Semileptonic decay of the neutral B meson into a D^* meson: measurement of the branching fraction

Doctoral Dissertation

Cristiano Borean

Advisor: Professor Livio Piemontese

January 2003

Department of Physics
University of Ferrara

Semileptonic decay of the neutral B meson into a D^* meson: measurement of the branching fraction

Doctoral Dissertation

Cristiano Borean

Advisor: Professor Livio Piemontese

January 2003

Department of Physics
University of Ferrara

Dottorato di Ricerca in Fisica

XV Ciclo — Anni Accademici 2000–2002

This doctoral dissertation was written in American English, according to *Decreto Ministeriale n. 224 del 30 aprile 1999, Art. 6, Comma 4.*

To Sarah

"We shall not cease from exploration
and the end of all our exploring
will be to arrive where we started...
and know the place for the first time."
T.S. Eliot

Contents

Introduction

The *BABAR* experiment has been taking data on the $B\overline{B}$ system since 1999, at the PEP-II asymmetric collider, and has collected at the end of year 2002 about 82×10^6 $B\overline{B}$ events. The main purpose of the experiment was to study the CP violation of the b quark: one important result is the discovery of CP violation in the neutral B mesons system [77], perfomed by the *BABAR* Collaboration at the PEP-II asymmetric B Factory and also by the Belle Collaboration.

CP violation physics is only a part of B physics reach. There are many important measurements that can be performed with the large amount of data collected. Both precision measurements and searches for new B decay channels can now be performed. The subject of this thesis is one such precision measurement: the branching fraction of the semileptonic neutral B decay $B^0 \to D^{*-}\ell^+\nu_\ell$.

The semileptonic decays of a B meson are a very interesting and important topic. The measured branching fraction of the exclusive decay $B^0 \to D^{*-}\ell^+\nu_\ell$, where ℓ is either an e or a μ, accounts for roughly 10% (adding the $B^0 \to D^{*-}e^+\nu_e$ and $B^0 \to D^{*-}\mu^+\nu_\mu$ exclusive modes) of neutral B deacys. A precise understanding of this decay mode therefore improves our overall knowledge of the B branching fractions. The $B^0 \to D^{*-}\ell^+\nu_\ell$ decay mode is also of particular physical interest, since the study of its differential decay rate can be used to measure the CKM matrix element $|V_{cb}|$ [3], one of the fundamental parameters of the standard model [1].

We perform this measurement via full reconstruction of the B^0 meson decay products, which entails observing the lepton and all the decay products of the D^{*-}, but not the weakly interacting neutrino. We reconstruct the charged D^* meson only when it decays to a D^0 and a charged pion. The D^0 meson is then reconstructed in 4 different decay channels, i.e. $D^0 \to K\pi$, $D^0 \to K\pi\pi\pi$, $D^0 \to K\pi\pi^0$ and $D^0 \to K_S\pi\pi$.

The branching fraction is defined as:

$$\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell) = \frac{\Gamma(B^0 \to D^{*-}\ell^+\nu_\ell)}{\Gamma(B^0)} = \frac{N_{B^0\to D^{*-}\ell^+\nu_\ell}}{N_{B^0}} \ ,$$

where $\Gamma(B^0 \to D^{*-}\ell^+\nu_\ell)$ is the amplitude of the $B^0 \to D^{*-}\ell^+\nu_\ell$ decay, $\Gamma(B^0)$ is the total decay amplitude of the B^0 meson, $N_{B^0\to D^{*-}\ell^+\nu_\ell}$ is the total number of $B^0 \to D^{*-}\ell^+\nu_\ell$ events (efficiency corrected) in the data sample used in the analysis and N_{B^0} is the total number of neutral B mesons in the same data sample. The latter is calculated from a total number of 38×10^6 $\Upsilon(4S) \to B\overline{B}$ decays collected in *BABAR* 2001 data taking period, . In order to extract the total number of $B^0 \to D^{*-}\ell^+\nu_\ell$ events we correct for apparatus efficiency discrepancies.

We calculate the contributions of all backgrounds and then perform background subtraction. Most of the backgrounds are evaluated directly from data with control samples, i.e. data samples enriched in a specific background we want to study and selected with criteria different from the ones applied to the $B^0 \to D^{*-}\ell^+\nu_\ell$ signal we want to extract.

Among all backgrounds there is a poorly-understood one, the $B \to D^{*-}\ell^+\nu_\ell X$ events, where X is a pion (either charged or neutral). The experimental evidence of the exclusive B decays contributing to this background is poor and the parametrization of these exclusive decays is known with large uncertainties. This makes the Monte Carlo predictions of this background unrealiable.

In order to extract the $B^0 \to D^{*-}\ell^+\nu_\ell$ branching fraction we perform a fit to the distribution of $\cos\theta_{B^0,D^*\ell}$, where $\theta_{B,D^*\ell}$ is the angle between the parent B^0 meson and the system $D^*\ell$ of its observed decay products in the $\Upsilon(4S)$ frame of reference, calculated assuming only a massless neutrino missing in the reconstructed event. This fit separates the signal $B^0 \to D^{*-}\ell^+\nu_\ell$ events from the $B \to D^{*-}\ell^+\nu_\ell X$ background.

The analysis is complex since it requires the study of many systematic uncertainties that are not common to all signal channels used to extract $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$, i.e. tracking efficiencies systematic uncertainties are very different depending on the D^0 mode used in the charged D^* reconstruction.

With the large amount of data collected by *BABAR*, and analyzed in this thesis, we obtain the most precise $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ single measurement so far.

The next step of this analysis consists in the measure of the $|V_{cb}|$ parameter of the CKM matrix and is underway.

This thesis is structured as follows.

In chapter 1 we present an overview of the theoretical description of this measurement and the status of both the existing $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ and $|V_{cb}|$ experimental results.

In chapter 2 we describe the *BABAR* detector features and performances.

In chapter 3 we give a detailed description of the Silicon Vertex Tracker detector layout. We describe also a part of the assembling procedure we got involved during graduate studies.

In chapter 4 we describe the event selection and reconstruction. We describe all sources of background for this analysis and show the selection criteria adopted to select signal and background control samples.

In chapter 5 we show how we perform background subtraction and extract the yields, i.e. the total number of signal and background events in a given data sample.

In chapter 6 we show the measurement of the $B^0 \to D^{*-}\ell^+\nu_\ell$ branching fraction and how we assign the systematic uncertainties of this measurement. We also discuss our results comparing them with the other $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ published measurements.

We then conclude summarizing the results obtained in this analysis.

Chapter 1

Phenomenology of the semileptonic $B^0 \to D^{*-}\ell^+\nu_\ell$ decay

The analysis we performed aims to measure the semileptonic B meson branching ratio $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$, where the charged lepton ℓ is either an electron or a muon. Together, the electron and the muon modes account for roughly 10% of neutral B decays. A precise understanding of this decay mode therefore improves our overall knowledge of the B branching fractions. The $B^0 \to D^{*-}\ell^+\nu_\ell$ decay mode is also of particular physical interest, since the study of its differential decay rate can be used to measure the CKM matrix element $|V_{cb}|$ (see 1.1) , one of the fundamental parameters of the standard model [1]. Although $|V_{cb}|$ has been measured with reasonably good precision, further improvements are highly desirable, because some tests of the CKM framework are highly sensitive to the value of $|V_{cb}|$. For example, to determine the constraint in the $\rho - \eta$ plane (see below) arising from measurements of the CP violating quantity $|\epsilon_K|$ measured in the neutral K decays, one needs the quantity $A^4 \propto |V_{cb}|^4$ [2]. This dependence on $|V_{cb}|$ means that the relative uncertainty on $|V_{cb}|$ is magnified by a factor four.

In figure 1.1 we can see a diagram for the semileptonic decay of a $\overline{B}$ meson. In this decay, a b quark is converted to a c quark through the emission of a virtual W^- boson. The CKM matrix element $|V_{cb}|$ is associated with this vertex. The c quark recombines with the spectator $\overline{d}$ quark to form a D^* vector meson. The virtual W^- decays to a negatively-charged lepton plus an antineutrino. The invariant mass of this system is q^2, the square of the four-momentum transferred. The large number of gluons is meant to suggest the complexity of non-perturbative strong-interaction effects. Figure 1.1 is valid for all three lepton generations. However, we only study

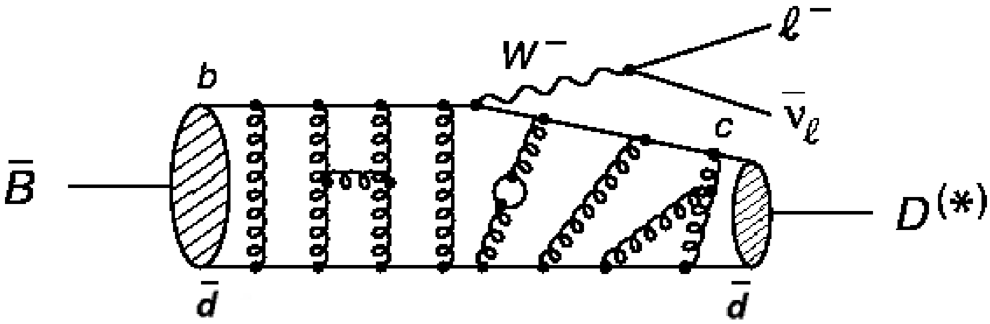


Figure 1.1: Semileptonic decay of a B meson. The complicated effects of the strong interactions can be parametrised by a small number of form factors, which are functions of q^2, the square of the mass of the virtual W boson.

those $B^0 \to D^{*-}\ell^+\nu_\ell$ decays in which the final-state charged lepton is either an electron (e) or a muon (μ). Either particle is light enough that we can neglect its mass, unlike the heavier tau (τ). Furthermore, we can observe electrons and muons directly with the *BABAR* detector, whereas taus must be reconstructed from their decay products. All τ decays involve a tau neutrino (ν_τ), which we cannot observe. The missing energy of the ν_τ makes the analysis of $B^0 \to D^{*-}\tau^+\nu_\tau$ difficult. Here we use the symbol ℓ to refer to either an electron or muon.

We refer to the decay $B^0 \to D^{*-}\ell^+\nu_\ell$ as a shorthand for both this decay and its charge-conjugate mode $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}$. Throughout this thesis, all references to particular decay modes imply the charge-conjugate mode as well.

1.1 The CKM Matrix

Flavor-changing quark transitions in the standard model are described by the coupling of the charged weak current

$$\mathcal{J}^\mu = \overline{u}_i \gamma^\mu \frac{1}{2}(1-\gamma_5)V_{ij}d_j \tag{1.1}$$

to the W boson via the interaction Lagrangian

$$\mathcal{L} = -\frac{g}{\sqrt{2}}\mathcal{J}^\mu W^+_\mu + \text{h.c.}\,. \tag{1.2}$$

The indices i and j in Equation 1.1 refer to the three quark generations so that $(u_1, u_2, u_3) \equiv (u, c, t)$, the annihilation operators for the three up-type quarks. Likewise, $(d_1, d_2, d_3) \equiv (d, s, b)$. The unitary matrix V_{ij} is the Cabibbo-Kobayashi-Maskawa (CKM) matrix [3]. This matrix is not diagonal for the familiar quark mass eigenstates. This feature gives rise to the generation-changing quark transitions in the standard model. The CKM matrix can be thought of as a rotation between the down-type mass eigenstates d, s, and b to a set of weak interaction eigenstates which couple to the up-type quarks only within each quark generation. This rotation can be written:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{1.3}$$

While a general unitary 3×3 matrix has 9 independent parameters, the phase invariance of the Lagrangian can be exploited to remove 5 relative phases between quark fields. These relative phases are not physically measurable quantities. A standard general parametrisation [4] writes the CKM matrix in terms of three real angles θ_{12}, θ_{13}, and θ_{23} and one phase δ_{13},

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} . \tag{1.4}$$

Here $c_{ij} \equiv \cos\theta_{ij}$ and $s_{ij} \equiv \sin\theta_{ij}$. The presence of the complex phase δ_{13} allows the existence of *CP* violation in the standard model: as there is a complex phase in the CKM matrix, the rates of charge-conjugate processes need not be exactly the same which implies *CP* violation. The values of the CKM matrix elements have been measured in a variety of experiments. Assuming the matrix is unitary, the bounds on the magnitude of the CKM matrix elements at the 90% confidence level are [5]:

$$|V| = \begin{pmatrix} 0.9741 \text{ to } 0.9756 & 0.219 \text{ to } 0.226 & 0.0025 \text{ to } 0.0048 \\ 0.219 \text{ to } 0.226 & 0.9732 \text{ to } 0.9748 & 0.038 \text{ to } 0.044 \\ 0.004 \text{ to } 0.014 & 0.037 \text{ to } 0.044 & 0.9990 \text{ to } 0.9993 \end{pmatrix} . \tag{1.5}$$

A hierarchy can be discerned in the values of the CKM matrix elements. The on-diagonal elements are close to unity, while the off-diagonal elements follow a pattern,

$$|V_{ud}|, |V_{cs}|, |V_{tb}| \gg |V_{us}|, |V_{cd}| \gg |V_{cb}|, |V_{ts}| \gg |V_{ub}|, |V_{td}| \tag{1.6}$$

Each successive set of elements is reduced from the previous by a factor of roughly 0.2. This hierarchy is exploited by the Wolfenstein parameterisation [6], in which one defines

$$s_{12} \equiv \lambda \; , \tag{1.7}$$

$$s_{23} \equiv A\lambda^2 \; , \tag{1.8}$$

$$s_{13}e^{-i\delta_{13}} \equiv A\lambda^3(\rho - i\eta) \; . \tag{1.9}$$

This parameterisation has been chosen such that λ accounts for the scale of the matrix elements, while the parameters A, ρ, and η are of order unity. In this scheme, the CKM matrix is written

$$V = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \; . \tag{1.10}$$

Unitary matrices have the property that each column is orthogonal to the complex conjugate of every other column; the same is true of the rows. That is, for $i \neq k$,

$$\sum_j V_{ij}V^*_{kj} = 0 \; , \tag{1.11}$$

$$\sum_j V_{ji}V^*_{jk} = 0 \; . \tag{1.12}$$

As each term in the sum is a complex number, these identities can be viewed as triangles in the complex plane.

It is instructive to draw three triangles (namely $\sum_i V_{id}V^*_{is} = 0$, $\sum_i V_{is}V^*_{ib} = 0$ and $\sum_i V_{id}V^*_{ib} = 0$), knowing the experimental values (within errors) for the various $|V_{ij}|$. In the first two triangles (figure 1.2), one side is much shorter than the other two, and so they almost collapse to a line. Since *CP* violation takes origin from the interference of amplitudes related to different sides of the triangle, then either the interference term is very small or the *CP* violating phase is very small. This would give an intuitive understanding of why *CP* violation is small in the leading K decays (the first triangle) and in the leading B_s decays (the second triangle).

The most exciting physics of *CP* violation lies in the B system, related to the third triangle. The openness of this triangle predicts large *CP* asymmetries in B decays as it has been confirmed by latest experimental results [7].

In terms of the Wolfenstein parameterisation, two of the three such triangles are near-degenerate, in the sense that one side is of higher order in λ than the other

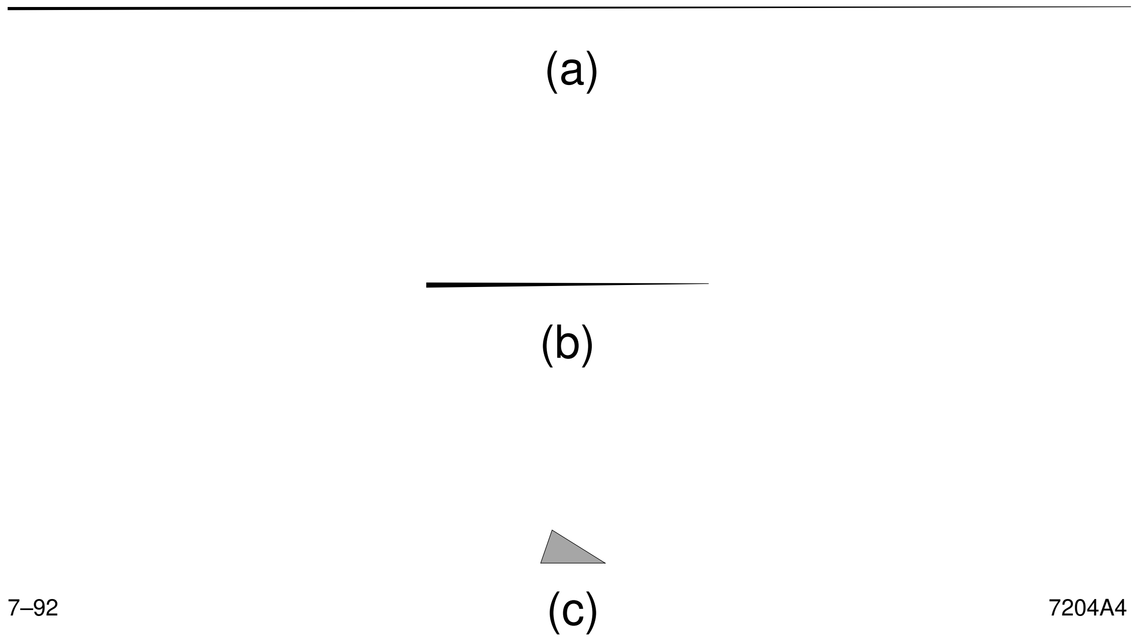


Figure 1.2: The three unitary triangles a)$\sum_i V_{id}V_{is}^* = 0$, b) $\sum_i V_{is}V_{ib}^* = 0$ and c) $\sum_i V_{id}V_{ib}^* = 0$, drawn to a common scale

two. The remaining triangle is commonly called the "Unitary Triangle", defined by applying Equation 1.12 to the first and third columns of the CKM matrix:

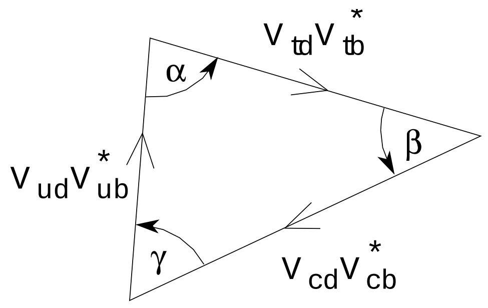
$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0\ . \tag{1.13}$$

The Unitarity Triangle is illustrated in figure 1.3. Figure1.3(a) shows the graphical representation of Equation 1.13. The angle α, β, and γ are also shown. The angles of this triangle (figure 1.3) can be expressed in terms of the CKM matrix elements as follows:

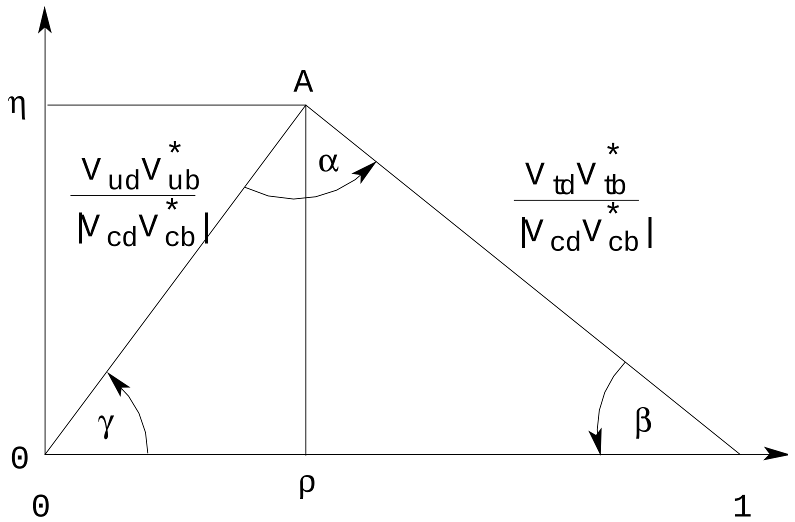
$$\alpha \equiv \arg\left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right]\ ,\ \beta \equiv \arg\left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right]\ ,\ \gamma \equiv \arg\left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right]\ . \tag{1.14}$$

The measurement of all this angles in the *BABAR* experiment will allow the over-determination of this triangle performing a consistency check of the standard model pattern for *CP* violation.

In the Wolfenstein parameterisation, to leading order in λ each side of this triangle is proportional to the product $A\lambda^3$. If we scale each side of the triangle down by this factor, we obtain a triangle with unit base length and an apex at the point (ρ, η). Experimental constraints on the unitarity triangle are commonly displayed in this representation, which I illustrate in figure 1.3(b).



(a)



(b)

Figure 1.3: The Unitary Triangle. Figure (a) is the representation of equation 1.13 in the complex plane with the length of the sides indicated in terms of the complex CKM matrix elements. Figure (b) show the Unitarity Triangle with all sides scaled by $V_{cb}^* V_{cd}$ with vertices (0,0), (1,0) and (ρ, η).

The precise measurement of the Unitarity Triangle is one of the central goals of heavy-flavor physics. This program is not motivated solely by a desire to determine the standard model parameters as well as possible, but to test the standard model itself. If the standard model is not a complete description of the physical world - for example if there is a fourth generation of quarks - then the Unitarity Triangle may not be a triangle at all. Alternatively, if experiments that study CKM parameters are affected by new physics, the Unitarity Triangle may appear not to close. Thus, the program of measurements of the Unitarity Triangle is also a search for signs of physics beyond the standard model. An examination of figure 1.3 shows the importance of the matrix element $|V_{cb}|$ in the study of the unitarity triangle. In figure 1.3(a), $|V_{cb}|$ sets the size of the base of the triangle. In terms of the Wolfenstein parameters, $|V_{cb}| \simeq A\lambda^2$ defines the size of the parameter A, as the parameter λ is known to the 1% level based on measurements of kaon decay. Thus, measurements of $|V_{cb}|$ sets the scale of the unitarity triangle via the scaling of figure 1.3(b).

1.2 Heavy-Quark Effective Theory

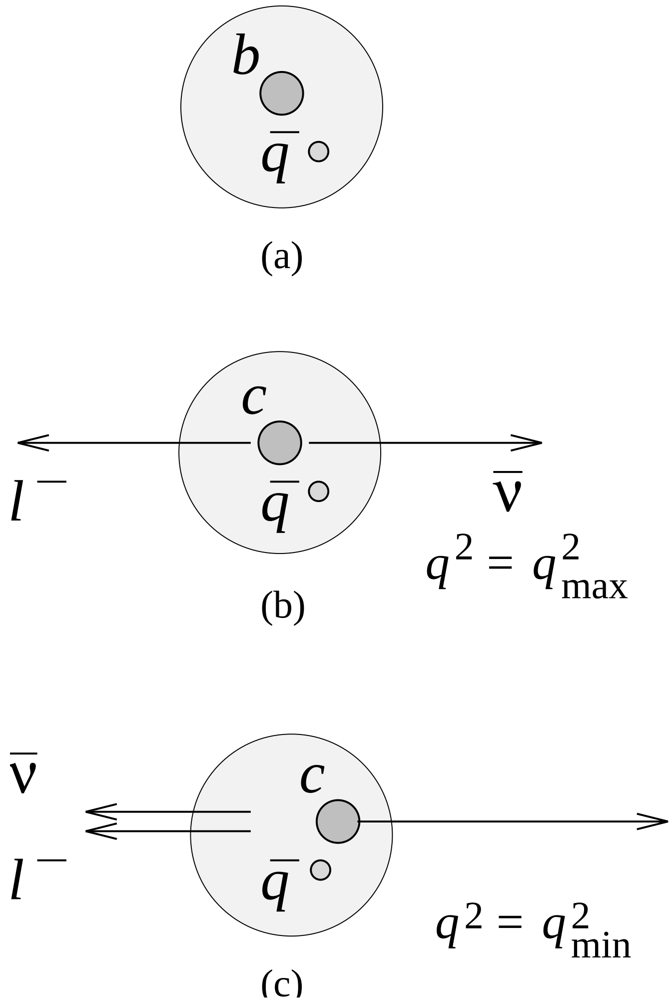


Figure 1.4: Kinematic configuration for the semileptonic decay of a B meson: (a) B meson before decay; (b) decay configuration for $q^2 = q^2_{\mathrm{max}}$ (zero recoil or $w = 1$, where the form factors are largest for producing a D^* vector meson in the final state; (c) decay configuration for $q^2 = q^2_{\mathrm{min}}$, where the form factors are smallest.

Heavy-quark effective theory [8], or HQET, is a theoretical approach used to describe meson systems that contain one heavy quark (b or c) and one light quark

(u, d or s). One can view such systems as analogues of the hydrogen atom. The hydrogen atom can be treated as an electron moving in the central electromagnetic field of the much heavier proton. Likewise, one can approach a heavy-light meson as a light quark moving in the central color field of the heavy quark. This description is illustrated for the case of a B meson in figure 1.4(a). As for the electron in the hydrogen atom, the light quark wavefunctions in the heavy-light meson are independent of the heavy quark mass and spin orientation, up to hyperfine corrections. Such corrections manifest themselves, for example, as the mass splittings between pseudoscalar and vector mesons, such as $\delta m = m_{D^*} - m_{D^0}$. To the extent that the light-quark wavefunctions do not depend on the heavy-quark properties, they must be the same in systems with different heavy-quark masses. This property is known as heavy-quark symmetry. This symmetry is exact in the limit that the heavy-quark mass is infinite, in particular, when it is much larger than Λ_{QCD}, the characteristic energy scale of QCD quark binding. This energy scale is on the order of 100 MeV.

A decay process involving hadrons has several possible final states depending of the final state quarks produced in the decay chain. A key ingredient in the parametrization of hadronic decays is the form factor. The form factor of a decay is simply a function that describes the probability of a particular final state configuration. In the decays of heavy-light mesons, the light quark starts in the central color field of the heavy meson. The heavy quark then decays, and in the process it generally acquires some motion relative to its initial state. Figure 1.4(c) illustrates an extreme case, in which the minimum possible q^2 has been transferred to the lepton-neutrino system in the semileptonic B decay. For a given q^2, the form factor tells us how likely it is that the light quark and the moving heavy quark will reassemble into a particular final state meson. The form factor is similar to an overlap integral between the wavefunction of the light quark in the initial, static color field and the wavefunction in the final, moving color field. A key result of HQET is that, in the heavy-quark symmetry limit, all decays from one heavy-light meson state to another are describe by a single, common form factor. As the initial and final-state light-quark wavefunctions do not depend on the heavy quark, the form factor cannot either. It can depend only on the motion of the final-state heavy quark relative to the heavy quark in the initial state. The universal form factor is known as the Isgur-Wise function [9], and is written $\xi(w)$. The Isgur-Wise function is written in terms of the variable w, which is defined as

$$w \equiv v \cdot v' = \frac{m_B^2 + m_{D^*}^2 - q^2}{2 m_B m_{D^*}} \, . \tag{1.15}$$

In this equation, v and v' are the four-velocities of the initial and final mesons in units of the speed of the light. In the heavy-quark symmetry limit, these meson four-velocities are identical to the heavy-quark four-velocities. HQET typically works in terms of the four-velocities of particles, rather than their four-momenta. The variable w explicitly refers to the relative motion of the initial and final heavy mesons, but it factors out the masses of these particles in such a way that w values across different decay modes can be related. The shape of $\xi(w)$ is not predicted by HQET. However, its value at $w = 1$ is well-defined. This value corresponds to a special kinematic point in HQET, at which $v = v'$: the final-state meson is at rest relative to the initial meson. It also corresponds to the maximum possible value of q^2. I show this configuration in figure 1.4(b). In this configuration, the color field of the final heavy-quark is stationary relative to the initial heavy quark. In the heavy-quark symmetry limit, the overlap between the initial and the final wavefunction is therefore perfect. The light quark is insensitive to the heavy-quark decay at $w = 1$. As a result, the Isgur-Wise function at $w = 1$ must be equal to unity, that is, $\xi(1) = 1$. Under the assumption of heavy-quark symmetry, the differential decay rate for $B^0 \to D^{*-}\ell^+\nu_\ell$ can be written in terms of $\xi(w)$ as

$$\begin{aligned} \frac{d\Gamma(B^0 \to D^{*-}\ell^+\nu_\ell)}{dw} &= \frac{G_F^2}{48\pi^3} m_{D^*}^3 (m_B - M_{D^*})^2 \sqrt{w^2-1}(w+1)^2 \\ &\times \left[1 + \frac{4w}{w+1}\frac{1-2wr+r^2}{(1-r)^2}\right] |V_{cb}|^2 \left[\xi(w)\right]^2 \ , \end{aligned} \tag{1.16}$$

where $r = m_{D^*}/m_B$. Outside the heavy-quark symmetry limit, this equation is no longer valid. The Isgur-Wise function $\xi(w)$ is replaced by general function $\mathcal{F}(w)$, which takes into account the hyperfine corrections in powers of $\Lambda_{\rm QCD}/m_b$ and $\Lambda_{\rm QCD}/m_c$. Unlike $\xi(1)$, the value of $\mathcal{F}(1)$ is shifted away from unity by these corrections. However, an important advantage of the $B^0 \to D^{*-}\ell^+\nu_\ell$ decay mode is that the first-order corrections to $\mathcal{F}(1)$ vanish as a consequence of Luke's theorem [10]. The leading corrections are of order $(\Lambda_{\rm QCD}/m_b)^2$, $(\Lambda_{\rm QCD}/m_c)^2$. As a result, $\mathcal{F}(1)$ is known theoretically to about the 5% level [11]:

$$\mathcal{F}(1) = 0.913 \pm 0.042 \tag{1.17}$$

Recently a new result on $\mathcal{F}(1)$ [12] has been obtained with lattice QCD calculations and shows a result compatible with equation 1.17 but with a smaller uncertainty [1]. The latest CLEO analysis [13] used this result to extract $|V_{cb}|$.

[1] $\mathcal{F}(1) = 0.919^{+0.030}_{-0.035}$ after applying a QED correction of +0.007 which reduces the uncertainty to 4%.

The basic approach used to determine $|V_{cb}|$ is to measure the differential cross section of $B^0 \to D^{*-}\ell^+\nu_\ell$, and then to fit the data with a simple functional form for $\mathcal{F}(w)$. The value of $\mathcal{F}(1)|V_{cb}|$ can then be inferred from the fit function at $w = 1$.

1.3 The $B^0 \to D^{*-}\ell^+\nu_\ell$ Decay

In this section, we provide parametrizations of the differential decay rate of $B^0 \to D^{*-}\ell^+\nu_\ell$ in terms of several form factor. This decay can be fully described by a set of three independent form factors, and there are a number of different ways to define such sets. We will mention a few of these sets and define the relationships between them. Finally, I will specify two HQET parameterisations in terms of these form factors. One is the " linear HQET " parameterisation assumed in a previous CLEO analysis [14] of the $B \to D^*\ell\nu$ form factors. The other is the theoretical result of Caprini, Lellouch, and Neubert [15], derived in the HQET framework through the application of dispersive bounds and unitarity constrains.

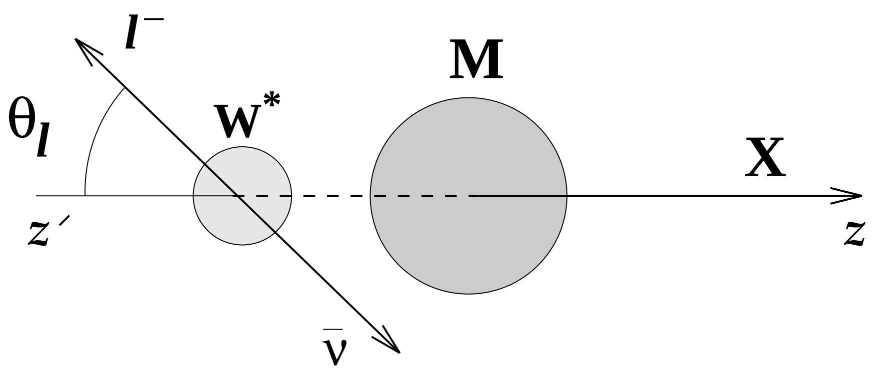


Figure 1.5: The polar angle θ_ℓ is defined in the rest frame of the virtual W (denoted here as W^*), in which the charged lepton and the neutrino are back-to-back. The angle is measured with respect to the axis z' pointed opposite to the momentum vector of the daughter meson X (in our case the D^*). In the M rest frame (M here stands for the B pseudoscalar meson) $p_{\vec{W}} = -p_{\vec{X}}$. Due to the Lorentz boost between the W and $M(orB)$ rest frames, leptons with small values of θ_ℓ have higher energy in the B rest frame than leptons with large values of θ_ℓ, at fixed q^2.

The general expression for semileptonic decays of pseudoscalar to vector mesons like $B^0 \to D^{*-}\ell^+\nu_\ell$ can be fully described by the four kinematic variables q^2, $\cos\theta_\ell$, $\cos\theta_{D^*}$ and χ. The first, q^2, is the invariant mass of the virtual W boson, or equivalently that of the lepton-neutrino system. The angle θ_ℓ is well illustrated in figure 1.5. The other two are decay angles and are illustrated in figure 1.6 together

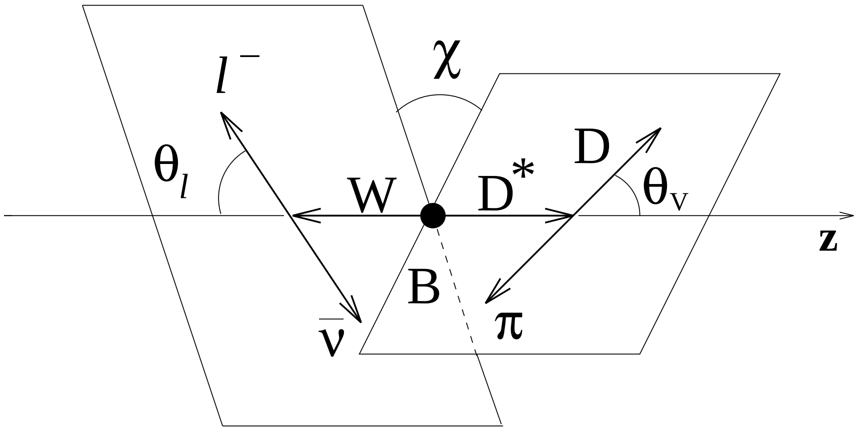


Figure 1.6: Definitions of the angles θ_ℓ, θ_V ($\equiv \theta_{D^*}$) and χin the decay $B^0 \to D^{*-}\ell^+\nu_\ell$. The lepton and the neutrino are drawn back to back because they are shown in the W^* rest frame. Similarly, the D and the π are shown in the D^* rest frame. The angle θ_ℓ is thus measured in the W, while θ_V is measured in the D^* rest frame. The azimuthal angle χ is measured between the W and the D^* decay planes.

with θ_ℓ. The angle θ_ℓ is that between the charged lepton in the W frame and the direction of motion of the W in the B rest frame. Similarly, θ_{D^*} is the angle between the D in the D^* rest frame and the direction of the D^* in the B rest frame. Finally, χ is the angle between the two decay planes; specifically, it is the angle in the B rest frame between the projections of the ℓ and the D onto the plane perpendicular to the D^*- W decay axis.

These definitions represent a particular choice between the two pseudoscalar daughters of the D^*. One can consistently replace the D in the definitions above with the slow π. In that case $\theta_{D^*} \to \pi - \theta_{D^*}$ and $\chi \to \pi - \chi$. All the formulae we give below are invariant under such transformations.

The differential decay rate for $B^0 \to D^{*-}\ell^+\nu_\ell$ can be written in terms of these four kinematic variables as [8]:

$$
\begin{aligned}
\frac{d\Gamma(B^0 \to D^{*-}\ell^+\nu_\ell; D^{*-} \to \overline{D}^0\pi^-)}{dq^2\, d\cos\theta_\ell\, d\cos\theta_{D^*}\, d\chi} = \frac{3G_F^2|V_{cb}|^2 p_{D^*} q^2}{8(4\pi)^4 m_B^2}\mathcal{B}(D^{*-} \to \overline{D}^0\pi^-) \\
\times \{ (1-\cos\theta_\ell)^2 \sin^2\theta_{D^*}[H_+(q^2)]^2 \\
+(1+\cos\theta_\ell)^2 \sin^2\theta_{D^*}[H_-(q^2)]^2 \\
+4\sin^2\theta_\ell \cos^2\theta_{D^*}[H_0(q^2)]^2 \\
-4\sin\theta_\ell(1-\cos\theta_\ell)\sin\theta D^* \cos\theta_{D^*}\cos\chi H_+(q^2)H_0(q^2) \\
+4\sin\theta_\ell(1+\cos\theta_\ell)\sin\theta D^* \cos\theta_{D^*}\cos\chi H_-(q^2)H_0(q^2) \\
-2\sin^2\theta_\ell \sin^2\theta_{D^*}\cos 2\chi H_+(q^2)H_-(q^2) \} \,.
\end{aligned}
\tag{1.18}
$$

This formula is fairly general. It assumes only the $(V-A)$ structure of the lepton current in the standard model and that the lepton mass may be neglected. As the B has spin zero, the W and D^* must have the same helicity in the decay. The functions $H_\pm(q^2)$, $H_0(q^2)$ are simply the helicity amplitudes for the three possible helicity states of virtual W. These helicity amplitudes, which are one complete set of $B^0 \to D^{*-}\ell^+\nu_\ell$ form factors, must be specified by a particular theory or model.

In figure 1.7 we can see the contribution to the differential decay rate from the different W helicities as computed by Neubert [8].

In equation 1.18, the quantity p_{D^*} is the magnitude of the three-momentum of the D^* meson in the B rest frame. It is related to w by the equation:

$$
p_D^* = m_{D^*}\sqrt{w^2-1} \tag{1.19}
$$

As w depends solely on q^2 and the masses of the B and D^* mesons (equation 1.15), p_D^* depends only on q^2, and not on any of the other kinematic variables.

Another common set of form factors is $A_1(q^2)$, $A_2(q^2)$ and $V(q^2)$. These functions are defined by writing the matrix element for the $B^0 \to D^{*-}\ell^+\nu_\ell$ decay in the form:

$$
\begin{aligned}
\langle D^*(p_{D^*},\epsilon)|(V^\mu - A^\mu)|B(p_B)\rangle &= \frac{2i\epsilon^{\mu\nu\alpha\beta}}{m_B+m_{D^*}}\epsilon^*_\nu(p_{D^*})_\alpha(p_B)_\beta V(q^2) \\
&- (m_B+m_{D^*})\epsilon^{*\mu}A_1(q^2) \\
&+ \frac{\epsilon^*\cdot q}{m_B+m_{D^*}}(p_B+p_{D^*})^\mu A_2(q^2) \,.
\end{aligned}
\tag{1.20}
$$

Here ϵ is the polarisation four-vector of the D^{*-}. This formula is a general one that assumes, again, only a $(V-A)$ current and that the lepton mass is negligible. The helicity amplitudes $H_\pm(q^2)$, $H_0(q^2)$ are related to the form factors $A_1(q^2)$, $A_2(q^2)$

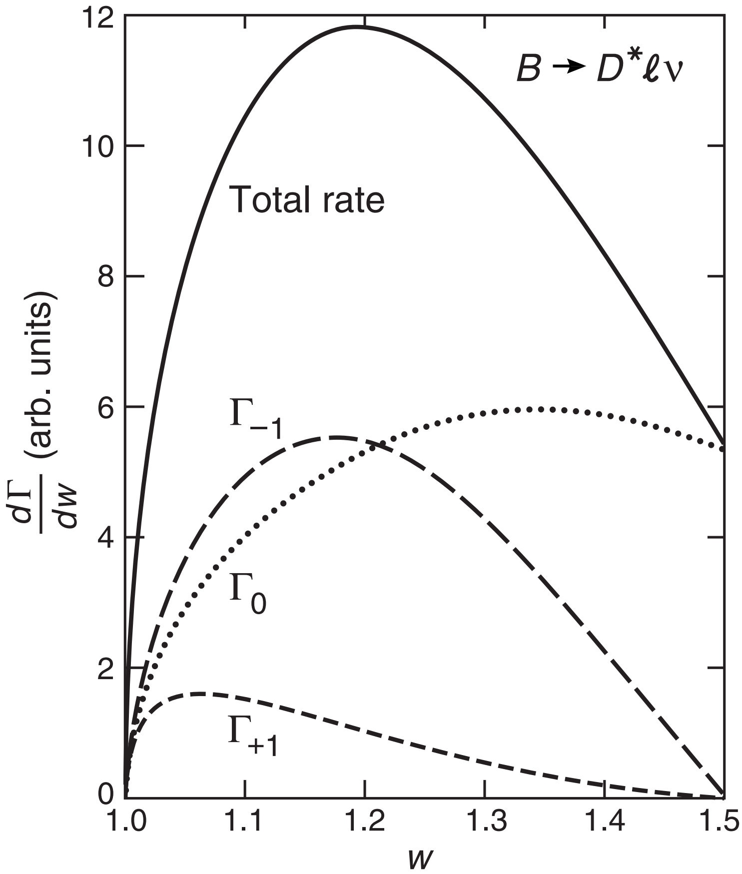


Figure 1.7: Contributions to the rate in $B \to D^*\ell\nu$ from different W (or D^*) helicities. The zero recoil configuration is at $w = 1$ ($q^2_{\rm max}$, while maximum recoil occurs at $w \approx 1.5$ ($q^2 \approx 0$). At intermediate values of w, the $V - A$ coupling produces more rate into the negative (-1) helicity state than into the positive (+1) state.

and $V(q^2)$ by the expression:

$$H_0(q^2) = \frac{1}{2m_{D^*}\sqrt{q^2}}\left[(m_B^2 - m_{D^*}^2 - q^2)(m_B + m_{D^*})A_1(q^2) - \frac{4m_B^2 p_{D^*}^2}{m_B + m_{D^*}}A_2(q^2)\right] , \tag{1.21}$$

$$H_\pm(q^2) = (m_B + m_{D^*})A_1(q^2) \mp \frac{2m_B p_{D^*}}{m_B + m_{D^*}}V(q^2) . \tag{1.22}$$

Equation 1.18 for the full differential cross section can be integrated over $\cos\theta_\ell$,

$\cos\theta_{D^*}$ and χ to give the differential cross section versus w. This result is generally written in a form parallel to 1.16 as

$$\begin{aligned}\frac{d\Gamma(B^0 \to D^{*-}\ell^+\nu_\ell)}{dw} &= \frac{G_F^2}{48\pi^3} m_{D^*}^3 (m_B - M_{D^*})^2 \sqrt{w^2} - 1(w+1)^2 \\ &\times \left[1 + \frac{4w}{w+1}\frac{1-2wr+r^2}{(1-r)^2}\right] |V_{cb}|^2 \left[\mathcal{F}(w)\right]^2 . \end{aligned} \quad (1.23)$$

In this equation, $\mathcal{F}(w)$ is a function of form factors given by:

$$\mathcal{F}(w) = h_{A_1}(w)\left[1 + \frac{4w}{w+1}\frac{1-2wr+r^2}{(1-r)^2}\right]^{-\frac{1}{2}} \sqrt{\left[\tilde{H}_0(w)\right]^2 + \left[\tilde{H}_+(w)\right]^2 + \left[\tilde{H}_-(w)\right]^2} , \quad (1.24)$$

where $\tilde{H}_0(w)$ and $\tilde{H}_\pm(w)$ are related to the form factor ratios $R_1(w)$ and $R_2(w)$, defined in appendix A, by:

$$\tilde{H}_0(w) = 1 + \left(\frac{w-1}{1-r}\right)\left[1 - R_2(w)\right] \quad (1.25)$$

and

$$\tilde{H}_\pm(w) = \frac{\sqrt{1-2wr+r^2}}{1-r}\left[1 \mp \sqrt{\frac{w-1}{w+1}} R_1(w)\right] ; \quad (1.26)$$

where $r = m_{D^*}/m_B$.

Comparison of equations A.8 and A.9, defined in appendix A, with equations 1.25 and 1.26 gives the relations between the $H_{0,\pm}(w)$'s and the $\tilde{H}_{0,\pm}(w)$'s :

$$H_{0,\pm}(w) = (m_B - m_{D^*})\sqrt{\frac{m_B m_{D^*}}{q^2(w)}}(w+1)h_{A_1}(w)\tilde{H}_{0,\pm}(w) . \quad (1.27)$$

Within the framework of HQET, one can define a particular parametrisation for $B^0 \to D^{*-}\ell^+\nu_\ell$ by simply writing down the forms of $h_{A_1}(w)/h_{A_1}(1)$, $R_1(w)$ and $R_2(w)$ with some number of free parameters. A simple "linear HQET" parametrisation is [16]:

$$\begin{aligned}\frac{h_{A_1}(w)}{h_{A_1}(1)} &= 1 - \rho^2_{h_{A_1}}(w-1) , & (1.28)\\ R_1(w) &= R_1(1) , & (1.29)\\ R_2(w) &= R_2(1) . & (1.30)\end{aligned}$$

This description has 3 free parameters: $\rho^2_{h_{A_1}}$, $R_1(1)$ and $R_2(1)$. It assumes that $h_{A_1}(w)$ is linear in w, and the the form factor ratios $R_1(w)$ and $R_2(w)$ are constant. This parametrisation was assumed in the CLEO measurement [14] of the $B^0 \to D^{*-}\ell^+\nu_\ell$ form factors.

Caprini, Lellouch and Neubert [15] provide a more theoretically driven parametrisation based on techniques involving unitarity constraints and dispersive bounds. They obtain the forms:

$$\frac{h_{A_1}(w)}{h_{A_1}(1)} = 1 - 8\rho^2_{h_{A_1}} z + (53\rho^2_{h_{A_1}} - 15)z^2 - (231\rho^2_{h_{A_1}} - 91)z^3 \,, \tag{1.31}$$

$$R_1(w) = 1.27 - 0.12(w-1) + 0.05(w-1)^2 \,, \tag{1.32}$$

$$R_2(w) = 0.80 + 0.11(w-1) - 0.06(w-1)^2 \,. \tag{1.33}$$

In this parametrisation, the form factor $h_{A_1}(w)$ is written in terms of the variable z, which is related to w by

$$z = \frac{\sqrt{w+1} - \sqrt{2}}{\sqrt{w+1} + \sqrt{2}} \,. \tag{1.34}$$

The slope $\rho^2_{h_{A_1}}$ in the expression for $h_{A_1}(w)$ has the same meaning in both HQET parametrisations. It is the slope of the function $h_{A_1}(w)/h_{A_1}(1)$ versus w, evaluated at the point $w = 1$.

The linear HQET parametrisation is less sophisticated than that of Caprini *et al.*, but it does contain the right basic structure. The w dependence of $R_1(w)$ and $R_2(w)$ in equation 1.31 is clearly very weak. The theoretical uncertainties [16] in the predictions for $R_1(1)$ and $R_2(1)$ are about 0.1 and 0.2 respectively, so theory and experiment are in excellent agreement. In the w region allowed for $B^0 \to D^{*-}\ell^+\nu_\ell$ decays the $R_1(w)$ and $R_2(w)$ theoretical dependence is much smaller than the experimental uncertainty hence the assumption of constant R_1 and R_2 is clearly a small error. We will see if this assumption still holds when the analyses with the huge amount of data collected by the B Factory *BABAR* and Belle will be available. Right now the only published result in literature is the CLEO one [14]. The greater difference between the parametrisations is in the shape of $h_{A_1}(w)$.

Figure 1.8 compares the shape $\mathcal{F}(w)$ between the parametrisation of Equation 1.31 and a linear form similar to that used in the 1995 CLEO $|V_{cb}|$ analysis [17]. The two forms have similar values over most of the w region, and differ only near the ends of the distributions. Due to the limited phase space available, few events populate these high and low w regions.

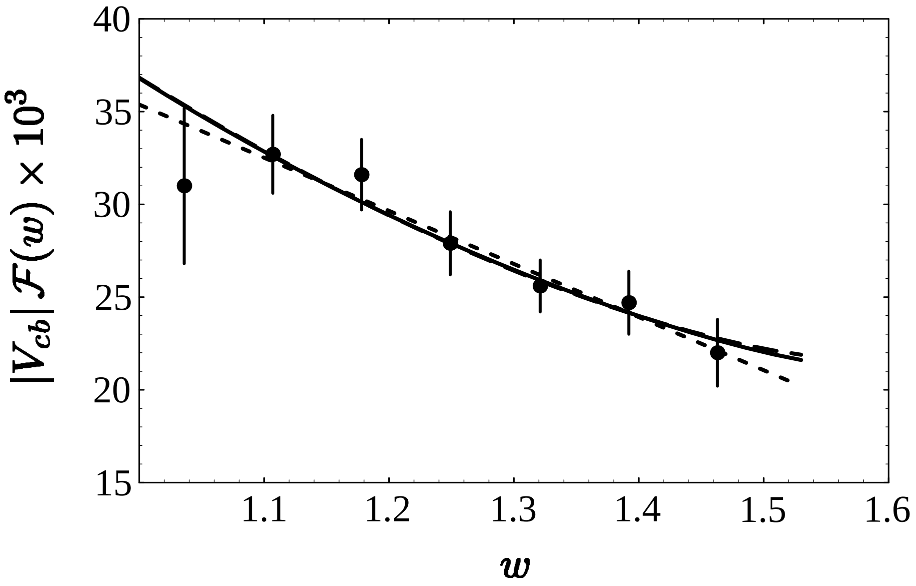


Figure 1.8: A comparison of a linear form of $\mathcal{F}(w)$ to one consistent with the parametrisation of Caprini *et al.* (Equation 1.31). The solid curve is the Caprini *et al.* shape, while (barely visible) curve in long dashes is a second parametrisation by the same authors. The dashed line is a straight line fit. All three curves are *ad hoc* fits to the points, which derive from the 1995 CLEO $B \to D^*\ell\nu$ analysis [17].

Thus, for simply describing the bulk of the events, the difference between the two parametrisation is small. However, in order to reliably extract $|V_{cb}|$, it is critical that the extrapolation in the low-w region be as correct as possible, and the linear shape appears to underestimate the true value of $|V_{cb}|$. The fits of Caprini *et al* [15] suggest that the linear fits systematically underestimate the extrapolation at zero recoil, hence $|V_{cb}|$, by roughly 3%. Most of this effect is related to the curvature of the form factor, and does not depend strongly upon the details of the non-linear shape chosen [18].

Both shapes in figure1.8 are *ad hoc* fits directly to the points, which represent the same data used for the 1995 CLEO analysis [17]. This approach is, of course, not how such fits are done in practice. Operationally, one first convolves a form factor parametrisation with a smearing matrix. The smearing matrix includes both the reconstruction efficiency as a function of w and the probabilities that events will be reconstructed in different w bins. Such a matrix is calculated with Monte Carlo events. The convolved shape is compared directly to the event yields from the data as a function of w, and the parametrisation adjusted to obtain the best fit.

