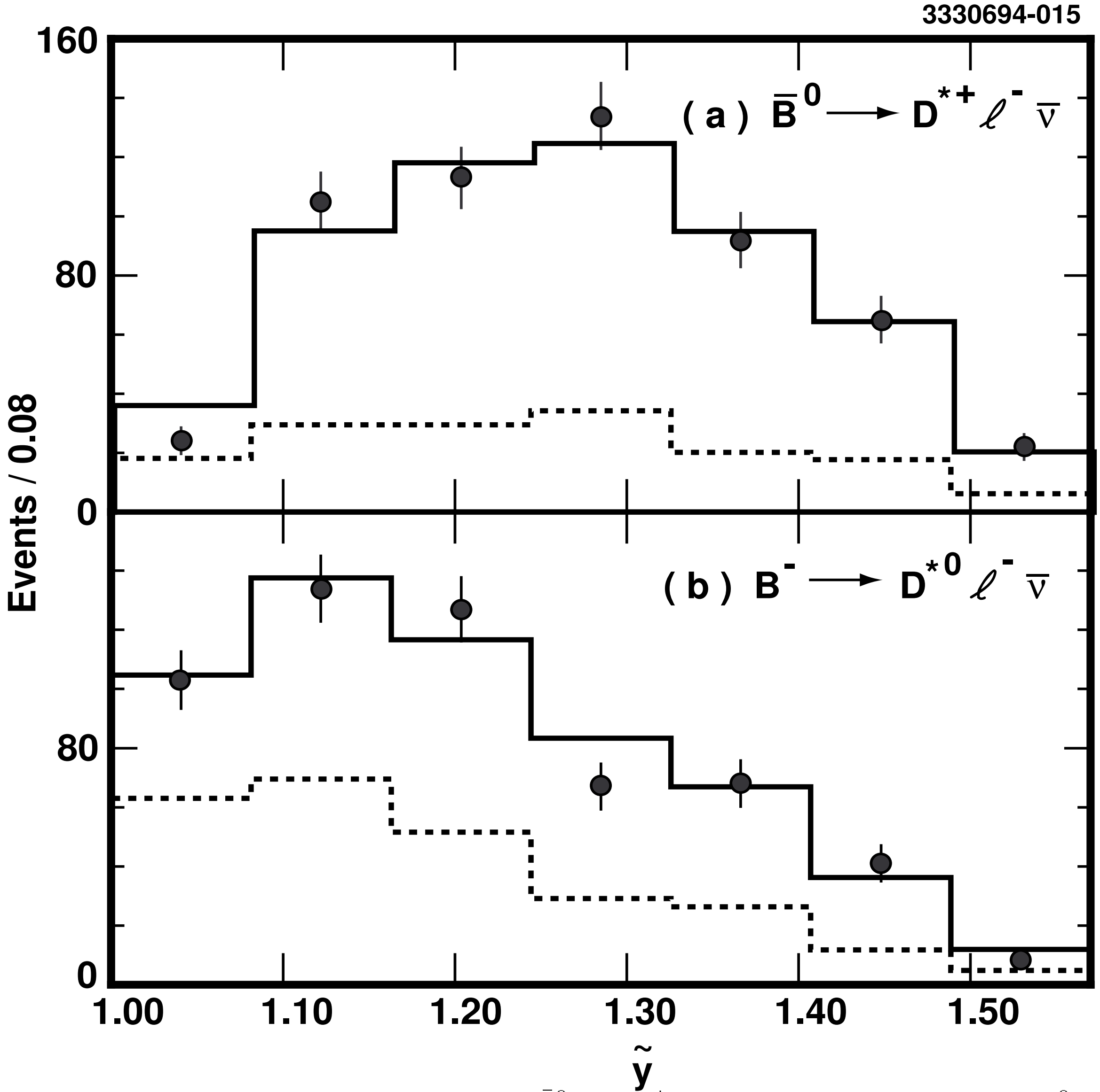

FIG. 10. Differential yield, $dN/d\tilde{y}$, for (a) $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ events and (b) $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ events in the data, with the projections of the unbinned fits superimposed. The solid histogram represents the result of the fit using $\mathcal{F}(y) = 1 - a^2(y-1)$. The dashed histogram shows the level of the background from all sources.