1.4 Experimental results on $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ and $|V_{cb}|$

The present $|V_{cb}|$ analyses are taking advantage of the additional knowledge of the form factors contained in equation 1.31.

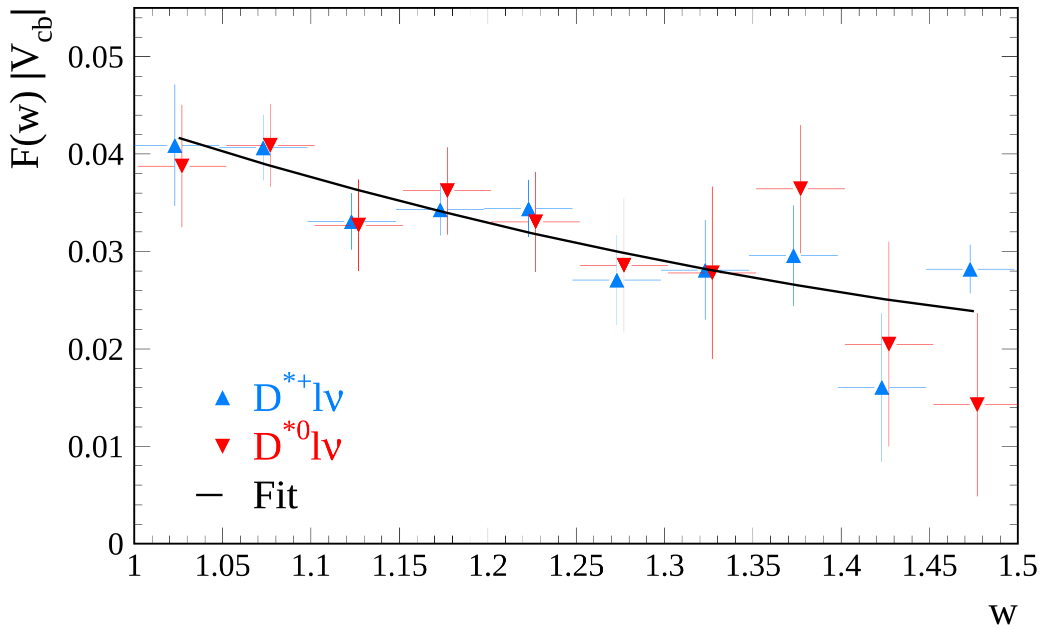


Figure 1.9: CLEO 2002 latest result on $\mathcal{F}(1) \cdot |V_{cb}|$ as we can see they used both the $B^0 \to D^{*-}\ell^+\nu_\ell$ (blue triangles with error bars) and the $B^\pm \to D^{*0}\ell^\pm\nu_\ell$ (red triangles with error bars) data sample. The result of their fit is $\mathcal{F}(1) \cdot |V_{cb}| = (4.31 \pm 0.13(\text{stat}) \pm 0.18(\text{syst})) \times 10^{-2}$. Integrating the w spectrum they obtained $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell) = (6.09 \pm 0.19(\text{stat}) \pm 0.40(\text{syst})) \times 10^{-2}$ and using the value of $\mathcal{F}(1)$ as in [12] they extracted $|V_{cb}| = (4.69 \pm 0.14(\text{stat}) \pm 0.20(\text{syst}) \pm 0.18(\text{theor})) \times 10^{-2}$.

Measurements of the product $\mathcal{F}(1) \cdot |V_{cb}|$ have been performed at the $\Upsilon(4S)$ by CLEO [13] and Belle [19]. At LEP results are available from ALEPH [20], DELPHI [21] and OPAL [22]. Figures 1.9 and 1.10 show the latest CLEO and Belle measurements of $|V_{cb}|$ and $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$.

To combine the published data, the central values and errors of $\mathcal{F}(1) \cdot |V_{cb}|$ and ρ^2 ([2]) are rescaled to the same set of form factors input parameters and their quoted uncertainties [5]. The $\mathcal{F}(1) \cdot |V_{cb}|$ values used for this average are extracted using the parametrisation adopted by the latest $B \to D^*\ell\nu$ CLEO analysis [13]

[2] defined as the slope of the form-factor at zero recoil

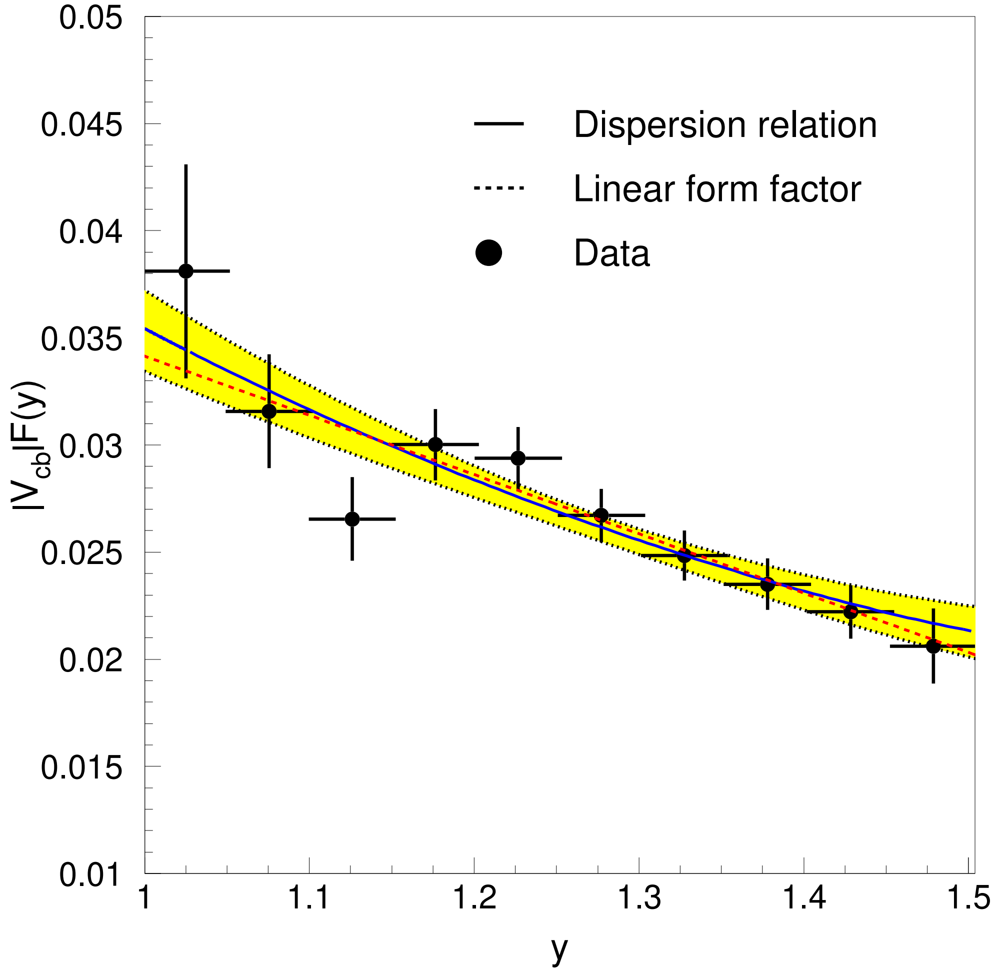


Figure 1.10: Belle results on $\mathcal{F}(1) \cdot |V_{cb}|$ using only the decay $\overline{B}^0 \to D^{*+}e^-\overline{\nu}_e$ and its charge-conjugate. The result of their fit is $\mathcal{F}(1) \cdot |V_{cb}| = (3.358 \pm 0.19(\mathrm{stat}) \pm 0.18(\mathrm{syst})) \times 10^{-2}$. Integrating the w spectrum they obtained $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell) = (4.59 \pm 0.23(\mathrm{stat}) \pm 0.40(\mathrm{syst})) \times 10^{-2}$ and using the value of $\mathcal{F}(1)$ as in 1.17 they extracted $|V_{cb}| = (4.19 \pm 0.45(\mathrm{stat}) \pm 0.53(\mathrm{syst}) \pm 0.30(\mathrm{theor})) \times 10^{-2}$. Their notation adopt y instead of w.

based on the CLEO experimental determination of the form factor ratios R_1 and R_2 [14]. The LEP data, which originally used theoretical values for these ratios, are rescaled accordingly. The averaging procedure [23] takes into account statistical and systematic correlations between $\mathcal{F}(1) \cdot |V_{cb}|$ and ρ^2.

Averaging the corrected measurements the PDG [5] quotes:

$$\mathcal{F}(1) \cdot |V_{cb}| = (38.3 \pm 1.0) \times 10^3$$

and

$$\rho^2 = 1.5 \pm 0.13 \ .$$

The error ellipses for the corrected measurements and for the world average are shown in figure 1.11.

The value of $|V_{cb}|$ obtained from exclusive semileptonic $B \to D^*\ell\nu$ decays is[3]:

$$|V_{cb}| = (4.21 \pm 0.11 \text{exp} \pm 1.9 \text{theor}) \times 10^{-2} \ . \tag{1.35}$$

.

[3]PDG 2002 [5] explains they did not use the last QCD lattice calculation for $\mathcal{F}(1)$. In that case the theoretical error would be 25% less than the one in 1.35.

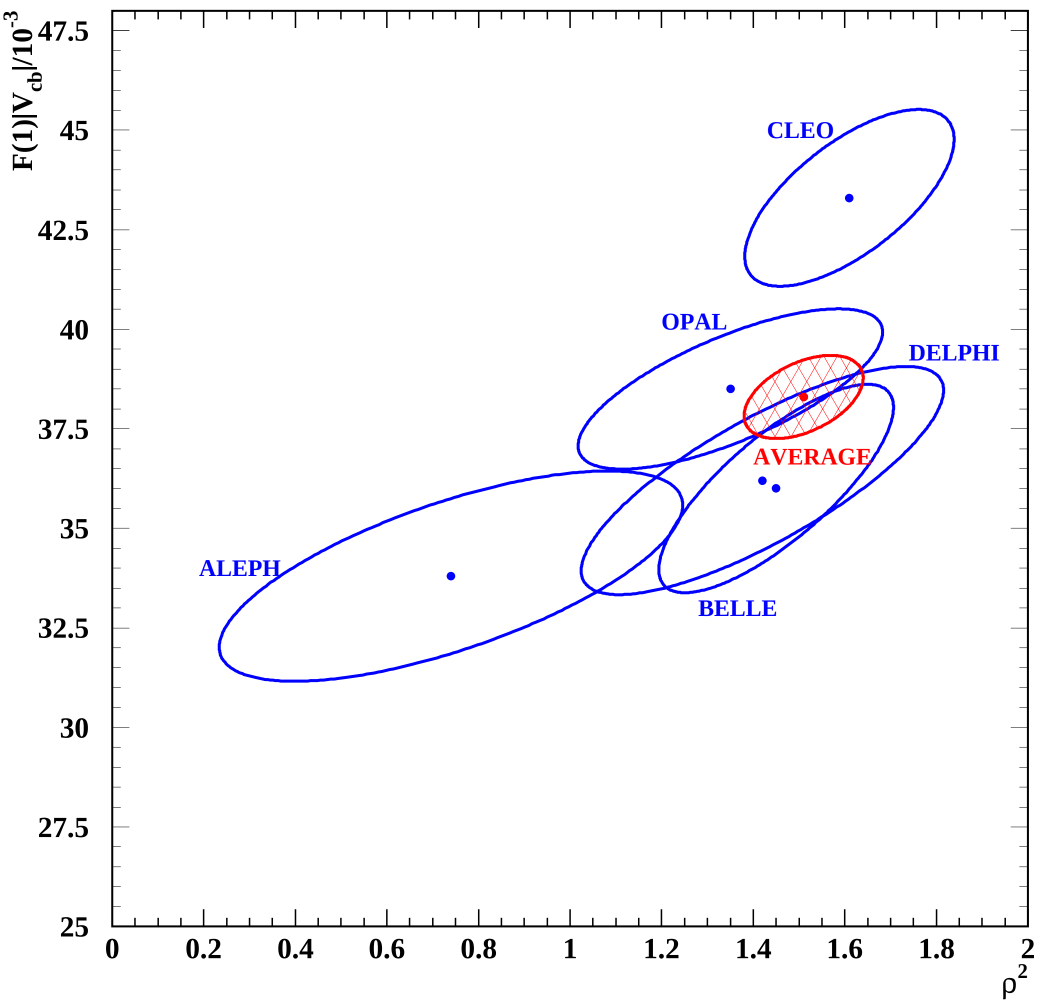


Figure 1.11: The error ellipses for the corrected measurements and world average for $\mathcal{F}(1) \cdot |V_{cb}|$ versus ρ^2. The ellipses are obtained from product between the 1 σ error of $\mathcal{F}(1) \cdot |V_{cb}|$, ρ^2 and the correlation between the two.

Chapter 2

The BABAR detector

2.1 Introduction

The PEP-II B Factoryis an asymmetric e^+e^- collider designed to operate at a luminosity of 3×10^{33} cm^{-2}s^{-1} and beyond at a center-of-mass energy of 10.58 GeV, the mass of the $\Upsilon(4S)$ resonance (see figure 2.1). This resonance decays almost exclusively into $B^0\overline{B}^0$ or B^+B^- pairs and thus provides an ideal laboratory for the study of B mesons. In PEP-II, the electron beam of 9.0 GeV collides head-on with the positron beam of 3.1 GeV resulting in a boost to the $\Upsilon(4S)$ resonance of $\beta\gamma = 0.56$. This boost makes it possible to reconstruct the decay vertices of the two B mesons and to determine their relative decay times. One can therefore measure the decay time dependence of decay rates. The crucial test of CP invariance is a comparison of these rates for B^0 and $\overline{B}^0$ to a CP eigenstate. Experimentally this requires events in which one B meson decays to a CP eigenstate and is fully reconstructed and the other B meson is tagged as B^0 or a $\overline{B}^0$ by its decay products, a charged lepton, a charged kaon or a D^* decay. The very small branching ratios of B masons to CP modes, typically 10^{-4}, and the need for full reconstruction of final states with two or more charged particles and several π^0 plus the tagging of the second B place stringent requirements on the detector.

2.1.1 Physics requirements

To achieve the goal of performing accurate CP violation measurements there are many requirements:

- a large and uniform acceptance, in particular down to small polar angles rel-

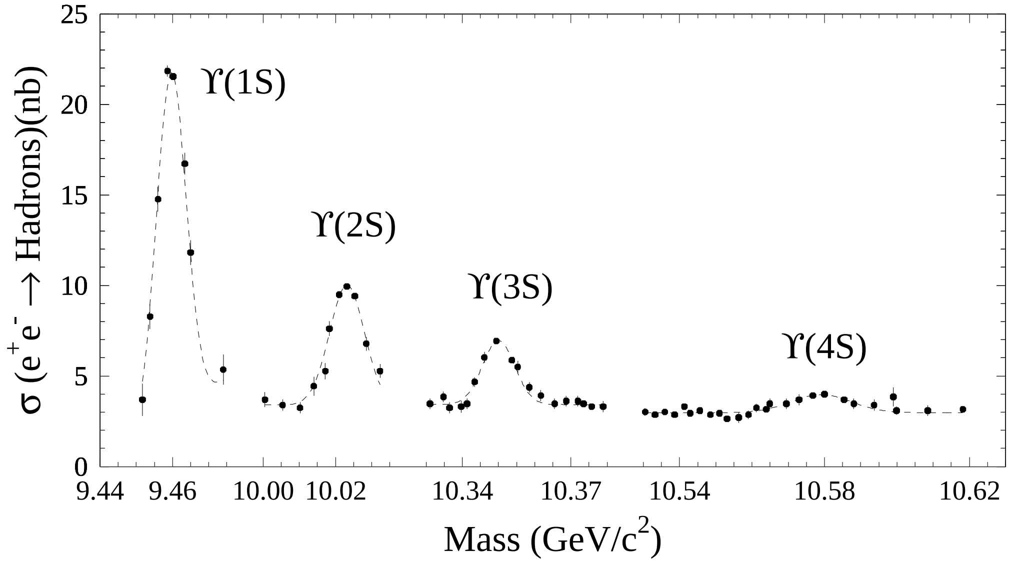


Figure 2.1: The first four S-wave Υ resonances shown with the hadronic cross section versus center-of-mass energy/c^2 in the Υ mass region. The $\Upsilon(4S)$, at 10.58 GeV/c^2, is the third radial excitation of the ground state. It has sufficient mass to decay into $B^0\overline{B}^0$ and B^+B^- pairs, which results in its larger width. Note the difference in resonance amplitude and the significant continuum background between the $\Upsilon(4S)$ and the lower resonances. The continuum events underneath the $\Upsilon(4S)$ typically have a two-jet topology, which allows them to be distinguished from the much more isotropic distribution of tracks in the $\Upsilon(4S) \to B\overline{B}$ decays.

ative to the boost direction, to avoid particle losses;

- excellent detection efficiency for charged particles down to 60 MeV/c and for photons to 25 MeV;
- high momentum resolution to separate small signals from background;
- excellent energy and angular resolution for the detection of photons from π^0 and radiative B decays in the range from 25 MeV to 4 GeV;
- very good vertex resolution, both transverse and parallel to the beam;
- identification of electrons and muons over a range of momentum, primarily for the detection of semi-leptonic decays used to tag the B flavor and for the study of semi-leptonic and rare decays;

- identification of hadrons over a wide range of momentum for B flavor tagging as well as for the separation of pions from kaons in decay modes like $B^0 \to K^\pm\pi^\mp$ and $B^0 \to \pi^+\pi^-$ as well as in charm meson and τ decays;
- a highly efficient, selective trigger system with redundancy so as to avoid significant signal losses and systematic uncertainties;

2.1.2 Factory requirements

- low noise electronics and data acquisition systems of high flexibility and operational stability;
- high degree of reliability of components and frequent monitoring and automated calibrations, plus control of the environmental conditions to assure continuous and stable operation;
- an on-line computing and network system that can control, process, and store the expected high volume of data;
- detector components that can tolerate significant doses of radiation and operate under high background condition.

To reach the desired sensitivity for the most interesting measurements, but not for the $|V_{cb}|$ measurement which is systematics limited, data sets of some 10^8 B mesons will be needed. For the peak cross section at the $\Upsilon(4S)$ of 1.1 fb^{-1} this will require an integrated luminosity of some $200 - 500$ fb^{-1} which is more than twice the actual *BABAR* recorded luminosity as shown in figure 2.2. Targeting to collect such a huge amount of data requires the machine to work with few interruptions or shutdown periods.

2.2 PEP-II

PEP-II is an e^+e^- storage ring system designed to operate at a center of mass energy of 10.58 GeV, corresponding to the mass of the $\Upsilon(4S)$ resonance. The parameters of these energy asymmetric storage rings are presented in table 2.1. PEP-II has surpassed its design luminosity, both in terms of the instantaneous and the integrated daily luminosity, with significantly fewer bunches than anticipated. PEP-II typically operates in a series of 40 minute fills during which the colliding beams coast. At the

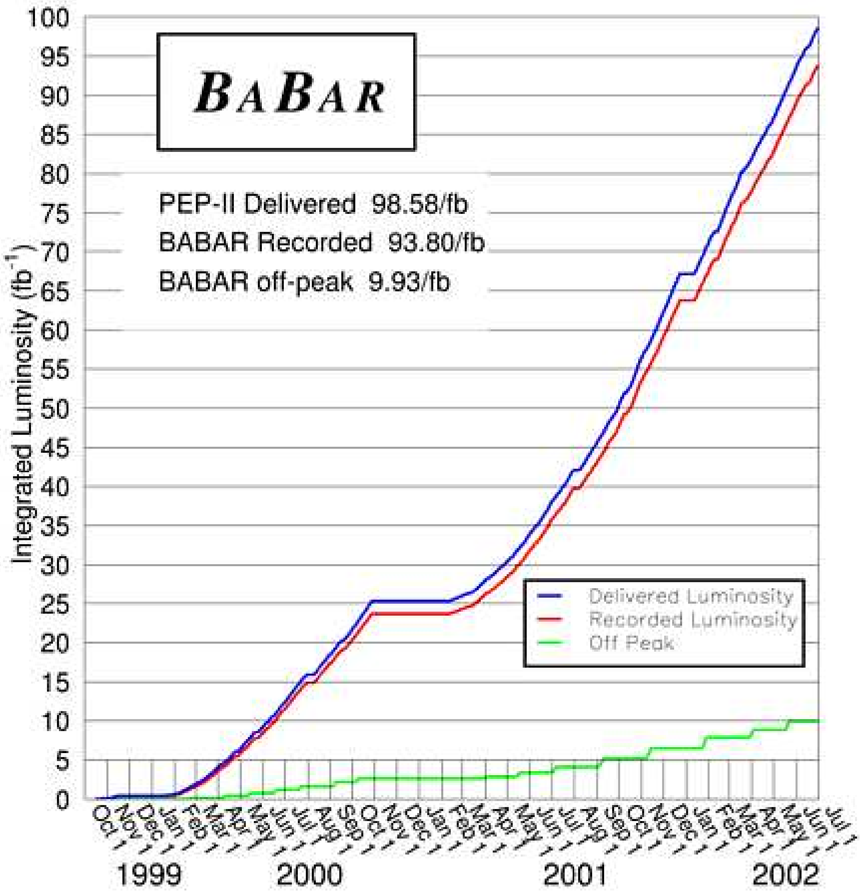


Figure 2.2: PEP-II delivered and *BABAR*-recorded integrated luminosity in RUN1 and RUN2 (from october 1999 to June 2002).

end of each fill, it takes about three to five minutes to replenish the beams for the next fill. After a loss of the stored beams, it takes approximately 15 minutes to refill the two beams. *BABAR* divides the data into runs, defined as periods of about two hour duration or less during which beam and detector conditions are judged to be stable. While most of the data are recorded at the peak of the $\Upsilon(4S)$ resonance, some 11% are taken 40 MeV below to allow for studies of the non-resonant background [1].

2.2.1 Interaction region

PEP-II collides the bunches head-on to avoid the coupling of transverse and longitudinal modes caused by the crossing angle which can limit its luminosity. The energy asymmetry allows to separate the beams magnetically but is not possible to

[1]mostly from $e^+e^- \to q\overline{q}$ or $e^+e^- \to \ell^+\ell^-$.

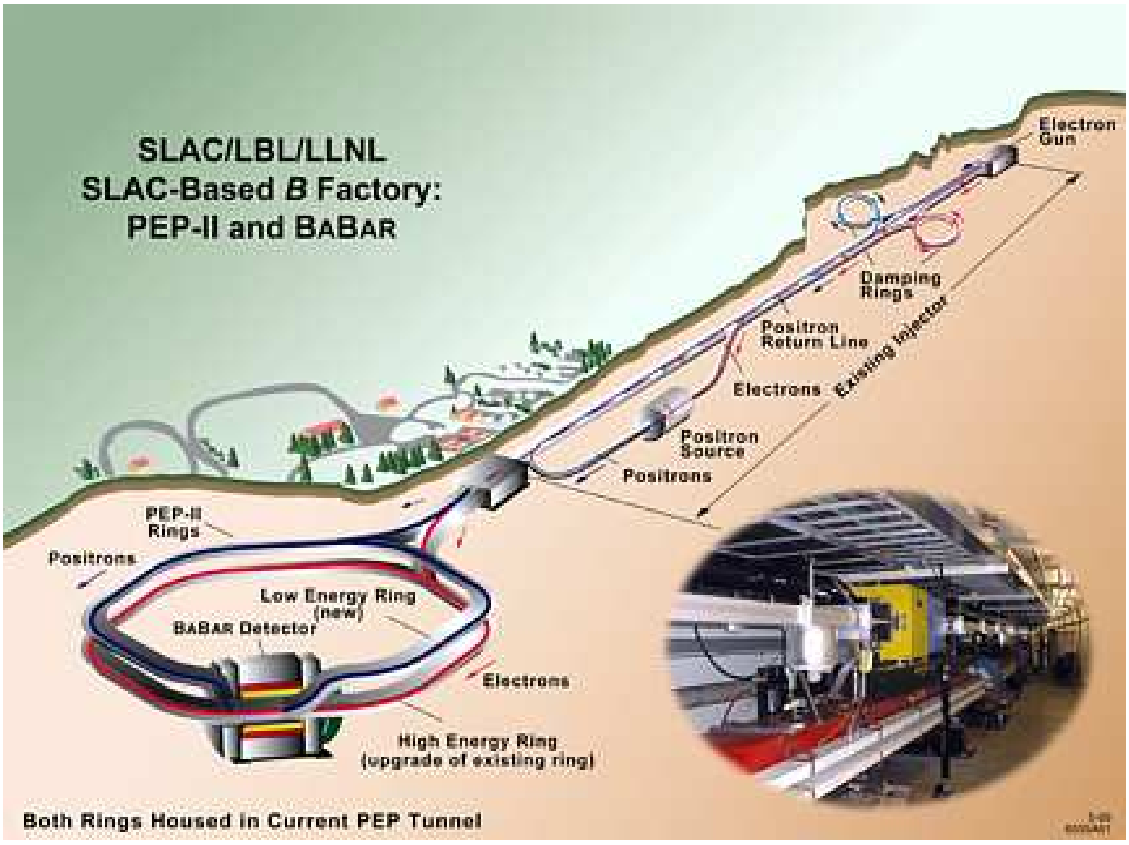


Figure 2.3: PEP-II overview.

Parameters	Units	Design	Typical
Energy HER/LER	GeV	9.0/3.1	9.0/3.1
Current HER/LER	A	0.75/2.15	1.0/1.7
# of bunches		1658	553-829
Bunch spacing	ns	4.2	6.3-10.5
σ_x	μm	110	147
σ_y	μm	3.3	5.6
σ_z	mm	9	9
Luminosity	pb$^{-1}/day$	135	240

Table 2.1: *PEP-II beam parameters; values are given both for design and typical colliding beam operation in the first year. HER and LER refer to the high energy* e^- *and low energy* e^+ *ring, respectly.* σ_x, σ_y *and* σ_z *refer to the horizontal, vertical, and longitudinal r.m.s. size of the luminous region.*

use different focusing optics for the two beams. The higher energy beam (HER) has indeed more focusing than the other with equal currents, that would result in an asymmetric focusing. To compensate this and obtain a symmetric focusing, the lower energy beam has been designed to have a higher current. It is usually more difficult [25] to obtain a high current electron beam with respect to a positron beam because the positive ions that remain in the beam pipe are attracted by the electron beam and can generate orbit instabilities. This is indeed the reason why the electron beam has been chosen to have the lower current (and the higher energy). However PEP-II experts [25] claim that with the present working conditions the positron beam (the one with higher current) becomes instable at high currents because of a multiplication effect of electrons extracted from the walls of the beam pipe. The large beam currents and the necessity to separate closely-spaced bunches as close as possible to the interaction point (IP), tightly couple the issues of detector design, interaction region layout, and remediation of machine induced background. The bunches are separated magnetically in the horizontal plane by a pair of dipole magnets (B1), followed by a series of offset quadrupoles. The tapered B1 dipoles, located at $\pm$ 21 cm on either side of the IP, and the Q1 quadrupoles operate inside the field of the *BABAR* solenoid, while Q2, Q4, and Q5, are located outside or in the fringe field of the solenoid.

The interaction region is closed by a water-cooled beam pipe consisting of two thin layers of beryllium (0.83 mm and 0.53 mm) with a 1.48-mm water channel in between. To attenuate synchrotron radiation, the inner surface of the pipe is gold-plated (approximately 4 μm). In addition, the beam pipe is wrapped with 150 μm Ta foil on either side of the IP, i.e., beyond $z = +10.1$ cm and $z = -7.9$ cm. The total thickness of the central beam pipe section at normal incidence corresponds to 1.06 % of a radiation length. The beam pipe, the permanent magnets and the Silicon Vertex Tracker (SVT) are assembled and aligned and then enclosed in a 4.4-m long support tube. This rigid structure is inserted into the *BABAR* detector, spanning the IP.

2.3 Luminosity, Beam Energies and Position

The beam parameters that are most critical for the *BABAR* data analysis are the luminosity, the energies of the two beams and the position and size of the luminous region.

2.3.1 Luminosity

While PEP-II uses a high-rate luminosity sampling radiative Bhabha scattering to provide a fast relative measurement of the luminosity for operations, *BABAR* derives the absolute luminosity off-line from QED processes. The best result is obtained from $\mu^+\mu^-$ pairs. For a sample of 1 fb^{-1}, the statistical error is 1.3 % compared to a systematic error of 0.5 % on the relative, and 1.5 % on the absolute value of the luminosity. This error is currently dominated by uncertainties in the Monte Carlo generator and the simulation of the detector. It is expected that with a better understanding of the efficiency the overall systematic error on the absolute value of the luminosity can be reduced by a factor two [26].

2.3.2 Beam Energies

During operation, the mean energies of the two beams are calculated from the total magnetic bending strength (including the effects of off-axis quadrupole fields, steering magnets, and wigglers) and the beam orbits. While the systematic uncertainty in the PEP-II calculation of the absolute beam energies is estimated to be $5-10$ MeV [25], the relative energy setting for each beam is accurate and stable to about 1 MeV [25]. The energy spread of the LER and HER is 2.3 MeV [25] and 5.5 MeV [25], respectively. To ensure that data are recorded close to the peak of the $\Upsilon(4S)$ resonance, the observed ratio of $B\overline{B}$ enriched hadronic events to lepton pair production is monitored on-line. At the peak of the resonance, a 2.5 % change in the $B\overline{B}$ production rate corresponds to a 2 MeV change in the c.m. energy, a value that is close to the tolerance to which the energy of PEP-II can be held. However, a drop in the $B\overline{B}$ rate does not distinguish between energy settings below or above $\Upsilon(4S)$ peak. The sign of the energy change must be determined from other indicators. The best monitor and absolute calibration of the c.m. energy is derived from the measured c.m. momentum of fully reconstructed B mesons constrained with the known B meson mass. An absolute error of 1.1 MeV can be obtained for an integrated luminosity of 1 fb^{-1}. This error is equally limited by the knowledge of the B mass [27] and the detector resolution.

2.3.3 Beam Size and Position

The size and position of the luminous region are critical parameters for the time-dependent analyses[2] and their values are monitored continuously on-line and off-line. The design values for the size of the luminous region are presented in table 2.1 above. The vertical size is too small to be measured directly. It is inferred from the measured luminosity, the horizontal size, and the beam currents; it varies typically by 1-2 μm. The transverse position, size and angles of the luminous region relative to the *BABAR* coordinate system are determined by analysing the distribution of the distance of closest approach to the z-axis of the tracks in well measured two-track events as a function of the azimuth ϕ. The longitudinal parameters are derived from the longitudinal vertex distribution of the two tracks. The uncertainties in the average beam position are of the order of a few μm in the transverse plane and about 100 μm along the collision axis. Run-by-run variations in the beam position are comparable to these measurement uncertainties, indicating that the beams are stable over the period of a typical run. The measured horizontal and longitudinal beam sizes, corrected for tracking resolution, are consistent with those measured by PEP-II.

2.4 Detector overview

The *BABAR* detector (figure 2.4, figure 2.5) has been designed and fabricated by a collaboration of 600 physicists of 75 institutions from 9 countries and installed at SLAC in Fall 1998. First data with PEP-II colliding beams were collected in May 1999. The *BABAR* superconducting solenoid which produces a field of 1.5 T axial magnetic field contains a set of nested detectors: a five layers Silicon Vertex Tracker (SVT), a central Drift Chamber (DCH) for charged particles detection and momentum measurement, a quartz-bar Cherenkov radiation detector (DIRC) for particle identification, and a CsI crystal electromagnetic calorimeter for detecting photons and electrons. The calorimeter has a barrel and an endcap which extends it asymmetrically into the forward direction (e^- beam direction), where many of the collision products emerge. All the detectors located inside the magnet have full acceptance in azimuth. The flux return outside the cryostat is composed of 18 layers of steel, which increase in thickness outwards, and are instrumented with 19 layers of planar RPCs in the barrel and 18 in the end-caps. The RPCs allow the separation

[2] or whenever a primary decay vertex is evaluated with the beam-spot constraint.

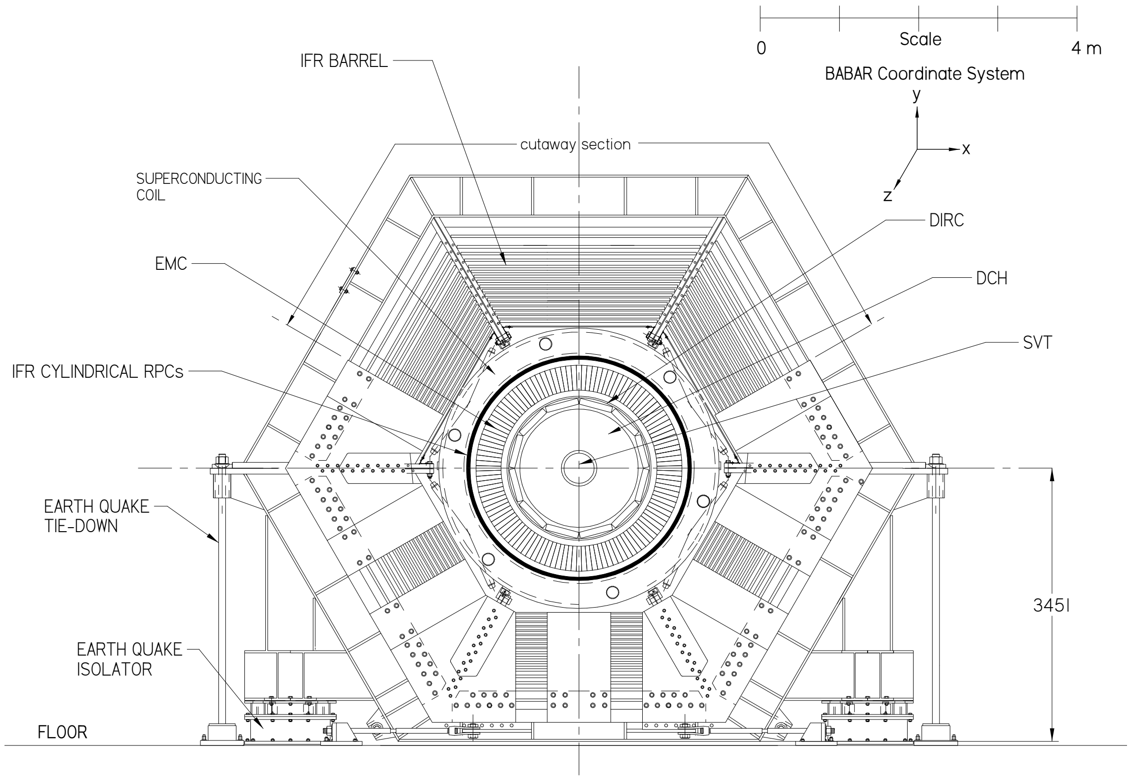


Figure 2.4: *BABAR* detector front view.

of muons and charged hadrons, and also detect penetrating neutral-hadrons. Since the average momentum of charged particles produced in B meson decay is less than $1\,\mathrm{GeV}/c$, the errors on the measured track parameters are dominated by multiple Coulomb scattering, rather than the intrinsic spatial resolution of the detectors. Similarly, the detection efficiency and energy resolution of low energy photons are severely impacted by material in front of the calorimeter. Thus, special care has been given to keep the material in the active volume of the detector to a minimum. Figure 2.6 shows the distribution of material in the various detector systems in units of radiation lengths. Specifically, each curve indicates the material a particle traverses before it reaches the first active element of a specific detector system.

2.5 Silicon Vertex Tracker

The Silicon Vertex Tracker (SVT) is the fundamental detector for the reconstruction of the B decay vertex and the secondary vertices, as well as for the tracking of low

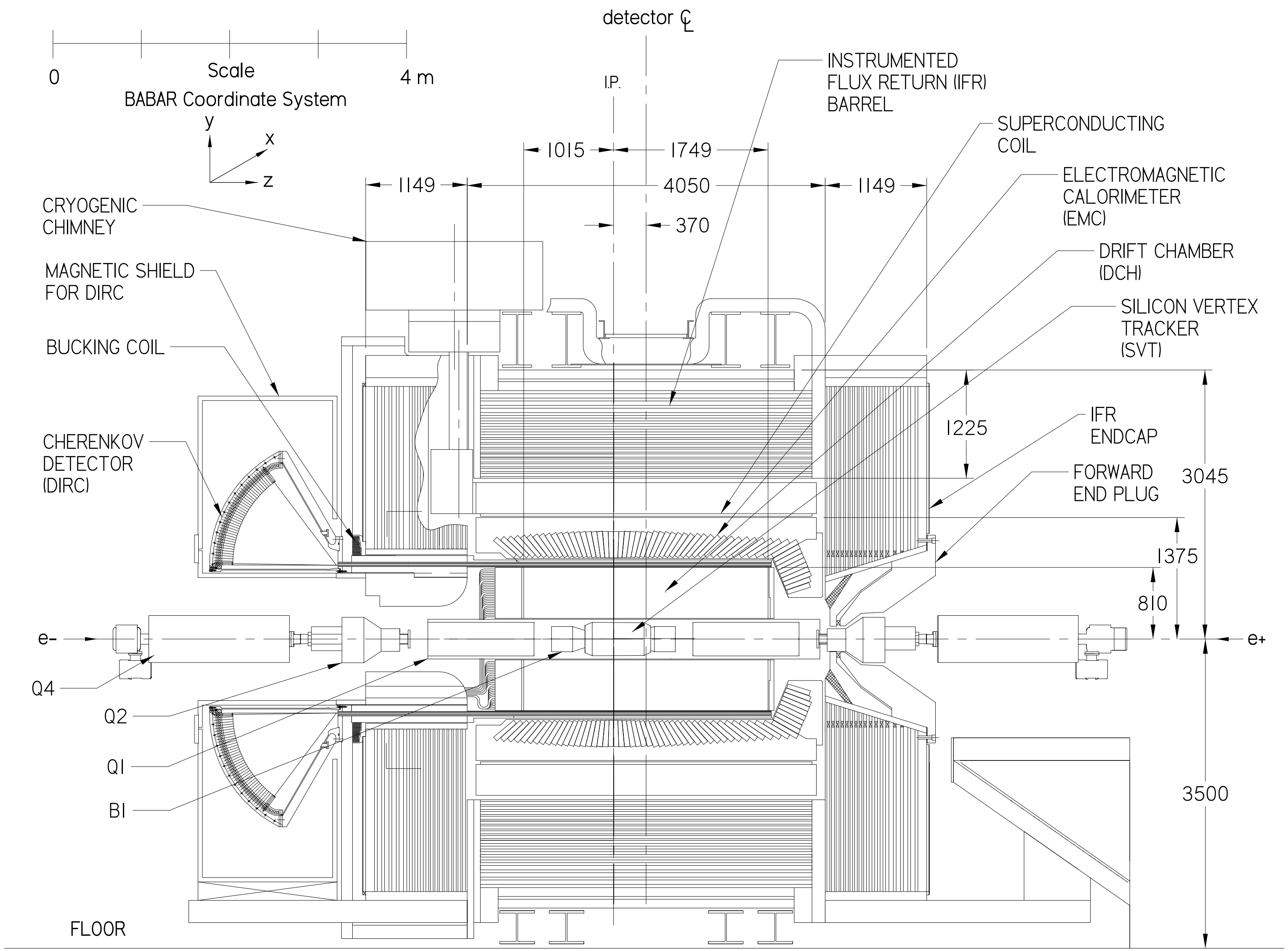


Figure 2.5: *BABAR* detector side view.

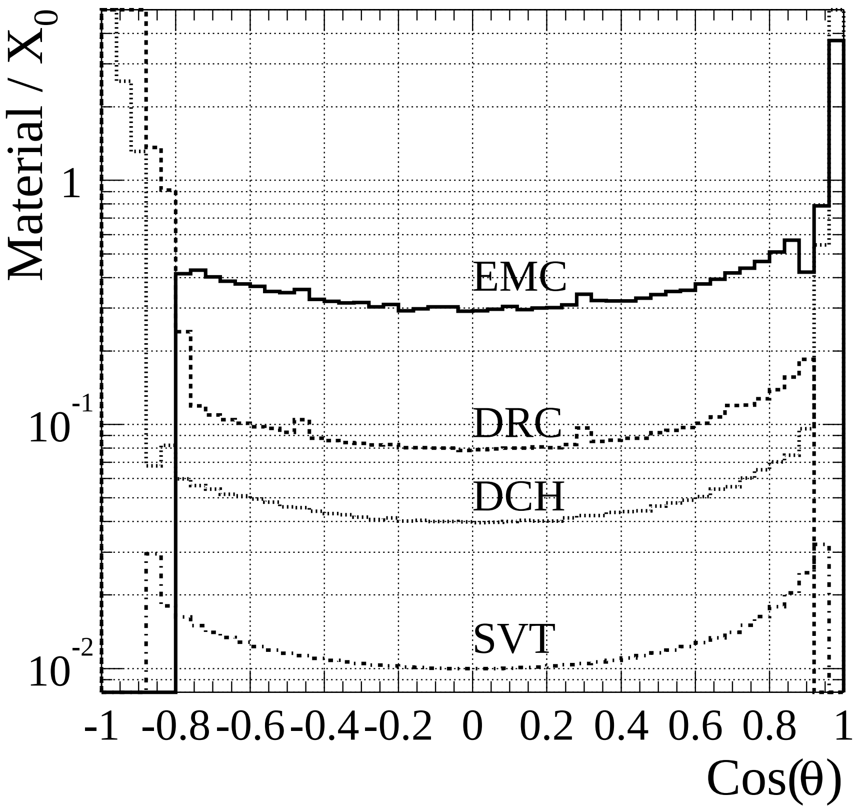


Figure 2.6: Amount of material (in units of radiation lengths) which a high energy particle, originating from the center of the coordinate system at a polar angle θ, traverses before it reaches the first active element of a specific detector system.

momenta particles such as the soft pion coming from the $D^{*\pm} \to D^0\pi^\pm$ decay of the $D^{*\pm}$ in the $B \to D^{*\pm}\ell^\mp\nu_\ell$ decay.

A separate chapter has been dedicated to this detector.

2.6 Drift Chamber

The Drift Chamber (DCH) is the main tracking device for charged particles with transverse momenta p_T above $\sim 120\,\mathrm{MeV}/c$, providing the measurement of p_T from the curvature of the particle traversing the 1.5 T magnetic field.

2.6.1 Purpose and design requirements

The DCH is designed to provide excellent momentum resolution in order to reconstruct exclusive final states with minimal background. For almost all momenta, multiple scattering is a significant, if not dominant, limitation on track parameter resolution. In order to reduce this contribution to the chamber p_T measurement, material within the chamber volume has been minimised using low-mass helium-based gas. A thin inner wall improves the contribution of the high-precision measurement in the outer layer of the SVT to the p_T resolution, and minimises backgrounds due to photon conversions in the chamber wall. The DCH trigger requires one and one-half tracks, i.e. one track traversing the full chamber radius together with one track traversing at least half the chamber radius.

For low momentum particles, the DCH provides particle identification by measurement of ionization loss ($\mathrm{d}E/\mathrm{d}x$). For the helium-isobutane gas mixtures under consideration, a resolution of around 7 % has been attained for the $\mathrm{d}E/\mathrm{d}x$ measurement, allowing π/K separation up to 700 MeV/c. This capability is complementary to the DIRC in the barrel region, while in the extreme backward and forward directions where there is no dedicated particle identification device, it is the only available measurement.

The DCH also allows the reconstructions of secondary vertices, such as decays of K_S^0 outside the silicon detector volume. For this purpose, the chamber is able to measure not only the transverse coordinate, but also the longitudinal position of tracks with good ($\sim 1\,\mathrm{mm}$) resolution. Good z resolution also aids in matching DCH and SVT tracks, and in projecting tracks to the DIRC and the calorimeter.

2.6.2 Mechanical design and overview

The final design adopted for the Drift Chamber, illustrated in figure 2.7, consists of a 280 cm-long cylinder with an inner radius of 23.6 cm and an outer radius of 80.9 cm.

DCH is located within the volume inside the DIRC and outside the PEP-II support tube. The center of the chamber is displaced forward by 36.7 cm to improve the forward track length, given PEP's asymmetric boost.

The drift system design consists of 7104 hexagonal cells (figure 2.8), approximately 1.8 cm wide by 1.2 cm high, arranged in 40 concentric layers between a radius of 25.3 and 79.0 cm. The active volume provides charged particle tracking

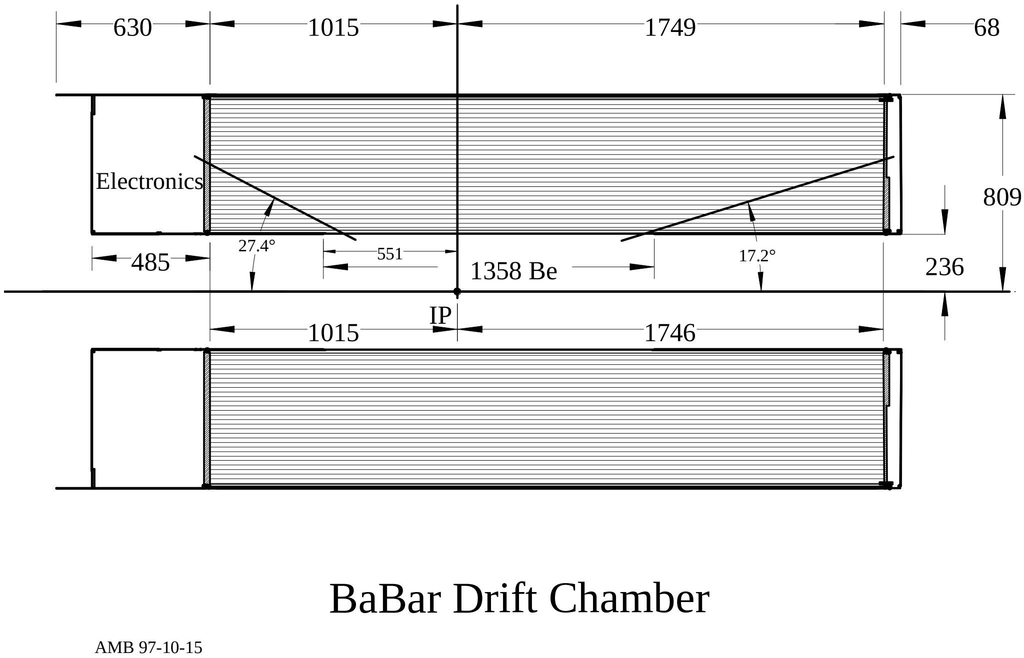


Figure 2.7: *BABAR* Drift Chamber side view.

over the polar angle range $-0.92 < \cos\theta_{lab} < 0.96$. The 40 layers are grouped into 10 superlayers[3] of 4 layers each, organised with the same orientation for sense and field wires within a given superlayer. The superlayer structure facilitates fast local segment finding as the first step in pattern recognition. This arrangement is particularly important for L1 trigger decisions. For measuring also the z coordinate superlayers alternate in orientation, first axial, then plus a small stereo angle, followed by minus a small stereo angle. All superlayers participate in the L1 track finding.

Individual cells are formed by either $3 \times 120\,\mu$m grounded aluminium field wires (inner cells in a superlayer), or 2 field wires and $1 \times 80\,\mu$m aluminium cathode wire at about 350 V (outer cells in a superlayer), per each 20 μm rhenium-tungsten sense wire operating at a nominal 1900-1930-1960 V. The sense wires have been working at 1930 V since the beginning of 2001 (which embed the dataset used in this analysis). Table 2.2 lists a summary of the main geometric features of the design. The counting gas is a low-mass 80 % : 20 % mixture of He:Isobutane, with

[3]Cylindrical layers of cells are grouped four at a time into superlayers. The first 16 layers are shown in figure 2.8. The plus signs, open circles, filled circles and crosses denote sense wires, field wires, guard wires and clearing wires, respectively. Lines have been added between field wires to aid in visualisation of the cells. The numbers on the right side give the stereo angles (mrad) of sense wires in each layer. A 1 mm thick beryllium inner cylinder is shown inside the first layer.

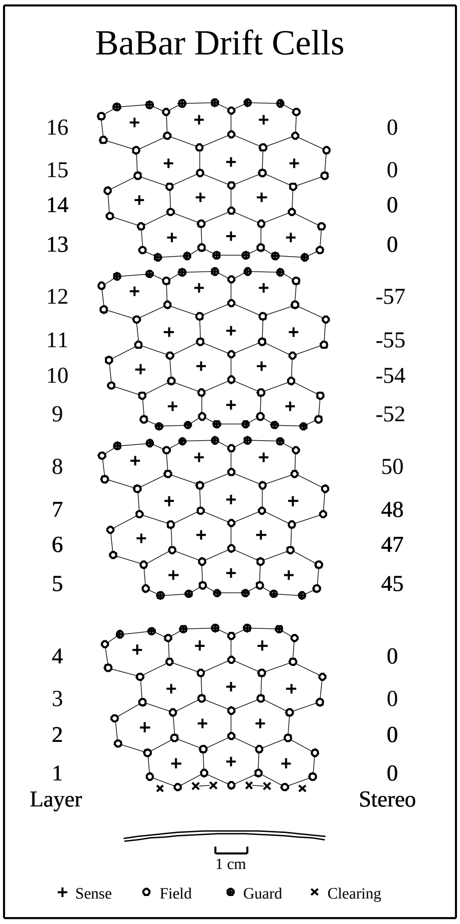


Figure 2.8: *BABAR* Drift Cell configuration.

a small amount (3000 ppm) of water vapor to prolong the lifetime of the chamber in a high rate environment and to prevent electrical discharge. The combination of low-mass gas and minimisation of material in the field cage itself, is designed to reduce the contribution of multiple scattering to the p_T resolution for the typically soft tracks produced in a $\Upsilon(4S)$ event.

2.6.3 Tracking resolution

For the purpose of most physics analyses, charged tracks are defined by five parameters ($d_0,\ \phi_0,\ \omega,\ z_0,\ \tan\lambda$) at the track's point of closest approach (POCA) to the z axis, and the associated error matrix. d_0 and z_0 are the distances from the origin to this POCA in the transverse (x, y) plane and along the z axis respectively. ϕ_0 is the angle between the transverse component of the track tangent vector at this

Superlayer	Inner Radius [cm]	Cells/Layer	Stereo Angle [mrad]
1	26.04	96	0
2	31.85	112	+[44.9-50.0]
3	37.05	128	-[52.3-57.4]
4	42.27	144	0
5	48.08	176	+[55.6-59.7]
6	53.32	192	-[62.8-66.9]
7	58.34	208	0
8	64.30	224	+[65.0-68.5]
9	69.52	240	-[72.1-75.8]
10	74.72	256	0

Table 2.2: *Layer arrangement for the BABAR Drift Chamber.*

POCA and the x axis. λ is the angle between the transverse plane and the track tangent vector at this POCA (the so called "dip" angle). ω is the curvature of the track. The charge of the track is incorporated in the signing convention for ω while the sign of d_0 is determined from the angular momentum of the track w.r.t. the x axis. d_0, ϕ_0, z_0 and $\tan\lambda$ resolutions are dominated by the resolution of the SVT while DCH makes a significant contribution to the p_T measurement.

Track resolution is monitored in online prompt reconstruction with Bhabha and $\mu^+\mu^-$ events, and is further investigated after the data is fully reconstructed with tracks in hadronic events and reconstructed cosmic ray events.

Figure 2.9 shows the d_0 and z_0 resolutions as a function of p_T as determined from tracks in hadronic events. For this study, the tracks not being examined were vertexed together, and the d_0 and z_0 with respect to this vertex were calculated. These were accumulated and a Gaussian fit was made to the negative part of the distribution (particle lifetime can contribute to the positive part of both the the d_0 and z_0 distributions). Requirements were also made on event thrust and vertex χ^2 to reduce the effect of decay lifetimes on this measurement. The contributions from the vertex errors were accumulated and fit for, and removed from the measured resolutions in quadrature. The d_0 and z_0 resolutions so measured are about 25 and 40 μm respectively at $p_T = 3$ GeV/c. These agree well with resolutions measured in other studies.

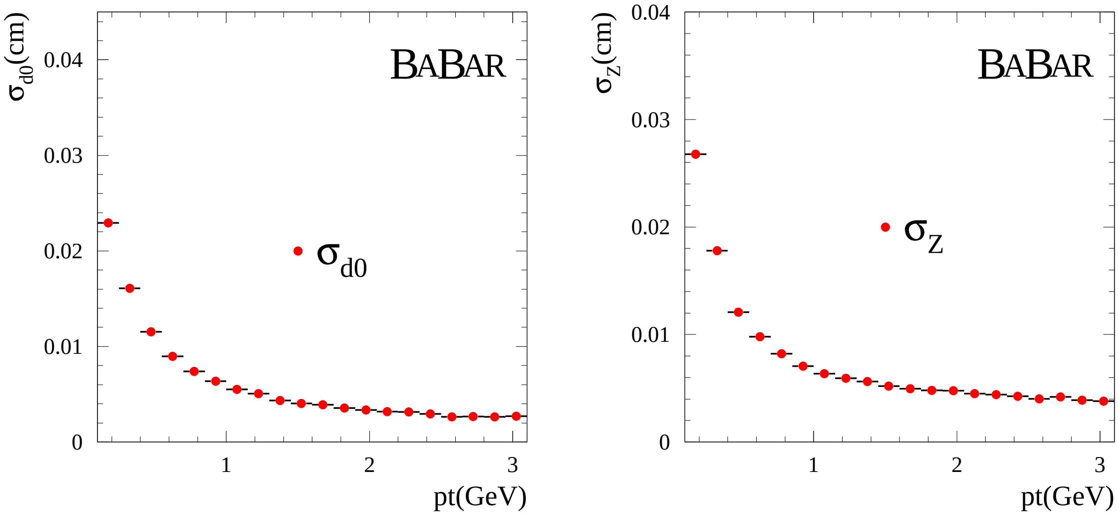


Figure 2.9: Impact parameter resolution of tracks, reconstructed using all available tracking information,in the xy plane as a function of transverse momentum (left) and for the z projection (right).

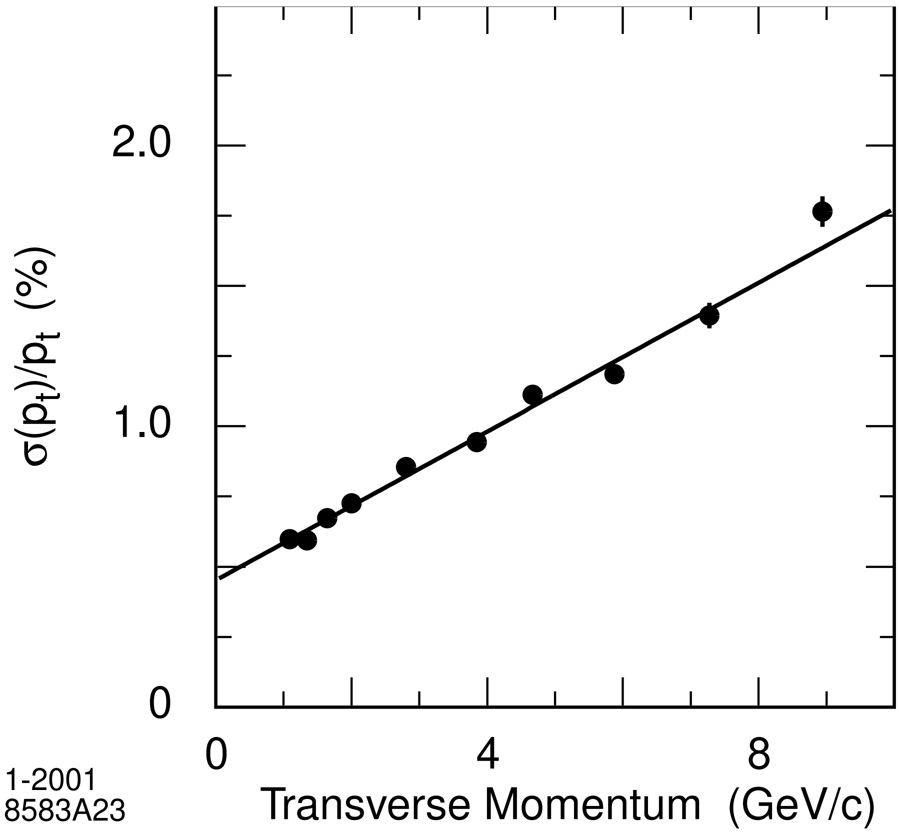


Figure 2.10: p_T resolution determined from cosmic ray muons.

The ϕ_0 and $\tan\lambda$ resolutions are measured in the cosmic ray studies to be 0.43 mrad and 0.53×10^{-3}, respectively. The p_T resolution, directly related to the ω resolution, is measured in the cosmic ray studies as a function of p_T (figure 2.10). The data are well represented by a linear function:

$$\frac{\sigma_{p_T}}{p_T} = (0.13 \pm 0.01)\% \cdot p_T + (0.45 \pm 0.03)\% \, ,$$

where p_T is measured in GeV/c. For the cosmic rays studies, cuts were made on the d_0, z_0 and $\tan\lambda$ of the track to make sure the track passed close to the nominal beam interaction point in d_0 and z_0 and populated the parts of the detector normally populated with tracks in beam interaction events. The upper and lower halves of the cosmic ray were fit as two separate tracks and the resolutions were derived by comparing the values of the track parameters for the two tracks at their POCAs to the z axis.

2.6.4 dE/dx Resolution

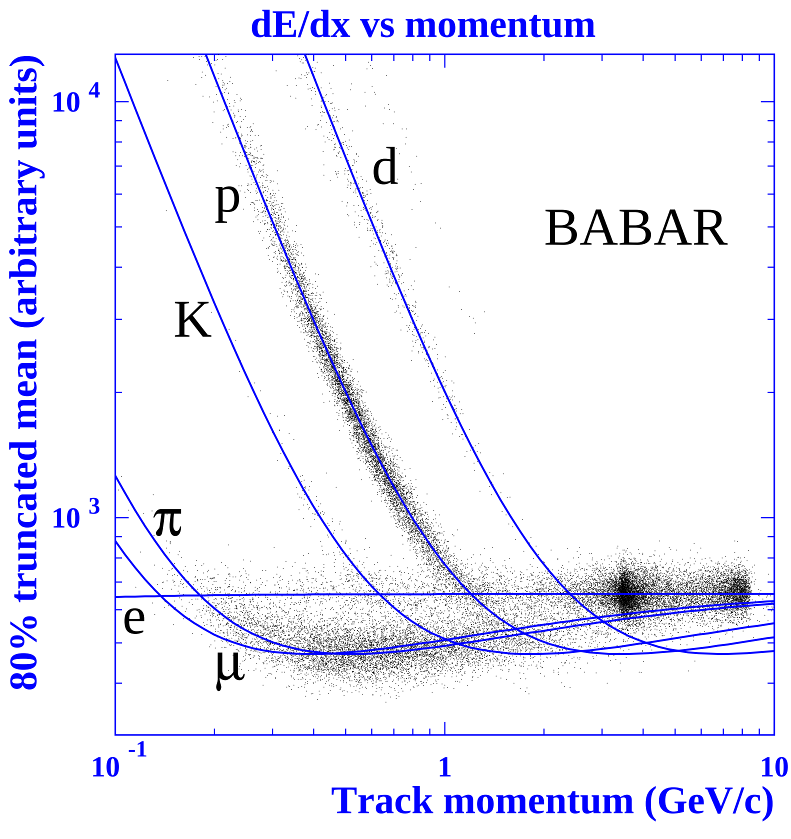


Figure 2.11: Reconstructed dE/dx as a function of track momenta.

Figure 2.11 shows the distribution of the reconstructed and corrected dE/dx from the drift chamber as a function of track momenta. The superimposed Bethe-Block curves for the individual particle species have been determined using various particle control samples.

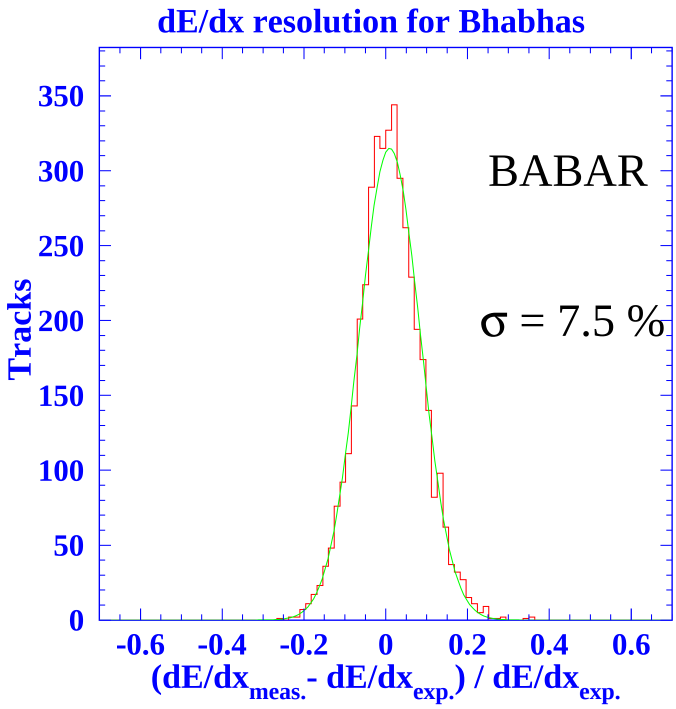


Figure 2.12: Distribution of the difference between the measured and expected dE/dx for Bhabha while operating at 1900 V.

The resolution achieved to date is typically about 7.5 % for Bhabha, as shown in figure 2.12. This is close to the expected resolution of 7 %, which we expect to achieve with further refinement or addition to the applied corrections.

Independent dE/dx of charged tracks is a less important, but non-negligible, function of the SVT. A truncated mean algorithm is used, keeping 60 % of the hits for tracks with at least 4 associated SVT hits. The resulting SVT dE/dx distribution as a function of momentum is shown in figure 2.13.

2.7 Cherenkov light detector

The DIRC (figure 2.14), the Detector of Internally Reflected Cherenkov radiation, is a novel concept and technology employed primarily for the separation of pions and kaons from about 500 MeV/c to the kinematic limit of 4 GeV/c.

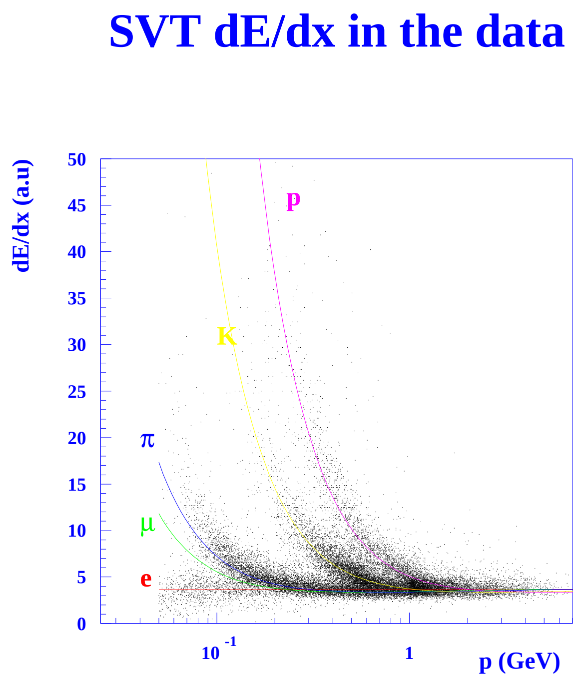


Figure 2.13: Energy loss per unit length (dE/dx) as measured in the SVT as a function of momentum. The enhancement of protons is due to beam-gas interactions. The vertical scale is arbitrary.

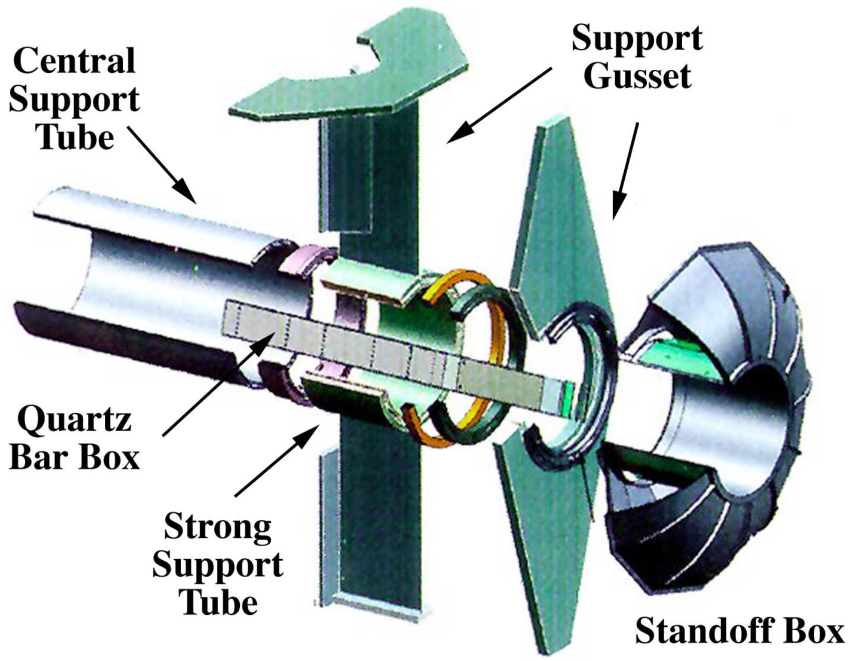


Figure 2.14: Schematics of the DIRC mechanical support structure.

2.7.1 Principles and layout

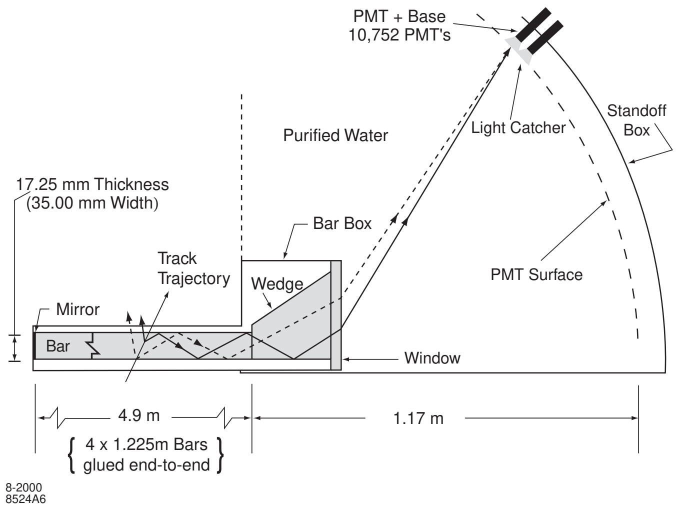


Figure 2.15: Schematics of the DIRC fused silica radiator bar and imaging region.

The DIRC [28, 29] is a novel type of ring-imaging Cherenkov detector, based on the principle that the magnitudes of angles are maintained upon reflection from a flat surface. Figure 2.15 shows a schematic of the DIRC geometry that illustrates the principles of light production, transport, and imaging. The radiator material of the DIRC is synthetic fused silica (also referred to as "quartz") in the form of long, thin bars with regular rectangular cross section. These bars serve both as radiators and as light pipes for the portion of the light trapped in the radiator by total internal reflection. A charged particle with velocity v, traversing the fused silica bar with index of refraction $n = 1.473$, generates a cone of Cherenkov photons of half-angle θ_c with respect to the particle direction, where $\cos\theta_c = 1/\beta n$. For particles with $\beta \approx 1$, some photons will always lie within the total internal reflection limit, and will be transported to either one or both ends of the bar, depending on the particle incident angle. To avoid having to instrument both bar ends with photon detectors, a mirror is placed at the forward end, perpendicular to the bar axis, to reflect incident photons to the backward (instrumented) bar end. Once photons arrive at the instrumented end, most of them emerge into a water-filled expansion region, called the stand-off box. A fused silica wedge at the exit of the bar reflects photons at large angles and thereby reduces the size of the required detection surface. The photons are detected by an array of densely packed photomultiplier tubes (PMTs), each surrounded by reflecting "light catcher" cones [30]to capture light which would otherwise miss the PMT active area. The PMTs are placed about 1.2 m from the bar end. The expected Cherenkov light pattern at this surface is essentially a conic section, whose cone opening-angle is the Cherenkov production angle modified by refraction at the exit from the fused silica window. The time taken for the photon to travel from the point of origin to the PMT is also related to the photon propagation angle $(\alpha_x, \alpha_y, \alpha_z)$ with respect to the bar axis. As the track position and angles are known from the tracking system, these three α angles can be used to (over-)determine the Cherenkov angle θ_c. This over-constraint on the angles is particularly useful in dealing with high background rates.

2.7.2 Cherenkov angle reconstruction

An unbinned maximum likelihood formalism incorporates all information provided by the space and time measurements from the DIRC. The emission angle and the arrival time of the Cherenkov photons are reconstructed from the observed space-time coordinates of PMT hits, transformed into the Cherenkov coordinate system. The known spatial position of the bar through which the track passed and the PMTs hit within the readout window of $\pm 300\,\mathrm{ns}$ of the trigger signal is used to calculate the three-dimensional vector pointing from the bar-end center to the center of each tube. This vector is then extrapolated into the quartz bar using Snell's law. This procedure defines, up to the top/bottom, left/right, forward/backward and wedge/no-wedge ambiguities, the Cherenkov angles θ_c and ϕ_c (azimuth) of a photon. The DIRC time measurement represents the third dimension of the photomultiplier hit reconstruction. The timing resolution is insufficiently precise to be competitive with the positional information for Cherenkov angle reconstruction, but is used to suppress background hits from the accelerator and, more importantly, other tracks in the same event. It is also used to resolve the forward-backward and wedge ambiguities in the hit-to-track association. The relevant observable to distinguish between signal and background photons is the difference between the measured and expected photon time, δt_γ.

It is calculated for each photon using the track-time of the PMT and the photon propagation time within the bar and the water filled standoff box. Applying the time information substantially improves the correct matching of photons with tracks and reduces the number of accelerator induced background hits by approximately a factor 40, as can be seen in figure 2.16. The reconstruction routine currently provides a likelihood value for each of the 5 particle types (e, μ, τ, K, p) if the track passes through the active volume of the DIRC. These likelihoods are calculated in an iterative process by maximizing the likelihood value for the entire event while testing different hypotheses for each track.

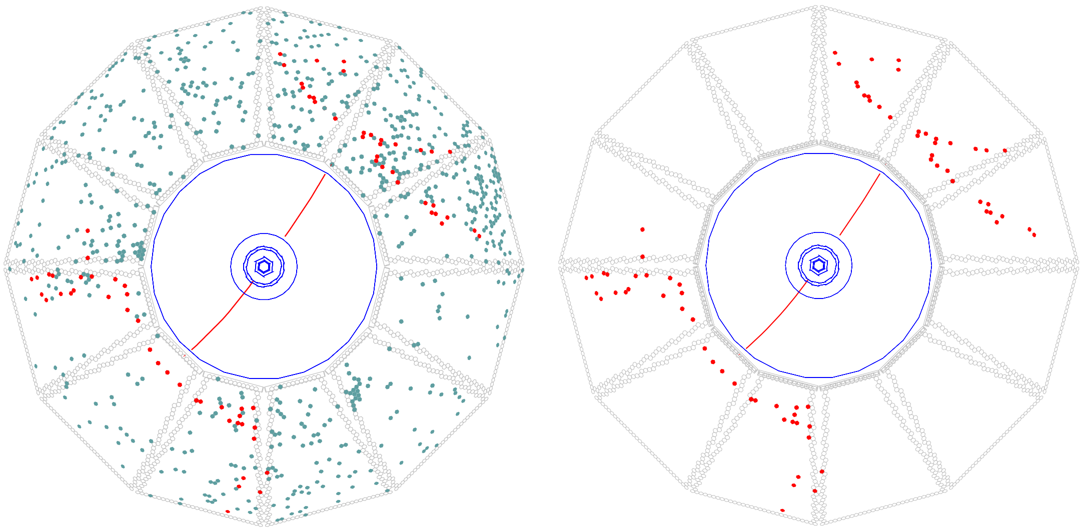


Figure 2.16: Display of one $e^+e^- \to \mu^+\mu^-$ event reconstructed in *BABAR* with two different time cuts. On the left, all DIRC PMTs that were hit within the ± 300 ns trigger window are shown. On the right, only those PMTs that were hit within 8 ns of the expected Cherenkov photon arrival time are displayed.

2.7.3 Performance

Figure 2.17 shows the number of photoelectrons as a function of the polar angle. It varies from a minimum of about 20 for small polar angles at the center of the barrel to well over 50 at large polar angles. This is in good agreement with the value expected from the Monte Carlo simulation at all angles. The bump at $\cos\theta = 0$ is a result of the fact that for tracks at small angles internal reflection of the Cherenkov photons occurs in both the forward and backward direction. This spectrum also demonstrates a very useful feature of the DIRC in the *BABAR* environment, namely, the performance improves in the forward direction, as is needed to cope with the angle-momentum correlation of particles from the boost. The small decreasing of the number of photons from the backward direction to the forward one is a consequence of the photon absorption along the bar before reaching the stand-off box in the backward end.

With the present alignment, the typical track Cherenkov angle resolution for di-muon events is shown in Figure 2.18 to be about 2.4 mrad. This is in good agreement with the simulation.

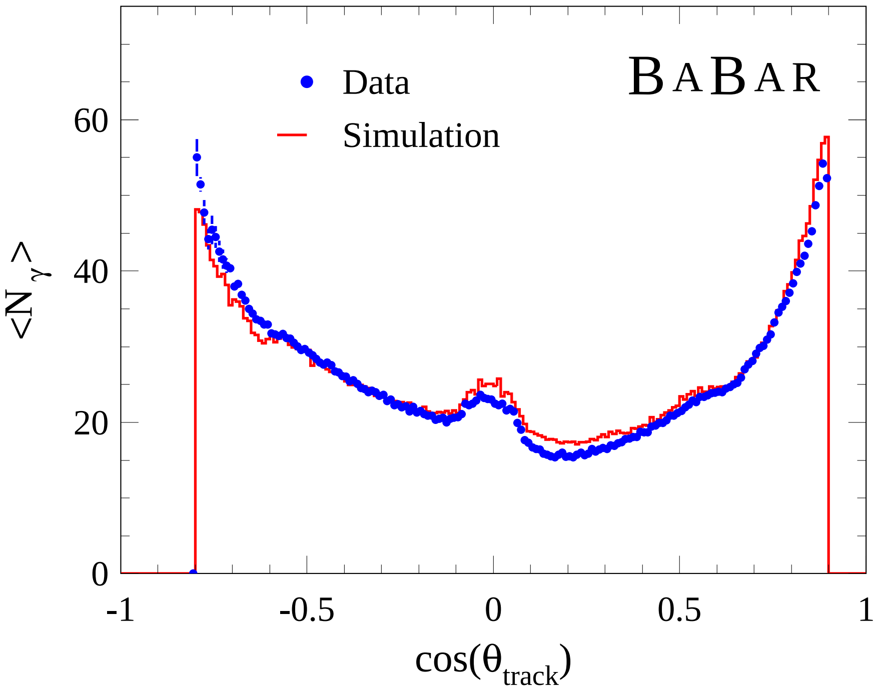


Figure 2.17: Number of detected photoelectrons versus track polar angle for reconstructed di-muon events in data and simulation.

From the measured single track resolution *vs.* momentum in di-muon events and the difference between the expected Cherenkov angles of charged pions and kaons, the pion-kaon separation power of the DIRC can be inferred. As shown in figure 2.19, the separation between kaons and pions at 3 GeV/c is about 4.3 σ, which agrees with the predicted separation using Monte Carlo simulation.

Figure 2.20 shows an example of the use of the DIRC for sample selection. The $K\pi$ invariant two body mass spectra are shown without (a), and with (b) the use of the DIRC for kaon identification. The mass peak corresponds to the decay of the D^0 particle. The efficiency for correctly identifying a charged kaon that hits a radiator bar and the probability to wrongly identify a pion as kaon are determined using D^0 decays selected kinematically from inclusive D^* production and are shown as a function of track momentum in figure 2.21. The mean kaon selection efficiency and pion mis-identification are 96.2 ± 0.2 %(stat) and 2.1 ± 0.1 %(stat), respectively [29].

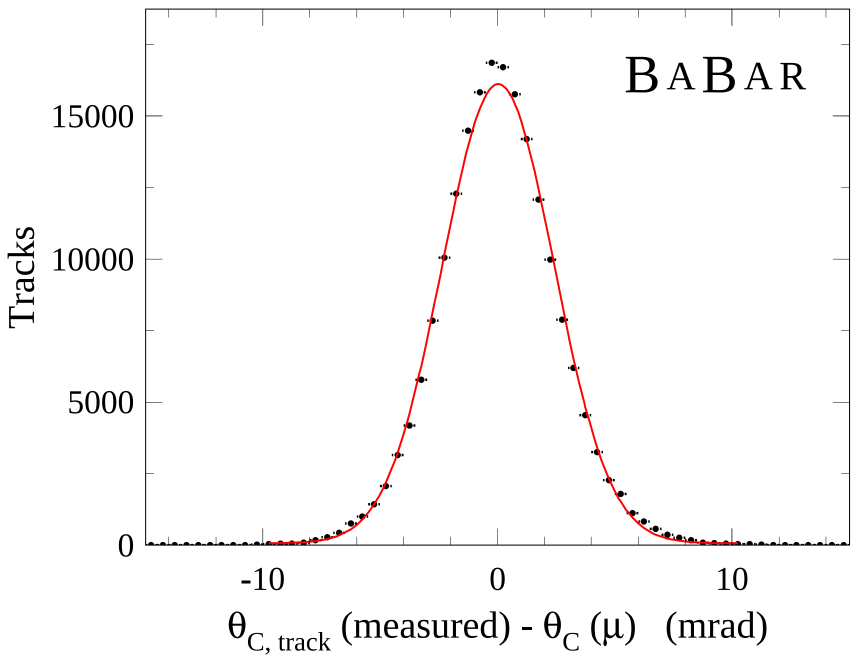


Figure 2.18: Resolution of the reconstructed Cherenkov polar angle per track for di-muons.

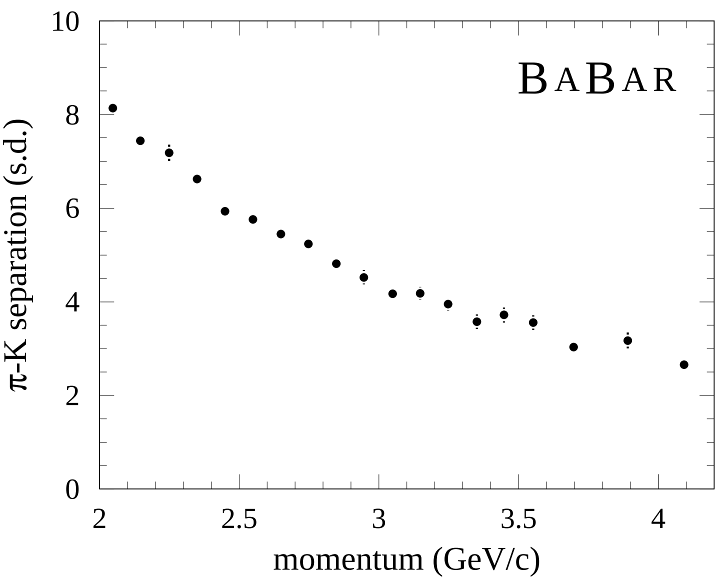


Figure 2.19: $\pi - K$ separation in $B^0 \rightarrow \pi^+\pi^-$ events versus tracks momentum inferred from the measured Cherenkov angle resolution and number of Cherenkov photons per track in di-muon events.

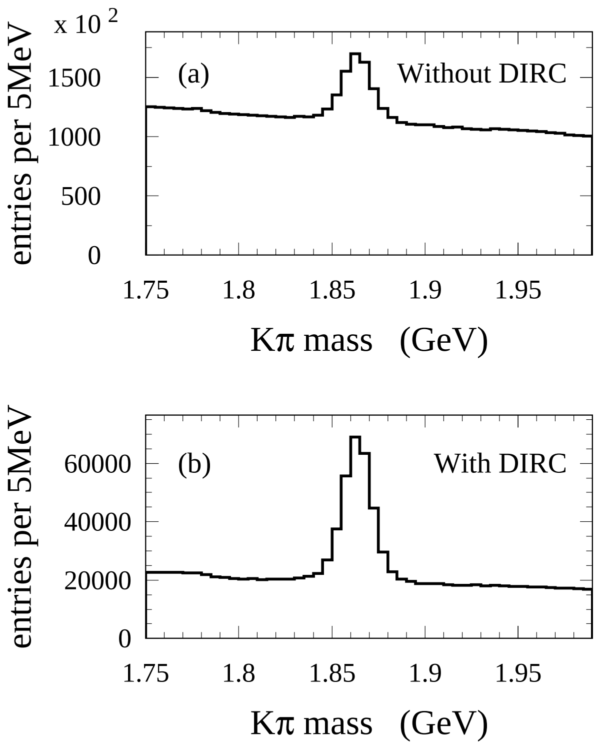


Figure 2.20: Invariant $K\ \pi$ inclusive mass spectrum (a) without and (b) with the use of the DIRC for kaon identification. The mass peak correspnds to the decay of the D^0 particle.

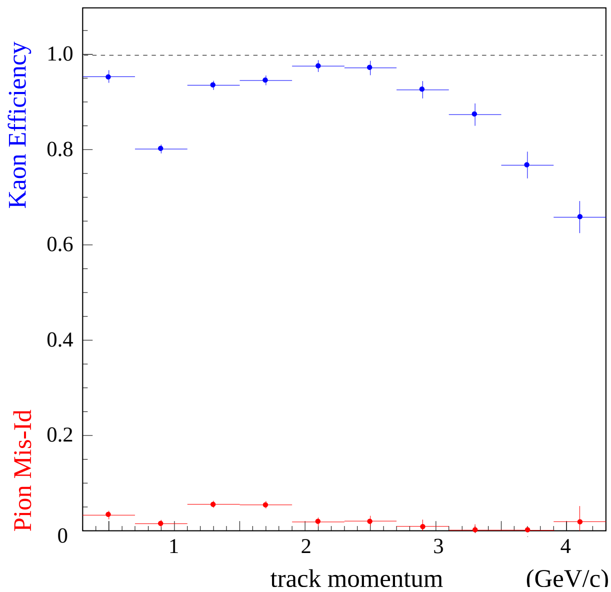


Figure 2.21: Efficiency and mis-identification probability for the selection of charged kaons as a function of track momentum determined using D^0 decays selected kinematically from inclusive D^* production.

2.8 Electromagnetic calorimeter

2.8.1 Design requirements

The *BABAR* electromagnetic calorimeter is required to detect low-energy photons with high efficiency, and to precisely measure their energy and direction. It is also the primary sub-detector providing electron-hadron separation. The efficient reconstruction of extremely rare decays of B mesons containing π^0s (e.g. $B^0 \rightarrow \pi^0\pi^0$) pose the most stringent design requirements on energy resolution of order 1% while excellent photon efficiency at low energy ($\sim$ 20 MeV) is required for efficient reconstruction of B meson decays containing multiple π^0 and η [31, 29]. Similar precision is required for efficient separation of electrons and hadrons with purities required at the 0.1% level for momentum as low as 500 MeV/c. The upper bound on the energy range required is set by the highest energy particles produced. At the $\Upsilon(4S)$ these are 9 GeV electrons from Bhabha events. The π^0 mass resolution is dominated by the energy resolution at low energies (less than 2 GeV) and by the angular resolution at high energies. The angular resolution is required to be a few milliradians in order to maintain good π^0 resolution at all energies. The need for high efficiency requires hermetic coverage of the acceptance region while excellent resolution can only be achieved by minimising the material in front of and between the active detector elements.

2.8.2 Detector layout

The electromagnetic calorimeter (figure 2.22) contains 6580 CsI crystals doped with thallium iodide at about 1000 ppm. Each crystal is a truncated trapezoidal pyramid and ranges from 16 to 17.5 radiation lengths in thickness. The front faces are typically about 5 cm in each dimension. The crystals are arranged quasi-projectively in a barrel structure of 48 θ rows by 120 crystals in azimuth (ϕ), with an inner radius of 90 cm. The forward end is closed by a separable end-cap capable of holding nine additional rows.

At present, eight are filled with crystals and the innermost with shielding absorber. The polar angle coverage of the calorimeter is $-0.78 < \cos\theta_{lab} < 0.96$. Each crystal is wrapped with a diffuse reflective material (TYVEK) and housed in a thin eggcrate-like carbon fiber composite mechanical structure. There are 280 such modules in the barrel (7 types, 40 of each type) and 20 identical end-cap modules.

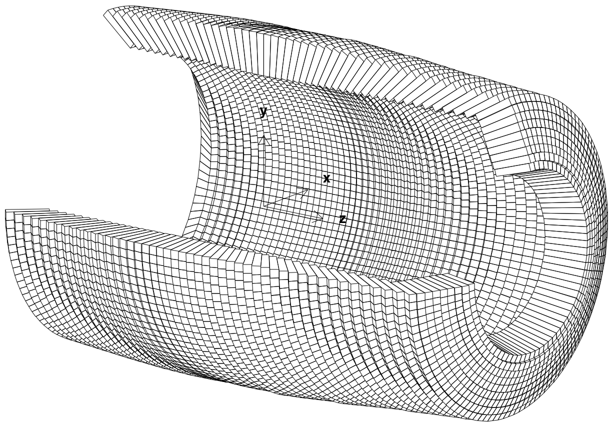


Figure 2.22: The crystal geometry of the Electromagnetic Calorimeter.

Crystals are read out with two independent 2 cm^2 large PIN photodiodes epoxied to their rear faces.

2.8.3 Crystals description

Radiation Length	1.86 cm
Molière Radius	3.8 cm
Density	4.53 g/ cm^3
Light Yield	50000 γ/ MeV
Light Yield Temperature Coefficient	0.28 %/°C
Peak Emission λ_{max}	565 nm
Refractive Index (λ_{max})	1.79

Table 2.3: *Properties of CsI(Tl).*

The high light yield and small Molière radius of CsI(Tl) gives the excellent energy and angular resolution required, while the short radiation length guarantees complete shower containment at *BABAR* energies with a relatively compact design. Furthermore, the high light yield and peak of the emission spectrum permit an efficient use of a silicon photodiode readout.

The limiting energy resolution of a homogeneous calorimeter is determined by fluctuations in the electromagnetic shower propagation and in the case of a crystal detector is empirically described by equation 2.1. The resolution as a function of energy E is given by σ_E parametrised as the quadratic sum of a stochastic term σ_1 and a constant term σ_2.

$$\frac{\sigma_E}{E} = \sigma_1 \ E^{-\frac{1}{4}} \oplus \sigma_2 \oplus \frac{\sigma_3}{E} \tag{2.1}$$

The stochastic term, σ_1 which is dominant at low energies (less than 100 MeV), arises from fluctuations in photon statistics, while electronic noise in the readout chain is takne into account in σ_3.

The constant term, dominant at higher energies (above 1 GeV), arises from several effects of which the main are fluctuations in shower containment and albedo.

Non-uniformities in the crystals result in a non-linear energy response as mean shower depth varies logarithmically with incident energy. Some can be reduced to negligible levels initially, by careful quality assurance of the crystals and the wrapping procedures during construction. However, radiation damage during operation causes a degradation in response that changes with time. This requires a novel energy dependent calibration and monitoring system which employs a radioactive source, a light pulser, and reconstructed data. These calibrations must be updated every few days during operation, and be fully automated to track time dependent radiation-damage effects.

The angular resolution improves as the transverse size of the crystal decreases. However the energy resolution is degraded by the summing of electronic noise from each crystal. The transverse size of the crystals is chosen to be comparable to the Molière radius: this allows to achieve the required angular resolution[4] at low energies. The total number of crystals and readout channels is so limited to an acceptable noise and cost level.

2.8.4 Performance

Energy resolution

The energy resolution is $\sigma_E/E = 5.0 \pm 0.8$ % at 6.13 MeV and $\sigma_E/E = 1.9 \pm 0.07$ % at 7.5 GeV. Figure 2.23 shows the energy resolution extracted from a variety of data as a function of energy. The mass resolution of π^0 and η candidates in which the two photons in the decay have approximately equal energy can be used to infer the energy resolution at an energy less than 1 GeV. The process $\chi_{c1} \rightarrow J/\psi \ \gamma$ provides

[4]few milliradians

an additional measurement at 500 MeV.

The expectation from Monte Carlo, where the contributions from electronic noise and low energy beam backgrounds are included, is shown as a dashed curve. The data fit curve given by equation 2.1 with $\sigma_1 = 2.32 \pm 0.30$ % and $\sigma_2 = 1.85 \pm 0.12$ %.

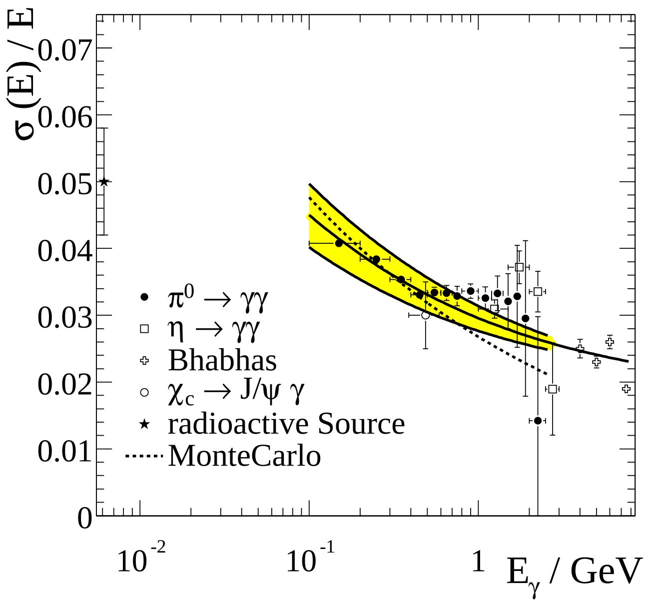


Figure 2.23: The energy resolution for the electromagnetic calorimeter measured using photon and electron candidates. The solid curve is a fit to equation 2.1 and the shaded area denotes the one sigma error on the fit.

Angular resolution

Figure 2.24 shows the angular resolution measured as a function of energy. The decays of π^0 and η candidates in which the two photons in the decay have approximately equal energy are used to infer angular resolution. It varies between about 12 milliradians at low energies and 3 milliradians at high energy. The expectation from Monte Carlo is shown in the dashed curve. The data fit the empirical parametrisation:

$$\frac{\sigma_{\theta,\phi}}{E} = \frac{\sigma_1}{\sqrt{E}} + \sigma_2 \tag{2.2}$$

where $\sigma_1 = 3.87 \pm 0.07$ mrad and $\sigma_2 = 0.00 \pm 0.04$ mrad. There are many more π^0 than η decays. However the statistical significance of the disagreement between data and Monte Carlo is similar for both samples.

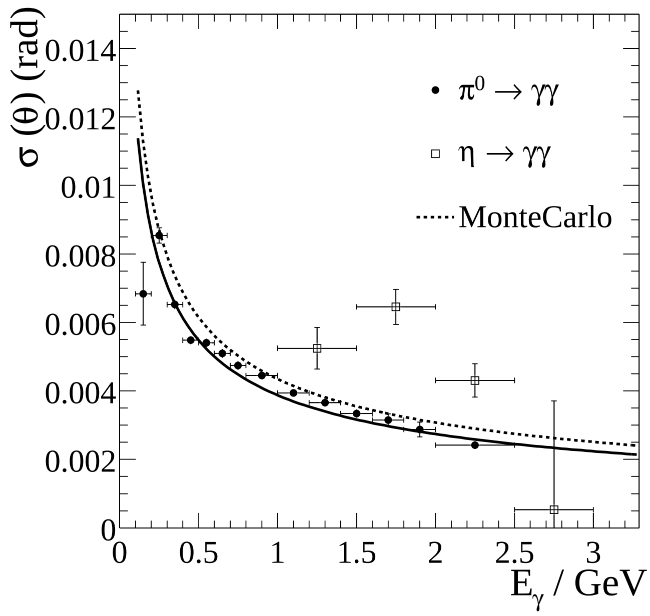


Figure 2.24: The angular resolution for the electromagnetic calorimeter measured using photon and η candidates. The solid curve is a fit to equation 2.2.

π^0 Mass and Width

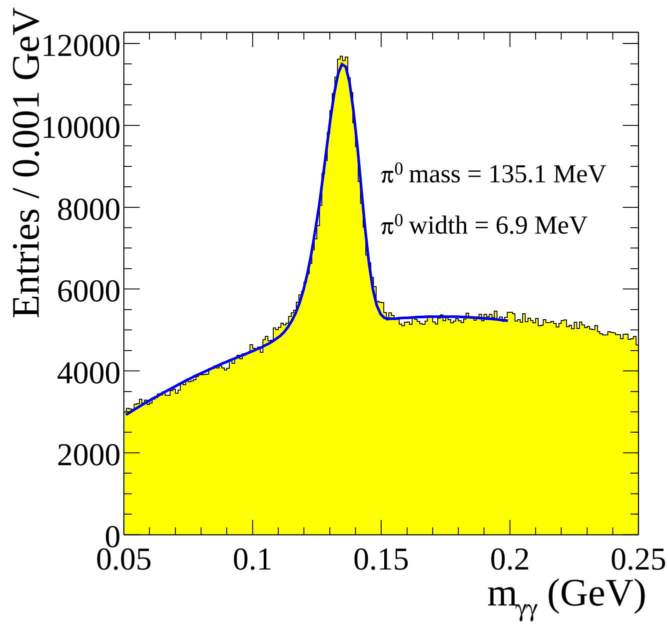


Figure 2.25: The π^0 mass peak reconstructed from photon candidates in hadronic events. The photon candidates are required to have an energy of at least 30 MeV and the energy of the π^0 must be greater than 300 MeV to reduce combinatoric backgrounds. The solid line is a fit to the data [29].

Figure 2.25 shows the two-photon invariant mass for π^0 candidates. The π^0 candidates are taken from hadronic B meson decays. The invariant mass is stable to less than 1 % over the full photon energy range. The width of 6.9 MeV/c^2 compares to a Monte Carlo estimate of 6.8 MeV/c^2 in hadronic B meson events. In low occupancy $\tau^+\tau^-$ events the width is observed to be 6.5 MeV/c^2 for π^0 energies below 1 GeV.

Electron-Hadron separation

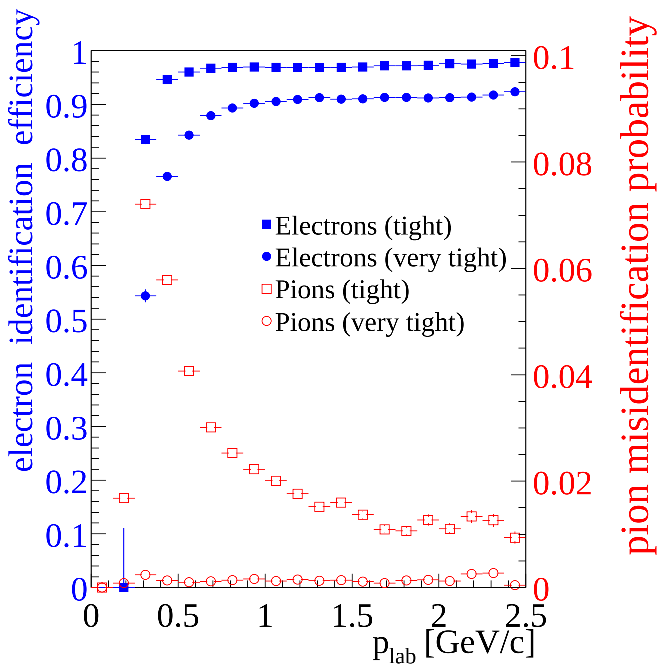


Figure 2.26: The electron efficiency and pion mis-identification rate for different momenta.

Electron-hadron separation is accomplished by use of the shower energy, lateral shower moments (see 4.5) and incident track parameters. The comparison of shower energy and incident momentum (E/p) is the most significant separation variable. Figure 2.26 shows the electron efficiency and pion misidentification rate for different momenta using a selection algorithm. The efficiency of electron identification is measured using electrons from radiative Bhabhas and $\gamma\gamma \rightarrow e^+e^-$ events. The pion misidentification probability is measured in three prong τ decays. The very tight selector has an efficiency of 90.8 % with an average pion misidentification of 0.2 % The electron-hadron separation power meets design expectations.

2.9 Instrumented flux return

2.9.1 Physics Requirements and Goals

The Instrumented Flux Return (IFR) was designed to identify muons with high efficiency and good purity and to detect neutral hadrons (primarily K_L^0 and neutrons) over a wide range of momenta and angles. Muons are important for tagging the flavor of neutral B mesons via semi-leptonic decays, for the reconstruction of vector mesons, like the J/ψ, and the study of semi-leptonic and rare decays involving leptons from B and D mesons and τ leptons. K_L^0 detection allows for the study of exclusive B decays, in particular CP eigenstates, it could also help in vetoing charm decays and improve the reconstruction of neutrinos. The principal requirements for IFR are large solid angle coverage, good efficiency and high background rejection for muons down to momenta below 1 GeV/c. For neutral hadrons, high efficiency and good angular resolution are most important.

2.9.2 Overview and RPC Concept

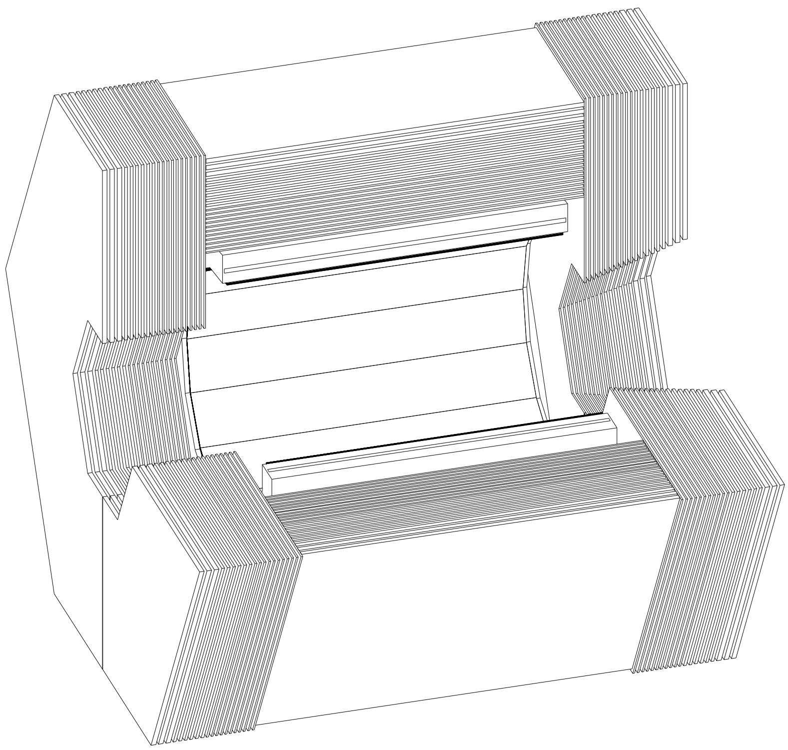


Figure 2.27: Overview of the IFR Barrel sectors and forward and backward end-doors; the shape of the RPC modules and the way they are stratified is shown.

The IFR uses the steel flux return of the magnet as muon filter and hadron absorber. Single gap resistive plate chambers (RPC) [32, 29], with two-coordinate readout have been chosen as detectors. The RPC are installed in the gaps of the finely segmented steel of the six barrel sectors and the two end-doors of the flux return, as illustrated in figure 2.27. The steel segmentation has been optimised on the basis of Monte Carlo studies of muon penetration and charged and neutral hadron interactions. The steel is segmented into 18 plates, increasing in thickness from 2 cm of the inner 9 plates to 10 cm of outermost plates for a total 65 cm. In addition, two layers of cylindrical RPCs are installed between the EMC and the magnet cryostat to detect particles exiting the EMC.

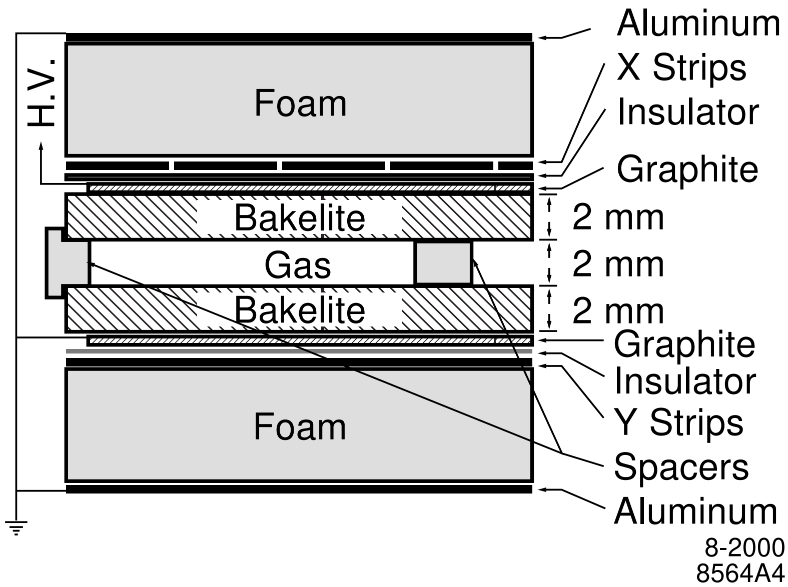


Figure 2.28: Cross section of a planar RPC with the schematics of the HV connection.