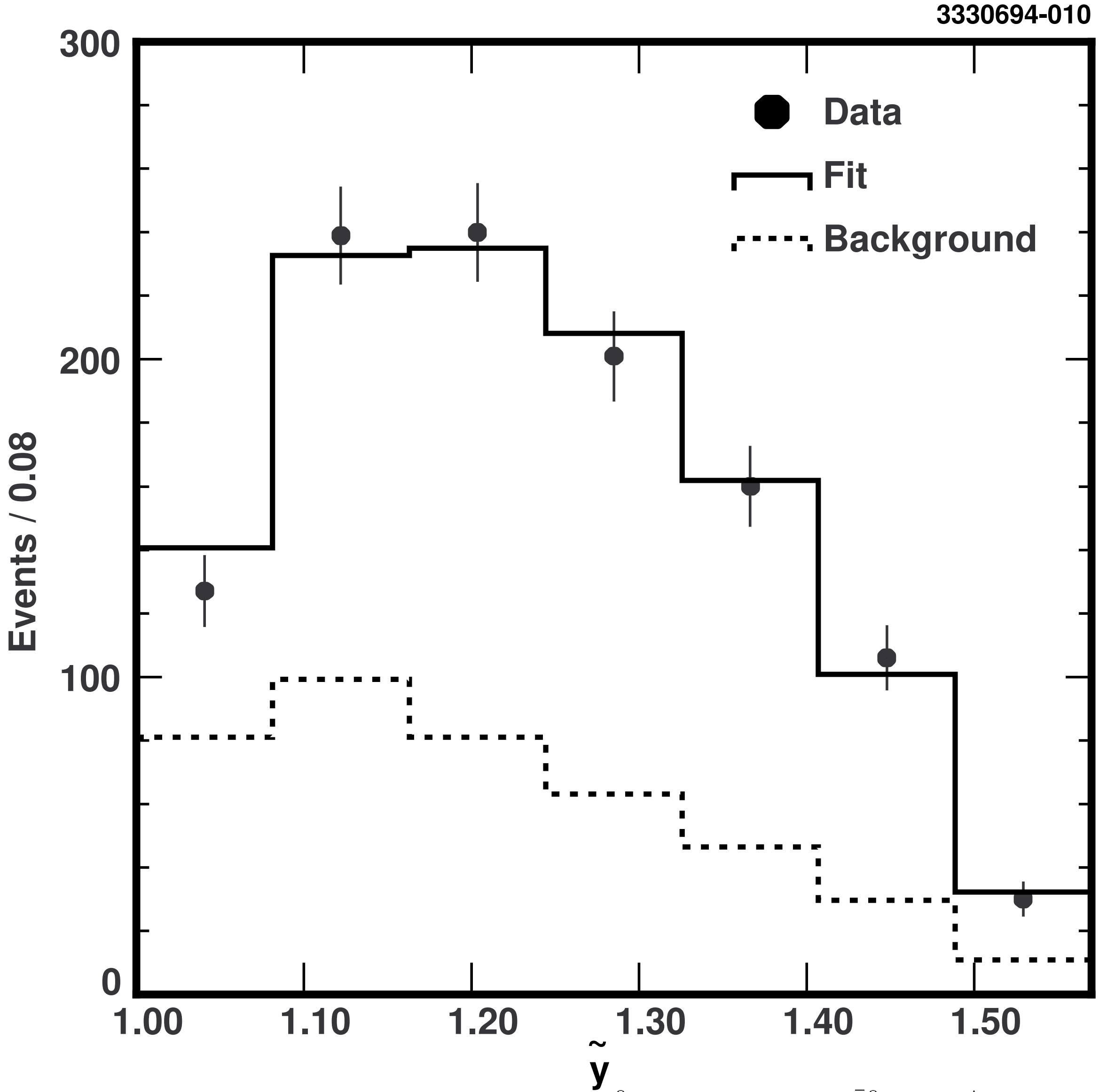

FIG. 11. Differential yield, $dN/d\tilde{y}$, for $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ events and $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ events combined with the projection of the unbinned simultaneous fit superimposed. The solid histogram represents the result of the fit using $\mathcal{F}(y) = 1 - a^2(y-1)$. The dashed histogram shows the sum of the background levels.