RPCs detect streamers from ionising particles via capacitive readout strips. They offer the advantage of simple and low cost construction. Further benefits are large signals and fast response allowing for simple and robust front end electronics and good time resolution, typically 1-2 ns. The position resolution depends on the segmentation of the readout; few millimeters are achievable. A cross section of an RPC is shown schematically in figure 2.28. The planar RPCs consist of two bakelite (phenolic polymer) sheets, 2 mm thick and separated by a gap of 2 mm. The bulk resistivity of the bakelite sheets has been especially tuned to $10^{11} - 10^{12}\,\Omega\,\mathrm{m}$. The external surfaces are coated with graphite to achieve a surface resistivity of approximately 100 $k\Omega/square$. These two graphite surfaces are connected to high voltage (approximately 8 kV) and ground, and protected by an insulating mylar film. The bakelite surfaces facing the gap are treated with linseed oil to improve performance. The modules are operated in limited streamer mode and the signals are read out capacitively, on both sides of the gap, by external electrodes made of aluminium strips on a mylar substrate.

2.9.3 Efficiency Measurements and Performance

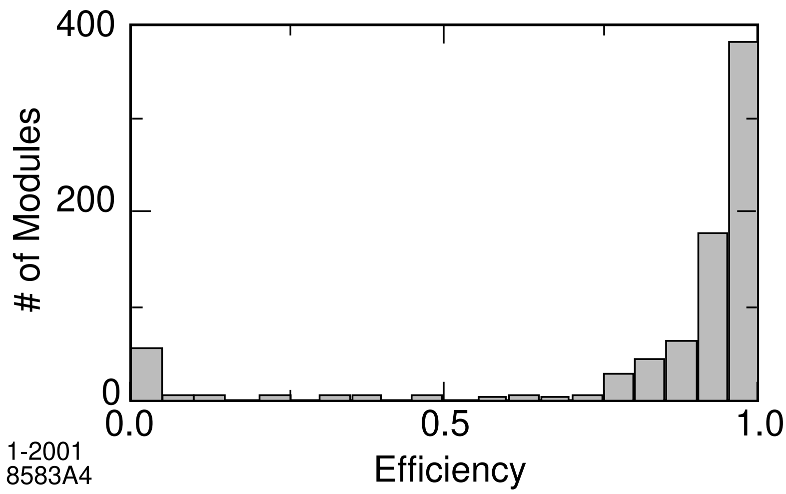


Figure 2.29: Distribution of the efficiency for all RPC modules measured with cosmic rays in June 1999. Some 50 modules were not operational at that time.

The efficiency of the RPCs is evaluated for both normal collision data and dedicated cosmic rays. Every week cosmic ray data are recorded at different voltage settings and the efficiency is measured chamber by chamber as a function of the applied voltage. The absolute efficiency at the nominal working voltage (tipically 7.6 kV) is stored in the database for use in the event reconstruction software. To calculate the efficiency in a given chamber, nearby hits in a given layer and hits in different layers are combined to form clusters. One-dimensional IFR clusters are defined as a group of adjacent hits in one of the two readout coordinates. The cluster position is defined as the centroid of the strips in the cluster. Two-dimensional clusters are formed by joining one-dimensional clusters (of the same readout coordinate) in different layers, provided the distance between their coordinate centroids is less than a given value. In each sector, two-dimensional clusters in different coordinates are combined to three-dimensional clusters as long as there are less than three layers missing one of the two coordinates. Charged tracks reconstructed by the drift chamber are extrapolated into the IFR. Clusters which are less than 12 cm from the extrapolated track are combined to form three-dimensional or two-dimensional clusters. The residual distributions from straight line fits to two-dimensional clusters typically have an r.m.s. width of less than 14 cm. An RPC is considered efficient if a signal is detected at a distance of less than 10 cm from the fitted straight line in either of the two readout planes. Following the installation and commissioning of the IFR system all the RPC modules were tested with cosmic rays and their efficiency was measured. The results are presented in figure 2.29. Of the RPC active modules 75 % exceeds an efficiency of 90 %.

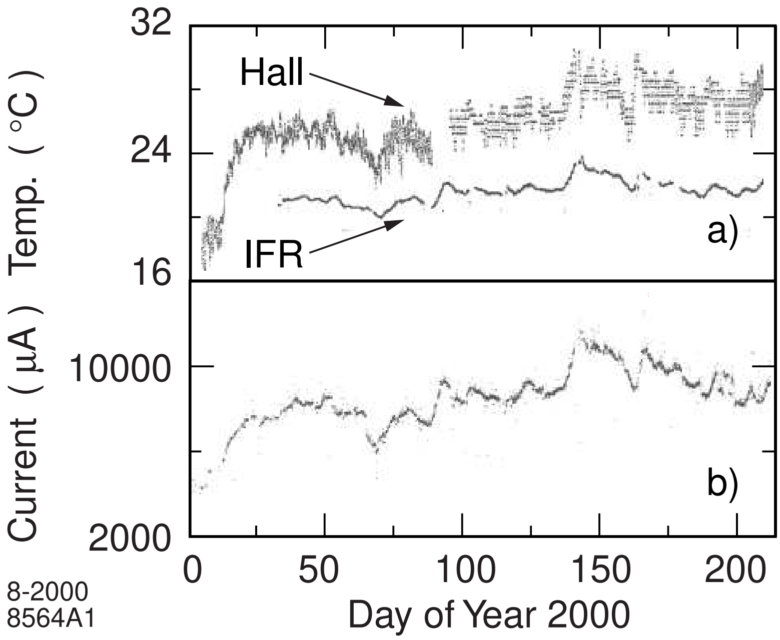


Figure 2.30: History of the temperature and dark current in the RPC modules in the first part of the year 2000. The top figure shows the temperature in the IR-2 hall and in the IFR backward end-door. The bottom figure shown the total dark current in the backward end-door.

Early tests indicated that the RPC dark current was very temperature dependent, specifically, the current increases 14-20 % per °C. Since the IR experimental hall does not have temperatures regulation this presents a serious problem. The FECs that are installed in the steel gaps each produce 3 W, adding up to a total power dissipation of 3.3 kW in the barrel and 1.3 kW in the forward end-door. During the first summer of operation the daily average temperature in the IR hall was about 28 °C and the maximum hall temperature frequently exceeded 31 °C. The temperature inside the steel rose to more than 37 °C and the dark currents in many modules exceeded the capabilities of the HV system and RPCs had to be temporarily disconnected. To overcome this problem water cooling was installed on the barrel and end-door steel, removing approximately 10 kW of heat and stabilizing the temperature to 20-21 °C in the barrel, 22 °C in the backward and 24 °C in the forward end-doors. Figure 2.30 shows the history of temperature in the hall and temperature and total dark current in the backward end-door. While the current closely follow the temperature variations, the range of change is now limited to a few degrees.

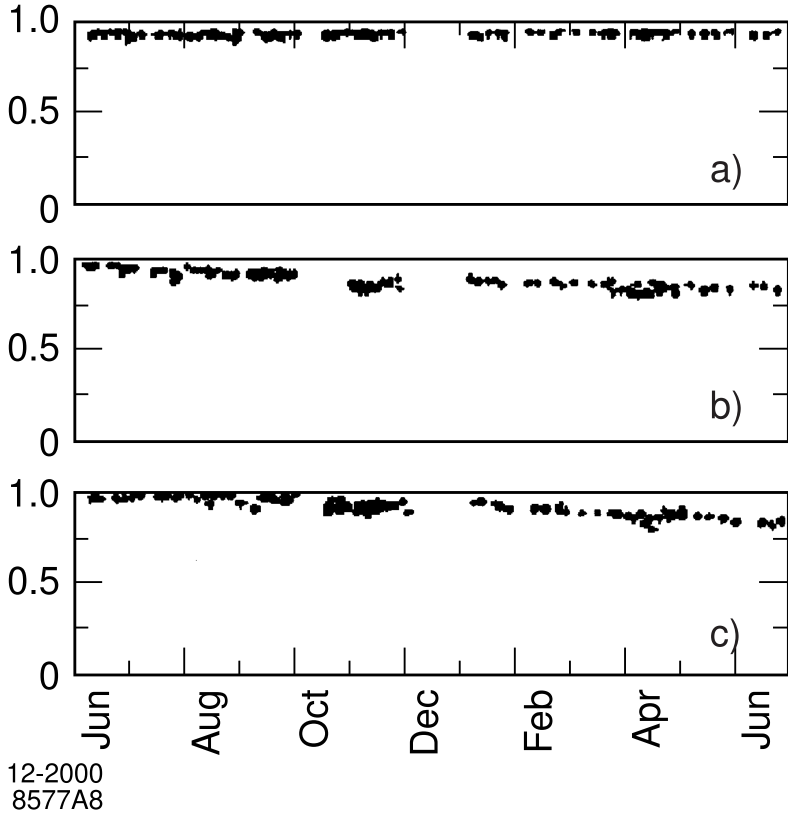


Figure 2.31: Efficiency history for 12 months starting in June 1999 for RPC modules showing different performance. Top: highly efficient and stable. Middle: slow continuous decrease in efficiency. Bottom: faster decrease in efficiency.

During the operations at high temperatures a large fraction of the RPCs (more than 50 %) showed not only very high dark currents but also a reduction in efficiency [33]. After the cooling was installed and the RPCs were reconnected, some of them continued to deteriorate while others remained stable, some of them (more than 30 %) at full efficiency. (see figure 2.31). Detailed efficiency studies revealed large regions of very low efficiencies in the modules, but no clear pattern was identified. The cause of the efficiency loss remains under investigation. Several possible causes have been included as the primary source of the problem, such as the change in the bakelite bulk resistivity, loosened spacers, gas flow or composition. A number of prototype RPCs developed similar efficiency problems after being operated above a temperature 36 °C for a period of two weeks. In some of these modules evidence was found that the linseed oil had failed to cure and had accumulated at various spots under the impact of the high voltage[5].

Muon Identification

While muon identification relies almost entirely on the IFR, other detector systems provide complementary information. Charged particles are reconstructed in the SVT and DCH and muon candidates are required to meet the criteria for minimum ionising particles in the EMC, i.e. tracks depositing large amounts of energy are rejected. Charged tracks that are reconstructed in the tracking systems are extrapolated to the IFR taking into account the non-uniform magnetic field, multiple scattering and the average energy loss. The projected intersections with the RPC planes are computed and for each readout plane all clusters detected within a maximum distance from the predicted intersection are associated with the track. For each cluster in the IFR associated with a charged track a number of variables are defined to discriminate muons from charged hadrons:

- the total number of interaction lengths traversed from the IP to the last RPC layer with an associated cluster
- the difference between this measured number of interaction lengths and the number of interaction lengths predicted for a muon of the same momentum and angle
- the average number and the r.m.s. of the distribution of RPC strips per layer
- the χ^2 the for geometric match between the projected track and the centroids of clusters in different RPC layers
- the χ^2 of a polynomial fit to the 2-dimensional IFR clusters.

[5] In order to solve some of the inefficiency problems an improvement program is currently under development. At the moment by November 2002 we should end the substitution of the external chamber RPCs in the forward end-cap region with new ones based on the same base concept but with an improved linseed oil essication technique [34]. The modules in the barrel region are more complicated to substitute and the eventual replacement is foreseen in two years from now [35].

At present, cuts on individual variables are employed to optimise the efficiency and purity of muons. The performance of muon selection has been tested on samples of muons from $\mu\mu ee$ and $\mu\mu\gamma$ final states and pions from 3-prong τ decays and $K^0_S \rightarrow \pi^+\pi^-$ decays. The selection of these control samples is based on kinematic variables, and not on variables used for muon selection. We will illustrate later the achieved muon detection efficiency and the hadron misidentification using Very-Tight selection cuts on the specific sample used for this analysis.

K^0_L and Neutral Hadron Detection

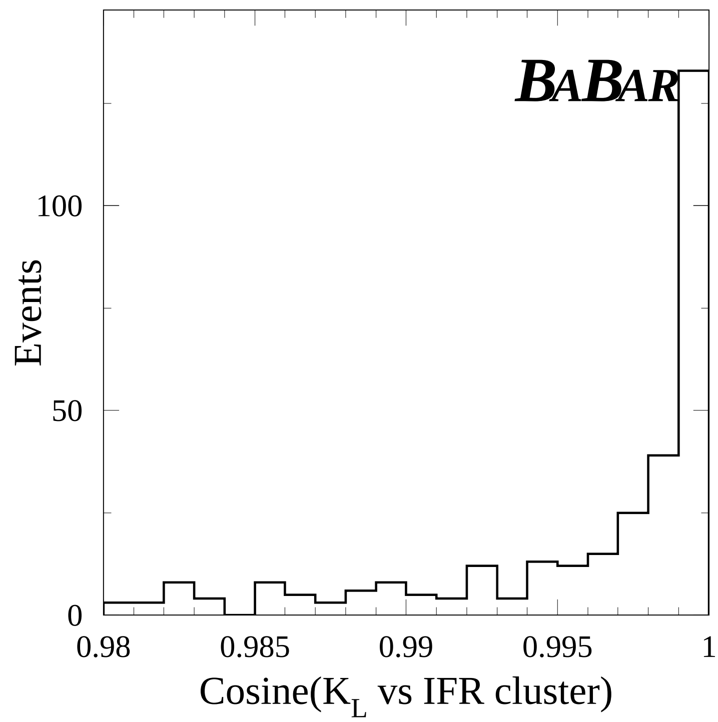


Figure 2.32: Angular difference, $\cos\Delta_\theta$, between the direction of the missing momentum and the closest neutral IFR cluster for a sample of ϕ mesons produced in the reaction $e^+e^- \rightarrow \phi\gamma$ with $\phi \rightarrow K^0_L K^0_S$.

K^0_L and other neutral hadrons interact in the steel of the IFR and can be identified as clusters that are not associated with a charged track. Monte Carlo simulations predict that about 64 % of K^0_L's above momentum of 1 GeV/c produce one cluster in the cylindrical RPC, and/or a cluster with hits in two or more planar RPC layers. The centroid of the 3-dimensional IFR cluster is taken as the interaction point of the neutral hadron, and thus defines its production angle. Since a significant fraction of the hadrons interacts before reaching the IFR, information from the EMC and the cylindrical RPCs is combined and the production angle of the neutral hadron is taken from the first interaction point in the detector. Neutral showers in the EMC are associated with the neutral hadrons detected in the IFR, based on a match in production angles. For a good match a χ^2 probably of ≥ 1 % is required, assuming uncorrelated angular errors of $\sigma_\phi = 180\,\mathrm{mrad}$ and $\sigma_\theta = 150\,\mathrm{mrad}$ for the polar and azimuthal angles. An estimate of the angular resolution of the neutral hadron cluster can be inferred from figure 2.32 indicating that the resolution is of the order of 50 mrad.

2.10 Electronics, Trigger, Data Acquisition and Online Computing

The electronics, trigger, data acquisition and on-line computing represent a collection of tightly coupled hardware and software systems. These systems were designed to maximise the physics data acceptance, maintainability and reliability while minimising complexity, dead-time, down-time and cost.

- *Front-End Electronics* (FEE) assemblies are physically located on the detector and consist of signal processing and digitisation electronics along with a standard digital transport mechanism into the data acquisition system

- A data reduction, transport and event building system, *Online Dataflow* (ODF), handles digitised data from FEE, performs initial data reduction and formatting[6] and builds complete events from the various data sources within the detector

- A farm of commercial Unix processors and associated software, *Online Event Processing* (OEP), provides the real-time environment within which events are processed by the *Level 3* trigger algorithms, completely reconstructed for the monitoring of a carefully selected subsample of events, and recorded in an intermediate raw data format to mass storage

- In a second farm of processors, *Online Prompt Reconstruction* (OPR), completely reconstructs all physics events, performs monitoring and constants gen-

[6] also known as *feature extraction*

eration in near real-time. Physics event data are transferred to an object database and are thus available for further analyses

- A *Fast Control and Timing System* (FCTS) generates needed beam and trigger-related signals along with timing fiducials

- A robust and flexible two-level trigger handles the full accelerator collision rate, one in hardware, *Level 1* (L1), the other one in software, *Level 3* (L3).

- An *Online Run Control* (ORC) system handles the logic for managing the state of the detector systems, starting and stopping runs, and performing calibrations. A user interface provides shifters and experts an easy access to detector operations

- A system to control and monitor the detector and its support system, *Online Detector Control* (ODC), is based upon the EPICS (Experimental Physics Industrial Control System) toolbox. This system includes communications links with the PEP-II accelerator complex. User interface provide operator access and displays.

2.10.1 Electronics

A common electronics architecture is shared by all *BABAR* detector subsystems; common vendors were chosen for many of its components such as VME crates and power supplies. Event data from the detector flows through the FEE circuitry, while monitoring and control signals are handled by a separate, parallel system.

Each FEE chain consists of an amplifier, digitiser, a trigger latency buffer for storing data during the L1 trigger processing, and a de-randomizing event buffer for storing data between the L1 accept and subsequent transfer to the data acquisition system (figure 2.33).

The digital L1 latency buffers function as fixed length data pipelines and are managed by common protocol signals generated by the FCTS. All de-randomizing event buffers function as FIFOs capable of storing a fixed number of events. During normal operations, analog signal processing, digitization and data readout occur continuously and simultaneously.

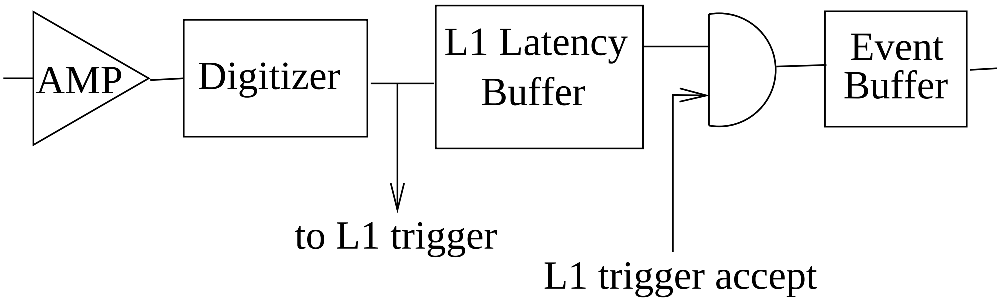


Figure 2.33: Front-End Electronics scheme. Analog signals arrive from the left, proceed conditionally through the indicated steps and are injected into the remainder of the data acqusition system.

2.10.2 Trigger

The trigger system operates as a sequence of two independent stages, the second conditional upon the first. The L1 trigger is performed first at the machine crossing rate. Its goal is to sufficiently reduce that rate to a level acceptable for the L3 software trigger which runs on a farm of commercial processors.

The L1 trigger is optimised for simplicity and speed. It consists of a pipelined hardware processor. It is designed to provide an output trigger rate of the order of 2 kHz or less. The L1 trigger selection is based on a reduced data set from the DCH, EMC and IFR. Its maximum L1 response latency for a given collision is 12 μs. Based on both the complete event and L1 trigger information, the L3 software algorithms select events of interest allowing them to be transferred to mass storage for further analysis. The L3 output rate is administratively limited to 140 Hz [26] so as not to overload the downstream processing. Dedicated L1 trigger processors receive data which is continuously clocked in from the DCH, EMC and IFR detector subsystems. The L1 trigger processor produces a clocked output to the FCTS at 30 MHz, the time granularity of resultant L1-Accept signals. The arrival of a L1-Accept signal by the data acquisition system causes a window of each subsystem's L1 latency buffer to be read out. This window is about 500 ns wide for the silicon vertex tracker and $4 \div 16$ μs wide for the calorimeter. Absolute timing information for the event, i.e. associating an event with a particular beam crossing, is determined offline. The ingredients for this calculation include the timing distributions for drift chambers hits within a track segment, waveforms from the electromagnetic calorimeter and accelerator timing fiducials.

2.10.3 Data Acquisition and Online Computing

The data acquisition and computing systems are responsible for the transport of event data from the detector FEE circuits all the way to mass storage with a minimum of dead-time and is shown schematically in figure 2.34. These systems interface with the trigger system and enable specialised modes of operation, such as calibration and testing. Other parts of these systems provide for the control and monitoring of the detector and supporting facilities.

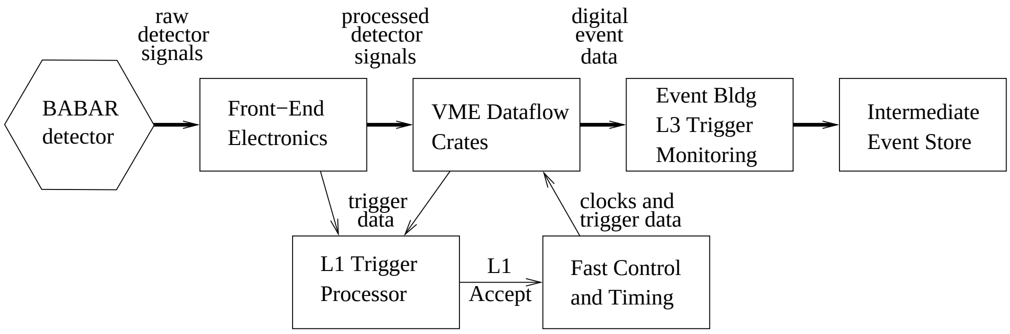


Figure 2.34: Data acquisition schematic diagram.

Online Dataflow

The Online Dataflow software connects, controls and manages the flow of data in the data acquisition hardware while incurring almost no dead-time. This code is divided between embedded processors in the ReadOut Modules (ROMs) which run the VxWorks real-time operating system [37] and UNIX processors running the Solaris operating system [38]. Dataflow provides:

1. configuration and readout of the FEE over fiber links to the ROMs
2. data transport, buffering and event building from the ROMs to the OEP farm
3. masking and prescaling of L1 triggers
4. partitioning into multiple, independent data acquisition systems

Additional code in ROMs, supplied by the detector systems, performs gain and pedestal correction, feature extraction and data formatting. Data from electronics calibrations are accumulated in the ROMs, the channel response functions are evaluated, the results are compared to reference data and subsequently applied in feature extraction. Calibration data are stored in a *Conditions Database*.

Online Event Processing

The Online Event Processing receives and processes data from the Dataflow event builder on each of the 60 UNIX processors. In particular, OEP orchestrates the following tasks:

1. L3 trigger algorithms
2. fast monitoring to assure data quality
3. logging the selected events to disk

Logging involves a process which merges the multiple data output streams to a single file.

Online Prompt Reconstruction

Bridging the online and offline worlds is OPR [39, 29]. This system reads raw data recorded to disk by OEP and, operating on a farm of 150 LINUX processors [40] , selects physics events, performs complete reconstruction, performs a "rolling calibration", collects extensive monitoring data for quality assurance and writes the result into an object database. A "rolling calibration" is the generation of reconstruction constants during normal event processing which are then applied to the processing of subsequent data.

Online Detector and Run Control

The ODC system controls and extensively monitors the electronics and environment of the detector and assures the safety of the detector. Its implementation is based on EPICS [41], providing detector-wide standardisation for the control and monitoring, diagnostics and alarm handling. ODC also provides communication with PEP-II and the magnet control system. Monitoring data are achieved in the *Ambient Database*.

The ORC system ties together the various components of the online system and provides the operator with a single graphical interface to control the detector operation. Complex configurations are stored in a *Configuration Database*; the key to the configuration used for any run is stored along with the data. The Event, Ambient, Conditions and Configurations Databases are implemented in an object database [42].

Chapter 3

Silicon Vertex Tracker

The SVT (Silicon Vertex Tracker) is the most relevant detector for the measurement of time dependent *CP* asymmetries in *BABAR*. It has been designed to provide precise reconstruction of decay vertexes and charged particle reconstruction, taking into account physics requirements and the constraints imposed by the PEP-II interaction region.

3.0.4 Physics requirements

The finite precision in the reconstruction of the z difference must degrade the error on the *CP* asymmetry by less than 10 %. In order to achieve this goal, the resolution is required to be better than 80 μm on each B decay vertex.

A resolution in the transverse plane of about 100 μm is required in order to separate secondary decay vertexes such as D^0 meson decays.

The SVT must provide stand-alone tracking for particles with transverse momentum less than 120 MeV/c, the minimum that can be reliably measured in the drift chamber alone. This feature is fundamental, especially for this analysis, for the identification of slow pions from D^* meson decays: a tracking efficiency of 70 % or more is required for tracks with a transverse momentum in the range 50-120 MeV/c where almost all the soft pion tracks have only SVT hits. Another important requirement to study semileptonic B^0 decays is the capability to reconstruct the B decay vertex with a resolution of less than 80 μm. This also means that the material traversed by particles must be minimized.

3.0.5 PEP-II constraints

Additional constraints are imposed by the storage rings. The geometrical layout of the SVT is constrained by the limited amount of space available. The SVT is located inside the long support tube, that extends all the way through the detector. To maximize the angular coverage, the SVT must extend down to 17.2° in polar angle, where the B1 permanent magnets sit. In the backward direction, it is sufficient to extend the SVT sensitive area down to 30°.

The SVT must withstand 2 MRad of ionizing radiation. A radiation monitoring system capable of aborting the beams is required. The expected radiation dose at the design luminosity is 240 kRad/yr in the horizontal plane immediately outside the beam pipe (where the highest radiation is concentrated), and 33 kRad/yr on average otherwise.

Reliability and robustness are essential: all components of the SVT inside the support tube must have long mean-time-to-failure, because the time needed for any replacement is estimated to be 4-5 months. Redundancies must be built in whenever possible.

The SVT must be cooled to remove the heat generated by the electronics. In addition, it must operate in a magnetic field.

The relative position of the individual silicon sensors must be stable over long time periods. The assembly must allow for relative motion of the support structures with respect to the B1 magnets. This feature is also necessary to sustain an earthquake of moderate intensity.

The choice of a Silicon Vertex Tracker made of five layers of double-sided silicon strip sensors with spatial resolution, for perpendicular tracks, of 10-15 μm in the three inner layers and about 40 μm in the two outer layers fulfills the physics requirements: the inner three layers perform the impact parameter measurements, while the outer ones are necessary for pattern recognition and low p_T tracking.

3.1 Silicon sensors

The SVT sensors [43] are made of 300 μm thick double-sided silicon strip devices. They were designed at INFN Pisa and Trieste (Italy) and fabricated commercially. A schematic representation of a silicon detector is shown in figure 3.1 They are built on high-resistivity (6-15 k$\Omega \cdot$ cm) n-type substrates with p^+ strips and n^+ strips

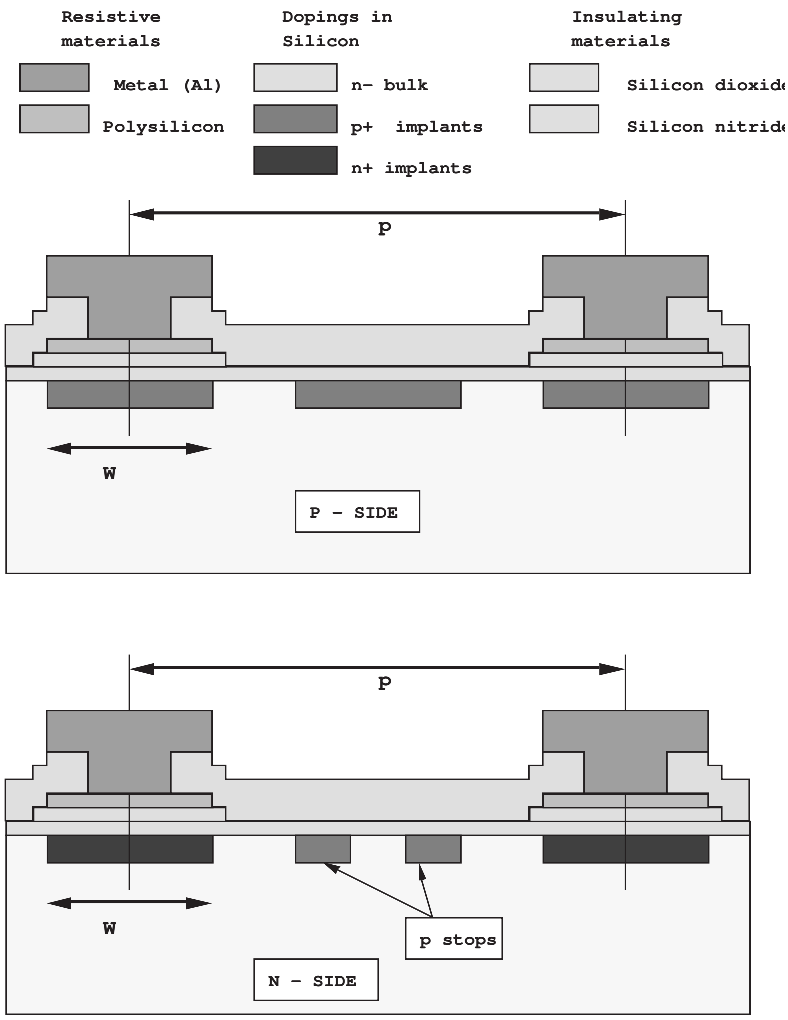


Figure 3.1: Schematic view of a silicon detector adopted for *BaBar* SVT.

on the two opposite sides. The insulation of the n^+ strips is provided by individual p-stops, in such a way as to achieve an inter-strip resistance greater than 100 MΩ at operating bias voltage, normally about 10 V above the depletion voltage. Typical depletion voltage is in the range 25-35 V. On both sides the strips are biased with polysilicon resistors (4-20 MΩ) to insure the required radiation hardness, keeping the voltage drop across resistors and the parallel noise as low as possible. Strips are AC-coupled to electronics via integrated decoupling capacitors, whose capacitance depends on the sensor shape, but is always greater than 14 pF/ cm. The sensors were designed to maximize the active area, which extends to within 0.7 mm of the physical edges. Another design goal was to control the inter-strip capacitance, values between 0.7 pF/ cm and 1.1 pF/ cm were obtained for the various sensor shapes. To achieve the required spatial resolution, while keeping the number of readout channels as low as possible, most of the modules have a floating strip between two readout strips.

Layer/ view	C_{input} (pF)	R_{series} (Ω)	Noise calculated (elec)	Noise measured (elec)
1 z	3.2	100	1	40
1 ϕ	3.2	50-100	0-1	82
2 z	4.0	100	1	48
2 ϕ	4.0	55-110	0-1	88
3 z	5.4	100	1	70
3 ϕ	5.4	110	1	128
4 z	9.1-12.7	210	1	104
4 ϕ	9.1-12.7	100	1	224
5 z	11.4-14.4	210	1	104
5 ϕ	11.4-14.4	100	1	265

Table 3.1: *Electrical parameters of the SVT, shown for the different layers and view. C_{input} refers to the total input capacitance, R_{series} is the series resistance. The amplifier peaking time is 200* ns *for layers 1-3 and 400* ns *for layers 4-5.*

The leakage currents, because of the excellent performance of the manufacturing process, were as low as 50 nA/ cm^2 on average, measured at 10 V above depletion voltage. The silicon sensor parameters have been measured after irradiation with ^{60}Co sources. Apart from an increase in the inter-strip capacitance of about 12 % during the first 100 kRad, the main effect was an increase of the leakage current of 0.7 μA/ cm^2 per MRad. However, in a radiation test performed in a 1 GeV electron beam, an increase in leakage current of about 2 μA/ cm^2 per MRad, and a significant shift in the depletion voltage, dependent on the initial dopant concentration, were observed. A shift of about 8-10 V was seen for irradiation corresponding to a dose of approximately 1 MRad. These observations indicate significant bulk damage caused by energetic electrons. Due to the change in depletion voltage, the SVT sensors could undergo type inversion after about 1-3 MRad. Preliminary tests show that the sensors continue to operate properly after type inversion [44]. Studies of the behavior of SVT modules as a function of radiation dose continue.

3.2 SVT Layout

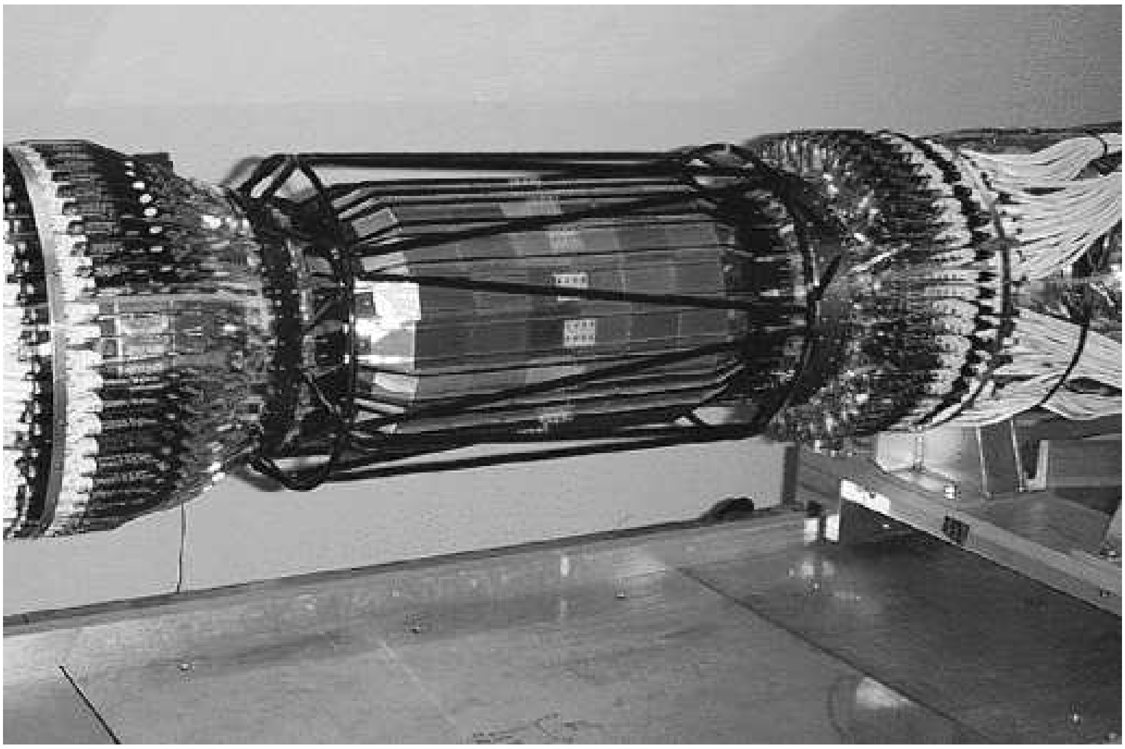


Figure 3.2: Fully assembled Silicon Vertex Tracker. The silicon sensors of the outer layer are visible, as is the carbon-fiber space frame (black structure) that surrounds the silicon.

The SVT detector is organized in five layers of double-sided silicon strip sensors: a photograph is shown in figure 3.2. The strips on the two sides are oriented orthogonally to each other: the ϕ strips run parallel to the beam and the z strips are oriented transversely to the beam axis.

The modules are made of several rectangular silicon sensors, of different sizes, that are assembled in a plane. The modules are divided electrically into two half-modules, which are read out at the ends. Two identical trapezoidal sensors must be added (one each at the forward and backward ends) to form the arch modules. To satisfy the different geometrical requirements of the five SVT layers, six different sensors shapes are required (five rectangular and one trapezoidal). The half-modules are given mechanical stiffness by means of the carbon fiber/Kevlar ribs, which are visible in figure 3.3; ϕ strips of the same half-module are electrically connected with wire bonds. This results in a total strip length up to 14 cm in the inner layers and up to 24 cm in the outer layers.

The modules of the inner 3 layers are straight, while the modules of layers 4 and

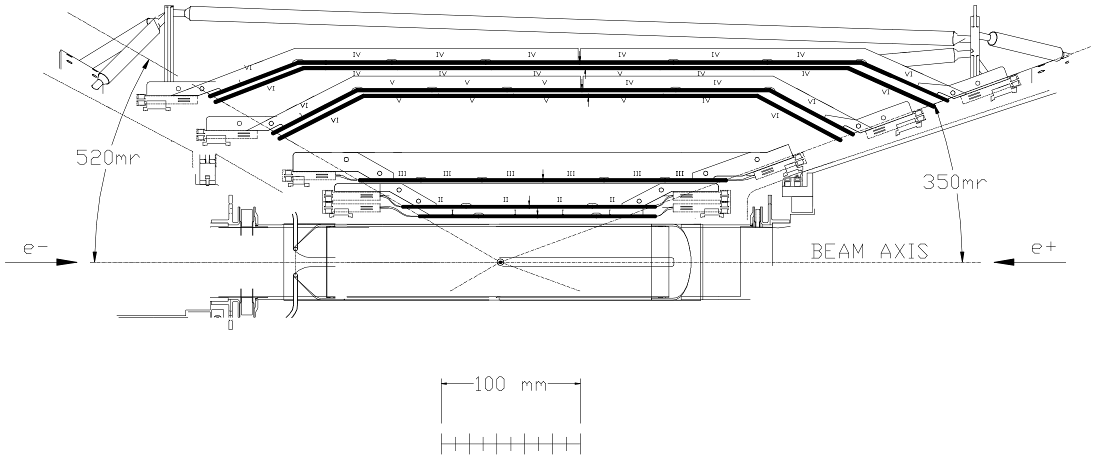


Figure 3.3: Schematic view of SVT: longitudinal section.

5 are arch-shaped (figure 3.3). This arch design was chosen to minimize the amount of silicon required to cover the solid angle, while increasing the crossing angle for particles near the edges of acceptance. A photograph of an outer layer arch module is shown in figure 3.5. The signals from the z strips are brought to the readout electronics using fanout circuits consisting of conducting traces on a thin (50 μm) insulating Upilex substrate. In layers 1 through 3 each z strip is connected to its own preamplifier channel, while in layers 4 and 5 the number of z strips on a half module exceeds the number of electronics channels available, requiring that pairs of z strips are electrically connected (ganged) to a single electronics channel. The length of a z strip is about 5 cm (no ganging) or 10 cm (two strips connected). The ganging introduces an ambiguity on the z coordinate measurement, which must be resolved in the pattern recognition procedure. The total number of readout channels is approximately 145,000.

The inner modules are tilted in ϕ by 5°, allowing an overlap region between adjacent modules, a feature that is very useful for the alignment. The outer modules cannot be tilted, because of the arch geometry. To avoid gaps and to have a suitable overlap in ϕ, layers 4 and 5 are divided into two sub-layers (4a, 4b, 5a, 5b) placed slightly different radii. The relevant geometrical parameters of each layer are summarized in table 3.2.

In order to minimize the material in the detector acceptance region, the readout electronics are mounted entirely outside the active detector volume. The forward electronics must be mounted in the 1 cm space between the 350 milliradians stay-

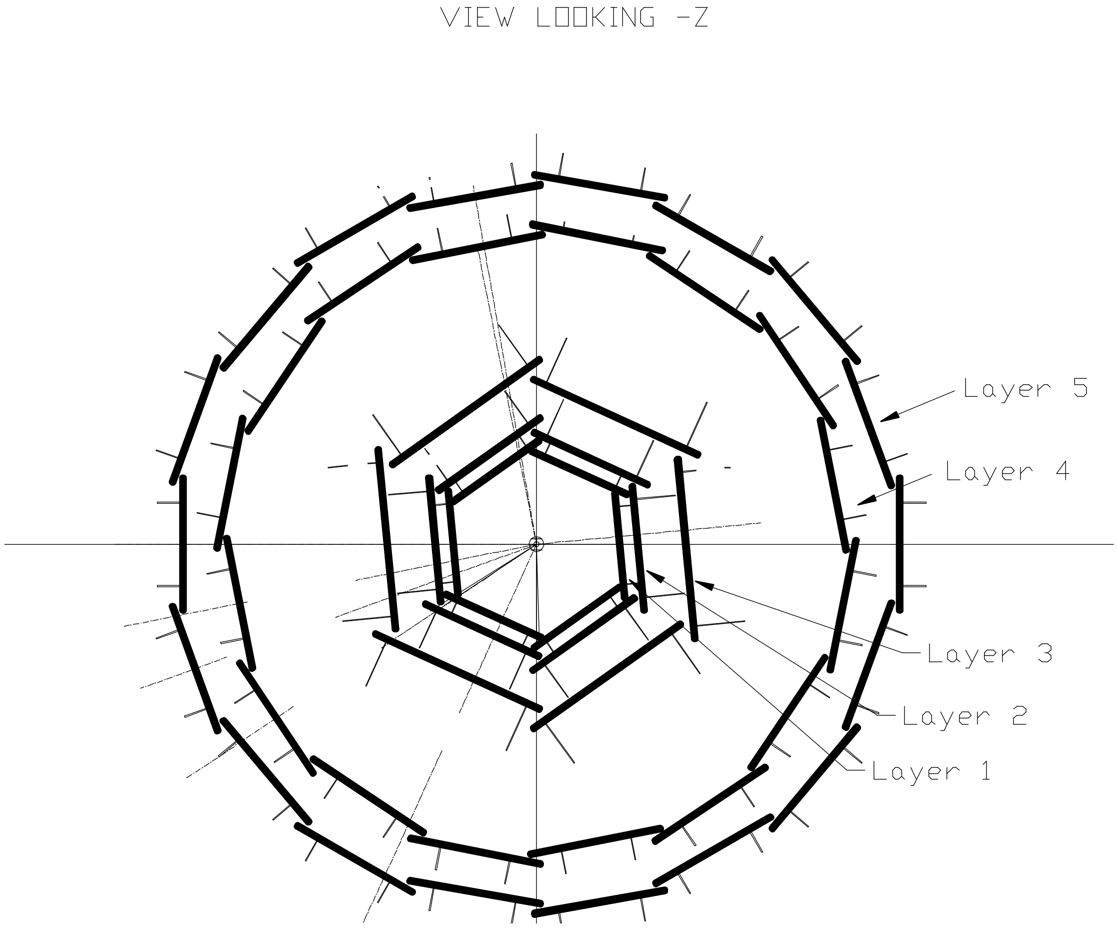


Figure 3.4: Schematic view of SVT: transverse section.

clear space and B1 magnet. This implies that the hybrids carrying the front-end chip must be positioned at an angle of 350 mrad relative to the sensor for the inner layers, and at an even larger angle for the outer ones (figure 3.4). In the backward direction the available space is larger and the inner layer electronics can be placed in the sensor plane, allowing a simplified assembly procedure.

The SVT support structure (figure 3.2) is a rigid body made from the carbon-fiber cones, connected by a “space-frame”, also made of carbon-fiber epoxy laminate. This structure is attached to the B1 magnets by a set of gimbal rings, in such a way as to follow for relative motion of the two B1 magnets while fixing the position of the SVT relative to the forward B1 and the orientation relative to the axis of both B1s. Note that the structure composed by SVT, B1 magnet and support tube is supported independently of the rest of the *BABAR* apparatus, leading to the possible movements between the SVT and the rest of *BABAR*. Precise position monitoring of

Layer/ view	Radius (cm)	Readout pitch (μm)	Floating strips	Strip length (mm)
1 z	3.2	100	1	40
1 ϕ	3.2	50-100	0-1	82
2 z	4.0	100	1	48
2 ϕ	4.0	55-110	0-1	88
3 z	5.4	100	1	70
3 ϕ	5.4	110	1	128
4 z	9.1-12.7	210	1	104
4 ϕ	9.1-12.7	100	1	224
5 z	11.4-14.4	210	1	104
5 ϕ	11.4-14.4	100	1	265

Table 3.2: *Geometrical parameters of the SVT, for each layer and view. Floating strips refers to the number of strips between readout strips. Note that part of the ϕ sides of layers 1 and 2 are bonded at 100 μm and 110 μm pitch, respectively, with one floating strip. Strip length for layers 4 and 5 z includes ganging. The radial range for layers 4 and 5 considers the radial extent of the arched sections.*

the beam interaction point is necessary and has been described in section 2.3.3.

The total active silicon area is 0.96 m^2 and the material traversed by particles is minimal. The geometrical acceptance of SVT is 92 % of the solid angle in the center-of-mass system.

The module assembly and the mechanics are quite complicated, especially for the arch modules, and are described in detail in [43].

In order to perfrom precision measurements, the resolution required for vertexing has to be of the order of few tens of microns. To achieve this goal, precision of few microns are needed for each step of the module assembly. For example, when assembling the barrel detectors, each silicon sensors has to be head-to-head glued with micrometric precision and each readout channel of a silicon sensor has to be wirebonded with the corresponding channel of the head-to-head glued silicon sensor. Micrometric wirebonding is done by a soldering robot controlled by an operator. I had an active role in the Detector-Fanout-Assembly (DFA) which I briefly report below.

Figure 3.5: Photograph of an arch module.

3.2.1 Detector-Fanout-Assembly

The Detector-Fanout-Assembly (DFA) consists of the silicon wafers and the ϕ and z fanouts for a half-module. The DFA assembly procedures are:

- ϕ fanout gluing to the first silicon wafer
- head-to-head gluing of the barrel detectors
- z fanout gluing

In the glue deposition task the accidental squirting of the glue on the bonding pads must be carefully prevented. For the fanouts a gluing robot is used, with a glue dispenser and a calibrated tip under controlled pressure. The layer of glue is $50\,\mu$m thick. The ϕ fanout is put on the chuck and aligned with respect to the first wafer using a system of micrometric screws, incorporated in the fanout gluing jig. After the deposition of the glue, the vacuum on the fanout chuck is removed and the fanout drops under the force of gravity onto the detector plane. Head-to-head gluing is performed only for the wafers of the barrel part of the half-module. A groove of depth $50\,\mu$m is inscribed into a teflon surface and the glue is applied to the wafers by dipping them into this glue loading "pot". Then they are transferred on vacuum towers controlled by micrometric screws and aligned with a gap of $75\,\mu$m between them. The alignment of the barrel section with the wedge detectors is performed using the precision machined towers as in the head-to-head gluing. The aligned objects are then transferred together to the z fanout gluing jig. The z fanout chuck allows the deposition of the glue line and the dropping of the fanout on the aligned detectors. The DFA is completed and is stored in the storage box ready to be wirebonded with micrometric soldering. The fanouts are equipped with a testing extension that allows the final testing of the DFA to detect new defects. The most important defects are:

- pinhole: a break in the AC coupling capacitor that causes a short between the aluminum strip and the strip implant. Pinholes are in general created during the detector fabrication, but can also be created in the assembly phase if too much power is used in the wirebonding or if insufficient electrostatic protection is employed. The fraction of channels with pinholes is at the level of 1-2%. These channels are not connected to the electronics;

- p-stop-short: a short between the aluminum strip and the p-stop on the ohmic side. This problem affects 0.3% of the inner three layers but is almost absent in the outer layers due to the larger strip pitch;

- shorts: in addition to manufacturing defects, any mechanical damage due to improper handling may cause scratching of the detectors and fanout;

- un-bondable channels: a pad on the detector or on the fanout that has been damaged or obstructed such that wirebonding is no longer possible.

On a DFA the sum of the manufacturing defects and defects that occur during assembly is 2-4%.

3.3 SVT Components

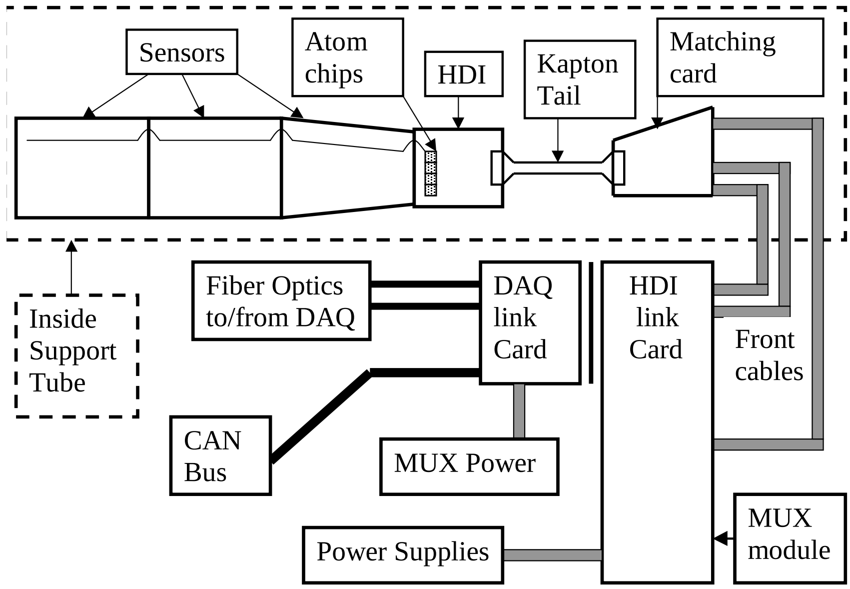


Figure 3.6: Schematic block diagram showing the different components of the SVT.

A block diagram of SVT components is shown in figure 3.6. The basic components of the detector are the silicon sensors, the fanout circuits, the front-end electronics and the data transmission system.

3.3.1 Fanout circuits

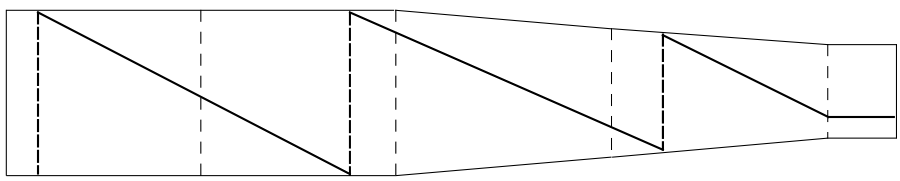


Figure 3.7: Schematic drawing of fanout connection to z strips showing ganging.

The fanout circuits, which route the signal from the strips to the electronics, have been designed to minimize the series resistance and the inter-strip capacitance. As described in [45], a trace on the fanout has a series resistance of about 1.6 $\Omega/$ cm, an inter-strip resistance greater than 20 MΩ, and an inter-strip capacitance lower than 0.5 pF/ cm. The electrical parameters of the final assembly of sensors and fanouts (known as DFAs which stands for Detector Fanout Assemblies) are summarized in table 3.1. Due to the different strip lengths, there are large differences between the inner and the outer layers. Smaller differences are also present between the forward and backward halves of the module, that are of different lengths.

While the ϕ fanout circuits are just one-to-one connections a few centimeters long, the z fanouts are more complicated. They must extend the full lenght of the detector modules in order to connect all the z strips and to provide ganging in cases where the number of available readout channels exceeds the number of z strips, such as in outer layers. Figure 3.7 depicts the concept of ganging. Two or three z strips (dashed lines) are connected in series by the fanout traces (solid lines), bringing the signals to the end of the detector module where they can be wire-bonded to the front-end electronics IC.

3.3.2 Front-end electronics and Data transmission

The electrical parameters of a DFA and the general *BABAR* requirements are the basic inputs that drove the design of the SVT front-end chip: the ATOM[1]. In particular, the front-end chip had to cope with the following requirements:

- signal to noise ratio bigger than 15 for minimum ionizing particles (MIP) signals for all modules;
- signals from all hit strips must be retained, in order to improve the spatial resolution through interpolation, while keeping the number of transmitted hits as low as possible. Hit means a deposited charge bigger than 0.95 fC, corresponding to 0.25 times the average charge deposited by a MIP;
- the amplifier must be sensitive both to negative and positive charge;
- the peaking time must be programmable, with a minimum of 100 ns (in layers 1 and 2, because of the high occupancy), up to 400 ns (outer layers, with high capacitance);
- capability to accept random triggers with a latency up to 11.5 μs and a jitter up to $\pm 1\,\mu$s, without dead-time;
- radiation hardness;

[1] A Time-Over-threshold Machine.

- small dimensions: 128 channels in a 6.2 mm wide chip.

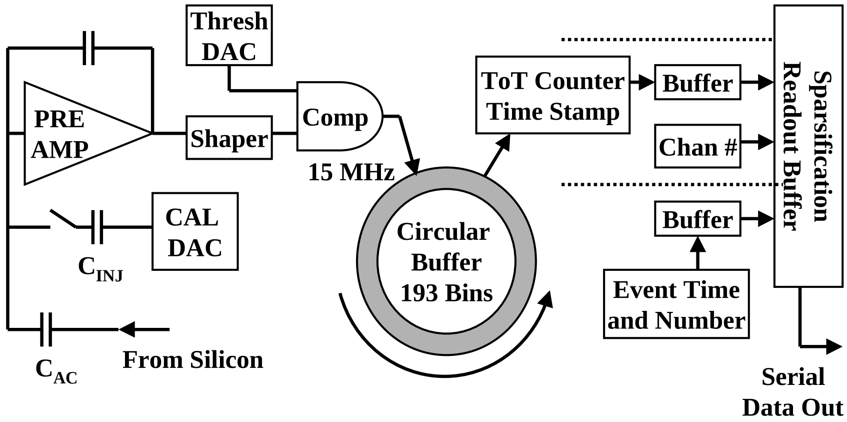


Figure 3.8: Schematic diagram of the AToM front end IC.

These requirements are fully satisfied by the ATOM chip [46], which is described schematically in figure 3.8. The linear analog section consists of a charge-sensitive preamplifier followed by a shaper. Signals are presented to the Time-Over-Threshold (ToT) comparator, whose threshold is generated and controlled by a digital-to-analog converter(DAC). The signals above threshold induce at the comparator output a logic signal whose width, the ToT, is related to the collected charge by a quasi-logarithmic relationship. This design choice effectively retains the analog charge information with only 4 bits. This is not, however, the limiting factor to the measurement of the spatial resolution. The ToT is digitized by comparison with a reference clock (30 MHz) and the resulting stream of pulses is stored into a 193 location buffer, to compensate for the trigger latency and jitter. Upon receipt of a first level trigger, the ToT is retrieved from the latency buffer, associated to a 5 bit time stamp and stored into one of four buffers.

The output data (4 bits of the ToT value, 5 bits for the time stamp and 7 bits for the strip address) are formatted, serialized and delivered upon the receipt of a readout command.

The typical noise behavior of the ATOM is described by the equivalent noise charge (ENC) of the linear analog section and is given in table 3.3.

Radiation hardness was measured up to 2.4 MRad of ^{60}Co irradiation. At that dose, a decrease of 20% in the charge gain has been observed, with a noise increase

ATOM peaking time	Noise Charge at zero Capacitance	Noise slope
100 ns	380 e^-	40.9 e^-/pF
200 ns	280 e^-	33.9 e^-/pF
400 ns	220 e^-	25.4 e^-/pF

Table 3.3: *ATOM chip Equivalent Noise Charge (ENC) parameters at different peaking times.*

of less than 15% [29].

With the parameters of tables 3.3 and 3.1, it is possible to perform a calculation of the expected noise in the various SVT modules, taking into account that the noise due to the sensors leakage current fluctuations is negligible. The results of such a calculation are shown in table 3.1, together with the measured average values for the selected modules. The agreement is satisfactory and the maximum average noise is 1600 electrons, leading to a signal to noise ratio bigger than 15.

The ATOM chips are housed on thick-film double-sided hybrid circuits, known as High Density Interconnects (HDIs), which are fabricated on 1.2 mm thick aluminum-nitride substrate [47]. The chips and the other components are mounted on both sides of the hybrid substrate to read out the ϕ and z sides of the half-modules. The electronics on each side of the HDI is biased through a floating power supply system, in such a way to guarantee a small voltage drop (less than 1 V) across the detector coupling capacitors.

The digitized signals are transmitted from the ATOM chips through a thin kapton cable[2] to the matching cards, from where they are routed to conventional cables.

Just outside the *BABAR* detector, signals are multiplexed by the MUX modules, converted onto optical signals and transmitted to the ROMs. The MUX also receive digital signals from the DAQ via fiber optical connection. The SVT is interfaced to the *BABAR* ODC via the industry standard CAN-bus.

3.4 Data analysis and performances

This section describes the reconstruction of raw data into physics-related quantities such as trajectories of charged particles.

[2]known as tail.

3.4.1 Defects

Due to a series of minor mishaps incurred during the installation of the SVT, few SVT modules were damaged and are currently not functioning. There is no single primary failure mode, but rather a mixture of defective connectors, mis-handling during installation and not-fully-understood problems on the front-end electronics hybrid. There has been no module failure due to radiation damage. A total of 9 readout sections (which corresponds to one side of a half-module) out of 208 are not functioning[3].

It should be noted that due to the redundancy afforded by the five layers of the SVT, the presence of the defective modules has minimal impact on physics. In addition, it is estimated that there are individual channel defects, of various types, at a level of about 1 %. Calibrations reveal an increase in the number of defective channels at a rate of less than 0.1 %/year.

3.4.2 Cluster and hit reconstruction

The average occupancy, in a time window of 1 μs, of the SVT under normal running conditions is about 3 % for the inner layers, with a significant azimuthal variation due to beam-induced backgrounds, and less than 1 % for the outer layers, where noise hits dominate. Figure 3.9 shows the typical occupancy as a function of chip number (equivalent to azimuthal angle, in this case) for layer 1, ϕ side. In the inner layers, the occupancy is dominated by machine backgrounds, which are significantly higher in the horizontal plane, seen in the plot as the peaks near chip numbers 3 and 25, approximately.

The first step of the reconstruction program consists in discarding out-of-time channels. A time correction, the time between the particle crossing and the time the shaper exceeds threshold, is performed, after which hits with times more than 200 ns from the event time (determined by the drift chamber) are discarded. The loss of real hits from this procedure is negligible. The resulting in-time hits are then passed to the cluster finding algorithm. First, the charge pulse height (Q) of a single pulse in calculated from the ToT value. In a first pass, clusters are formed grouping adjacent strips with consistent times. A second pass is then performed, where clusters separated by just one strip are merged into one cluster. The two

[3]at the time of the data collected for this analysis. After the 2002 summer shutdown there are only 5 readout sections malfunctioning.

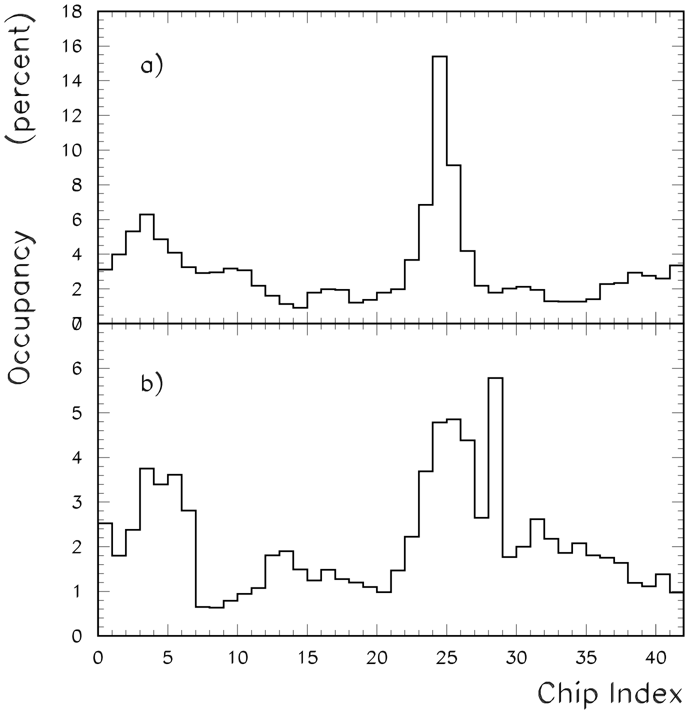


Figure 3.9: Typical occupancy in percent as a function of chip number in layer 1, ϕ side for a) forward half-modules and b) backward half-modules. Chip number corresponds to the azimuthal angle and the increased occupancy in the horizontal plane is visible near chips 3 and 25.

original clusters plus the merged cluster are made available to the pattern recognition algorithm, which chooses among the three.

The position x of a cluster formed by n strips is determined, using the "head-to-tail" algorithm:

$$x = \frac{(x_1 + x_n)}{2} + \frac{p}{2}\frac{(Q_n - Q_1)}{(Q_n + Q_1)}$$

where x_i and Q_i are the position and collected charge of i-th strip, respectively, and p is the readout pitch. This formula results in a cluster position that is always within $p/2$ of the geometrical center of the cluster. The cluster pulse height is simply the sum of the strip charges, while the cluster time is the average of the strip times. Clusters are then converted to hits, before being passed to the tracking algorithms. Hits differ from clusters in that the hit position is expressed in the global coordinates, taking into account the physical positions of the sensors.

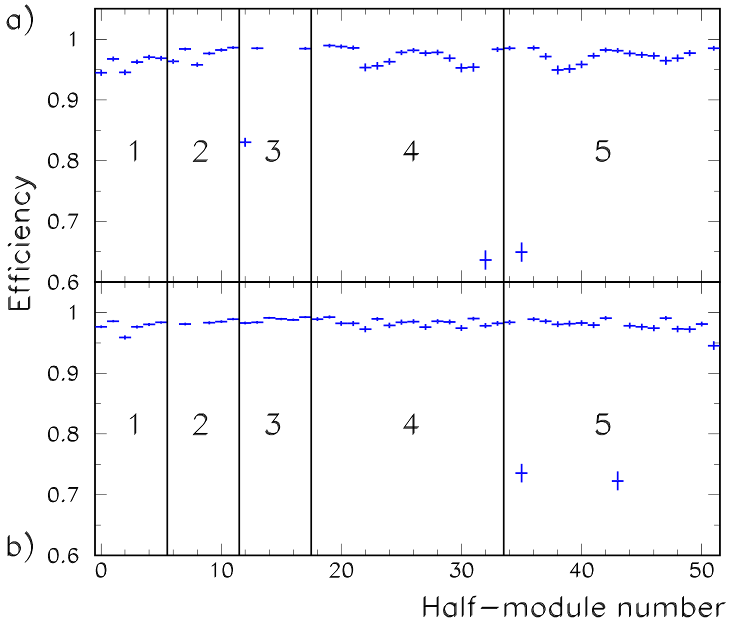


Figure 3.10: SVT hit reconstruction efficiency, as measured on $\mu^+\mu^-$ events for a) forward half-modules and b) backward half-modules. The plots show the probability of associating both a ϕ and a z hit to a track passing through the active part of the detector. The horizontal axis corresponds to the different modules, with the vertical lines separating the different layers as numbered. Missing values correspond to non-functioning half-modules.

3.4.3 Performance

The SVT hit efficiency can be calculated for each half-module by comparing the number of associated hits to the number of tracks crossing the module. As can be seen in figure 3.10, we typically measure combined hardware and software efficiencies of 98 %, excluding defective readout sections (9 out of 208[4]), but employing no special treatment for other defects, such as broken AC coupling capacitors or dead channels on front-end chips. Actually, since most of the defects affect a single channel, they do not contribute heavily to inefficiency, because most charge depositions involve two or more strips, due to track crossing angles and, to some degree, charge diffusion.

The spatial resolution of SVT hits is determined by considering the distance along the sensor plane between the track trajectory and the hit, using high-momentum

[4]after 2002 summer shutdown this number has been reduced from 9 to 5.

tracks in two prong events. The uncertainty due to the track trajectory is subtracted from the width of the residual distribution to obtain the hit resolution. Figure 3.11 shows the SVT hit resolution for z side and ϕ side hits as a function of track incident angle, for each of the five layers. The multiple scattering in the beam pipe and in the silicon itself sets a lower limit on intrinsic resolution, corresponding to a hit resolution of about 10-15 μm for mesurements made close to the IP (layer 1, 2 and 3) and 30-40 μm for the outer layers (4 and 5) [31]. We see from figure 3.11 that SVT meets Technical Design Report [31] specifications for perpendicular tracks.

Studies have shown that hit reconstruction efficiency and spatial resolution are effectively independent of occupancy for the occupancy levels observed so far [29].

Even though tracking resolution studies were discussed in detail in the previous chapter, it's important to stress that SVT dominates in the determination of the track angles.

Figure 3.12 shows the resolutions for the track impact parameter along the beam (σ_z) and in the plane perpendicular to the beam (σ_{xy}) for various momenta as a function of the dip angle λ measured relative to normal incidence. The resolution is dominated by multiple scattering and not by the intrinsic hit resolution of the silicon detectors, except for the highest momentum tracks from B decays. In this analysis the impact parameter resolution of the soft pion coming from from a D^* decay is well represented by the solid line in figure 3.12.

SVT standalone tracking efficiency

A key issue for a reliable measurement of the $\mathcal{B}(B^0 \rightarrow D^{*-}\ell^+\nu_\ell)$ is an efficient detection of the soft pion at low transverse momenta, where it has only SVT hits since it is not energetic enough to reach the DCH. This requires a detailed study of data and Monte Carlo agreement and systematics effects evaluation of soft pion tracking in the SVT.

The standalone SVT tracking algorithms have a high efficiency for tracks with low transverse momentum. This feature is crucial for the detection of D^* decays. To study this efficiency, decays $D^{*+} \rightarrow D^0\pi^+$ are selected by reconstructing events of the type $\overline{B} \rightarrow D^{*+}X$ followed by $D^{*+} \rightarrow D^0\pi^+ \rightarrow K^-\pi^+\pi^+$. As we already said, for the majority of these low momentum pions the momentum resolution is limited by multiple scattering and the production angle can be determined from the signals in the innermost layers of the SVT.

We will show later how the $\delta m = M(K^-\pi^+\pi^+) - M(K^-\pi^+)$ resolution is affected by the information of the soft pion tracks reconstructed only in the SVT or both in the SVT and the DCH. Clearly the δm resolution is much better in the latter case but the statistics is much smaller with respect to the case of "SVT tracks only".

The SVT standalone tracking allow to gain in efficiency mostly at very low momenta. The resolution is impacted by multiple scattering and by limited track length of the slow pions. To derive an estimate of the tracking efficiency for these low momentum tracks, a detailed Monte Carlo simulation was performed. Specifically, the pion spectrum was derived from simulation of the inclusive D^* production in $B\overline{B}$ events, and the Monte Carlo events were selected in the same way as the data. A comparison of the detected slow pion spectrum with the Monte Carlo prediction is presented in figure 3.13. Based on this very good agreement, the detection efficiency has been derived from the Monte Carlo simulation.

The SVT significantly extends the capability of the charged particle detection down to transverse momenta of about 50 MeV/c.

We will show in the next chapters how to evaluate the systematic uncertainty due to evaluation with Monte Carlo of the detection efficiency for soft pions when calculating the $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$.

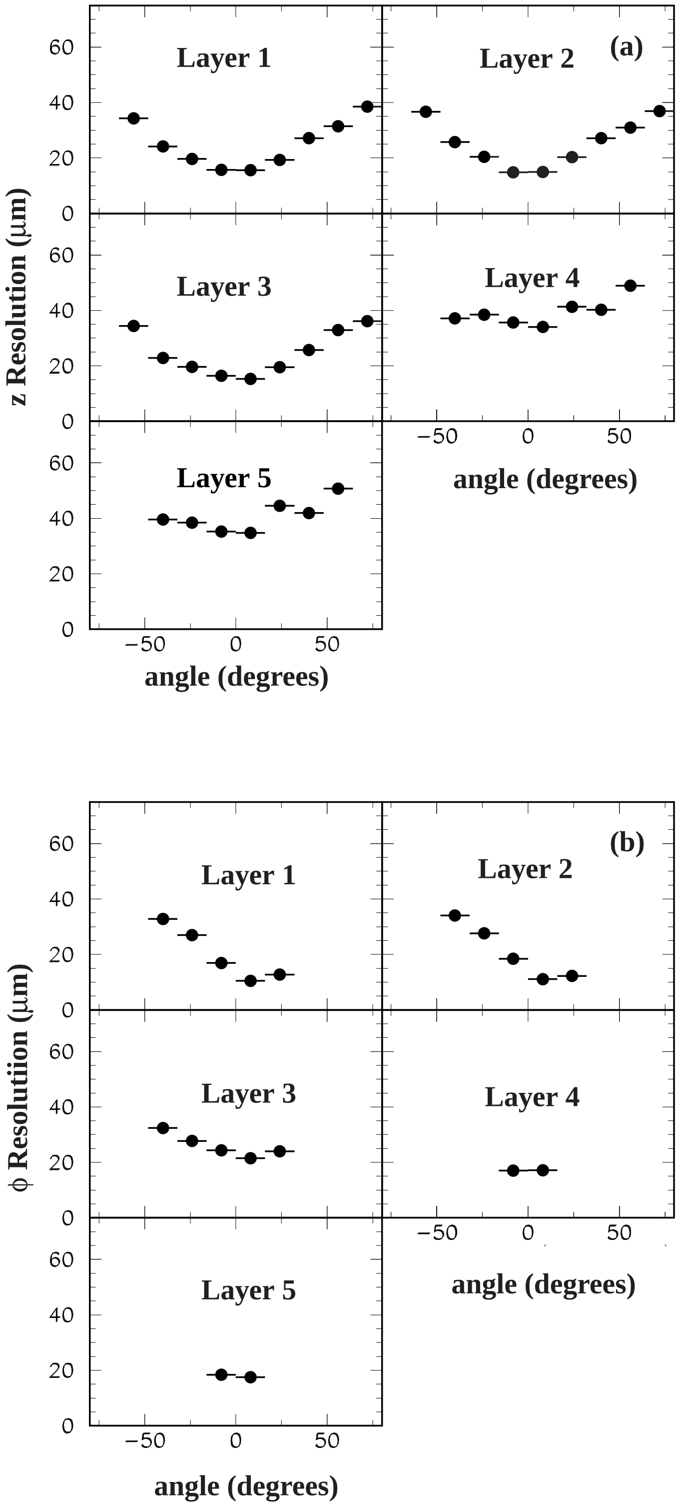


Figure 3.11: SVT hit resolution in the a) z and b) ϕ coordinate in microns, plotted as a function of track incident angle in degrees. Each plot shows a different layer of the SVT. The plots in the ϕ coordinate for layers 1-3 are asymmetric around $\phi = 0$ because of the pinwheel design of the inner layers. There are fewer points in the ϕ resolution plots for the outer layers as they subtend smaller angles than the inner layers.

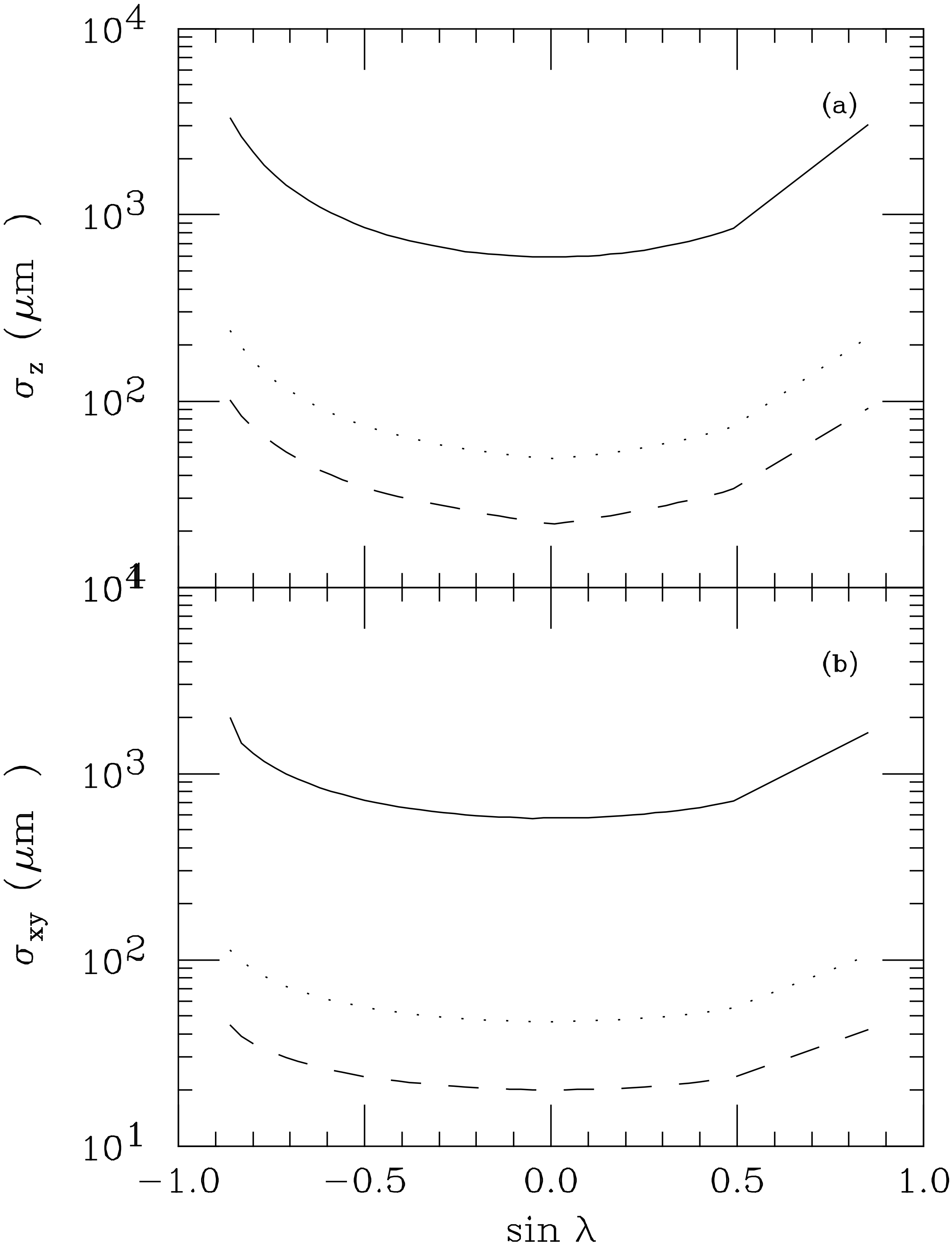


Figure 3.12: (a) Impact parameter resolution along the beam (σ_z) and (b) in the plane perpendicular to the beam (σ_{xy}) versus dip angle λ for three different momenta: 100 MeV/c (solid line), 1.0 GeV/c (dotted line) and 3.0 GeV/c (dashed line).

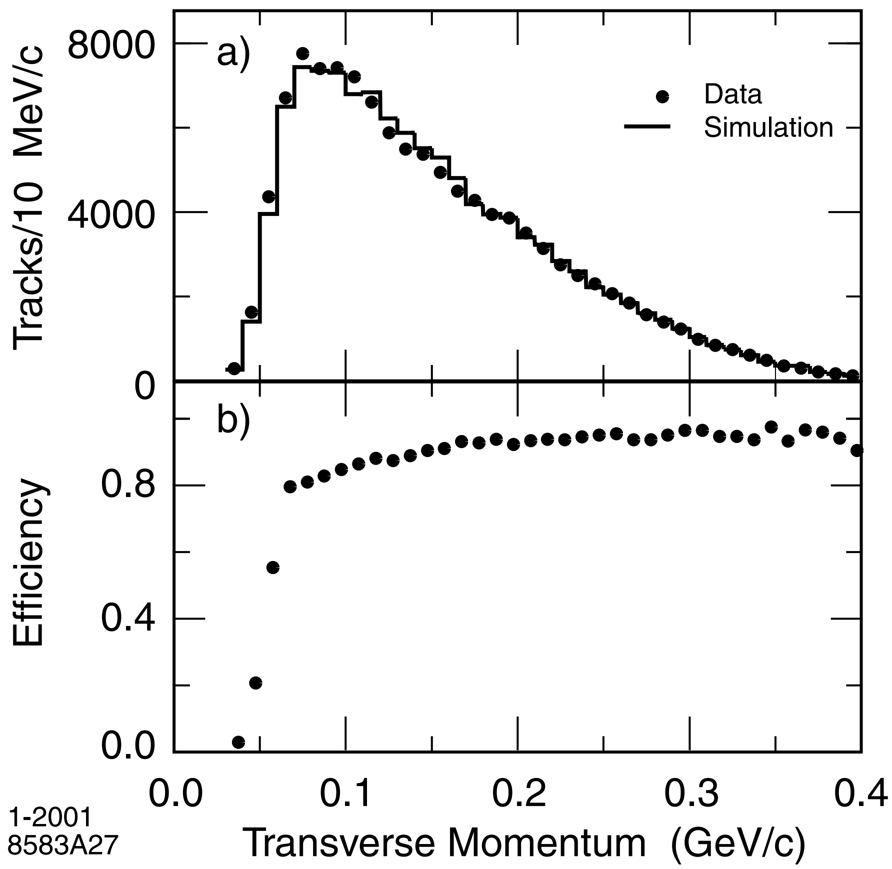


Figure 3.13: Monte Carlo studies of low momentum tracks in the SVT: a) comparison of data (contributions from combinatoric background and non $B\overline{B}$ events have been subtracted) with simulation of the transverse momentum spectrum of pions from $D^{*+} \rightarrow D^0\pi^+$ in $B\overline{B}$ events, and b) efficiency for slow pion detection derived from simulated events.

Chapter 4

Event selection and reconstruction: signal, backgrounds and control samples

In this chapter we describe the criteria adopted in the event selection for the study of the semileptonic decay $B^0 \to D^{*-}\ell^+\nu_\ell$ [1]. The signal is: $\Upsilon(4S) \to B^0\overline{B}^0$ where one of the neutral B mesons decays semileptonically to a charged D^* meson. We perform full reconstruction of the B^0 meson decay products, which entails observing the lepton and all the decay products of the charged D^*, but not the weakly interacting neutrino. We reconstruct the charged D^* meson only when it decays to a D^0 and a charged pion. The D^0 meson is then reconstructed in 4 different decay channels, i.e. $D^0 \to K\pi$, $D^0 \to K\pi\pi\pi$, $D^0 \to K\pi\pi^0$ and $D^0 \to K_S\pi\pi$. The backgrounds that are present in this analysis are described in section 4.1.

We show the distribution of the relevant kinematic variables of the decay process. Then we explain the criteria adopted for charged tracks reconstruction, particle identification, vertexing and for the D^0 and D^* reconstruction. We will need to know the efficiencies of all these procedures. A section is dedicated to the description of the background discriminating variables adopted in the analysis. The problem of background subtraction is complicated by the circumstance that several backgrounds have kinematic (for example the value of the D^*-D^0 mass difference δm) similar to the signal: they are called "peaking δm backgrounds" or, shortly, peaking backgrounds, as can be seen in figures 5.10 to 5.12.

[1] we recall, again, that throughout this thesis we refer to the decay $B^0 \to D^{*-}\ell^+\nu_\ell$ as a shorthand for both this decay and its charge-conjugate mode $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}$.

One of the starting points of this measurement is that we want to keep Monte Carlo dependence as small as possible; hence we study most of the backgrounds directly from data control samples as defined in section 4.3.

4.1 Description of signal and backgrounds

We define the following classification of the sources of signal and background that we expect to contribute to this sample:

1. events with a correctly reconstructed D^* candidate,
   (a) events originate from $B\overline{B}$ events,
      i. events with a correctly identified lepton candidate,
         A. **SIGNAL**: $B^0 \to D^* \, \ell \, \nu_\ell \, (n\gamma)$ decays,
         B. or, **CORRELATED BACKGROUND**: B^0 or $B^\pm \to D^*$ X, X$\to \ell$ Y,
         C. or, **UNCORRELATED BACKGROUND**: B^0 or $B^+ \to D^*$ X, and $\overline{B}^0$ or $B^- \to \ell$ X or (via mixing) $B^0 \to \ell$ X,
         D. or, **CHARGED B BACKGROUND**: $B^\pm \to D^* \, \ell \, \nu_\ell$ X, (via $B^\pm \to D^{**} \, \ell \, \nu_\ell$ or non resonant $B^\pm \to D^* \, \ell \, \nu_\ell \, \pi^\pm$ production)
         E. or, **NEUTRAL B BACKGROUND**: $B^0 \to D^* \, \ell \, \nu_\ell$ X, (via $B^0 \to D^{**} \, \ell \, \nu_\ell$ or non resonant $B^0 \to D^* \, \ell \, \nu_\ell \, \pi^0$ production)
      ii. or, **FAKE-LEPTON BACKGROUND**: events with a misidentified lepton candidate.
   (b) or, **CONTINUUM BACKGROUND**: $c\bar{c} \to D^*\ell$ X.
2. or, **COMBINATORIC BACKGROUND**: events with a misreconstructed D^* candidate.

4.2 Basic kinematic quantities for the semileptonic B^0 decay

Before introducing the choice of signal observables, we first review some quantities defined in the B^0 rest frame (denoted by the hat accent) of the semileptonic decay

$B^0 \to X^-\ell^+\nu_\ell$, where X represents the hadronic state (in our signal X is the D^* meson):

- $\hat{E}_\ell$, the lepton energy,
- $\hat{q}^2$, the W^+ invariant mass, calculated as $m_{B^0}^2 + m_X^2 - 2m_{B^0}\hat{E}_X$,
- $\hat{p}_X$, the magnitude of the hadronic momentum,
- $\cos\hat{\theta}_{X,\ell}$, the cosine of the angle between the directions of the hadronic and lepton momenta.

The variables $\hat{E}_\ell$ and $\hat{q}^2$ parametrize the phase space of the decay

$$\mathrm{d}\Phi\left(B^0 \to X^-\ell^+\nu_\ell\right) \propto \mathrm{d}\hat{E}_\ell\,\mathrm{d}\hat{q}^2 \ ,$$

and so are useful for disentangling the effects of phase space and dynamics. Figure 4.1 compares the distribution of these variables ("Dalitz plot") for pure three-body phase space and for $B^0 \to D^{*-}\ell^+\nu_\ell$ decays described by HQET form factors, respectively.

The variables $\hat{p}_X$ and $\cos\hat{\theta}_{X,\ell}$ are completely correlated with the Dalitz variables

$$\begin{aligned}
\hat{p}_X &= \frac{1}{2m_B}\sqrt{\left(m_B^2 - (\hat{q}+m_X)^2\right)\left(m_B^2 - (\hat{q}-m_X)^2\right)} \\
\cos\hat{\theta}_{X,\ell} &= \frac{\hat{q}^2 + m_\ell^2 - \hat{E}_\ell(m_B^2 + m_X^2 - \hat{q}^2)/m_B}{2\hat{p}_X\hat{p}_\ell}\ ,
\end{aligned}$$

and so cuts applied to more than two of these variables are not independent: $\hat{q}^2$, $\hat{E}_\ell$, $\hat{p}_X$, and $\cos\hat{\theta}_{X,\ell}$. Figure 4.2 shows the distribution of these variables, and Figure 4.3 shows their correlations with the lepton energy. Figure 4.2(a) shows that the hadronic momentum has a sharp edge at the upper kinematic limit, although the spectrum is significantly softened by the form factors and angular-momentum constraints. Figure 4.3(a) shows that the lepton energy and hadronic momenta are anti-correlated: a minimum cut on the lepton energy reduces the contribution from events with large hadronic momentum, and vice versa. Figure 4.3(c) demonstrates this effect on the hadronic momentum spectrum.

Figure 4.2(b) shows that the D^* and lepton tend to be back-to-back in the B decay frame, and that this effect is mostly due to the phase space constraints. Figure 4.3(b) shows that events with a soft lepton are essentially isotropic, but that events with a hard lepton are mostly back-to-back. As a result, the $\cos\hat{\theta}_{X,\ell}$

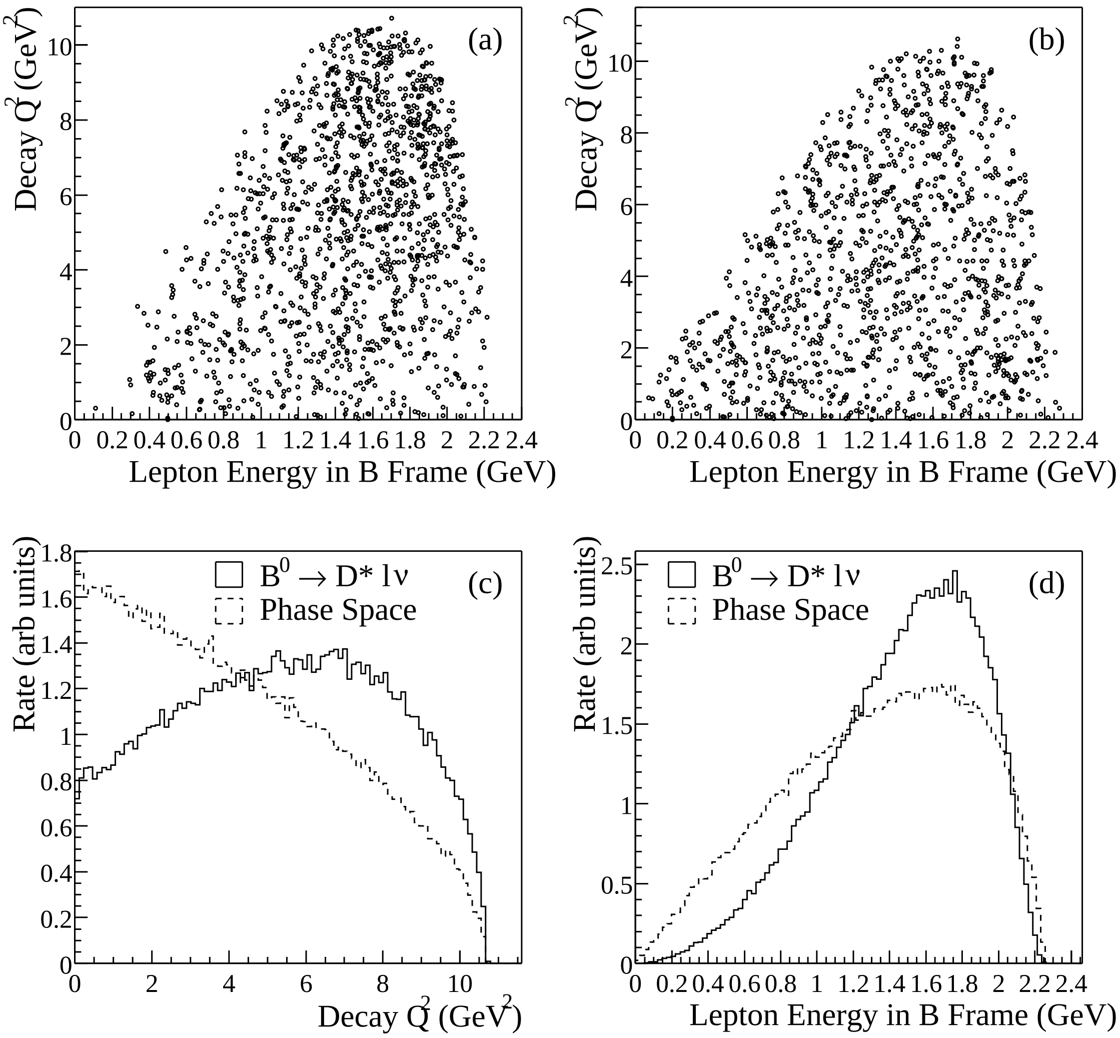


Figure 4.1: Dalitz distribution of $B^0 \to D^{*-}\ell^+\nu_\ell$ using (a) the form factors and angular-momentum correlations, or (b) three-body phase space. The lower plots show the normalized projections of these variables: (c) $\hat{q}^2$, and (d) $\hat{E}_\ell$.

distribution peaks more strongly at minus one as a tighter cut is applied to the minimum lepton energy, as shown in figure 4.3(d).

Although observables defined in the B^0 decay frame are the most natural for a theoretical description of semileptonic decays, we do not generally know the B^0 momentum in the lab frame as a vector, since we can infer only its modulus from $E^2 = m^2c^2 + |\vec{p}|^2c^4$. Instead, we use decay frame of the $\Upsilon(4S)$, which is determined by the beam energies

$$p(\Upsilon(4S)) = p(e^+) + p(e^-)\ .$$

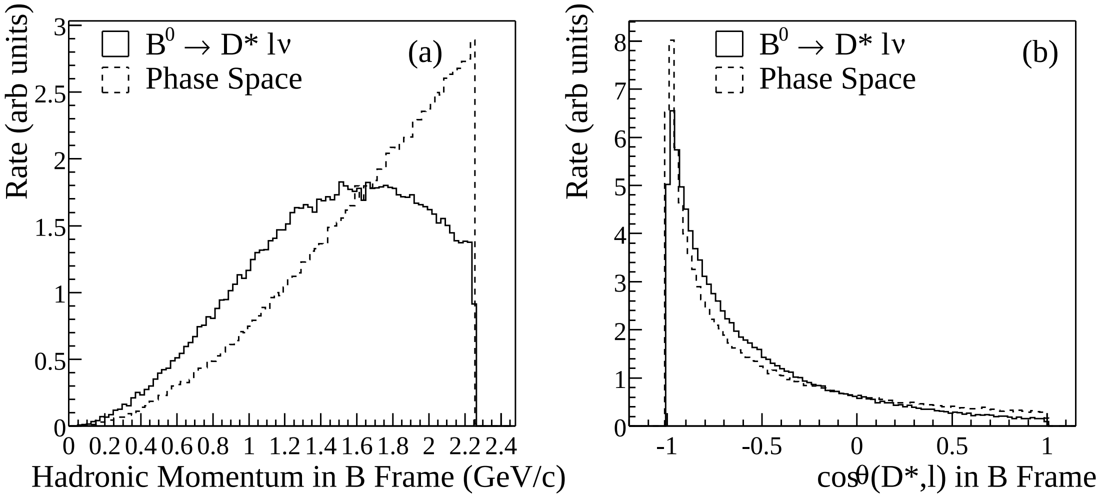


Figure 4.2: Normalized distributions of (a) the hadronic momentum, $\hat{p}_X$, and (b) the angle between the hadronic and lepton momenta, $\cos\hat{\theta}_{X,\ell}$, measured in the B^0 decay frame. The plots compare the distributions calculated using either three-body phase space, or else including the full decay dynamics.

The difference between the $\Upsilon(4S)$ and B^0 decay frames is due to the small B^0 decay momentum

$$p^*_{B^0} \sim 300\,\mathrm{MeV}/c\ ,$$

and so to a good approximation, the kinematic considerations made above are equivalent for these frames.

In order to distinguish the $\Upsilon(4S)$ rest frame from laboratory frame we use a star to indicate variables calculated in the $\Upsilon(4S)$ frame such E^*_ℓ, p^*_ℓ, p^*_X.

Some of the kinematic variables adopted in this analysis are evaluated both in the CM frame ($\Upsilon(4S)$ rest frame) and in the laboratory frame. In figure 4.4 we show, for example, the correlation between lepton momentum in CM and in laboratory frame and the effects, in the laboratory frame, of a cut on lepton momentum in the CM frame. The p^*_ℓ cut is quite important for event selection as we will show later.

4.3 Classification of signal and background control samples

In addition to the signal sample, we also select several control samples that are used to characterize most of the main sources of background.

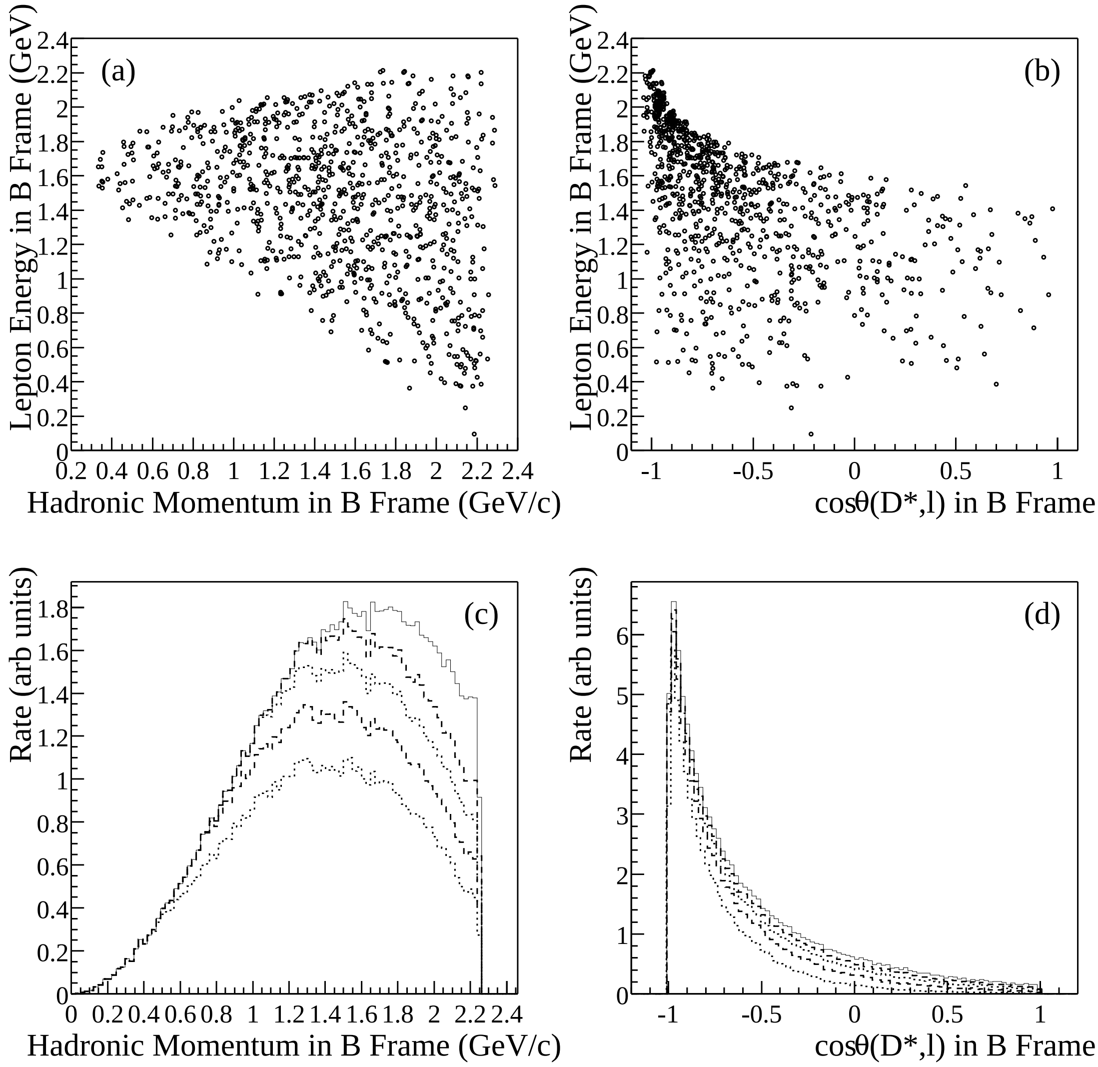


Figure 4.3: The upper plots show the correlations between the variables (a) $\hat{p}_X$ and $\hat{E}_\ell$, and (b) $\cos\hat{\theta}_{X,\ell}$ and $\hat{p}_X$, for $B^0 \to D^{*-}e^+\nu_e$ decays. The bottom plots show the effects of a minimum cut on the lepton energy on the distributions of (c) $\hat{p}_X$ and (d) $\cos\hat{\theta}_{X,\ell}$. The different histograms starting from the top dashed one correspond to $\hat{E}_\ell \geq 0$ GeV, 0.8 GeV, 1.0 GeV, 1.2 GeV, and 1.4 GeV.

We select about 209.000 events from a sample of about $34.54\,\text{fb}^{-1}$ recorded on the $\Upsilon(4S)$ resonance (on peak) and 6.500 events from a sample of about $3.79\,\text{fb}^{-1}$ recorded off resonance (about 40 MeV below $\Upsilon(4S)$ resonance) including events used to study the combinatoric background.

In order to separate real D^* from misreconstructed D^* background we use the variable δm, defined as $\delta m = m_{D^*} - m_{D^0}$. When studying the combinatoric background we keep δm values up to 165 MeV (in a misreconstructed D^* the δm

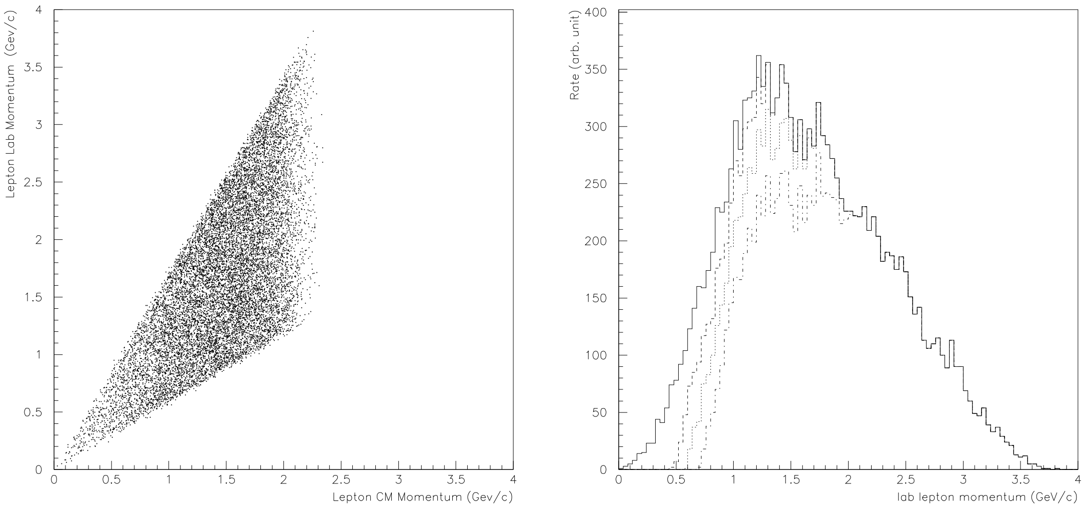


Figure 4.4: Distributions (left) of scatter plot for total momentum in lab and CM frame, and (right) total momentum for electrons and muons produced by B^0 decay into $D^{*-}\ell^+\nu_\ell$ with different CM momentum cuts. The histograms starting from the top one represents the lepton laboratory momentum with $p^*_\ell > 0$, $p^*_\ell > 0.8$ GeV/c, $p^*_\ell > 1.0$ GeV/c and $p^*_\ell > 1.2$ GeV/c. Both quantities are calculated from Monte Carlo truth.

distribution does not necessarily peak around m_π as requested for a true D^*).

We select about 99K events in the δm accepted range (see section 4.11), of which 38K are signal candidates and the remainder are in background control samples. Control samples are selection of events with enriched contribution of various backgrounds, and are listed in table 4.1, together with the quantities ("attributes") which are used to define them.

We distinguish between the signal and control samples using the first three attributes listed in table 4.1. Of the 2(`angCut`)×2(`onOffRes`)×3(`leptID`)= 12 possible combinations, 2 represent the signal sample and the remaining 10 are primary and secondary background control samples (see table 4.2). Signal candidates must pass the opposite-side angular cuts (see section 4.13.2), have lepton candidates identified as either electrons or muons (see section 4.5), and be recorded near the $\Upsilon(4S)$ resonance. Of the 10 control samples, 5 are primary samples that we use to directly study backgrounds that contribute to the signal sample, and the remaining 5 are secondary samples that we use to study backgrounds in other control samples.

In the reconstruction process we select only the D^* coming from $D^{*+} \to D^0\,\pi^+$ and charge-conjugate decay mode. We use the `svtDch` category to subdivide each of the 12 signal and control samples into up to 2 subsamples that can be studied independently. The `svtDch` category is useful to isolate the subsample (of about

Attribute	N_{cat}	Description
`angCut`	2	What are the angular correlations between the D^* and ℓ ?
		(1) candidates pass the "opposite-side" (OS) cuts ($\cos\theta_{D^*,\ell} < 0$)
		(2) candidates pass the "same-side" (SS) cuts ($\cos\theta_{D^*,\ell} \geq 0$)
`onOffRes`	2	What center-of-mass energy was the event recorded at ?
		(1) near the $\Upsilon(4S)$ resonance
		(2) below the $\Upsilon(4S)$ resonance ("continuum")
`leptID`	3	Is the kinematically-selected ℓ candidate lepton-like ?
		(1) candidate passes "very-tight" electron ID
		(2) candidate passes "very-tight" muon ID
		(3) candidate fails "loose" electron and muon ID
`Dmode`	4	How was the D^0 candidate reconstructed ?
		(1) $D^0 \to K\pi$
		(2) $D^0 \to K\pi\pi\pi$
		(3) $D^0 \to K\pi\pi^0$
		(4) $D^0 \to K_S\pi\pi$
`svtDch`	2	How was soft-pion candidate track reconstructed ?
		(1) in the SVT only
		(2) in the SVT and the DCH

Table 4.1: *Attributes used to subdivide the signal and control samples for the purposes of background characterization and extraction of the $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$. When using the $\cos\theta_{D^*,\ell}$ variable for SS and OS sample splitting, it is important to say that $\cos\theta_{D^*,\ell}$ is evaluated in the $\Upsilon(4S)$ rest frame.*

20%) of candidates for which the D^* mass is more precisely determined because DCH hits are available for the soft pion coming from the $D^* \to D^0\ \pi_{soft}$ decay(see figure 4.5, 4.16 and 4.17). For the purpose of this branching fraction measurement we introduce the `Dmode` categories to account for the different amounts and properties of the backgrounds (especially the combinatoric background) that we observe in these subsamples.

Table 4.3 shows the D^* and D^0 PDG2002 values for the modes adopted in this analysis.