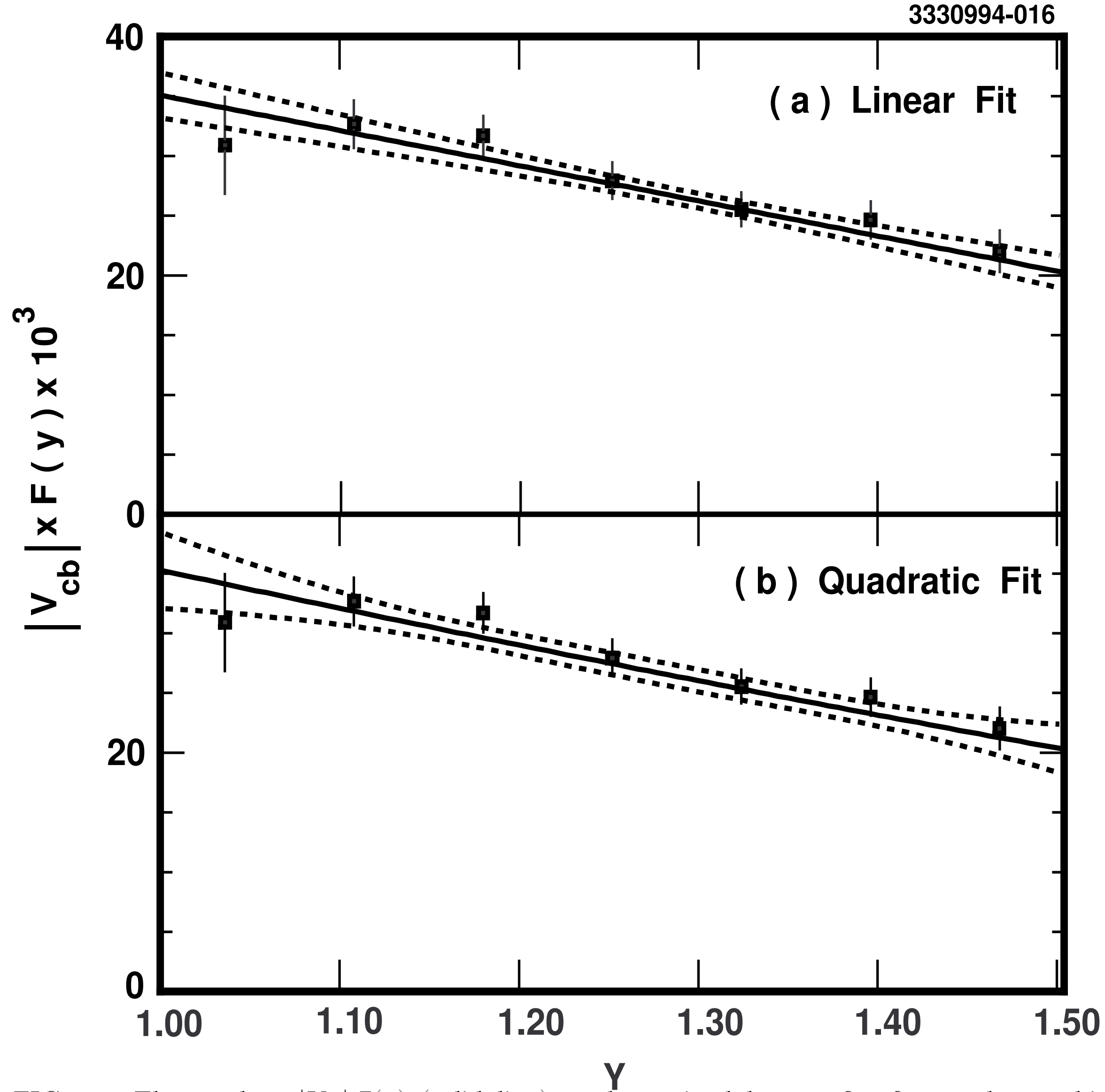

FIG. 12. The product $|V_{cb}|\mathcal{F}(y)$ (solid line) as determined by our fits fits to the combined $D^{*+}l^-$ and $D^{*0}l^-$ data, using (a) a linear expansion and (b) a quadratic expansion of the form factor $\mathcal{F}(y)$. The dotted lines show the contours for 1σ variations of the fit parameters. The points are data for the square root of the measured decay rate divided by the factors other than $\mathcal{F}(y)$ and V_{cb} of Eq. (32) (error bars are statistical only). As explained in the text, the data are binned in $y_A(\tilde{y})$, which is not an unbiased estimator of y. Therefore, the data points do not exactly correspond to the product $|V_{cb}|\mathcal{F}(y)$.