We analyzed signal Monte Carlo separately for each of the final reconstructed D^0 modes ($e^+e^- \to B^0(D^{*-}\ell^+\nu_\ell)\overline{B}^0(X)$ where $D^{*-} \to \overline{D}^0\pi^-$ and $\overline{D}^0$ is forced to decay to one of the 4 modes listed in table 4.3)[2], on generic $B^0\overline{B}^0$ ($e^+e^- \to B^0(X)\overline{B}^0(Y)$)

[2] X, Y means the particle is left free to decay in the simulation.

Sample	`angCut`	`onOffRes`	`leptID`	Sample enriched in...
1	OS	On	e	electron signal
2	OS	On	μ	muon signal
3	OS	On	$!e,!\mu$	fake BG to samples 1 and 2
4	OS	Off	e	continuum BG to sample 1
5	OS	Off	μ	continuum BG to sample 2
6	SS	On	e	uncorrelated BG to sample 1
7	SS	On	μ	uncorrelated BG to sample 2
8	OS	Off	$!e,!\mu$	fake BG to samples 4 and 5
9	SS	On	$!e,!\mu$	fake BG to samples 6 and 7
10	SS	Off	e	continuum BG to sample 6
11	SS	Off	μ	continuum BG to sample 7
12	SS	Off	$!e,!\mu$	fake BG to samples 10 and 11

Table 4.2: *Summary of the signal and control samples selected for the $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ analysis. Samples 1 and 2 are signal, 3 to 7 are primary background control samples, and 8 to 12 are secondary background control samples. By writing $!e$ or $!\mu$ we mean that lepton candidate fails loose particle identification criterium as defined in section 4.5.*

and B^+B^- ($e^+e^- \to B^+(X)B^-(Y)$) Monte Carlo and on continuum $c\bar{c}$ ($e^+e^- \to c\bar{c}$) Monte Carlo for the amount shown in table 4.4. We analyzed also specific $B \to D^{**}\ell\nu$ and non resonant $B \to D^*\pi\ell\nu$ to perform systematics checks as we show in table 4.8.

This section is meant to get acquainted with the rather complicated, but necessary, jargon adopted throughout this thesis.

Reconstructed particle	Decay Mode	Branching ratio
$D^{*\pm}(m_{D^{*\pm}} = 2010.0 \pm 0.5\,\text{MeV})$	$D^{*\pm} \to D^0\pi^\pm$	$67.7 \pm 0.5\%$
$D^0(m_{D^0} = 1864.5 \pm 0.5\,\text{MeV})$	$D^0 \to K^-\pi^+$	$3.83 \pm 0.09\%$
	$D^0 \to K_s\pi^+\pi^-$	$2.7 \pm 0.2\%$
	$D^0 \to K^-\pi^+\pi^0$	$13.9 \pm 0.9\%$
	$D^0 \to K^-\pi^+\pi^+\pi^-$	$7.49 \pm 0.31\%$

Table 4.3: *Decay modes reconstructed in this analysis for the D^* meson. The Branching ratios are referred to PDG 2002 [5].*

Mode	SP4 Monte Carlo events	Equiv. Lumi (fb^{-1})
$B^0\overline{B}^0$	58.548.713	106.45
B^+B^-	62.342.000	113.35
$c\overline{c}$	92.584.401	71.22
Kpi	84.000	40.82
K3pi	82.000	12.49
Kpipi0	126.000	13.23
Kspipi	21.000	11.30

Table 4.4: *SP4 (acronym for Simulation Production 4) Monte Carlo samples adopted in this analysis. To evaluate equivalent luminosities from number of events we use $\sigma_{b\overline{b}}$ = 1.10 nb and $\sigma_{c\overline{c}}$ = 1.30 nb.*

4.4 Track Reconstruction

In order to explain the recostruction criteria adopted in the selection of the charged particles coming from the decay $B^0 \rightarrow D^{*-}\ell^+\nu_\ell$ we want to study, it is necessary to give a short list of the jargon adopted in *BABAR* to describe different recontsruction categories. The charged tracks we are speaking about are: $e^\pm$, $\mu^\pm$, $\pi^\pm$ and $K^\pm$. We recall that the volume in which the electron and positron beam interaction is estimated to happen is called beamspot. It has an ellissoidal shape and its position is continuosly monitored and stored in the conditions database.

1. **ChargedTracks (CT):** Any reconstructed track.
2. **GoodTracksVeryLoose (GTVL):** Any track meeting the following criteria:
   - Momentum lower than 10 GeV/c;
   - Distance of closest approach to the origin of the beamspot in xy plane less than 1.5 cm;
   - distance of closest approach on the z coordinate to the beamspot center lower than 10 cm;
3. **GoodTracksLoose (GTL):** Any GoodTrackVeryLoose that also satisfies:
   - $p_T > 100$ MeV/c;
   - 12 or more DCH hits;
4. **GoodTracksTight (GTT):** Any GoodTrackLoose that also satisfies:

- Distance of closest approach to the origin of the beamspot in xy plane less than 1.0 cm;
- distance of closest approach on the z coordinate to the beamspot center lower than 3 cm;
- 20 or more DCH hits

It is important to use the charged tracks selected only with the criteria mentioned above since, as we will show later, there are dedicated studies and recipes available inside *BABAR* to evaluate tracking efficiencies and systematic uncertainties on them.

The charged tracks selected to reconstruct the D^0 are $\pi^{\pm}$ and $K^{\pm}$. The decay mode $D^0 \rightarrow K_S\pi\pi$ contains a K^0_S, which we decided to reconstruct only in the charged pion decay mode (the decay $K^0_S \rightarrow \pi^+\pi^-$ account for about $68.60 \pm 0.27\%$ of all K^0_S's, according to [5]). The decision not to reconstruct the $K^0_S \rightarrow \pi^0\pi^0$ comes from the the big systematic effect introduced in having two π^0 which practically makes useless the mode for a precision branching fraction measurement as we are targeting. The K^0_S reconstruction requires two charged pions selected from the CT criterium with the pion hypotheses. Given this, all the charged kaons and pions of the 4 D^0 modes used in this analysis are selected with the GTL criterion.

The selection adopted to reconstruct the soft pion from the $D^* \rightarrow D^0\ \pi_{soft}$ is the GTVL criterium. As can be seen in figure 4.5, the soft pion in about 75% of the selected $B^0 \rightarrow D^{*-}\ell^+\nu_\ell$ events has no DCH hits, making almost impossible the GTL requirement without decreasing dramatically selection efficiency. Another important physics motivation to select the SVT-hits only soft pions comes from the fact that for a $|V_{cb}|$ mesurements it is crucial to keep as many events with low w (defined in chapter 1.2) as possible in order to achieve the best sensitivity and to reduce uncertainties in the $|V_{cb}|$ fit technique.

We will show in section 6.2 how to account for all the different tracking efficiency corrections and systematic errors to be associated for each of the categories adopted for the soft pion (GTVL), the lepton (GTT) and charged D^0 tracks (GTL a part from the K^0_S charged pion daughters which are CT). This is one of the reasons which explains why we evaluate tracking corrections and systematic errors separately, when measuring $\mathcal{B}(B^0 \rightarrow D^{*-}\ell^+\nu_\ell)$, for each of the 8 channels identfied by the values of `Dmode` (4)$\otimes e/\mu$ (2).

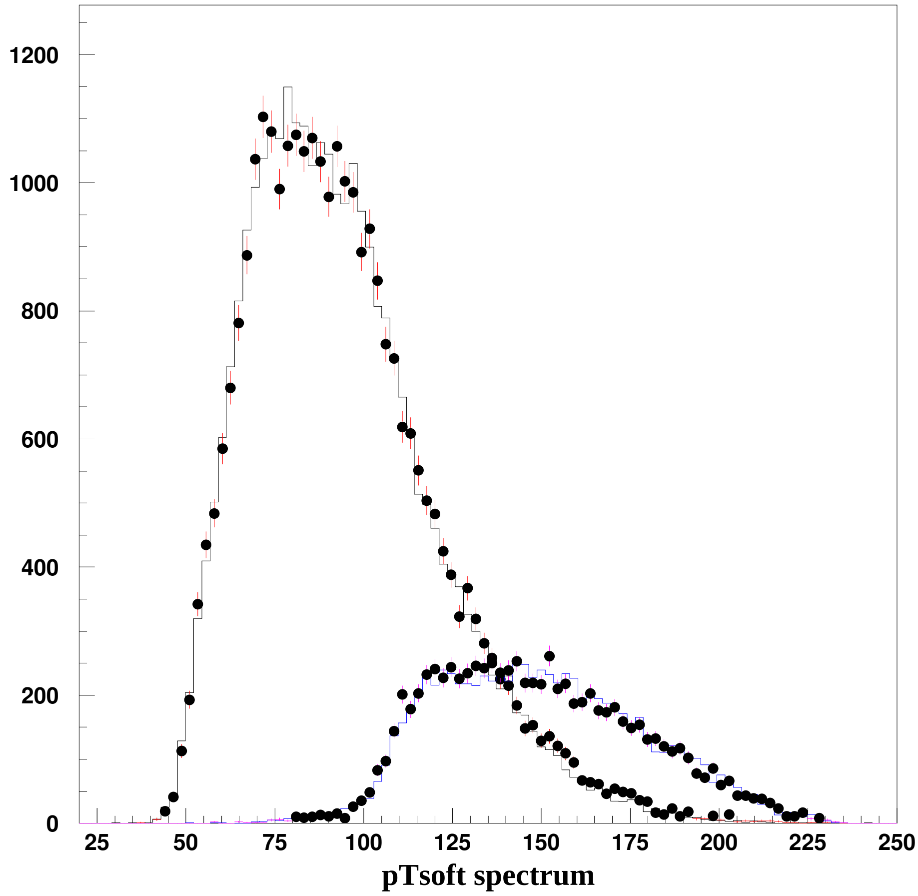


Figure 4.5: Soft pion p_T distribution in data (dots) and in MC (histograms) normalized w.r.t. data. We can distinguish among events with only SVT hits (low p_T peaked around 80 MeV and with both SVT and DCH hits (distributed above 100 MeV) for the soft pion.

4.5 Particle Identification

In order to reconstruct the $D^*\ell$ candidate, it is necessary to identify the leptons and all the D^0 meson decay products such as $K^{\pm}$, π^0 and K^0_S [3]. We adopted criteria and algorithms common to all *BABAR* physics analyses.

In *BABAR* there are several classification categories for each particle hypothesis, each one corresponding to a different level of sample purity. Selection efficiencies and misidentification probabilities have been measured for each one of these categories

[3] while $K^{\pm}$ can be identified with the DIRC that has a very good K/π separation, π^0 and K^0_S are reconstructed from a $\gamma\gamma$ and a $\pi^+\pi^-$ pair respectively. In this analysis we decided to reconstruct K^0_S only from charged pions pairs since the neutral mode $K^0_S \to \pi^0\pi^0$ is less efficient and introduces larger systematics effect in reconstruction.

in specific control samples. Radiative Bhabha events, for example, have been used to study electrons, while $e^+e^- \to e^+e^- \ \mu^+\mu^-$ and $e^+e^- \to \mu^+\mu^- \ \gamma$ control samples have been used to study muons. The efficiencies and misidentification rates are tabulated in bins of (p, θ, ϕ), measured in the laboratory frame.

Below we show identification criteria adopted in this analysis for $e^\pm$, $\mu^\pm$, $K^\pm$ and γ used for π^0 reconstruction.

4.5.1 Electron Identification

Electron identification is performed through selection criteria based on the following kinematic variables [63]: the E/p ratio between the energy deposited in the EMC and the candidate electron momentum; the number of crystals involved in the electromagnetic shower; the cluster shape in the calorimeter and the $\mathrm{d}E/\mathrm{d}x$ ionization energy loss in the SVT and DCH. When using the identification criteria which has best purity the Cherenkov angle in the DIRC is adopted in the selection too.

The E/p ratio is one of the principal discriminating variables for electrons with momentum greater than 500 MeV/c. All other charged muons or hadrons above this threshold are MIPs (Minimum Ionizing Particle). Some hadrons may interact in the EMC, giving rise to hadronic showers. The ratio E/p of an electron in the EMC is distributed around 1.

Electromagnetic showers in the EMC have a different signature than showers of hadronic interactions resulting in different longitudinal and lateral energy distributions. Two parametrizations of lateral shower shapes currently in use are:

- Lateral energy distribution LAT, first introduced at Argus [48]:

$$LAT = \frac{\sum_{i=3}^{n} E_i r_i^2}{\sum_{i=3}^{n} E_i r_i^2 + E_1 r_0^2 + E_2 r_0^2}, \qquad E_1 \geq E_2 \geq \ldots \geq E_n$$

  where the sum extends over all crystals in a shower, $r_0 = 5\,\mathrm{cm}$, the average distance between two crystal frontfaces, and r_i is the distance between crystal i and the shower center (calculated as the center of gravity with linear energy-weighting of every crystal).

  Electromagnetic showers in the EMC have most of their energy deposited in one or two crytals (narrow showers), hence LAT is smaller for electromagnetic showers than for hadronic showers.

- Furthermore, hadronic showers in the EMC tend to be more irregular than electromagnetic showers. This is exploited in the expansion of the lateral shower shape in terms of Zernike moments (first introduced at ZEUS [49]):

$$A_{nm} = \sum_{r_i \leq R_0}^{n} \frac{E_i}{E} \cdot f_{nm}(\frac{r_i}{R_0}) \cdot e^{-im\phi_i}, \qquad R_0 = 15\,\text{cm}$$

  with

$$f_{nm}(\rho_i \equiv \frac{r_i}{R_0}) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)! \rho_i^{n-2s}}{s!((n+m)/2-s)!((n-m)/2-s)!},$$

  with $n, m \geq 0$ integers, $n - m$ even and $m \leq n$. Zernike functions are used as they provide an expansion independent of the orientation of the lateral coordinate system (in which ϕ_i is measured).

  The azimuthal variation in shower shape enters only for moments with $m \geq 2$, like in $|A_{42}|$, the only Zernike-moment currently used in electron identification.

Shower shapes and E/p are complementary in so far as hadronic showers depositing a large energy amount (and hence a large E/p) tend to have large values for the shower shape variables (due to the irregularities of the hadronic shower).

There are differences of longitudinal shower development for electrons and hadrons: electrons shower shortly after entering the EMC, while interacting hadrons usually shower later. To take into account this different behaviour we use the variable:

$$\Delta\phi = q(\phi_{EMC} - \phi_{cluster}),$$

where q is the particle charge, ϕ_{EMC} is the ϕ coordinate of the track intersection with the EMC and $\phi_{cluster}$ is the azimuth of the cluster associated to the track. The $\Delta\phi$ variable for electrons is distributed around smaller values than for hadrons as can be seen in figure 4.6

The electrons $\mathrm{d}E/\mathrm{d}x$ distribution has a gaussian shape whose peak is centered around $< \mathrm{d}E/\mathrm{d}x >\sim 650$ with a width $\sigma = 50$. Depending on the level of purity required in the selection, different cuts are applied around the distribution mean value.

The classification category with the best purity, called *ElectronVeryTight*, is used in this analysis to select electrons in the signal sample. The candidates failing

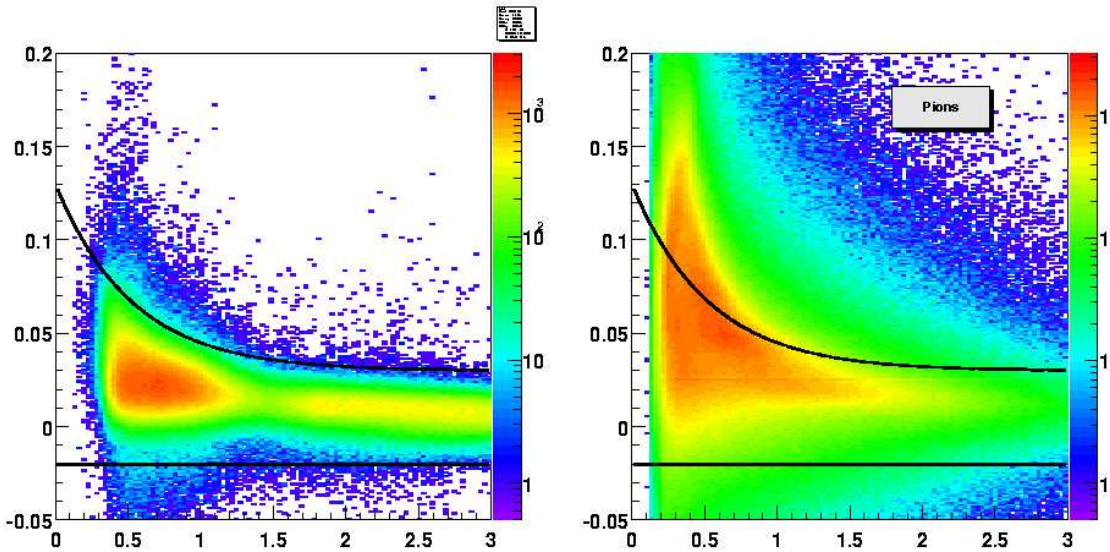


Figure 4.6: Discriminant power of $\Delta\phi$ variable. The left plot show $\Delta\phi$ as a function of electron p_T in GeV/c obtained from control samples. Selected events are inside the black curves. On the right there is the same plot done on a pion control sample.

a selection criteria with higher efficiency[4], called *ElectronLoose*, are used to characterize the fake-lepton control sample. Table 4.5 describes the cuts applied for the *ElectronVeryTight* and *ElectronLoose* selection criteria.

Crit.	dE/dx	N_{Xtal}	E/p	LAT	A_{42}	$\Delta\phi$	DIRC
ElectronLoose	500 ... 1000	3	0.65 ... 5.0	-10. ... 10.	-10. ... 10.	no	no
ElectronVeryTight	540 ... 860	3	0.89 ... 1.2	0. ... 0.6	-10. ... 0.11	yes	yes

Table 4.5: *Definition of the various electron selector criteria. Here N_{Xtal} is the number of crystals involved in the shower.*

The Cherenkov angle cut in the *ElectronVeryTight* category is applied only if the number of photons detected in the DIRC, associated to the track, is larger than 10. In such case the Cherenkov angle is required to be inside 3σ from the corresponing value of an electron hypothesis.

We will show later in section 5.2 the plots showing both electron and muon selector efficiencies and misidentification probabilities, which have been used to evaluate the fake-lepton background in our selected data sample.

[4]hence with a lower purity.

4.5.2 Muon Identification

We already mentioned muon identification in section 2.9.3. Here we give a quantitative description of the identification procedure.

Muon identification is performed through selection criteria based on the following discriminating variables [65]:

- the energy E_{cal} lost in the calorimeter;
- the number of interaction lenghts, λ_{meas}, of the track inside the detector;
- the difference, $\Delta\lambda = \lambda_{exp} - \lambda_{meas}$, among λ_{meas} and the expected number of interaction lenghts, λ_{exp}, evaluated with the muon hypothesis;
- the spatial distribution of the hits with respect to the tracks extrapolation inside the IFR;
- the continuity of the track crossing the iron in the IFR;
- the average number of hit strips for each IFR layer crossed.

E_{cal} is used mainly to discriminate muons from electrons. On the contrary, the same distribution evaluated for muons is peaked around 200 MeV since, at the *BABAR* typical energies, they are MIP. The separation of hadrons is less easy, because hadrons may have interactions with the CsI nuclei. For μ separation we study the penetration of the track in the iron of the IFR. The number of interaction lenghts, λ, of the track inside the detector is estimated through an algorithm which extrapolates the track inside the IFR. Starting from the momentum mesurement, done with the tracking system (SVT+DCH), one can evaluate λ_{exp} with the muon hypothesis. The $\Delta\lambda$ distribution is clearly peaked around zero for a muon, while it is almost flat for a pion. It is possible to discriminate among muons and pions only with $\Delta\lambda$ if particle momentum is lower than 1.5 GeV/c.

The hit spatial dispersion with respect to the track extrapolation in the IFR is described by the variable

$$\chi^2_{trk} = \sum_{i=1}^{N_{hit}} \frac{(x_i^{meas} - x_i^{extr})^2}{(\sigma_i^{meas})^2},$$

where x_i^{meas} and x_i^{extr} are the measured and extrapolated hit coordinates, respectively. The cluster related to a pion interaction presents larger residuals of the extrapolated track compared to a muon cluster, giving a larger value of χ^2_{trk}. This is due to pion hadronic showers. Another variable, χ^2_{fit}, is used to discriminate muons from pions. It represents the χ^2 of a polinomial fit performed on the positions of the strips contributing to the IFR cluster. It is evident from their definitions that χ^2_{trk} and χ^2_{fit} are correlated.

The continuity of the track crossing the iron in the IFR is defined by:

$$T_c = \frac{N_L}{L_h - F_h},$$

where N_L indicates the total number of layers where the signal is detected and F_h and L_h correspond to the first and last layer of the cluster, respectively. Muon tracks are characterized by more continuity, compared to hadron tracks (which have non neglibile neutral particles in the hadronic showers). As a consequence we expect values of T_c closer to 1 for muons than for hadrons.

The average number of strips per layer which detected a particle, $< m >$, and its standard deviation, σ_m are also used to discriminate muons from pions since $< m >$ distribution is narrower and centered around smaller values for muons than for pions.

Same as was done for electrons, in this analysis we adopted two selection categories for muons. Signal muon candidates are selected with the highest purity category, *MuonVeryTight*. In order to study fake-lepton background coming from muon selection, we select events which fail a looser selection criterium, *MuonLoose*.

Crit.	E_{cal}	$\Delta\lambda$	λ	χ^2_{trk}	χ^2_{fit}	T_c	$< m >$	σ_m
MuonLoose	< 0.5 GeV	< 2	> 2	< 7	< 4	> 0.2	< 10	< 6
MuonVeryTight	0.05...0.4 GeV	< 0.8	> 2.2	< 5	< 3	> 0.34	< 8	< 4

Table 4.6: *Definition of the various muon selector criteria used in this analysis. In both categories only candidates detected in a minimum number of at least two layers are selected.*

We will show later in section 5.2 the plots with the muon identification efficiencies and misidentification probabilities for all the selector adopted in this analysis.

4.5.3 Identification of charged Kaons

The selection of charged kaons in *BABAR* is mainly based on the information from the DIRC, the SVT and the DCH. From the truncated mean dE/dx measurement, a better than 2σ separation of pions and kaons is possible up to about 0.6 GeV/c

K^+ decay	relative rate [%]
$\mu^+\nu_\mu$	63.51±0.18
$\pi^+\pi^0$	21.16±0.14
$\pi^+\pi^+\pi^-$	5.59±0.05
$\pi^0\mu^+\nu$	3.18±0.08
$\pi^0 e^+\nu$	4.82±0.06
$\pi^+\pi^0\pi^0$	1.73±0.04

Table 4.7: *Dominant kaon decay rates [5].*

particle momentum in the SVT and 0.7 GeV/c in the DCH. For momenta above 1.5 GeV/c the DCH provides about 2σ dE/dx separation between kaons and pions thanks to the relativistic rise. The selection algorithm in both momentum regions is based on the principle of maximum likelihood ratio.

The signature of charged kaons in the different detectors is complicated because they can decay or interact with the material of the detector. Furthermore, the hadronic interactions have different rates for the two charged modes. About 20% of the kaons decay before they reach the DIRC which is located at a radius of about 81 cm. The dominant kaon decay modes are listed in table 4.7. About 94% of the charged kaon decays result in a single charged track, which by following the initial kaon direction may not be distinguished from the original kaon.

Kaons interact with the material of the detector. The total cross sections for $K^\pm N$ scattering are shown in figure 4.7 together with a smooth function, which is used to simplify further estimates. The data are taken from the PDG [5]. From the total cross sections one can estimate how many kaons are lost after interacting in the detector material.

In order to perform our analysis and measure $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$, we have to reconstruct charged kaons coming from the $D^0 \to K\pi$, $D^0 \to K\pi\pi\pi$ and $D^0 \to K\pi\pi^0$ decays. We show in figure 4.9 the momentum and polar angle distribution of true kaons, from Monte Carlo simulation, coming from the D^0 decays mentioned above.

The signature of kaons in the Cherenkov detector is quite complex. Their Cherenkov angle distribution from Monte Carlo after the likelihood fit is shown in figure 4.8. The fit steps through the different particle hypotheses fitting the Cherenkov angle and arrival time distribution of the photons associated to a track. The kaon band is labeled as region C in the figure. In addition the track impact lo-

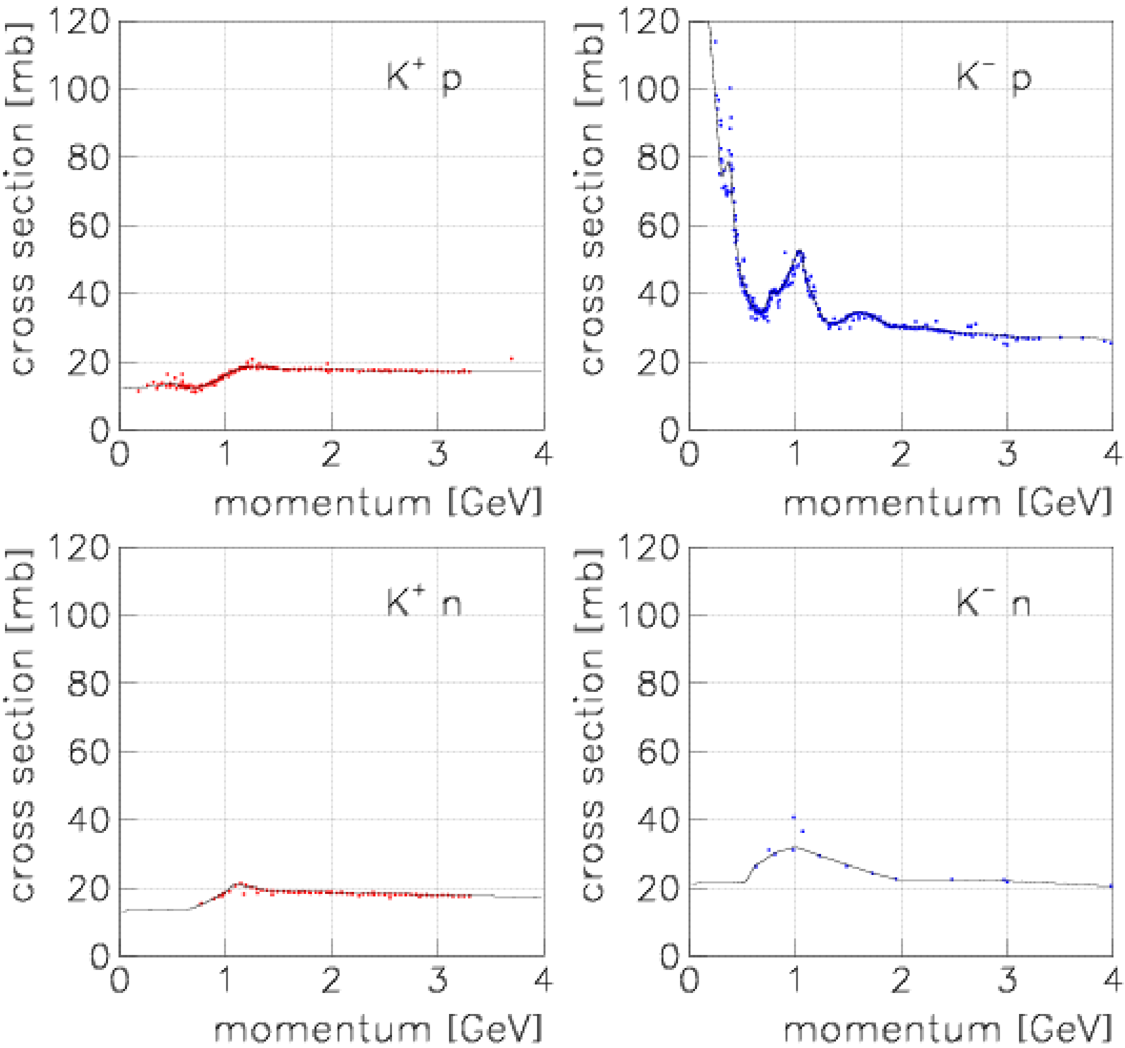


Figure 4.7: Total cross section for kaon scattering on protons and neutrons [5] for the different charges, together with a smooth function for estimates.

cation in the quartz bar is varied. This procedure leads to the gaussian distribution of measured Cherenkov angles with respect to the expceted angle for the different particle hypotheses. Region A in the figure is populated by tracks for which no Cherenkov angle fit is performed, because either they did not reach the DIRC (8% due to the solid angle coverage of the DIRC), or produced a small amount of photons compared to background photons near the particle threshold, or, for a small fractions of events, the fit failed to converge. Close to the kaon threshold, kaons do show up in the electron, muon and pion band with a rather wide spread, marked as region B in the figure. These kaons decayed or interacted before the DIRC and the reaction products did not deviate too much from their mother kaon direction. They are able to procuce light in the quartz radiators close to the expected range. The

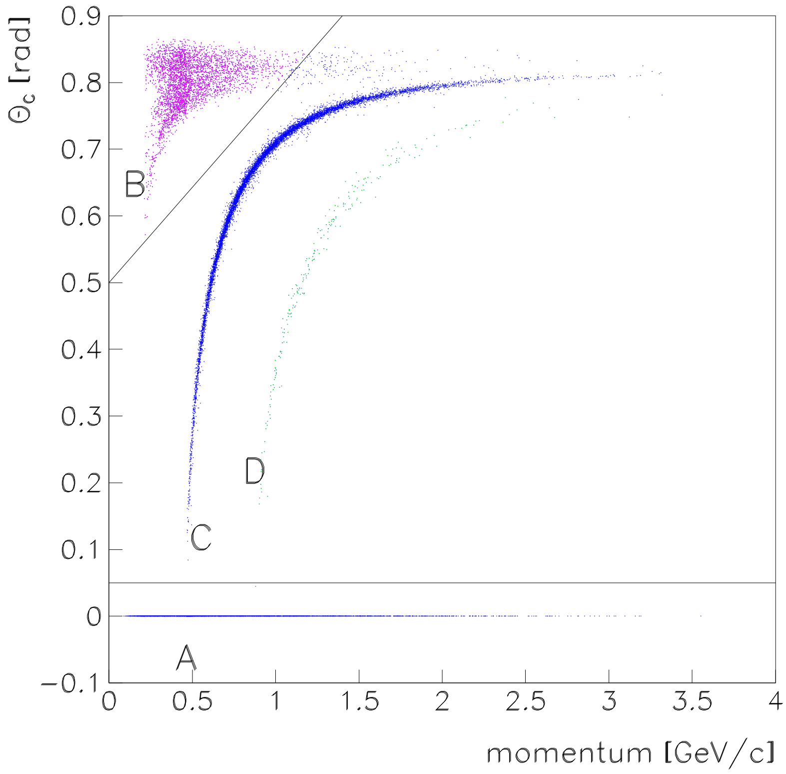


Figure 4.8: Distribution of the Cherenkov angle measurements versus track moemntua for simulated kaon tracks: the kaons appear in all signal regions of other particle types. Region A is populated by tracks for which no Cherenkov angle fit is performed. Region B is the electron, muon and pion band. Region C is the kaon band. Region D is the proton band. As can be seen from the figure, the simulated kaons show up in all the 4 regiond mentioned above.

imprecision in track-momentum and direction and Cherenkov-ring information reflects into the fit result. Region D in the figure correspond to the proton band which is also populated by kaons. They underwent hadronic interactions which set free real protons. A fraction of the tracks are misinterpreted by the fit due to a reduced number of detected photons or distorted track measurements. The selection based on a combination of likelihoods from the different detector components is sensitive to these kaon characteristics.

Once constructed the likelihoods, l_i, for each mass hypothesis, we perform cuts on each likelihood. The kaon selection efficiencies and misidentifications depend on the cuts. In this analysis we adopted two selection criteria[5] for kaon identification.

[5]all selection modes are tuned on a D^* decay sample.

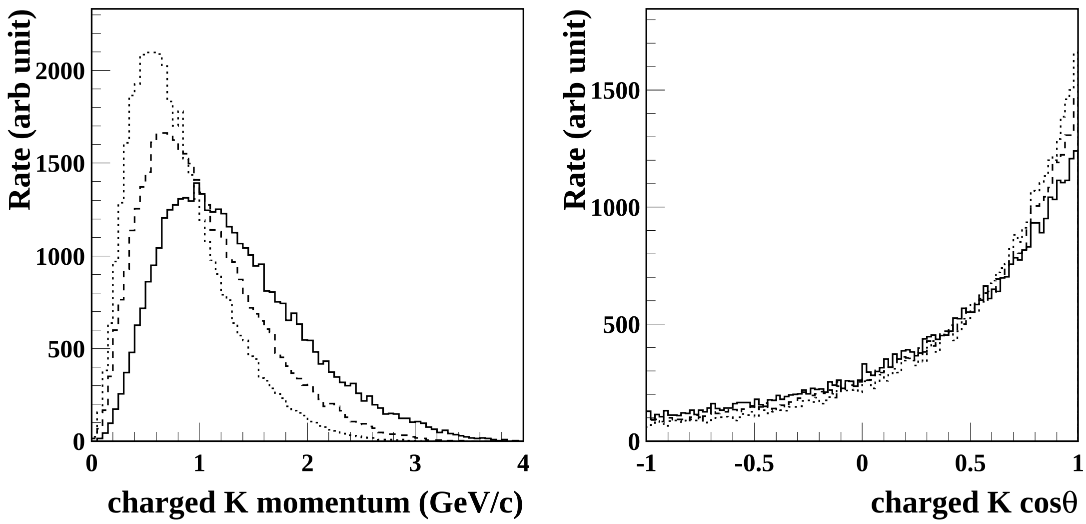


Figure 4.9: Distributions (right) of total momentum, and (left) cosine of the polar angle for charged kaons produced by D^0 decay into: $K^-\pi^+$ (solid curves), $K^-\pi^+\pi^0$ (dashed curves), and $K^-\pi^+\pi^+\pi^-$ (dotted curves). Both quantities are calculated from Monte Carlo truth in the lab frame. The plots for different decay modes are normalized to the same number of events.

The one called *KaonTight* is optimized to keep the misidentification below 5% up to momenta of 4 GeV/c. It provides SVT and DCH information up to momenta lower than 0.7 GeV/c. The DIRC information is used for momenta greater than 0.6 GeV/c. The likelihood cut are $l_K > l_p$ and $l_K > r_\pi l_\pi$, where $r_\pi = 1$ for momenta lower than 2.7 GeV/c, and $r_\pi = 80$ for momenta greater than 2.7 GeV/c. The *KaonTight* criterion is adopted in the recostruction of the $D^0 \to K\pi\pi\pi$ and $D^0 \to K\pi\pi^0$ modes. In the case of the two body $D^0 \to K\pi$ decay mode we adopted the *NotAPion* criterion, with a lower purity but an higher efficiency[6]. The basic concept of this kaon selector is rejecting pions keeping high efficiency. It provides SVT and DCH information up to momenta lower than 0.5 GeV/c and 0.6 GeV/c, respectively. The DIRC information is used for momenta greater than 0.6 GeV/c. The likelihood cut are $l_p > r_\pi l_\pi$ and $l_K > r_\pi l_\pi$ where $r_\pi = 0.1$ for momenta lower than 0.5 GeV/c and $r_\pi = 1$ for momenta greater than 0.5 GeV/c.

[6]this does not reduce the D^0 reconstruction purity in this mode since this decay is characterized by a smaller combinatoric background.

4.5.4 Photon identification

We call photons all energetic deposits in the calorimeter, which are consistently associated into clusters with proximity criteria and are isolated with respect to charged tracks extrapolated in the calorimeter. From the null mass hypothesis and the energy released in the calorimeter it is straightforward to evaluate the photon candidate momentum. Since the photons we are interested in this analysys are always pairs coming from the π^0 decay, we will give a quantitative description of the selection when presenting the π^0 reconstruction.

4.6 Vertex reconstruction

The vertexing algorithm used in this analysis is described in detail in Appendix B.

This section highlights the most important aspects of vertexing that relate to the decay vertices reconstruction for this analysis.

The vertex and momentum of the K_S^0 candidate are calculated by fitting two charged tracks to a common vertex, and the result forms a virtual track for the D^0 reconstruction. The π^0 candidate is form by combining two neutral clusters and is fit with a constraint so that the invariant mass of the π^0 candidate is equal to the true value [5]. The momentum of the π^0 candidate contributes to the calculation of the D^0 invariant mass, but does not contribute to the information for D^0 vertex. The daughter particles of the D^0 candidate are then fit to a common vertex, with a constraint so that the invariant mass of the D^0 candidate is equal to the world average [5]. The resulting momentum and position are combined with the slow pion and lepton candidate in a vertex fit that assumes the B meson decays within the beam colliding region (beam spot).

In section 4.13.1 we discuss the improvement in δm resolution coming from a refit of the $D^*\ell$ candidate including the D^0 decay lenght and all its correlations with D^0 momentum and soft pion momentum.

4.7 π^0 Reconstruction

Neutral pion candidates are formed from pairs of EMC bumps[7] with energy greater than 30 MeV and not matched with any track, that are assumed to be photons orig-

[7] we define bumps all local maxima of calorimeter energy deposit.

inating from the interaction point. At the level of final selection cuts, the invariant mass of the photon pair is required to be between 119.2 MeV and 150.0 MeV, which is about $\pm 2.5\sigma_{\gamma\gamma}$ of the nominal π^0 mass, where $\sigma_{\gamma\gamma}$ is the $\gamma\gamma$ mass resolution, with a minimum total energy of 200 MeV. Selected γ candidates are kinematically fitted so that $\gamma\gamma$ invariant mass is equal to the nominal π^0 mass. The χ^2 probability of the fit is required to be greater than 1%. Due mainly to the acceptance of the EMC, efficiencies for this selection vary from about 55% to 65% for π^0 energies from 0.3 to 2.5 GeV, the range of π^0 energies in B decays [71].

The $\gamma\gamma$ mass and energy distribution, the χ^2 probability of mass-constraint fit and the π^0 fitted energy versus the $\gamma\gamma$ energy scatter plot are shown in figure 4.10 for a typical data taking run.

4.8 K_S^0 Reconstruction

K_S^0 candidates are formed from pairs of charged tracks (CT), assumed to be pions. These two tracks are fitted to a common vertex. The χ^2 probability of the fit is required to be greater than 1%. At the level of final selection cuts, the $\pi^+\pi^-$ invariant mass, calculated at the vertex of the two tracks, is required to be within ± 15 MeV of the nominal K_S^0 mass [5].

The $\pi^+\pi^-$ mass and energy distribution evaluated at the vertex of mass-constraint fit are shown in figure 4.11 for a typical data taking run.

4.9 Event Shape

To reduce continuum background, we calculate the angle θ_{thrust} between the thrust of the $D^*\ell$ candidate and that of the rest of the event in the $\Upsilon(4S)$ center-of-mass frame. Events from continuum tend to be jet-like, therefore the angle between two thrusts tends to be close to zero or 180 degrees, so that the angle is peaked near $|\cos\theta_{\text{thrust}}| = 1$, while the decay topology of $B\overline{B}$ events is more spherical since two B mesons decay almost at rest, therefore the directions of the two trusts are not correlated and the angular distribution is flat. Figure 4.12 shows the $|\cos\theta_{\text{thrust}}|$ distributions of $B\overline{B}$ and continuum events, $e^+e^- \to c\overline{c}$, from our Monte Carlo sample. We select events with $|\cos\theta_{\text{thrust}}| < 0.85$. The efficiencies for signal and continuum events to satisfy this requirement are about 85% (signal distribution is flat) and 46% (about 1/2 of the continuum events are rejected), respectively.

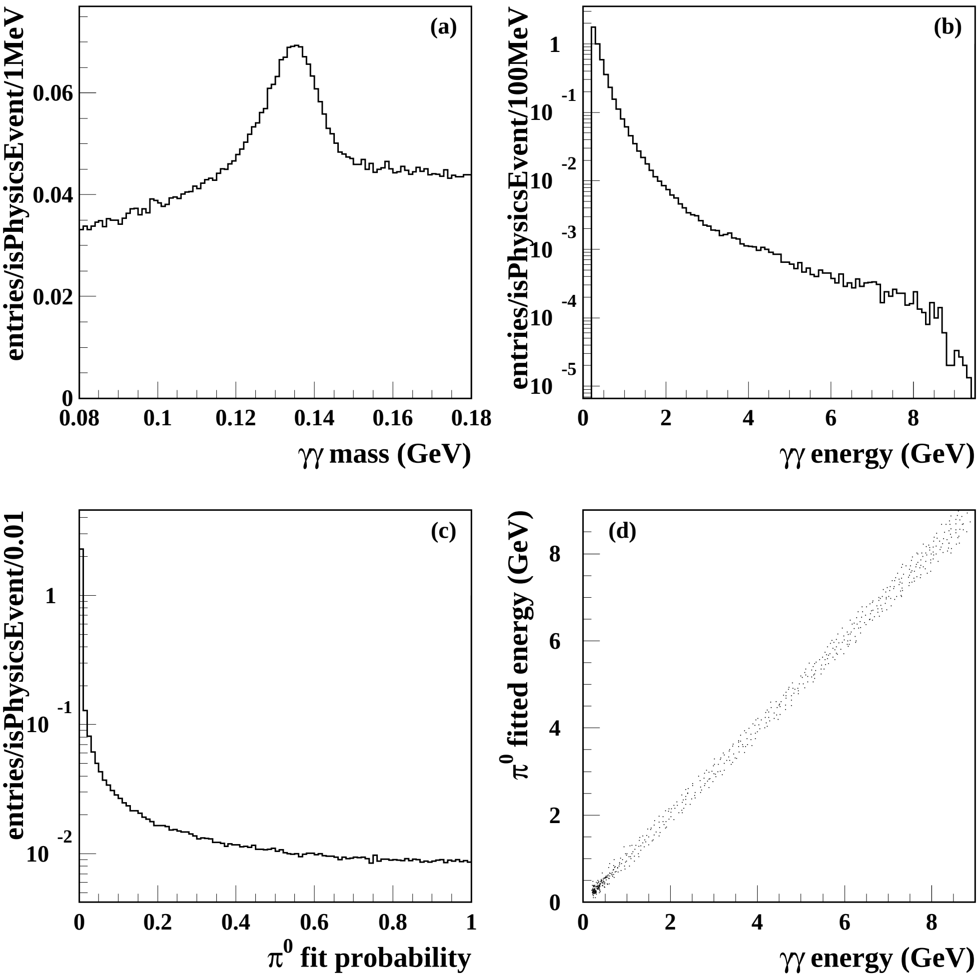


Figure 4.10: Distributions of (a) invariant mass, (b) total energy for pairs of neutral calorimeter clusters, (c) χ^2 probability of mass-constraint fit, and (d) a scatter plot of energies comparison from a typical run. The plots are normalized to show entries/event/bin and are shown before applying final selection cuts.

4.10 D^0 Reconstruction

The D^0 candidates are obtained from reconstructing $D^0 \to K^-\pi^+$, $D^0 \to K^-\pi^+\pi^-\pi^+$, $D^0 \to K^-\pi^+\pi^0$ and $D^0 \to K^0_S\pi^+\pi^-$ decay modes. The D^0 vertex is reconstructed as described in section 4.6. Charged kaon of $D^0 \to K\pi$ is required to pass *NotAPion* identification criterium, while charged kaons from $D^0 \to K\pi\pi\pi$ and $D^0 \to K\pi\pi^0$ modes are required to pass the *KaonTight* identification criterium. No particle iden-

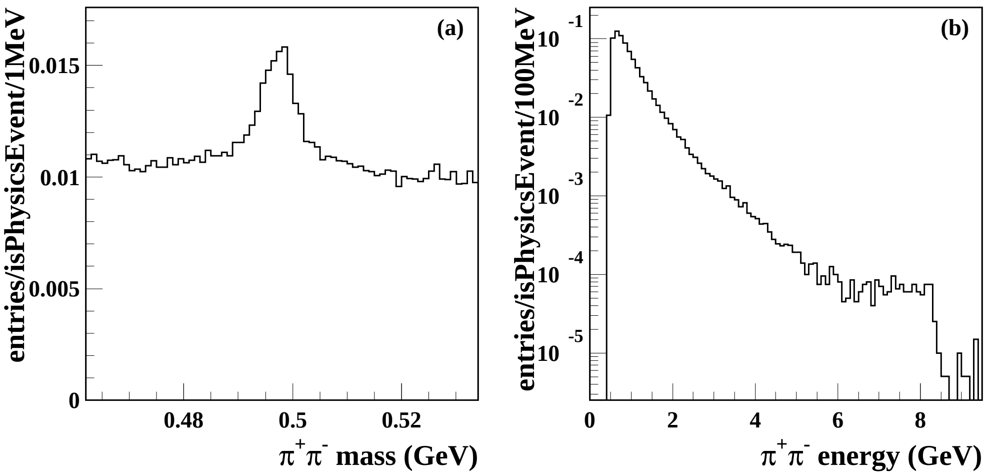


Figure 4.11: Distributions of (a) invariant mass, (b) total energy for pairs of CT, computed at the geometric vertex of the two tracks from a typical run. The plots are normalized to show entries/event/bin and are shown before applying final selection cuts.

tification is applied for $D^0 \to K_S\pi\pi$ mode.

The invariant mass distributions of D^0 candidates selected in Monte Carlo samples with the final selection cuts are shown in figure 4.13. The selection windows are chosen to be ± 34 and ± 17 MeV from nominal D^0 mass [5] for $D^0 \to K\pi\pi^0$ and other three modes, respectively, which are about 2.5 times the width of the D^0 mass distribution. We are forced to enlarge the mass window for the $D^0 \to K\pi\pi^0$ mode because of the presence of the π^0 in the decay chain which decrease the mass resolution. The D^0 candidates are then reconstructed with the invariant mass constrained to be the nominal D^0 mass. The χ^2 probability of the fit is required to be larger than 0.1%.

The $D^0 \to K^-\pi^+\pi^0$ and $D^0 \to K^0_S\pi^+\pi^-$ modes are known to have resonance substructure [72, 56]. The major resonances are $K^-\rho^+$, $K^{*-}\pi^+$ and $\overline{K}^{*0}\pi^0$ for $D^0 \to K\pi\pi^0$ mode and $K^{*-}\pi^+$,$\overline{K}^0\rho^0$, $\overline{K}^0 f_0$, $\overline{K}^0 f_2$, for $D^0 \to K_S\pi\pi$ mode. One can take advantage of these structures by selecting D^0 candidates near these resonances to increase signal to background ratios. Figure 4.14 shows the Dalitz plots for these two modes, in which the density is calculated based on the decay amplitudes and phases found in [56]. A detailed discussion about this method can be found in Appendix. C.

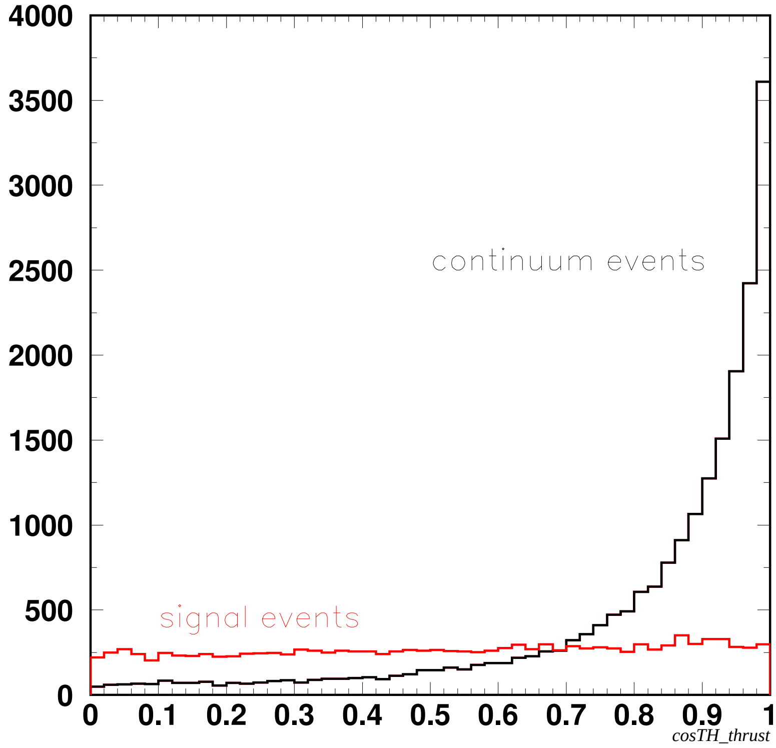


Figure 4.12: MC distribution of the $|\cos\theta^*_{\mathrm{thrust}}|$ variable in signal events (red flat histogram) and in continuum background events (black peaked histogram) for events which passed all cuts but the one on $\cos\theta^*_{\mathrm{thrust}}$. Both histograms are normalized to the same luminosity.

For each event, its expected density on the Dalitz plot is calculated. The event is kept if this value is larger than 10 % of the maximum value (within the red contours in plots (a) and (b) in figure 4.14). This cut is very effective against flat background.

It is worth noticing that the Dalitz plot for $D^0 \rightarrow K_S\pi\pi$ is not symmetric under the exchange of π^+ and π^-. It is due to the fact that a D^0 can produce the resonance $K^{*-}\pi^+$ but not $K^{*+}\pi^-$, and vice versa for $\overline{D}^0$. Therefore, we need to know whether the flavor of the mother particle is a D^0 or $\overline{D}^0$, (which cannot be determined from the final state $K^0_S\ \pi^+\pi^-$ alone), before the Dalitz density can be calculated correctly. We need to rely on the charge association between the D^* and soft pion, i.e., $D^{*+} \rightarrow D^0\pi^+$, and $D^{*-} \rightarrow \overline{D}^0\pi^-$ to determine the flavor of D^0.