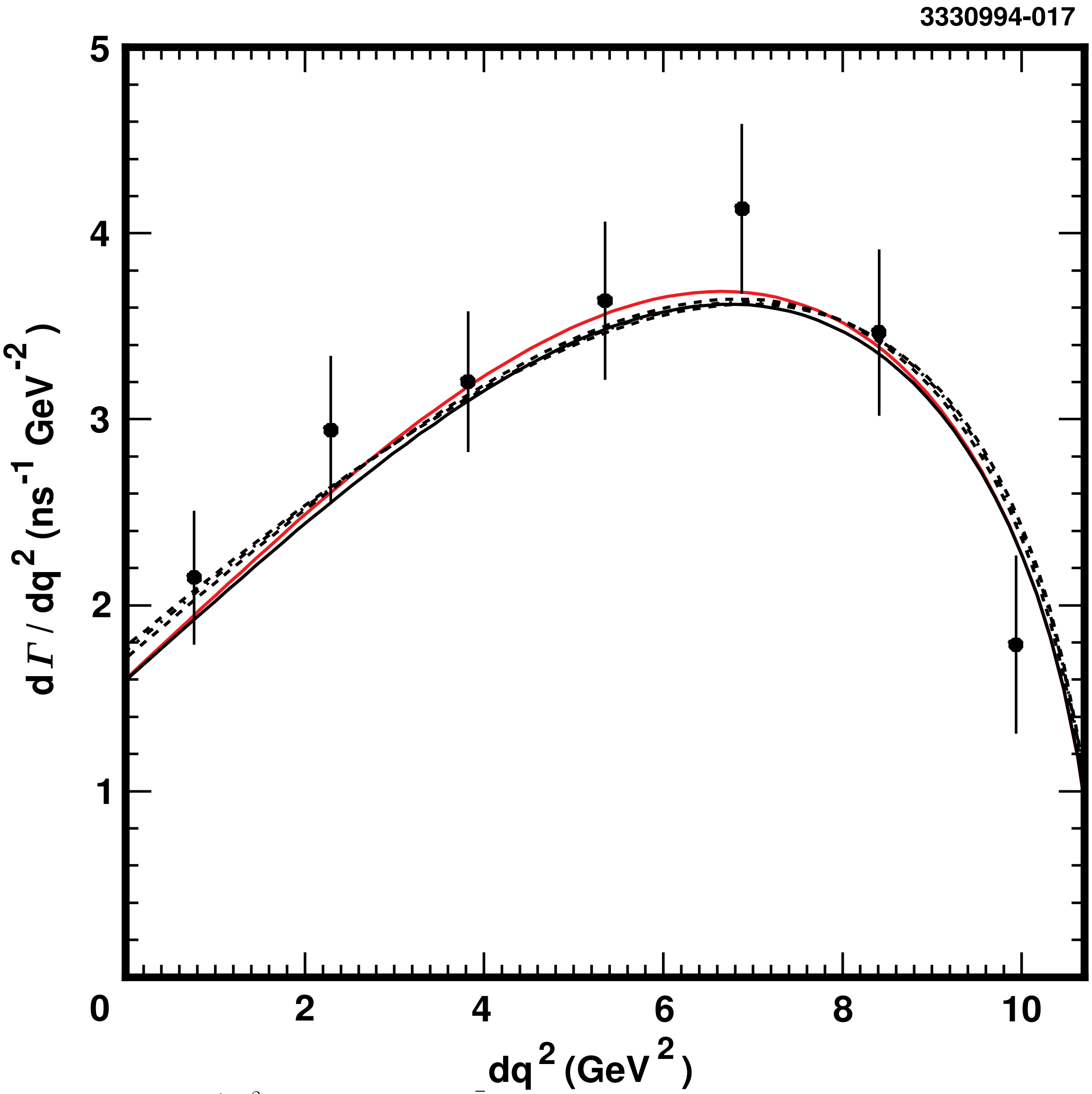

FIG. 13. The $d\Gamma/dq^2$ distribution for $\bar{B} \rightarrow D^* \ \ell \ \bar{\nu}$ decays. The upper solid line corresponds to the linear fit, and the lower solid line corresponds to the quadratic fit to our data. The dashed line shows $d\Gamma/dq^2$ for the ISGW model, the dotted line is the BSW model, the dot-dashed line is the pole-type $\xi(y)$. The data are shown as points. As explained in the text the data are binned in a smeared variable which is not an unbiased estimator of q^2. Therefore the data points do not exactly correspond to the fits shown.

REFERENCES

[1] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963); M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 653 (1973).

[2] ARGUS collaboration, H. Albrecht *et.al.*, Phys. Lett. B **249**, 359 (1990).

[3] CLEO collaboration, S. Henderson *et.al.*, Phys. Rev. D **45**, 2212 (1992).

[4] CLEO collaboration, J. Bartelt *et.al.*, CLEO Report No. CLEO CONF 93-19 (unpublished).

[5] ARGUS collaboration, H. Albrecht *et.al.*, Phys. Lett. B **197**, 452 (1987).

[6] CLEO collaboration, D. Bortoletto *et.al.*, Phys. Rev. Lett. **63**, 1667 (1989).

[7] CLEO collaboration, R. Fulton *et.al.*, Phys. Rev. D **43**, 651 (1991).

[8] ARGUS collaboration, H. Albrecht *et.al.*, Phys. Lett. B **275**, 195 (1992).

[9] ARGUS collaboration, H. Albrecht *et.al.*, Z. Phys. C **57**, 533 (1993).

[10] ARGUS collaboration, H. Albrecht *et.al.*, Phys. Lett. B **324**, 249 (1994).

[11] N. Isgur and M. Wise, Phys. Lett. B **232**, 113 (1989); N. Isgur and M. Wise, Phys. Lett. B **237**, 527 (1990); E. Eichten and B. Hill, Phys. Lett. B **234** 511 (1990); H. Georgi Phys. Lett. B **240** 447 (1990).

[12] We use the symbol ℓ and the term lepton to denote either an electron or a muon.

[13] M. Neubert, Phys. Lett. B **264**, 455 (1991).

[14] M. E. Luke, Phys. Lett. B **252**, 447 (1990); C. G. Boyd and D. E. Brahm, Phys. Lett. B **257**, 393 (1991).

[15] A. Falk, M. Neubert and M. Luke, Nucl. Phys B **388**, 3363 (1992); A. F. Falk and M. Neubert, Phys. Rev. D **47**, 2965 (1993); M. Neubert, Phys. Rev. D **47**, 4063 (1993).

[16] T. Mannel, CERN Report No. CERN-TH.7162/94 (unpublished).

[17] Throughout this paper, charge conjugate decay modes are implied.

[18] CLEO collaboration, F. Butler *et.al.*, Phys. Rev. Lett **69**, 2041 (1992).

[19] CLEO collaboration, D. S. Akerib *et.al.*, Phys. Rev. Lett **71**, 3070 (1993). CLEO reports two values for $\mathcal{B}(D^0 \to K^-\pi^+)$, a value obtained using efficiencies from a Monte Carlo the does not include final state radiation and a value obtained using efficiencies from a Monte Carlo that does include final state radiation. We use the former, because our Monte Carlo simulation of D^0 decays did not include final state radiation.