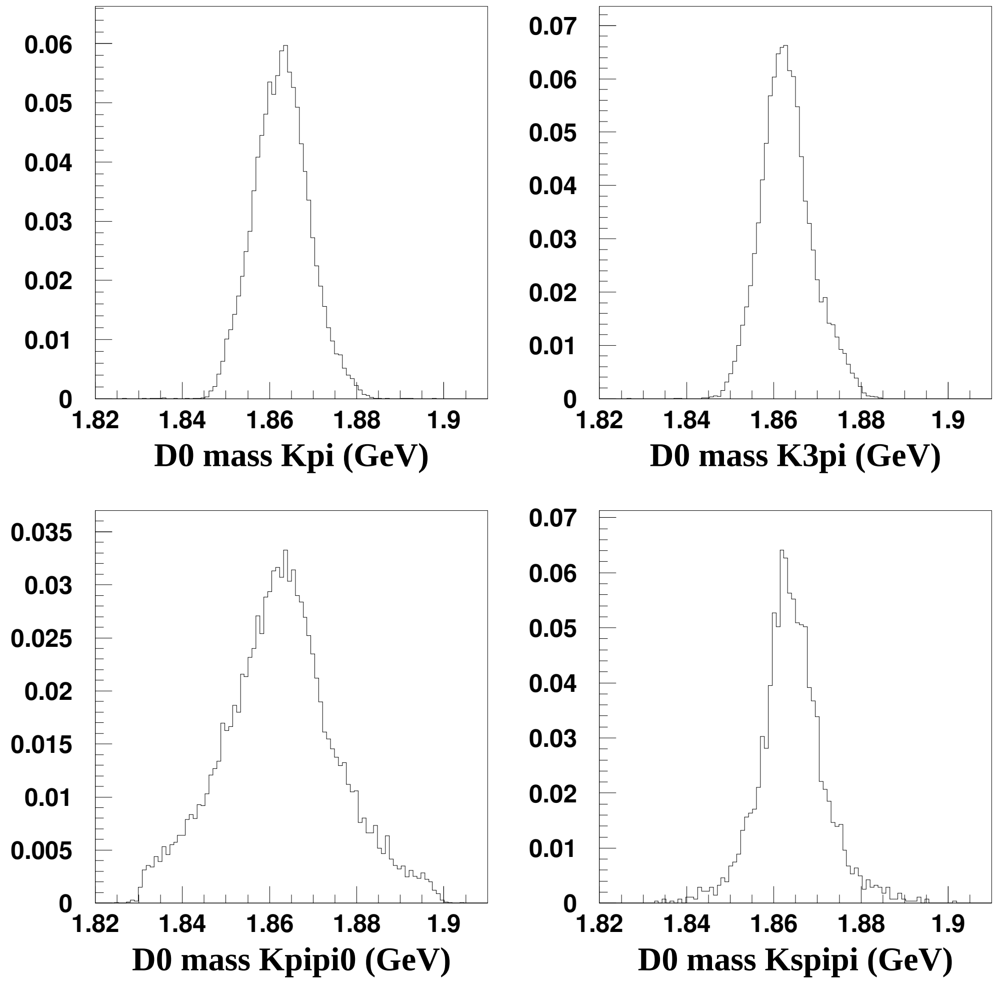


Figure 4.13: Reconstructed D^0 mass distributions in all D^0 modes adopted in this analysis. We show the mass spectra of the selected candidates in our data sample for $D^0 \to K\pi$ (top left), $D^0 \to K\pi\pi\pi$ (top right), $D^0 \to K\pi\pi^0$ (bottom left) and $D^0 \to K_S\pi\pi$ (bottom right). All histograms are normalized to have unity area. We recall that our selection, in case of multiple D^0 candidates, retains the candidate whose reconstructed mass, based on simple four-vector addition, is closer to the PDG2002 nominal mass [5].

4.11 D^* Reconstruction

D^* candidates are reconstructed by combining the D^0 candidate with a charged pion (known as the soft pion) that has transverse laboratory momentum, p_T, greater than 50 MeV/c and charge opposite to the charge of the kaon candidate. The cut on p_T is necessary to mantain tracking efficiency at an acceptable level since, owing

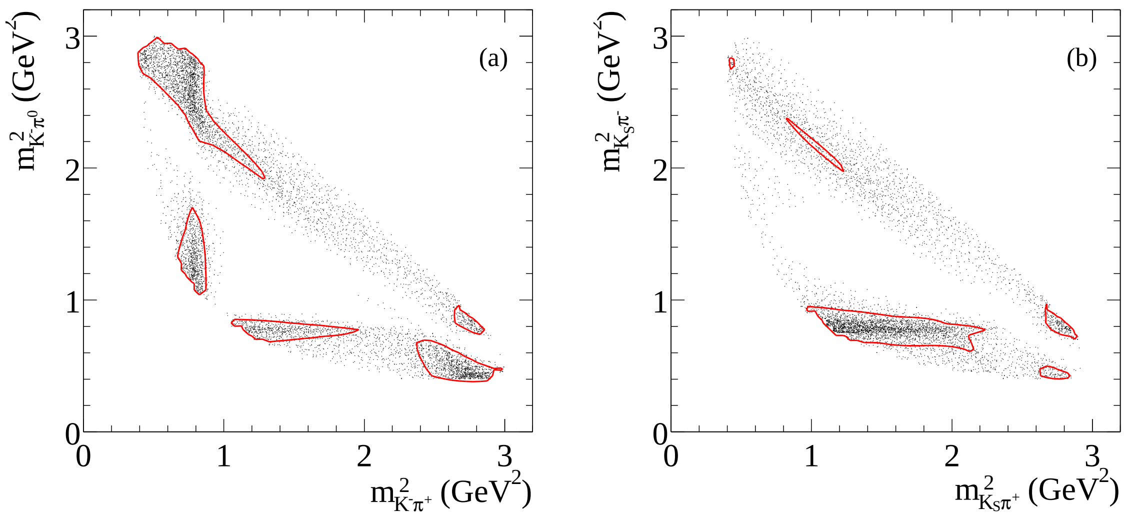


Figure 4.14: (a) Dalitz plot calculated with the amplitudes and phases measured by [56] for $D^0 \to K\pi\pi^0$. Red contours represent the regions where the density is higher than 10 % of the maximum value. (b)Same plot as (a) for $D^0 \to K_S\pi\pi$.

to uncertainties in the description of multiple scattering in the Monte Carlo, the lower the p_T of the soft pion the worse the tracking efficiency is. Another cut on the laboratory soft pion momentum, $p_\pi < 450$ MeV/c, is driven by kinematic constraints. For the $D^0 \to K_S\pi\pi$ mode, the charge of the soft pion is used to identify the flavor of D^0 candidates, so that the Dalitz density can be calculated correctly (see Section 4.10.).

D^* candidates from a B decay are required to have momentum, $p^*_{D^*}$, less than 2.5 GeV/c in the $\Upsilon(4S)$ rest frame. As shown in figure 4.15, this cut preserves all signal events but removes a significant amount of D^* from continuum events. We cut also on $p^*_{D^*} > 0.5$ GeV/c following standard *BABAR* D^* reconstruction cuts.

The spectrum that is used to determine whether a D^* is reconstructed correctly is the difference between the candidate $D^0\pi_{soft}$ mass with the candidate D^0 mass constrained to the true D^0 mass, and the true D^0 mass ($m_{D^0\pi_{soft}} - m_{D^0}$), which has to be smaller than 170 MeV at the first step of the selection. Subsequently, the soft pion is refitted with D^*-ℓ vertex and beamspot constraint (see section 4.13.1 for the details) and when the fit has converged we use the refitted $\delta m = m_{D^*} - m_{D^0}$ variable to select final candidates. The δm spectrum is peaked around 145.5 MeV with a width less than 1 MeV, depending on the soft pion reconstruction status as we will see in the next chapter.

We accept events in the window 143 MeV $< \delta m <$ 148 MeV window when the soft pion has hits only in the SVT and 144 MeV $< \delta m <$ 147 MeV when it has hits also in the DCH owing to the better δm resolution when the the soft pion is reconstructed also in the DCH.

Figure 4.16 shows the δm distribution for both signal and control samples events in data. All events with $\delta m < 165$ MeV will be used in the fit on control samples to extract the combinatoric (fake D^*) and peaking backgrounds (real D^*). In figure 4.17 we show the δm distribution in the accepted δm windows mentioned above; all the candidates are selected in the two signal samples as defined in table 4.2.

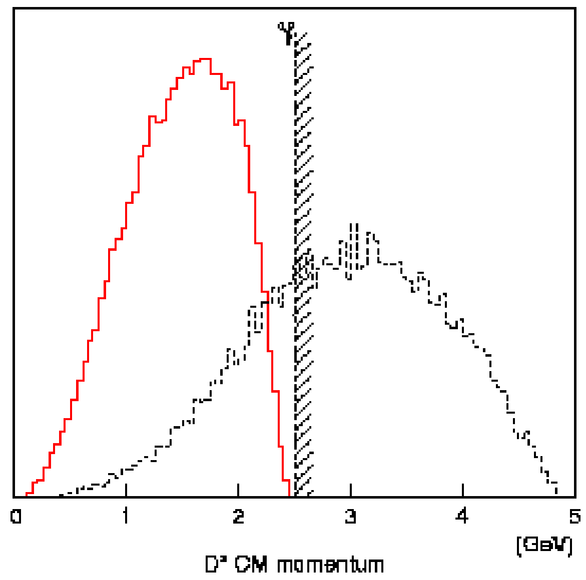


Figure 4.15: D^* meson momentum distribution in the center of mass frame in signal events (red continuum line) and in continuum background events (black dotted line) obtained from MC simulations.

4.12 Reconstruction of the $D^*\ell$ candidate

In order to reconstruct the $D^*\ell$ candidate a lepton cut $p^*_\ell > 1.2$ GeV/c in the $\Upsilon(4S)$ rest frame is applied. This cut is helpful in removing almost all the correlated background as can be seen from figure 4.18.

We already mentioned that in the final selection stage we calculate δm with the quantities evaluated after the refit of soft pion, D^0 and ℓ vertex (see 4.13.1). We require the fit χ^2 probability to be greater than 1% in the selection. In case of multiple D^0 candidates, we select the candidate with reconstructed mass, based on

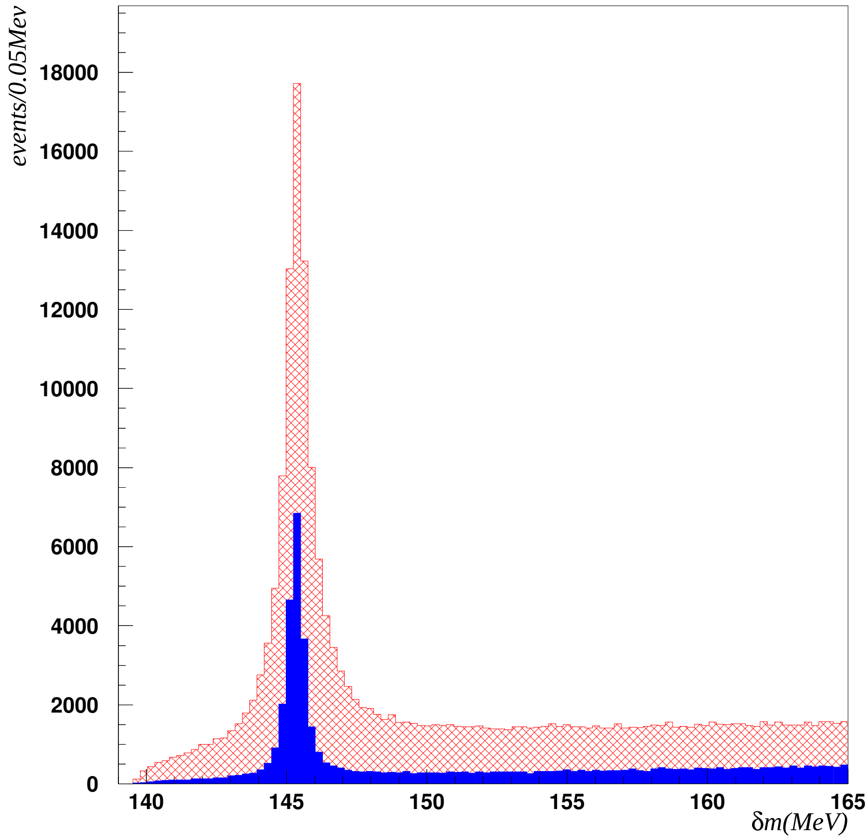


Figure 4.16: δm distribution of all selected candidates in data (both signal and control samples) together, after the refit described in section 4.13.1. The blue histogram shows the events where the soft pion candidate has both SVT and DCH hits. The red hatched histogram shows all the events selected. We show also the combinatoric region adopted for background studies which span $138\ \mathrm{MeV} < \delta m < 165\ \mathrm{MeV}$.

simple four-vector addition, closest to the PDG2002 mass [5], in each of the four possible reconstructed D^0 modes. If there are multiple $D^*\ell$ candidates with the same D^0 candidate, only the first candidate in the list is kept.

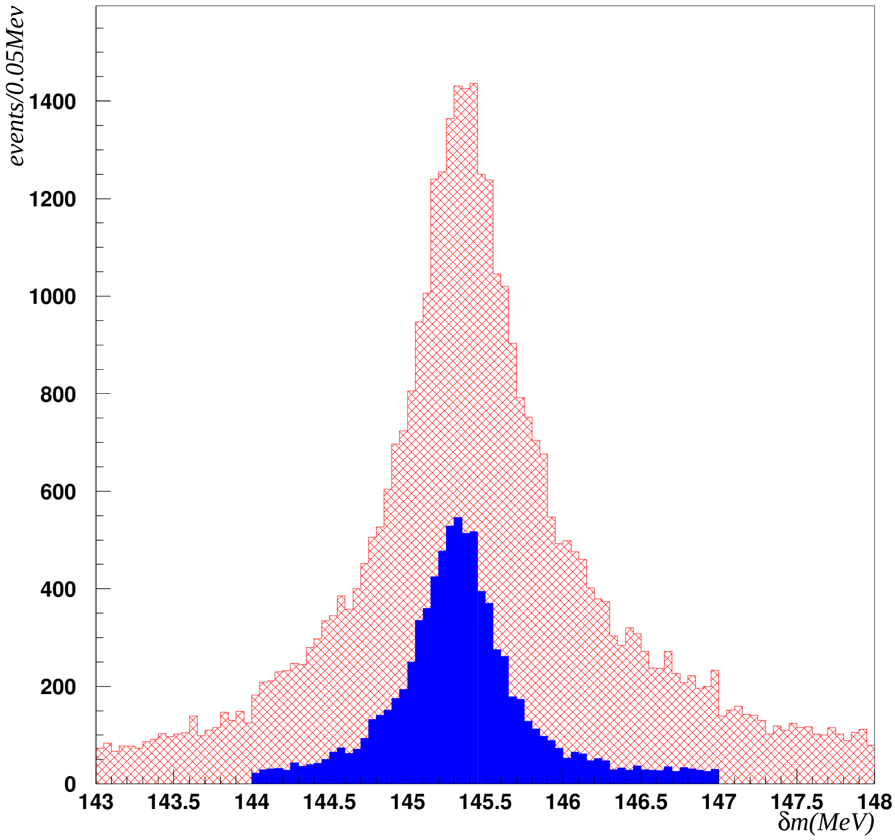


Figure 4.17: δm distribution of all selected (i.e. within the cuts in δm) candidates in data signal sample together, after the refit described in section 4.13.1. The blue histogram shows the events where the soft pion candidate has both SVT and DCH hits. The red hatched histogram shows all the events selected inside 143 MeV $< \delta m <$ 148 MeV.

4.13 Background discriminating variables

4.13.1 δm and $D^*\ell$ vertex (re)fitting

We use the δm variable to study and characterize most of the backgrounds. From figure 4.16 we see that there is a peak structure (events with a real D^* on top of an almost flat distribution (a misreconstructed D^*) which spans all the range from 138 MeV to 165 MeV. We characterize the backgrounds into two main categories:

1. peaking background which means a background where a real D^* is reconstructed. It is parametrized with two Gaussian distributions and

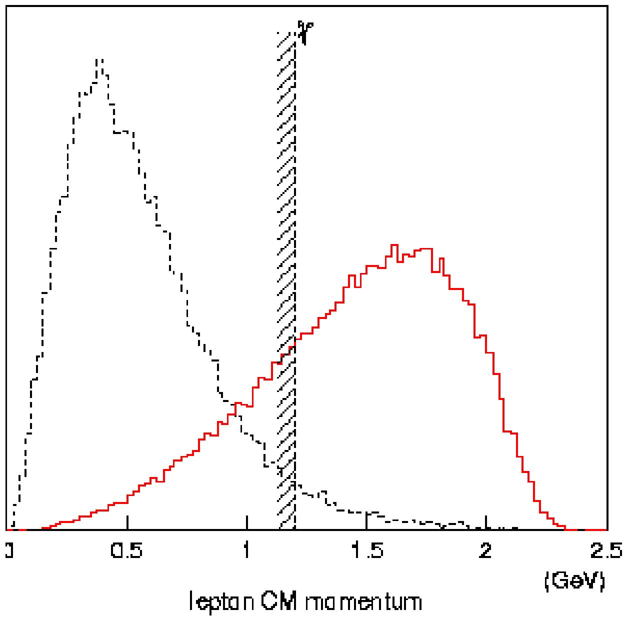


Figure 4.18: Lepton momentum distribution in the center of mass frame in signal events (red continuum line) and in correlated background events (black dotted line) obtained from MC simulations.

2. combinatoric background where the reconstructed D^* is not a real D^*. It is parametrized as a nearly flat background extending under and beyond the δm peak (see figure 4.16). Sidebands outside the δm peak are populated only by combinatoric background.

This analysis uses a constrained vertex fitting technique[8] that treats the ℓ, beamspot, D^* and its daughters simultaneously to find the best vertex. The algorithm explicitly incorporates a flight distance of the daughter D^0, correctly considers the correlated errors in the fit and provides an improved estimate of δm and the $D^*\ell$ vertex of the B. The basic algorithm was implemented in [59] and is available in the `VtxB0_DstarlnuAlgorithm` class of the *BABAR* VtxFitter package.

The implementation is based on a fitting algorithm described in [60]. The algorithm and its performance are described further in [61, 62]. The experimental fact that the refitting of the soft pion improves dramatically the δm resolution is a consequence of the correlation between the D^0 decay length, the D^0 momentum and the soft pion momentum. The list of parameters, unknowns and constraints is the following:

Parameters:

[8]see Appendix B for technical details.

- charged D^0 decay daughters
- slow pion
- lepton

Unknowns:

- D^0 decay vertex ($\vec{v}_{D^0}$)
- D^* decay vertex ($\vec{v}_{D^*}$)
- D^0 lifetime (τ_{D^0})

Constraints:

- xy vertex constraint at $\vec{v}_{D^0}$ for D^0 daughters
- z vertex constraint at $\vec{v}_{D^0}$ for D^0 daughters
- xy vertex constraint at $\vec{v}_{D^*}$ for D^0, π_{soft}, ℓ
- z vertex constraint at $\vec{v}_{D^*}$ for D^0, π_{soft}, ℓ
- $\vec{v}_{D^0} - \vec{v}_{D^*} = \frac{c\tau_{D0}}{m_{D0} \cdot \sum_i^{D^0 daughters} \vec{p}_i}$
- beamspot is applied as constraint on $\vec{v}_{D^*}$ and the lepton candidate is also considered in the vertex fit. With this χ^2 constraint the vertex position in the transverse plane (x, y) is constrained to be compatible with the beam-spot position
$$(x - x_{BS})^2 + (y - y_{BS})^2 = 0$$
As this is a χ^2 constraint, (x_{BS}, y_{BS}) are introduced as additional parameters in the fit, so their covariance matrix is used to account for the actual size of the beam.

The output of the algorithm is:

- improved versions of the inputs
- the values of the unknowns
- all correlations

There are two other ways of performing the vertex which are allowed by this fit method:

1. a D^* and ℓ refit without beamspot constraint;
2. a D^* with beamspot constraint refit.

Both are used to study the sytematic effects induced by the beamspot constraint and of the refit with the lepton. The vertexing method which uses the D^*, ℓ and imposes the beamspot constraint is the most efficient and the one with the smallest δm resolution of the gaussian peaking yield. After applying the refit δm resolution improves by about 15% [54].

4.13.2 Angular Variables

We define two angular quantities for each candidate to classify them into opposite-side and same-side samples. We calculate the following observables:

- the cosine of the angle between the D^* and the lepton, $\cos\theta_{D^*,\ell}$, calculated in the $\Upsilon(4S)$ rest frame,
- the cosine of the angle between the direction of the B^0 and the sum of the D^* and ℓ momenta, $\cos\theta_{B^0,D^*\ell}$, calculated in the $\Upsilon(4S)$ rest frame. We do not measure the angle directly since we do not have a handle on the B^0 direction. Instead, we use the following equation, coming from energy-momentum conservation relation applied to the $B^0 \to D^{*-}\ell^+\nu_\ell$ decay, assuming the only particle missed in our reconstruction is a massless neutrino,

$$\cos\theta_{B^0,D^*\ell} = \frac{-(m_{B^0}^2 + m_{D^*\ell}^2 - 2\,E_{B^0}^* E_{D^*\ell}^*)}{2\,|\vec{p}_{B^0}^{\,*}||\vec{p}_{D^*\ell}^{\,*}|}\,.$$

  The only information needed from B^0 is the amplitude of its C.M. momentum, which can be obtained from center-of-mass energy of $\Upsilon(4S)$ and B^0 mass. This quantity is sensitive to the presence of other hadronic states in the semileptonic decay of the B^0.

- the cosine of the same angle described in the previous bullet, calculated with the lepton momentum flipped ($\vec{p}_\ell^{\,*} \to -\vec{p}_\ell^{\,*}$), $\cos\theta_{B^0,D^*(-\ell)}$.

The $D^*\ell$ candidates are divided into two groups, so-called opposite-side ("OS") and same-side ("SS") candidates, based on these variables.

1. We define the opposite-side sample by requiring that $D^*\ell$ candidates have $\cos\theta_{D^*,\ell} < 0$ and $|\cos\theta_{B^0,D^*\ell}| < 10$[9]. These opposite-side candidates span the electron, muon, and fake-lepton control samples.

2. Same-side candidates are selected by requiring $\cos\theta_{D^*,\ell} \geq 0$. These events are the control sample for uncorrelated $D^*\ell$ background.

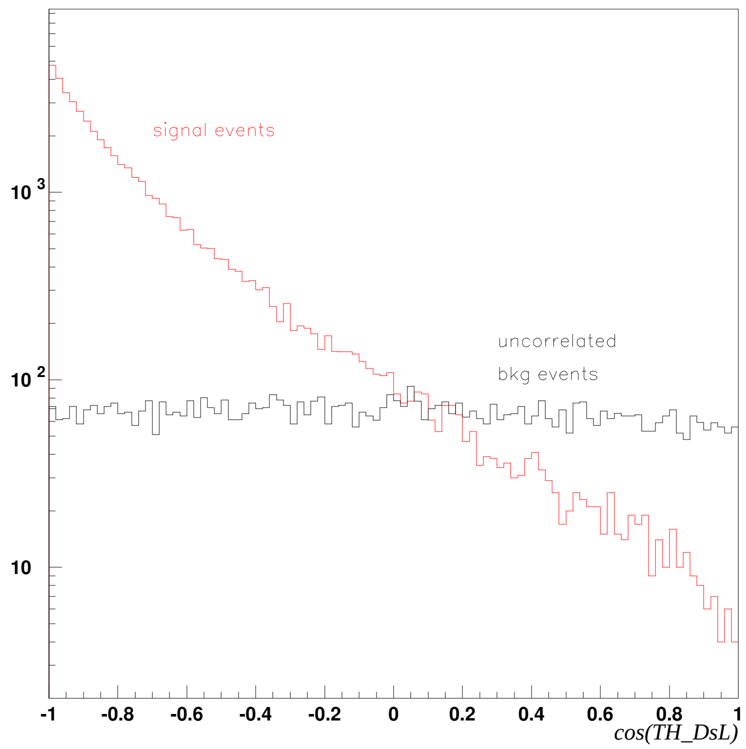


Figure 4.19: Monte Carlo distribution of the $\cos\theta_{D^*,\ell}$ variable in signal events (red peaked histogram) and in uncorrelated background events (black flat histogram).

The angular regions defining these control samples are shown in Figure 4.19 : we can see that the uncorrelated background is symmetric about $\cos\theta_{D^*,\ell} = 0$.

For these events we also compute the flipped version of $\cos\theta_{B^0,D^*\ell}$, i.e. $\cos\theta_{B^0,D^*(-\ell)}$. We use the distribution of flipped variable $\cos\theta_{B^0,D^*(-\ell)}$ to simulate events in the uncorrelated-lepton sample since, from Monte Carlo studies, the distribution of this flipped variable in the control sample is roughly similar to the distribution of the unflipped variable $\cos\theta_{B^0,D^*\ell}$ for uncorrelated background in the signal sample, whereas

[9] we select $|\cos\theta_{B^0,D^*\ell}| < 10$ since we want to include also the long $\cos\theta_{B^0,D^*\ell}$ tails characteristic of the $B \to D^*\ell\nu X$ background.

the distribution of $\cos\theta_{B^0,D^*\ell}$ in the control sample is systematically different since it is shifted to higher $\cos\theta_{B^0,D^*\ell}$ values.

We do not cut on the expected physical signal region $|\cos\theta_{B^0,D^*(-\ell)}| < 1.0$, we use the distributions to fit backgrounds to the larger range $|\cos\theta_{B^0,D^*(-\ell)}| < 10$.

In other words, we use the flipped variable to select a control sample that is enriched in the background we would like to study. To estimate the amount of δm uncorrelated peaking background we use the SS control samples and the relative efficiencies for a true signal or a true uncorrelated event to pass the OS or SS selection criteria from generic Monte Carlo. Then we extract the uncorrelated background shape in $\cos\theta_{B^0,D^*\ell}$ from Monte Carlo. We will show in the next chapter how to account for the correct amount and shape of all backgrounds in the $\cos\theta_{B^0,D^*\ell}$ distribution.

Considerations about $\cos\theta_{B^0,D^*\ell}$

As we have said before, several classes of backgrounds have peaks in the δm distribution, same as the signal events. There are reliable ways to simulate their distributions since most of these backgrounds are well understood. The exception is the background from $B \to D^*\ell\nu X$ events in which we reconstruct the D^* and the ℓ but we do not observe a final state particle X. This background includes both non-resonant $B \to D^*\ell\nu\pi$ and resonant $B \to D^{**}\ell\nu$ decays, in which some higher excited D resonance (generally indicated here as D^{**}) decays to a D^* plus another particle (usually a pion). Many of the modes that may contribute to this background have not been measured so far and therefore Monte Carlo simulation is not sufficiently precise to evaluate such background. Furthermore, our knowledge of the form factors that govern these decays is minimal. A detailed description of D^{**} resonaces is given in section 4.13.3.

To handle this background we use the $\cos\theta_{B^0,D^*\ell}$ variable which has very different distributions for signal events and for $B \to D^*\ell\nu X$ background. As $\cos\theta_{B^0,D^*\ell}$ is a physical quantity, it should lie between -1 and 1. The reconstructed value of $\cos\theta_{B^0,D^*\ell}$ falls outside this physical region only in a small fraction of signal events mainly due to the final state radiation of the electron which pulls the $\cos\theta_{B^0,D^*\ell}$ to negative values (less than -1). The situation is different for $B \to D^*\ell\nu X$ events. Here, when evaluating $\cos\theta_{B^0,D^*\ell}$ we neglect the particle X. To correct for this we should replace the $\cos\theta_{B^0,D^*\ell}$ definition above with:

$$\cos\theta_{B^0,D^*\ell} = \frac{-(m_{B^0}^2 + m_{D^*\ell}^2 - 2\,E_{B^0}E_{D^*\ell}) + m_{X\nu}^2}{2\,|\vec{p}_{B^0}||\vec{p}_{D^*\ell}|}\,.$$

where $m_{X\nu}$ is the invariant mass of the $X-\nu$ system, a quantity that is always positive. If we neglect this term when evaluating $\cos\theta_{B^0,D^*\ell}$ for such a background event, the result will be less than the true value and it can fall well below the physical region giving rise to long tails in $\cos\theta_{B^0,D^*\ell}$ distribution. No $B\to D^*\ell\nu X$ events fall in the $\cos\theta_{B^0,D^*\ell}>1$ region but events of several other backgrounds do.

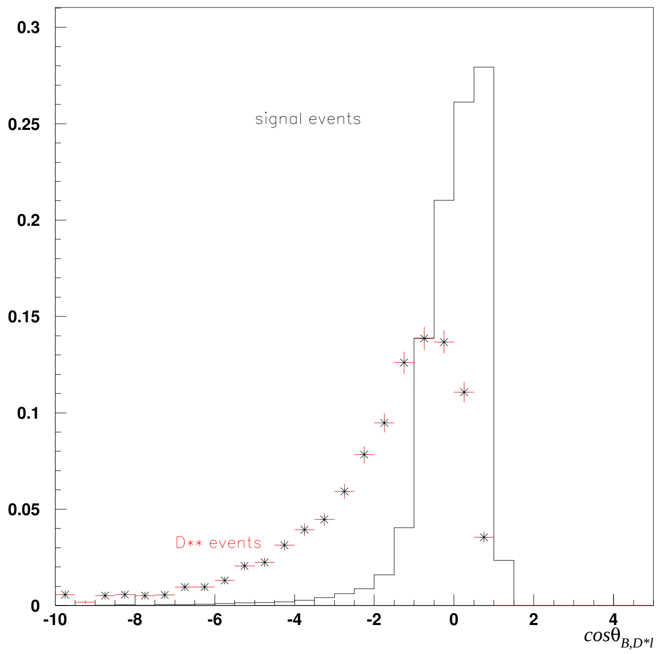


Figure 4.20: Comparison among thedistributions of $\cos\theta_{B^0,D^*\ell}$ for signal and $B\to D^*e\nu_eX$ background obtained from Monte Carlo simulations. We can see the radiative tail due to final state electron radiation in the signal distribution. Both distributions are normalized to unity.

In Figure 4.20 and 4.21 we show plots summarizing what we said above, separately for the electron and for the muon.

In addition to the control samples described in section 4.3, which are used to extract the fractions of combinatoric and most of peaking $\delta m = m_{D^*} - m_{D^0}$ backgrounds, we use Monte Carlo information to extract the distributions of all these backgrounds in the $\cos\theta_{B^0,D^*\ell}$ variable. We use the $\cos\theta_{B^0,D^*\ell}$ distribution to fit for the D^{**} and non resonant $D^*\pi\ell\nu_\ell$ composition in order to extract the

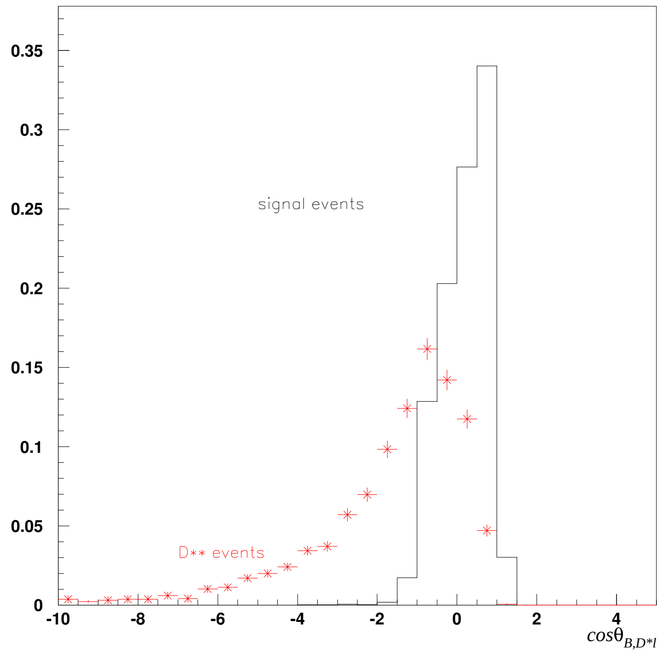


Figure 4.21: Comparison among the distributions of $\cos\theta_{B^0,D^*\ell}$ for signal and $B \to D^*\mu\nu X$ background obtained from Monte Carlo simulations. We can see in this case that there is almost no radiative radiative tail due to final state muon radiation in the signal. Both distributions are normalized to unity.

signal yield and measure $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$. We perform the fit in each of the 4(`Dmode`)×2(`leptID`)=8 different signal samples. Allowing the level of $B \to D^*\ell\nu X$ background to float in the fit we can extract the signal yield without knowledge of the various $B \to D^*\ell\nu X$ branching fractions. Furthermore, this fit allows us to measure $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ without a detailed knowledge of either the $B \to D^*\ell\nu X$ form factors or the relative contributions of different modes.

In table 4.8 we show the theoretical model adopted to generate the $B \to D^*\ell\nu X$ decays and the relative Branching Ratios adopted in the *BABAR* Monte Carlo.

4.13.3 Remarks on $B \to D^{*+}X\ell^-\overline{\nu}_\ell$ background

The *BABAR* Monte Carlo includes six D^{**} states. Four are orbitally excited states, with orbital angular momentum $L = 1$. The D_1 is a singlet $^1P_1$ state, while the D_0^*, D_1^* (also known as D_1' in the literature and adopted by *BABAR* `decay.dec` notation too) and D_2^* are three triplet $^3P_0$, $^3P_1$ and $^3P_2$ states. In HQET, the properties of

B decay mode	BR $\times 10^{-2}$	Model	D^{**} decay	BR	Overall BR $\times 10^{-4}$
$\bar{B}^0 \to D^{*+}\pi^0 l^- \bar{\nu}$	0.10	GR	no	no	10.
$\bar{B}^0 \to D_1^+ l^- \bar{\nu}$	0.56	ISGW2	$D_1^+ \to D^{*+}\pi^0$	0.33	18.48
$\bar{B}^0 \to D_1^{*+} l^- \bar{\nu}$	0.37	ISGW2	$D_1^{*+} \to D^{*+}\pi^0$	0.33	12.21
$\bar{B}^0 \to D_2^{*+} l^- \bar{\nu}$	0.37	ISGW2	$D_2^{*+} \to D^{*+}\pi^0$	0.103	3.81
$\bar{B}^0 \to D'^{+} l^- \bar{\nu}$	0.02	ISGW2	$D'^{+} \to D^{*+}\pi^0$	0.33	0.67
$\bar{B}^0 \to D^{*\prime +} l^- \bar{\nu}$	0.22	ISGW2	$D^{*\prime +} \to D^{*+}\pi^0$	0.17	3.74
$B^- \to D^{*+}\pi^- l^- \bar{\nu}$	0.20	GR	no	no	20.
$B^- \to D_1^0 l^- \bar{\nu}$	0.56	ISGW2	$D_1^0 \to D^{*+}\pi^-$	0.67	37.52
$B^- \to D_1^{*0} l^- \bar{\nu}$	0.37	ISGW2	$D_1^{*0} \to D^{*+}\pi^-$	0.67	24.79
$B^- \to D_2^{*0} l^- \bar{\nu}$	0.37	ISGW2	$D_2^{*0} \to D^{*+}\pi^-$	0.21	7.78
$B^- \to D'^{0} l^- \bar{\nu}$	0.02	ISGW2	$D'^{0} \to D^{*+}\pi^-$	0.67	1.32
$B^- \to D^{*\prime 0} l^- \bar{\nu}$	0.22	ISGW2	$D^{*\prime 0} \to D^{*+}\pi^-$	0.33	7.26

Table 4.8: *We show the decay modes used in our Monte Carlo simulation of the $B \to D^* X l\nu$. These values are based mainly on theoretical estimates. The theoretical models adopted to generate these events are also shown. For the non resonant $B \to D^*\ell\nu\pi$ we used the Goity and Roberts model (GR) while we adopted the Isgur and Wise model (ISGW2) for the $B \to D^{**}\ell\nu$ decays.*

these states should be dominated by the light-quark degrees of freedom, with small corrections from the coupling to the heavy-quark spin. The orbitally-excited D^{**} states can be organized into two doublets characterized by the quantum number j, the sum of the orbital momentum L with the light-quark spin s. The D_1 and D_2^* form a $j = 3/2$ doublet, while D_0^* and $D_1^{'}$ fall into a $j = 1/2$ doublet. The two $j = 3/2$ states can only undergo D wave decays, and are therefore fairly narrow. Both states have been observed with the expected quantum numbers [5]. As predicted by HQET, the mass difference is small, about 35 MeV, and the states are fairly narrow, with widths of the order of 20-30 MeV. In contrast, the $j = 1/2$ doublet states can decay via S wave, and are therefore expected to be very broad, with widths of the order of hundreds of MeV. The $j = 1/2$ states are therefore very hard to reconstruct. The only experimental evidence for these states is a preliminary CLEO observation of the $D_1^{\prime 0}$ [50].

The final two D^{**} states in the Monte Carlo are the two radially-excited $2S$ states. These states are a ${}^1S_1$, the D' (also known in literature as $D(2S)$), and a ${}^3S_1$ state, the $D^{*'}$ (also known in literature as $D^*(2S)$). These states are analogous to the D and the D^*. The DELPHI collaboration has reported evidence for a state they tentatively identify with the $D^{*'}$, although the quantum numbers of the observed state have not been confirmed. The DELPHI state is narrow, in accordance with some theoretical calculations [51]. The *BABAR* Monte Carlo assumes that the D' and $D^{*'}$ are very wide, a more common theoretical expectation. The branching fractions in table 4.8 are in certain case educated guesses from the several theoretical prediction available on the market. Little information exists about the total $B \to D^{(*)}X\ell^-\overline{\nu}_\ell$ branching fraction and the relative contributions of different modes, where by $D^{(*)}X$ we mean DX and $D^*\ X$ non resonant states as well as D^{**} resonances. There is an ALEPH measurement of $(2.26 \pm 0.29 \pm 0.33) \times 10^{-2}$, where the errors are statistical and systematic. The Monte Carlo $B \to D_1\ell^-\ \overline{\nu}_\ell$ branching fractions used come from a CLEO measurement [52][10]. There is little experimental evidence on the P wave D^{**} branching fractions, but these decay states are solid on theoretical footing. The dominant decay mode for these states is a pion emission to a $D^{(*)}\pi$ final state. By angular momentum conservation, the D wave D_1 decay and the S wave $D_1^{'}$ decay must be to a $D^*\ \pi$, and cannot be to a $D\pi$. Conversely, the S wave D_0^* decay can only be to a $D\pi$, not to $D^*\ \pi$. $B \to D_0^*\ell^-\ \overline{\nu}_\ell$ decays are therefore not a background to $B^0 \to D^{*-}\ell^+\nu_\ell$. Both D wave $D_2^* \to D\pi$ and $D_2^* \to D^*\pi$ transitions are allowed by angular momentum conservation, and both have been observed. The branching ratio of these two processes in the Monte Carlo is based on a CLEO measurement [53]. Otherwise, the P wave D^{**} branching fractions in `decay.dec` are based on isospin considerations. One expects, for example:

$$\frac{\mathcal{B}(D^{**+} \to D^{*+}\pi^0)}{\mathcal{B}(D^{**+} \to D^{*0}\pi^+)} = \frac{1}{2}. \tag{4.1}$$

These isospin considerations are the reason the $D^{*+}\ \pi^-\ \ell^-\ \overline{\nu}_\ell$ states typical of the $D^{*+}\ X\ell^-\ \overline{\nu}_\ell$ background from B^- decay occur twice as often as the $D^{*+}\ \pi^0\ \ell^-\ \overline{\nu}_\ell$ from $\overline{B}^0$ decay.

The radially-excited D' and $D^{*'}$ decays are handled in a similar manner. Due to parity and angular momentum conservation, the D' decays vis P wave to $D^*\ \pi$, while $D^{*'}$ decays via P wave to either $D\pi$ or $D^*\ \pi$. The *BABAR* Monte Carlo also includes $D\gamma$, $D^*\ \gamma$ and $D_1^{'}\ \pi$ final states at rate under 1%. It is clear that poor experimental evidence implies a lack of information about both the relative contribution of the different modes. From the theoretical side there is almost null knowledge of form factors in each mode giving an uncertainty in the shape distribution.

[10]CLEO quotes $\mathcal{B}(B^- \to D_1^0\ \ell^-\ \overline{\nu}_\ell)$=$(5.6 \pm 1.3 \pm 0.9) \times 10^{-3}$.

Chapter 5

Background subtraction

In this chapter we show the method used to perform background subtraction.

We evaluate separately the contributions of the backgrounds listed in section 4.3. We first evaluate -and subtract- the combinatoric background (i.e. a misreconstructed D^*), for each of the signal samples. After subtracting the combinatoric background we are left with $B^0 \to D^{*-}\ell^+\nu_\ell$ signal and $\delta m = m_{D^*} - m_{D^0}$ peaking backgrounds. The δm peaking backgrounds are composed of all the events with a D^* correctly reconstructed. We then subtract the fake-lepton background (i.e. all the events with a real D^* and a misidentified lepton) for each of the signal samples. After performing combinatoric and fake-lepton subtraction we subtract the continuum background (i.e. events with a real D^* and a real ℓ coming from an $e^+e^- \to c\bar{c} \to D^*\ell X$ decay and not from a $e^+e^- \to \Upsilon(4S)$ decay) for each of the signal samples. After these three steps of background subtraction we are left only with real $D^*\ell$ events coming from an $\Upsilon(4S) \to B\overline{B}$ decay. We then subtract the uncorrelated background (i.e. events where the D^* and the ℓ come from different B mesons) in each of the signal samples. After subtracting combinatoric, fake-lepton, continuum and uncorrelated backgrounds we then evaluate -and subtract- the correlated background (i.e. events with a real $D^*\ell$ candidate with both particles coming from the same B but with the ℓ that is not a direct product of the B decay; $B \to D^{*-}X\ ,\ X \to \ell^+Y$).

After performing the subtraction of all the backgrounds listed above we end up with the sum of $B^0 \to D^{*-}\ell^+\nu_\ell$ signal and the poorly-understood $B \to D^{*-}\ell^+\nu_\ell X$ background. We show in the next chapter how we separate the signal events from this background.

The method adopted to perform the subtraction of combinatoric, fake-lepton, continuum and uncorrelated backgrounds is based on a fit to the $D^* - D^0$ mass difference δm and has already been used in a previous *BABAR* analysis [74].

For the purpose of this study we merge together the $D^0 \to K_S\pi\pi$ and $D^0 \to K\pi\pi\pi$ `Dmode` since the combinatoric backgrounds in these two modes are similar [54] and also due to the fact that we have few $D^0 \to K_S\pi\pi$ events.

5.1 Combinatoric background

The tool adopted for extracting the fraction of combinatoric background that contributes to our signal sample is the fit of the distribution of the $D^* - D^0$ mass difference δm. We compute also some[1] of the fractions of the peaking backgrounds with this technique. In order to fit for a δm distribution we need to find a parametrization of the shape of the peak due to candidates with a correctly reconstructed D^* and also the shape of the combinatoric background due to candidates with an incorrectly reconstructed D^*. The output of the fit allows to extract the yields of peak and combinatoric components in each signal and control sample adopted for this analysis.

The general program for using these fits is to first determine the yields in different subsamples of the signal and control samples, and then to use these yields to calculate the fractions of different sources of "peaking background"; i.e., backgrounds containing a real D^*.

5.1.1 Modeling of Peak and Combinatoric Shapes

We model the shape of the peak in a δm distribution with a double Gaussian, which has five free parameters (since $f_1 + f_2 = 1$ because they represent the amount of peak events in the first and the second Gaussian, respectively):

$$F_{peak}(\delta m) = \sum_{k=1}^{2} \frac{f_k}{\sqrt{2\pi}\sigma_k} \exp\left(-\frac{(\delta m - \delta m_k)^2}{2\sigma_k^2}\right) .$$

We show in figure 5.7 the values of these 5 paramters in the signal samples used for the δm fit. We model the shape of the combinatoric background (i.e. events with a misreconstructed D^*) with a function

$$F_{comb}(\delta m) = \frac{1}{N}\left[1 - \exp\left(-\frac{\delta m - \delta m_0}{c_1}\right)\right]\left(\frac{\delta m}{\delta m_0}\right)^{c_2} ,$$

[1] fake-lepton, continuum and uncorrelated bkg.

where N is a normalization constant and δm_0 is the kinematic threshold equal to m_π. The values of c_1 and c_2 are shown in figures 5.1–5.6. The first step of the process of extracting yields consists in the determination of the shape parameters and can be done combining these models with one additional parameter f_{peak}:

$$F(\delta m) = f_{peak} \cdot F_{peak}(\delta m) + (1 - f_{peak}) \cdot F_{comb}(\delta m) \ .$$

The value of f_{peak} then corresponds to the fraction of peak events in the entire sample being fit (including the sideband).

The crucial point of all this method is that it is not a good idea to fit the entire sample at once to a single set of shape parameters since different subsamples have significantly different shapes (see figures 5.1–5.6). We have to account for these differences to improve our statistical sensitivity, and more important, to reduce possible systematic biases. At the other extreme, we could independently fit each of the 72 subsamples we obtain with the full breakdown by the attributes listed in table 4.1. We recall we end up with 72 categories since we merged together the $D^0 \to K_S\pi\pi$ and $D^0 \to K\pi\pi\pi$ modes.

This approach provides the maximum statistical sensitivity, in principle, but suffers from problems with fitting small samples.

The compromise we have chosen is to divide the sample into six groups that each share a common set of 5 peak shape parameters, `peakGroup` (6) = `svtDch` (2) $\otimes$ `Dmode` (3). We select this breakdown through a process of trial and error to determine which subdivisions of the sample result in statistically significant differences in δm distribution parameters. In particular, we find that the peak shape does not depend on whether an event is recorded on or off the $\Upsilon(4S)$ peak (`onOffRes`), whether a candidate passes or fails the angular cuts (`angCut`), or whether the lepton candidate passes or fails lepton ID requirements (`leptID`). But the shape depends, of course, on the D^0 decay channel and on wether the soft pion track has been measured by SVT alone or also by the DCH. Using the same peak shape (within a peak group) for the signal and its control samples simplifies the characterization of peaking backgrounds later.

Within each of the 6 peak groups, we further subdivide the sample into 4 subgroups that each share a common set of 2 combinatoric background shape parameters (c_1 and c_2). We again use the shape variations observed in data to determine an appropriate breakdown (see table 5.1.) We divide the large signal and fake control sample and do not distinguish between e and μ in the signal. We assign the smaller same-side (SS) and off-resonance control samples into their own groups with no further subdivision. Off-resonance same-side candidates are grouped with the smaller off-resonance group.

BG Group	`onOffRes`	`angCut`	`leptID`
Lept	On	OS	e,μ
Fake	On	OS	fake
SS	On	SS	*any*
OffRes	Off	*any*	*any*

Table 5.1: *Breakdown of each peak group (which shares a common set of the 5 peak shape parameters) into separate background groups that each share a common set of the combinatoric background shape parameters c_1 and c_2.*

In order to determine the 5 peak + 4 $\otimes$ 2 BG shape parameters for each peak group (in total, 78 parameters), we perform a simultaneous fit of each group using a likelihood

$$\mathcal{L} \equiv \prod_{k=1}^{4} \prod_{\text{evts } j} \left\{ f_{peak}^{(k)} \cdot F_{peak}(\delta m_j;\, m_1, \sigma_1, m_2, \sigma_2, f_1) + (1 - f_{peak}^{(k)}) \cdot F_{comb}(\delta m_j;\, c_1^{(k)}, c_2^{(k)}) \right\} \,,$$

where each candidate's contribution is calculated using the Probability Density Function (PDF) in this product corresponding to its background group k. With this approach we introduce 4 additional relative normalization parameters $f_{peak}^{(k)}$ for a total of 5 (peak) + 8 (BG) + 4 (fractions) = 17 parameters in each of the 6 simultaneous fits.

Figures 5.1–5.6 show the results of the simultaneous fits to each of the 6 peak groups. Note that these plots only list values for $c_1^{(k)}$, $c_2^{(k)}$, and $f_{peak}^{(k)}$ even though a common set of peak shape parameters was also allowed to vary in each fit (see Figure 5.7 for the values of the peak shape parameters).

In order to test our assumption of a common peak shape for the 6 samples shown in Figures 5.1–5.6, we compare the combined samples from each of these figures with the sum of the corresponding 4 models. Figure 5.7 shows the resulting plots over the narrow range of 143–149 MeV.

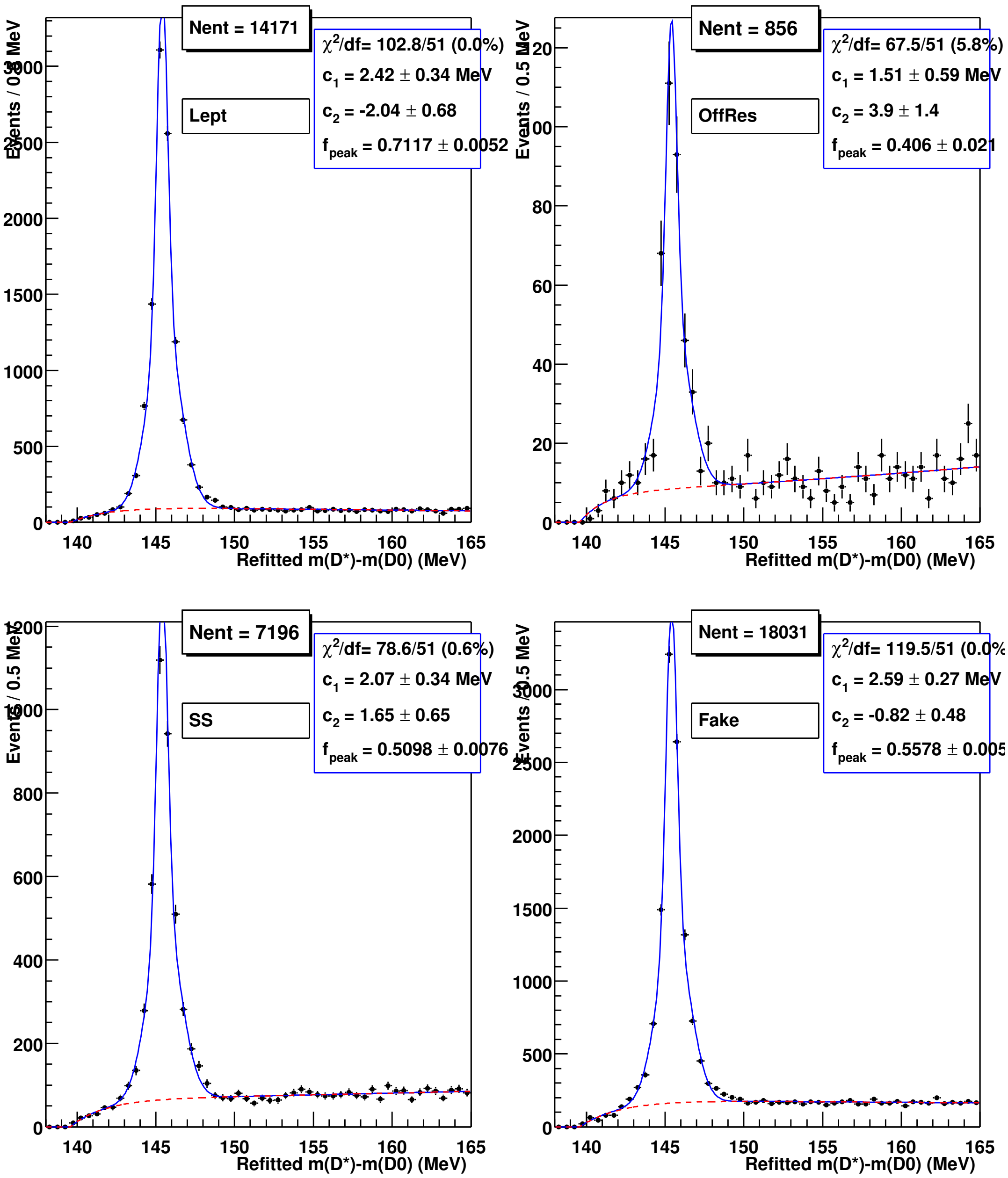


Figure 5.1: Results of the simultaneous fit described in the text to the sample consisting of candidates reconstructed as $D^0 \to K\pi$ and their soft pion reconstructed in the SVT only. The dashed curves show the fitted contributions of combinatoric background. The δm peak contains various backgrounds. For example the δm peak of the Off-Resonance sample contains mainly continuum background but also a small fraction of fake-lepton background.