[20] We investigated the feasibility of including other channels such as $D^0 \to K^-\pi^+\pi^+\pi^-$ and $D^0 \to K^-\pi^+\pi^0$, but the statistical significance of these modes is small due to the increased combinatoric backgrounds. Additionally, inclusion of these modes introduces new systematic errors, which offset the small statistical gain. Thus the overall significance of our result would not improve with more D^0 decay modes.

[21] CLEO collaboration, D. Bortoletto *et.al.*, Phys. Rev. D **45**, 21 (1992).

[22] The Signal Region (MM^2, C) boundaries are defined by the points $(0,0)$, $(-1.5,1.4)$ and $(1.1,1.1)$.

[23] This corresponds to about 1.6 million $e^+e^- \to \Upsilon(4S)$ events plus 4.5 million other hadronic (continuum) final states.

[24] CLEO collaboration, Y. Kubota *et.al.*, Nucl. Instr. and Meth. A **320**, pp 66-113 (1992).

[25] The calorimeter resolution quoted is for the region $|\cos\theta_\gamma| < 0.71$. The resolution is slightly worse for $0.71 < |\cos\theta_\gamma| < 0.8$.

[26] To be classified as hadronic, an event must have at least 3 charged tracks originating near the interaction point, and the total visible energy of the event must be greater

than 0.15 E_{CM}, where E_{CM} is the center of mass energy.
[27] G. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).
[28] The δ_m distributions are fit to Gaussian signal functions and a phase space background shape given by $a_1(\delta_m - m_\pi)^{1/2} + a_2(\delta_m - m_\pi)^{3/2}$, where a_1 and a_2 are free parameters and m_π is the charged or neutral pion mass. The signal widths from such fits are $\sigma =$ 1.01 MeV and $\sigma = 0.87$ MeV for D^{*+} and D^{*0} respectively. The sideband scale factors calculated from these fits are 0.40 ± 0.03 for D^{*+} and 0.36 ± 0.01 for D^{*0}.
[29] Particle Data Group, K. Hikasa *et.al.*, Phys. Rev. D **45**, No. 11, Pt. II (1992).
[30] CLEO collaboration, D. Bortoletto *et.al.*, Phys. Rev. Lett. **69**, 2046 (1992).
[31] N. Isgur, D. Scora, B. Grinstein, and M. B. Wise, Phys. Rev. D **39**, 799 (1989).
[32] The Correlated Background Region is defined by the (MM^2, C) points $(0.5, 0.4)$ and $(1.4, 1.3)$.
[33] The relative abundances used are $^1\mathrm{P}_1$:$^3\mathrm{P}_2$:$^3\mathrm{P}_1$ = 63:21:16; Daryl J. Scora, Ph.D. Thesis, University of Toronto, 1993 p. 60 (unpublished).
[34] P. Heiliger and L. M. Sehgal, Phys. Lett. B **229**, 409 (1989); ALEPH Collaboration, D. Buskulic *et.al.*, Phys. Lett. B **298** 479 (1992).
[35] The tag lepton is required to have momentum above 1.8 GeV. This momentum cut excludes leptons that were not directly produced by a B meson decay.
[36] F. Gilman and R. Singleton, Jr., Phys. Rev. D **41**, 142 (1990).
[37] R. Brun *et.al.*, GEANT v. 3.14, CERN Report No. CERN DD/EE/84-1 (unpublished).
[38] CLEO Collaboration, S. Sanghera *et al.*, Phys. Rev. D **47** 791 (1993).
[39] M. Neubert, SLAC Report No. SLAC-PUB-6263 (to appear in Physics Reports); Z. Ligeti, Y. Nir, and M. Neubert Phys. Rev. D **49** 1302 (1994).
[40] By the "Neubert model" we mean the estimate of the shape of the form factor $\mathcal{F}(y)$ from Ref. [39]. This involves the use of a theoretically-inspired form for the Isgur-Wise function with one free parameter, $\rho^2 \approx 0.8$, which is estimated using QCD sum rules. This is not regarded as a precision estimate. Note that HQET makes no predictions about the shape of the Isgur-Wise function. The expressions for the A_1, A_2 and V form factors in this model have been provided in Table VI.
[41] M. Bauer, B. Stech, M. Wirbel, Z. Phys. C **29**, 637 (1985).
[42] J. G. Körner and G. A. Schuler, Z. Phys. C **38**, 511 (1989).
[43] M. Garcia-Sciveres, Ph.D. thesis, Cornell University, 1994 (unpublished).
[44] J. G. Layter *et al.*, Phys. Rev. D **7**, 2565 (1973).
[45] D.Besson and T.Skwarnicki, Annu. Rev. Nucl. Part. Sci. **43** 333-378 (1993).
[46] G. P. Lepage, Phys. Rev. D **42**, 3251 (1990).
[47] W. Venus, in: Lepton and Photon Interactions, XVI International Symposium, eds. Persis Drell and David Rubin, AIP Press, New York (1994).
The averages obtained from these measurements and used throughout this paper are:
$$\tau_{B^0} = 1.53 \pm 0.09 \text{ ps}$$
$$\tau_{B^+} = 1.68 \pm 0.12 \text{ ps}$$
$$\tau_{B^+}/\tau_{B^0} = 1.10 \pm 0.11.$$
[48] Some previous measurements used the $D^0 \to K^-\pi^-\pi^+\pi^+$ decay channel as well as $D^0 \to K^-\pi^+$, but we have scaled all results by the shift in the $D^0 \to K^-\pi^+$ branching fraction.
[49] An additional form factor is required for non-zero lepton mass. The e and μ masses are

neglected relative to the B and D^* masses.

[50] Neubert writes the form factor $\mathcal{F}(y)$ as $\eta_A \hat{\xi}(y)$, where η_A is a calculable radiative correction for hard gluon processes ($\eta_A = 0.986 \pm 0.006$ [39]) and $\hat{\xi}(y)$ is the non-perturbative part of the form factor.

[51] D. M. Schmidt, R. J. Morrison, M. S. Witherell, Nucl. Instr. and Meth. A **328**, 547 (1993).

[52] D. G. Cassel *et.al.*, Phys. Rev. D **24**, 2787 (1981).

[53] In this parameterization, whenever $b = 0$, a^2 is equivalent to the slope, $\hat{\rho}^2$, as defined in Ref. [39] p. 163.

[54] The event selection for the unbinned likelihood fits use $M_{K\pi}$ cuts not present in background estimates listed in Tables I and II. We therefore scale the numbers from these tables according to the change in δ_m sideband subtracted raw yields caused by the $M_{K\pi}$ cuts. This is an effect of order 1%.

[55] The correlation coefficient from the linear fit is $\rho_{|V_{cb}|,a^2} = 0.90$. For the quadratic fit the correlations are: $\rho_{|V_{cb}|,a^2} = 0.92$, $\rho_{|V_{cb}|,b} = 0.84$, $\rho_{a^2,b} = 0.98$.

[56] The result for $|V_{cb}|\mathcal{F}(1)$ from the fit to the combined data scales as $1/\sqrt{\tau_{B^+} + \tau_{B^0}}$, as long as τ_{B^+}/τ_{B^0} remains constant. A change in the lifetime ratio has a non-trivial effect because it modifies the weights of the charged and neutral B components in the $d\Gamma/d\tilde{y}$ fit. By performing different fits with varying lifetime ratios within 2σ of the central value [47] (keeping $\tau_{B^+} + \tau_{B^0}$ fixed) we find, $\Delta|V_{cb}|\mathcal{F}(1) = 0.0012x - 0.003x^2$, where $x = \tau_{B^+}/\tau_{B^0} - 1.1$.

[57] M. Neubert, CERN Preprint CERN-TH.7395/94 (unpublished).

[58] M. Shifman, N. Uraltsev and A. Vainshtein, University of Minnesota Report No. TPI-MINN-94/13-T (unpublished).