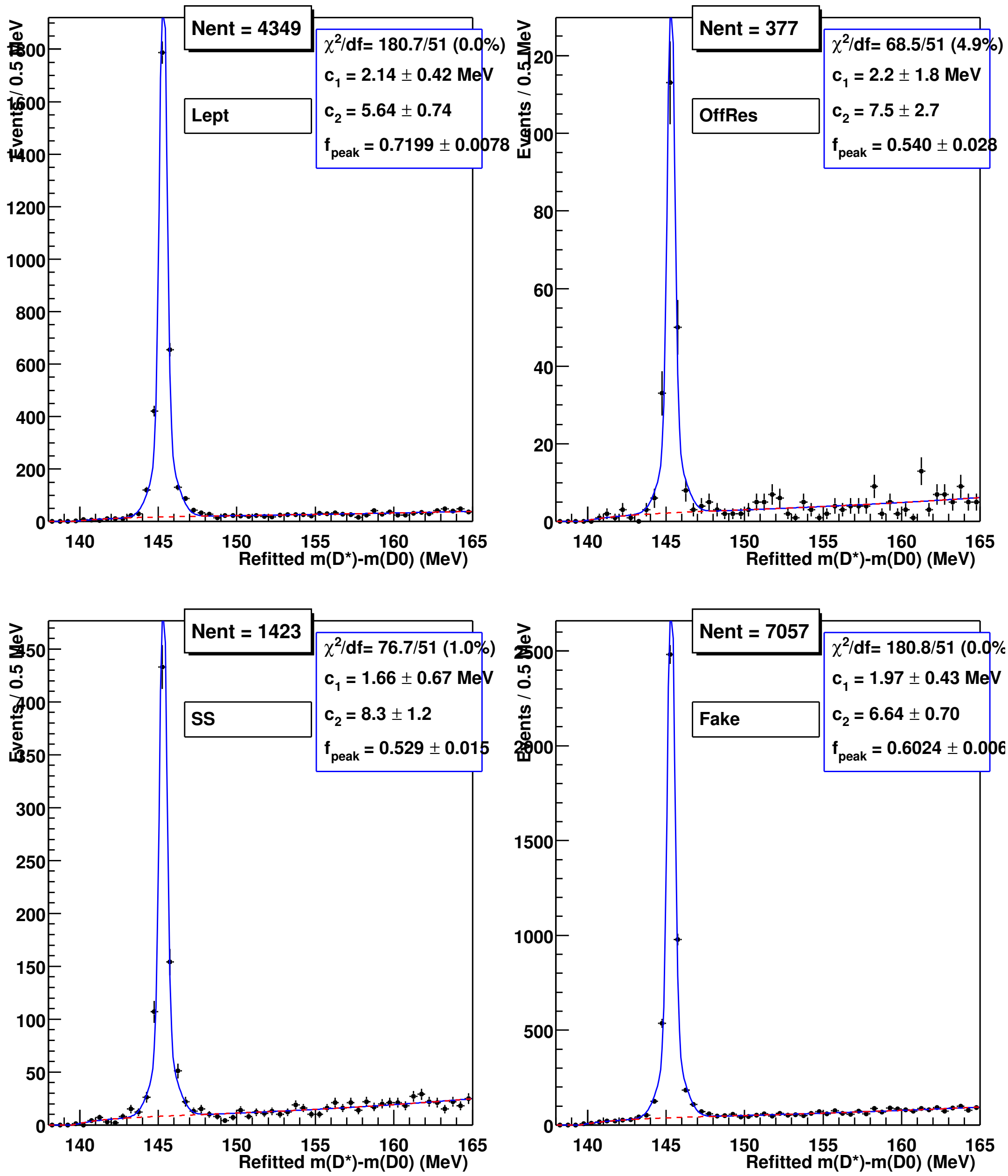


Figure 5.2: Results of the simultaneous fit described in the text to the sample consisting of candidates reconstructed as $D^0 \to K\pi$ and their soft pion reconstructed in the SVT and DCH. The dashed curves show the fitted contributions of combinatoric background. The δm peak contains various backgrounds. For example the δm peak of the Off-Resonance sample contains mainly continuum background but also a small fraction of fake-lepton background.

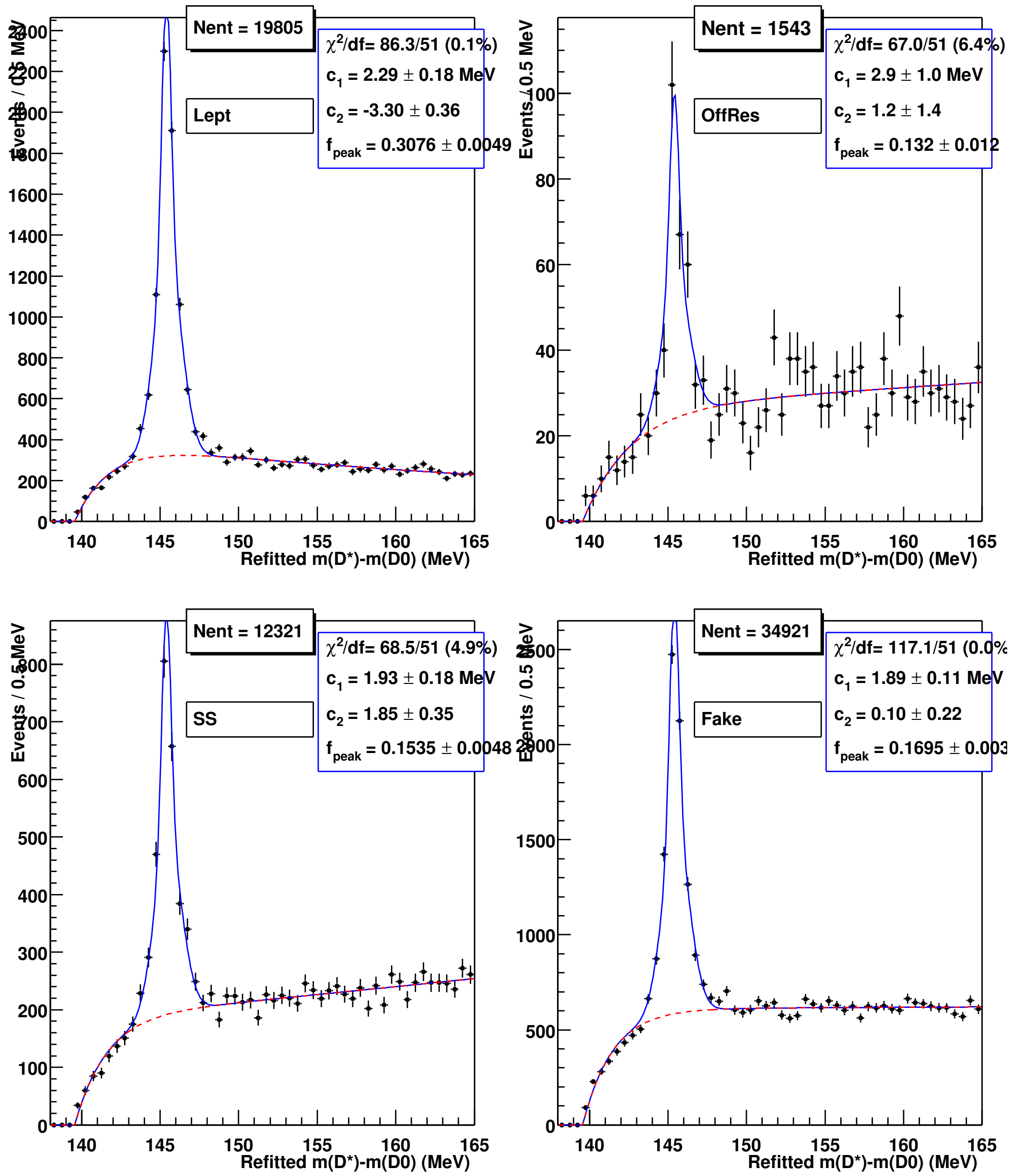


Figure 5.3: Results of the simultaneous fit described in the text to the sample consisting of candidates reconstructed as $D^0 \to K\pi\pi\pi$ or $D^0 \to K_S\pi\pi$ and their soft pion reconstructed in the SVT only. The dashed curves show the fitted contributions of combinatoric background. The δm peak contains various backgrounds. For example the δm peak of the Off-Resonance sample contains mainly continuum background but also a small fraction of fake-lepton background.

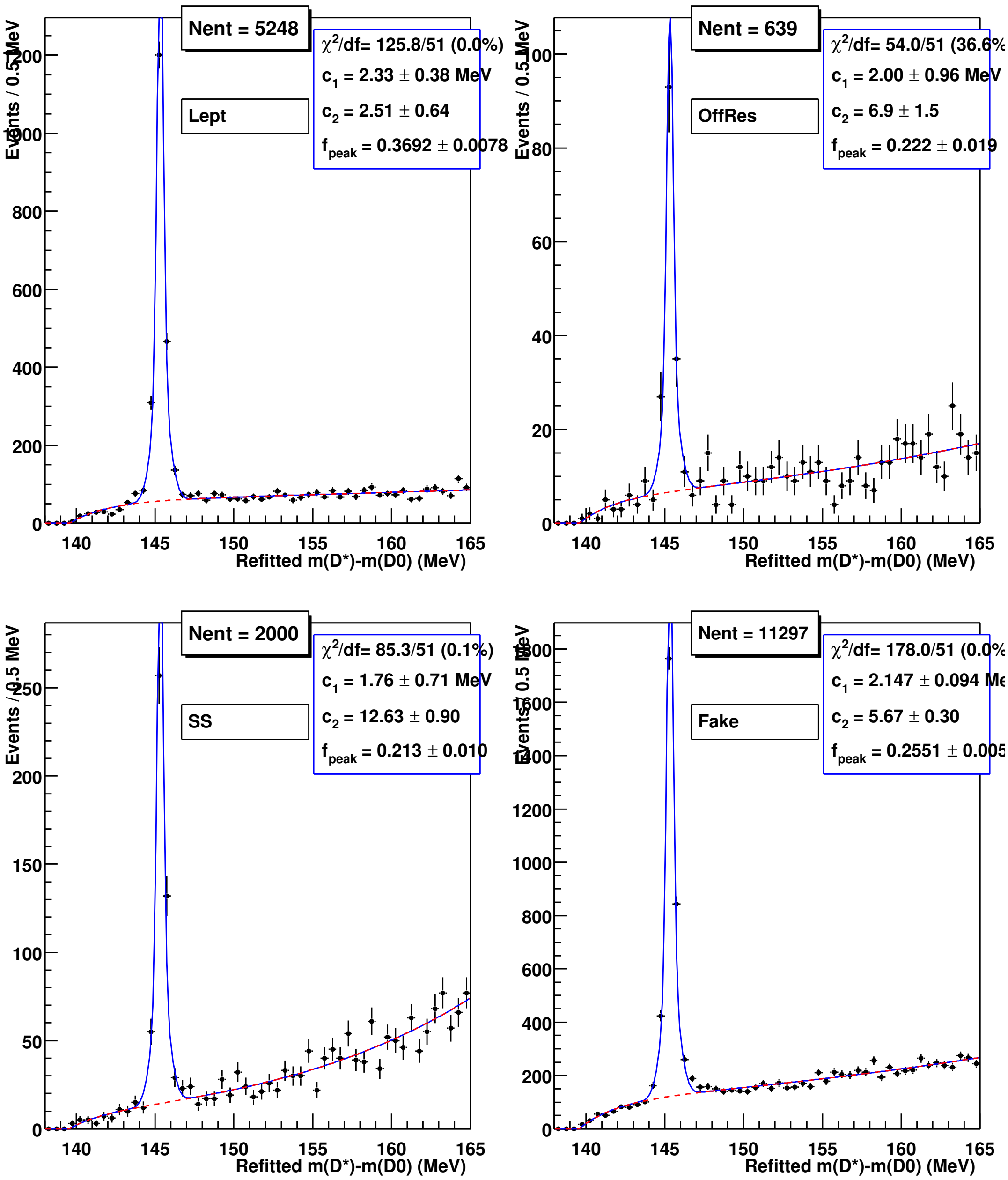


Figure 5.4: Results of the simultaneous fit described in the text to the sample consisting of candidates reconstructed as $D^0 \to K\pi\pi\pi$ or $D^0 \to K_S\pi\pi$ and their soft pion reconstructed in the SVT and DCH. The dashed curves show the fitted contributions of combinatoric background. The δm peak contains various backgrounds. For example the δm peak of the Off-Resonance sample contains mainly continuum background but also a small fraction of fake-lepton background.

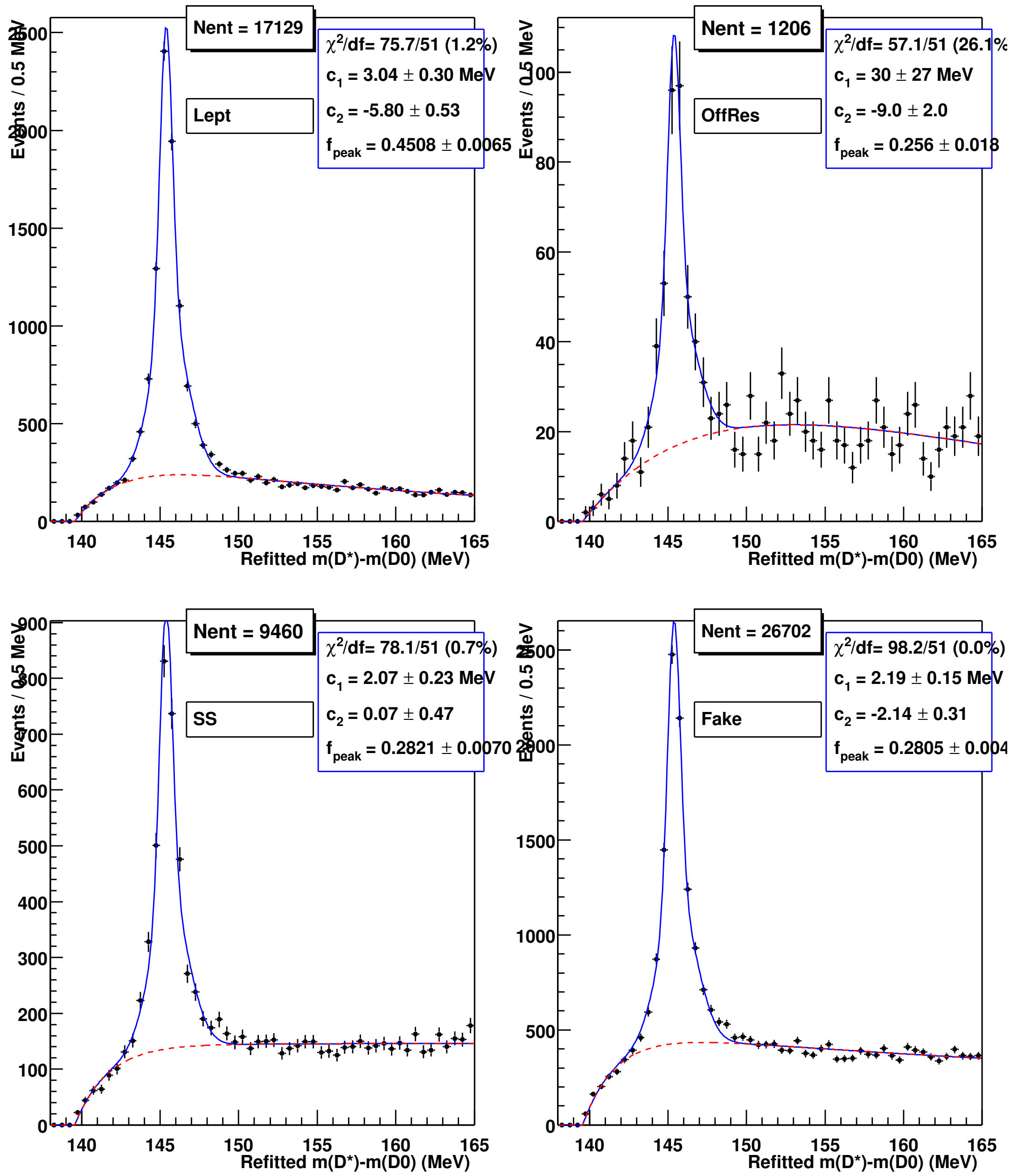


Figure 5.5: Results of the simultaneous fit described in the text to the sample consisting of candidates reconstructed as $D^0 \to K\pi\pi^0$ and their soft pion reconstructed in the SVT only. The dashed curves show the fitted contributions of combinatoric background. The δm peak contains various backgrounds. For example the δm peak of the Off-Resonance sample contains mainly continuum background but also a small fraction of fake-lepton background.

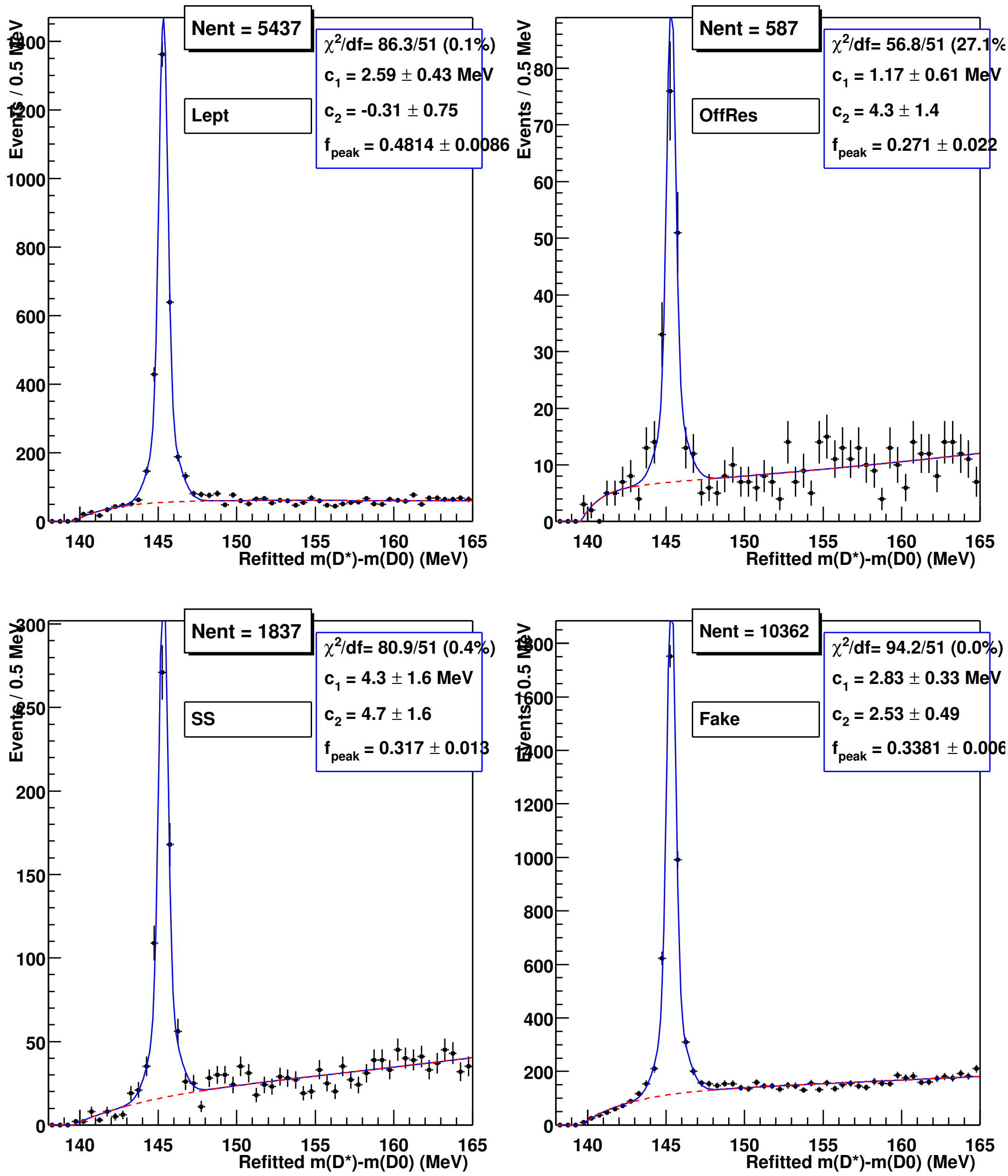


Figure 5.6: Results of the simultaneous fit described in the text to the sample consisting of candidates reconstructed as $D^0 \to K\pi\pi^0$ and their soft pion reconstructed in the SVT and DCH. The dashed curves show the fitted contributions of combinatoric background. The δm peak contains various backgrounds. For example the δm peak of the Off-Resonance sample contains mainly continuum background but also a small fraction of fake-lepton background.

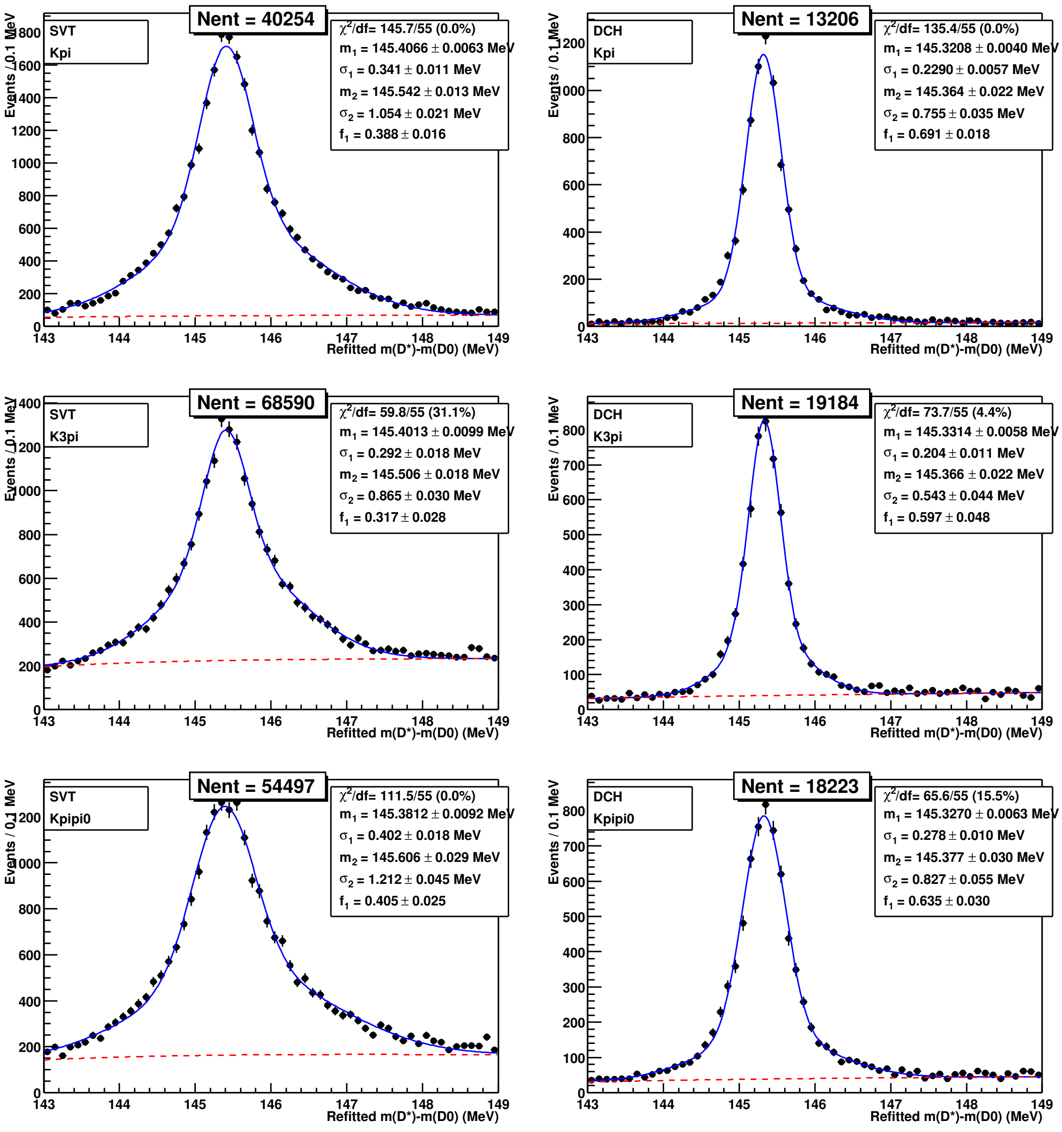


Figure 5.7: Results of the simultaneous fit described in the text to each of the 6 peak groups over the peak region 143–149 MeV. The dashed curves show the fitted contributions of combinatoric background.

5.1.2 Calculation of Yields

After finding all the peak and combinatoric shape parameters we can now evaluate the absolute yields of peak and combinatoric events within each of the 72 subsamples resulting from the most detailed decomposition by `svtDch` (2) $\otimes$ `Dmode` (3) $\otimes$ `onOffRes` (2) $\otimes$ `angCut` (2) $\otimes$ `leptID` (3). The main advantage of such a detailed breakdown is that it offers the maximum flexibility for later regrouping subsamples.

The method consists in combining peak and combinatoric distributions into an unbinned extended maximum-likelihood fit

$$\mathcal{L}(N'_{peak}, N'_{comb}) = \prod_{\text{evts } j} \left\{ \frac{N_{peak}}{N_{tot}} \cdot F_{peak}(\delta m_j) + \frac{N_{comb}}{N_{tot}} \cdot F_{comb}(\delta m_j) \right\} \cdot P(N_{tot}) \ ,$$

where $N_{tot} \equiv N_{peak} + N_{comb}$ and $P(N_{tot})$ is the Poisson probability of the observed number of events when N_{tot} are expected. The actual fit parameters in this likelihood, N'_{peak} and N'_{comb}, are yields calculated within a signal window $(\delta m_1, \delta m_2)$

$$N'_{peak} \equiv N_{peak} \cdot \int_{\delta m_1}^{\delta m_2} F_{peak}(\delta m) d(\delta m) \quad , \quad N'_{comb} \equiv N_{comb} \cdot \int_{\delta m_1}^{\delta m_2} F_{comb}(\delta m) d(\delta m) \ .$$

The choice of signal window for δm is 143–148 MeV for for candidates in which the slow pion is measured in the SVT only. For candidates in which the slow pion has SVT and DCH hits we use a narrower window of 144–147 MeV.

At the end of this procedure, we have for each sample the number of N'_{peak} events and N'_{comb} in the background under it. We are still left with the goal of evaluating and subtracting the "peaking" backgrounds due to fake-lepton, continuum, uncorrelated, correlated and $B \rightarrow D^{*-} \ell^+ \nu_\ell X$ background.

Appendix E contains the plots with each of the 72 two-parameter yield fits descibed above. All fits converge and we find good agreement between the models and the data.

Next come the subtraction to peak events of fake-lepton background: we subtract the fake-lepton background both in on and off-resonance samples, so that when after we subtract continuum background we do not perform a double subtraction of the continuum.

5.2 Lepton Identification and Fake Lepton determination

After selecting a charged D^* meson we study all the charged tracks with opposite charge with respect to the D^* meson that pass all $D^*\ell$ selection criteria except the lepton identification requirements.

We use the standard "very tight" electron and muon selectors to select the e and μ control samples. The tracks associated, when reconstructing the $D^*\ell$ candidate, to the charged D^* (with correct polarity) that fail both the "loose" electron and "loose" muon selectors are defined as the fake-lepton sample.

In order to measure the branching ratio, we need to correct the data for both background in the identification of leptons and its efficiency. For this we use tables of efficiencies $\omega_{x\to y}(p, \cos\theta, \phi)$ [63, 64, 65], as a function of momentum and angles (measured in the laboratory frame), of identifying as y an x particle. Then we compute -for each event sample- its own average efficiencies $\epsilon(x \to \ell)$ of identifying particle x as a lepton ℓ by averaging the $\omega_{x\to y}(p, \cos\theta, \phi)$ on the phase space -weighting with the appropriate distributions in $p, \cos\theta, \phi$ for the sample in question as

$$\epsilon(x \to y) \equiv \frac{1}{N}\int \mathrm{d}p\,\mathrm{d}\cos\theta\,\mathrm{d}\phi\,\frac{dN}{\mathrm{d}p\,\mathrm{d}\cos\theta\,\mathrm{d}\phi}\cdot\omega_{x\to y}(p, \cos\theta, \phi)\ , \tag{5.1}$$

where the function

$$\frac{1}{N}\frac{\mathrm{d}N}{\mathrm{d}p\,\mathrm{d}\cos\theta\,\mathrm{d}\phi}$$

describes the distribution of these variables in the relevant physics samples. The candidate events used to study the momentum and angular distributions for lepton identification efficiency evaluation are those with a true reconstructed D^* in the event, according to the Monte Carlo association.

The weighted PID efficiencies are shown in table 5.2.

The laboratory momentum distributions of lepton and hadron tracks (according to Monte Carlo) from generic $B\overline{B}$ and $c\overline{c}$ Monte Carlo are shown in Fig. 5.8.

For the overall rate of misidentifying a hadron as an electron or muon, $\epsilon(h \to \ell)$, we use the average of the misidentification rates for π, K, and p, weighted by the relative fraction of true π, K, and p candiates predicted by generic Monte Carlo for $B^0\overline{B}^0$ and B^+B^- and $c\overline{c}$:

$$\epsilon(h \to \ell) = f_\pi\,\epsilon(\pi \to \ell) + f_K\,\epsilon(K \to \ell) + f_p\,\epsilon(p \to \ell)\,.$$

sample	veryTight	Loose
ElecEffVTbha,ElecEffLSbha	0.90824 ± 0.00058	0.97703 ± 0.00041
ElecMFakeVTeemumu,ElecMFakeLSeemumu	0.0014 ± 0.0013	0.0016 ± 0.0013
ElecMFakeVTmumug2,ElecMFakeLSmumug2	0.0004 ± 0.0042	0.0010 ± 0.0042
ElecPIFakeVT,ElecPIFakeLS	0.001264 ± 0.000064	0.04811 ± 0.00035
ElecKFakeVT,ElecKFakeLS	0.00239 ± 0.00028	0.03499 ± 0.00096
ElecPFakeVT,ElecPFakeLS	0.01236 ± 0.00066	0.2755 ± 0.0045
MuonEffVTeemumu,MuonEffLSeemumu	0.6067 ± 0.0018	0.8394 ± 0.0016
MuonEffVTmumug2,MuonEffLSmumug2	0.6118 ± 0.0050	0.8406 ± 0.0045
MuonPIFakeVT,MuonPIFakeLS	0.01622 ± 0.00066	0.0620 ± 0.0012
MuonKFakeVT,MuonKFakeLS	0.00365 ± 0.00044	0.0573 ± 0.0012
MuonPFakeVT,MuonPFakeLS	0.00354 ± 0.00032	0.02083 ± 0.00070
KaonEffVTdstar,KaonEffLSdstar	0.8081 ± 0.0028	0.97265 ± 0.00093
KaonEFakeVT,KaonEFakeLS	0.00931 ± 0.00014	0.14556 ± 0.00073
KaonMFakeVTeemumu,KaonMFakeLSeemumu	0.0031 ± 0.0013	0.1499 ± 0.0017
KaonMFakeVTmumug2,KaonMFakeLSmumug2	0.0032 ± 0.0042	0.1420 ± 0.0046
KaonPIFakeVT,KaonPIFakeLS	0.01767 ± 0.00090	0.1597 ± 0.0016
KaonPFakeVT,KaonPFakeLS	0.03261 ± 0.00086	0.9608 ± 0.0013

Table 5.2: *Summary of angle- and momentum-weighted PID efficiencies for all selectors used to evaluate ℓ and K ID efficiencies.*

The relative fractions f_π, f_K, and f_p are given in table 5.3.

As will be shown later, we use the averaged efficiencies to both disentangle the hadron contribution in each control sample so that we can extract the correct amount of fake-lepton background in the signal samples and correct for lepton identification inefficiency.

Sample	π	K	p	total
BBbar	82168	7820	497	90485
CCbar	12094	1281	739	14114
Weighted Average	101784	9897.78	1695.66	113377.44
relative fraction	89.774 %	8.730 %	1.496 %	-

Table 5.3: *Relative contributions of π, K, and protons to hadron sample used for calculating misidentification probabilities. In the last two rows, the contributions have been weighted by the relative equivalent lumonosity of the $B^0\overline{B}^0$ B^+B^- and $c\overline{c}$ generic Monte Carlo samples.*

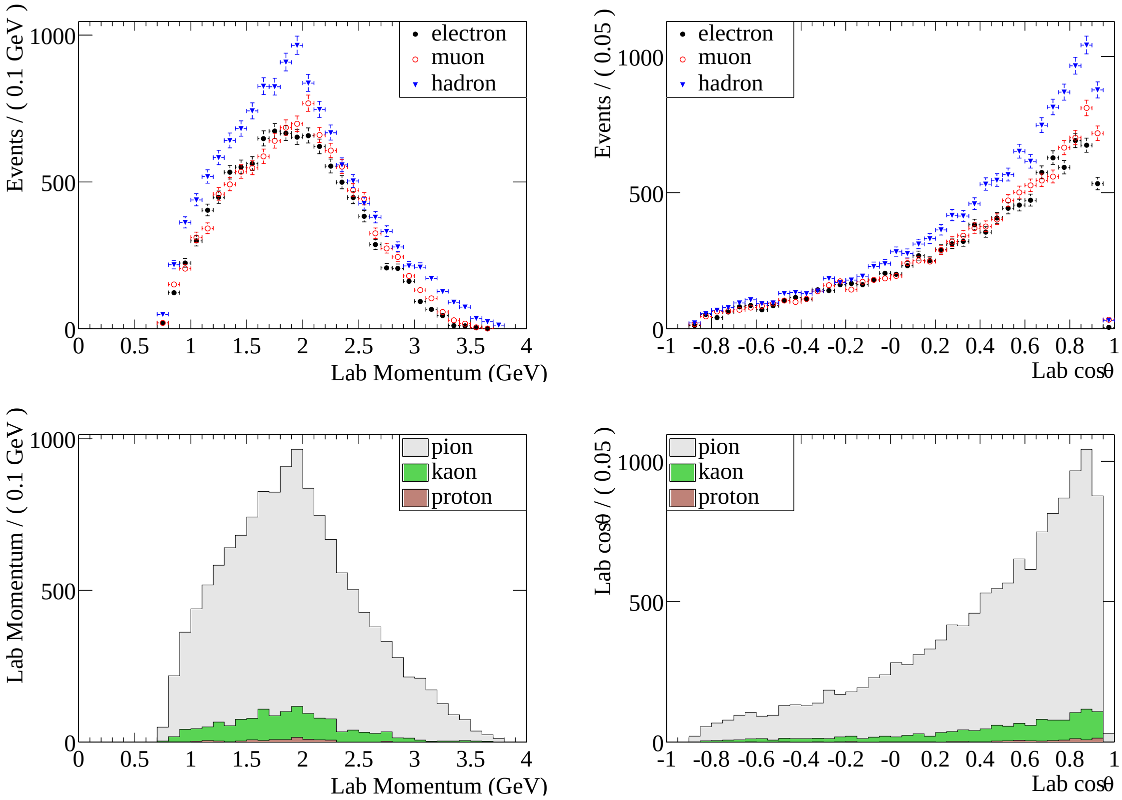


Figure 5.8: Upper plots: distributions of momentum (left) and polar angle (right) in the lab frame of real leptons and real hadrons (according to Monte Carlo) for the events that pass all the cuts except lepton identification, and the D^* is reconstructed correctly. Lower plots: The relative π, K and proton contributions to the hadron sample in the fake category of selected events.

For what regards electron efficiency evaluation, we use Bhabha control sample as default PID table. Other electron efficiency tables are for systematic checks. For muon efficiencies we use $ee\mu\mu$ PID tables.

Due to large variations in muon efficiencies during the 2001 data taking period the PID tables are evaluated dividing the 2001 run into 8 data sets. In Appendix F we show the particle identification efficiencies and the distribution of events, as a function of laboratory momentum, in the samples used to evaluate each kind of particle selector efficiency.

The treatment of the correction of efficiencies and misidentifications in the identification of leptons explained here is also applied to the identification of charged K and π used in the reconstruction of the charged D^* (see table 5.2).

The fake sample is required to fail both the Loose electron and Loose muon selectors. We denote the efficiency for species x to satisfy two criteria y and z based on the Very Tight (Loose) criteria as $\epsilon_x^{\mathrm{VT}}(y;z)$ ($\epsilon_x^{\mathrm{LS}}(y;z)$). The relations between these efficiencies and the individual selector efficiencies are

$$\begin{aligned}
\epsilon_e^{\mathrm{VT}}(e;!\mu) &= \epsilon^{\mathrm{VT}}(e\to e)(1-\epsilon^{\mathrm{VT}}(e\to\mu)) \simeq \epsilon^{\mathrm{VT}}(e\to e)\\
\epsilon_\mu^{\mathrm{VT}}(e;!\mu) &= \epsilon^{\mathrm{VT}}(\mu\to e)(1-\epsilon^{\mathrm{VT}}(\mu\to\mu)) \simeq 0\\
\epsilon_h^{\mathrm{VT}}(e;!\mu) &= \epsilon^{\mathrm{VT}}(h\to e)(1-\epsilon^{\mathrm{VT}}(h\to\mu))\\
\epsilon_e^{\mathrm{VT}}(!e;\mu) &= \epsilon^{\mathrm{VT}}(e\to\mu)(1-\epsilon^{\mathrm{VT}}(e\to e)) \simeq 0\\
\epsilon_\mu^{\mathrm{VT}}(!e;\mu) &= \epsilon^{\mathrm{VT}}(\mu\to\mu)(1-\epsilon^{\mathrm{VT}}(\mu\to e)) \simeq \epsilon^{\mathrm{VT}}(\mu\to\mu)\\
\epsilon_h^{\mathrm{VT}}(!e;\mu) &= \epsilon^{\mathrm{VT}}(h\to\mu)(1-\epsilon^{\mathrm{VT}}(h\to e))\\
\epsilon_e^{\mathrm{LS}}(!e;!\mu) &= (1-\epsilon^{\mathrm{LS}}(e\to e))(1-\epsilon^{\mathrm{LS}}(e\to\mu)) \simeq 1-\epsilon^{\mathrm{LS}}(e\to e)\\
\epsilon_\mu^{\mathrm{LS}}(!e;!\mu) &= (1-\epsilon^{\mathrm{LS}}(\mu\to e))(1-\epsilon^{\mathrm{LS}}(\mu\to\mu)) \simeq 1-\epsilon^{\mathrm{LS}}(\mu\to\mu)\\
\epsilon_h^{\mathrm{LS}}(!e;!\mu) &= (1-\epsilon^{\mathrm{LS}}(h\to e))(1-\epsilon^{\mathrm{LS}}(h\to\mu))
\end{aligned}$$

Here we assume that $\epsilon(e\to\mu)\simeq 0$ and $\epsilon(\mu\to e)\simeq 0$ for both the VeryTight and Loose selectors.

After subtracting combinatoric contributions, the remaining portion in the peak region of the δm distributions will contains events with a real D^* and a real or misidentified lepton candidate. The observed numbers of candidates in the D^* peak (after combinatoric background subtraction) in the lepton signal and fake control samples (N_e, N_μ, N_f) are related to the numbers of $D^*\ \ell$ candidates where the lepton candidate is a true lepton or a true hadron $(\eta_e, \eta_\mu, \eta_h)$ through a set of linear equations:

$$\begin{pmatrix} N_e \\ N_\mu \\ N_f \end{pmatrix} = \begin{pmatrix} \epsilon_e(e;!\mu) & \epsilon_\mu(e;!\mu) & \epsilon_h(e;!\mu) \\ \epsilon_e(!e;\mu) & \epsilon_\mu(!e;\mu) & \epsilon_h(!e;\mu) \\ \epsilon_e(!e;!\mu) & \epsilon_\mu(!e;!\mu) & \epsilon_h(!e;!\mu) \end{pmatrix} \begin{pmatrix} \eta_e \\ \eta_\mu \\ \eta_h \end{pmatrix}. \tag{5.2}$$

The true numbers of events $(\eta_e, \eta_\mu, \eta_h)$ can easily be found by inverting the matrix. The fake fractions in the signal and control samples are therefore

$$\begin{aligned} f_e &= \eta_h \epsilon_h(e;!\mu)/N_e \\ f_\mu &= \eta_h \epsilon_h(!e;\mu)/N_\mu \\ f_f &= \eta_h \epsilon_h(!e;!\mu)/N_f \end{aligned}$$

This calculation of fake fractions is done with the lepton and fake yields extracted from each of the 24 separate subdivisions of the data: `angCut` (2) $\otimes$ `onOffRes` (2) $\otimes$ `svtDch` (2) $\otimes$ `Dmode` (3).

The fractions of events in the δm peak that are due to fake leptons are shown as f_{fake} in tables 5.5 and Appendix G for various splittings of the data and also on all Monte Carlo samples we analyzed.

5.3 Continuum Background

The continuum background contribution can be estimated using off-resonance data. At this point fake lepton contributions are arlready subtracted from off-resonance data, and the remaining portions contain a real D^* and a real lepton. We scale the peak contribution by the ratio of on-resonance luminosity to off-resonance luminosity, which is approximately 9.113. The fractions of events in the δm peak that are due to continuum background are shown as f_{cont} in tables 5.5 and 5.6 and Appendix G for various splittings of the data and of the Monte Carlo we analyzed.

5.4 Uncorrelated Lepton Background

After subtracting the fake-lepton and continuum peaking background contributions, the remaining events in the δm peak in on-resonance data come from $B^0\overline{B}^0$ and B^+B^- events. To estimate the uncorrelated background, defined as a real D^* from one B and a real lepton from the other B, we use the same-side control samples and the relative efficiencies, computed from generic Monte Carlo and shown in table 5.4, for a true signal or a true uncorrelated event to pass the opposite-side or same-side selection criteria.

The efficiencies for SVT-only events and SVT+DCH events are significantly different, so we treat the two cases separately. This efficiency difference originates from the correlation between $\cos\theta_{D^*\ell}$ distribution and D^* momentum in the semileptonic decay process. As can be seen in figure 5.9 from Monte Carlo study, the higher the D^* center-of-mass momentum is, the stronger $\cos\theta_{D^*\ell}$ is peaked at -1. Since the available energy of D^* decay is very small (6 MeV), the soft pion p_t is highly correlated with the D^* momentum. As a result, $\cos\theta_{D^*\ell}$ is more strongly peaked at -1 for high p_t soft pion events, which leads to higher OS efficiency for SVT+DCH events.

Type	OS	SS
SVT-only;signal	0.97456 ± 0.00054	0.02544 ± 0.00054
SVT+DCH;signal	0.99561 ± 0.00038	0.00439 ± 0.00038
SVT-only;uncorr	0.5114 ± 0.0065	0.4886 ± 0.0065
SVT+DCH;uncorr	0.538 ± 0.018	0.462 ± 0.018

Table 5.4: *Relative efficiencies for passing opposite-side or same-side selection criteria, calculated from Monte Carlo.*

The measured yields of opposite-side (N_{OS}) and same-side (N_{SS}) events, after subtracting fake and continuum fractions, are related to the true number of signal (η_{sig}) and uncorrelated background (η_{unc}) events by

$$\begin{pmatrix} N_{\mathrm{OS}} \\ N_{\mathrm{SS}} \end{pmatrix} = \begin{pmatrix} \epsilon_{\mathrm{OS}}^{\mathrm{sig}} & \epsilon_{\mathrm{OS}}^{\mathrm{unc}} \\ \epsilon_{\mathrm{SS}}^{\mathrm{sig}} & \epsilon_{\mathrm{SS}}^{\mathrm{unc}} \end{pmatrix} \begin{pmatrix} \eta_{\mathrm{sig}} \\ \eta_{\mathrm{unc}} \end{pmatrix} . \tag{5.3}$$

Where:

$$\epsilon_{\mathrm{OS}}^{\mathrm{sig}} = \frac{N_{\mathrm{OS}}^{\mathrm{sig}}}{N_{\mathrm{OS}}^{\mathrm{sig}} + N_{\mathrm{OS}}^{\mathrm{unc}}}, \quad \epsilon_{\mathrm{OS}}^{\mathrm{unc}} = \frac{N_{\mathrm{OS}}^{\mathrm{unc}}}{N_{\mathrm{OS}}^{\mathrm{sig}} + N_{\mathrm{OS}}^{\mathrm{unc}}}, \quad \epsilon_{\mathrm{SS}}^{\mathrm{sig}} = \frac{N_{\mathrm{SS}}^{\mathrm{sig}}}{N_{\mathrm{SS}}^{\mathrm{sig}} + N_{\mathrm{SS}}^{\mathrm{unc}}}, \quad \epsilon_{\mathrm{SS}}^{\mathrm{unc}} = \frac{N_{\mathrm{SS}}^{\mathrm{unc}}}{N_{\mathrm{SS}}^{\mathrm{sig}} + N_{\mathrm{SS}}^{\mathrm{unc}}},$$

After solving for (η_{sig}, η_{unc}), the uncorrelated background fractions can be expressed as

$$\begin{aligned} f_{\mathrm{OS}} &= \epsilon_{\mathrm{OS}}^{\mathrm{unc}} \eta_{\mathrm{unc}} / N_{\mathrm{OS}} \\ f_{\mathrm{SS}} &= \epsilon_{\mathrm{SS}}^{\mathrm{unc}} \eta_{\mathrm{unc}} / N_{\mathrm{SS}} \ . \end{aligned}$$

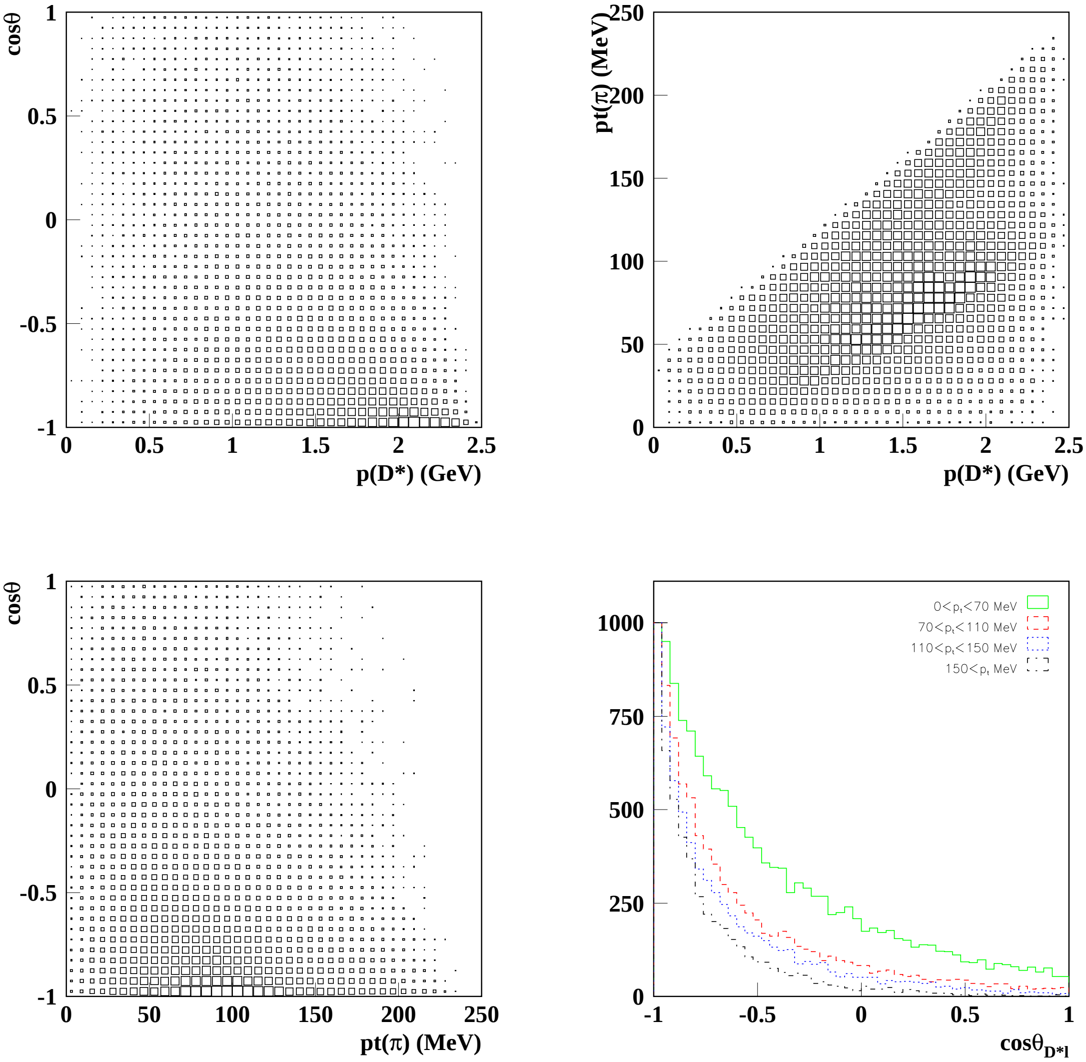


Figure 5.9: Distributions from event generator: upper-left, $\cos\theta_{D^*\ell}$ v.s. D^* CM momentum; upper-right, soft π p_T v.s. D^* CM momentum; bottom-left, $\cos\theta_{D^*\ell}$ v.s. soft π p_T; bottom-right, $\cos\theta_{D^*\ell}$ distriubtions with four ranges of soft π p_T. No cuts are applied on generated events.

Note that although we normalize the relative efficiencies for passing the OS or SS selections so that the sum of the efficiencies is 1, the resulting background fractions do not depend on this normalization. In other words, we are depending on the Monte Carlo for the relative efficiencies within each class of events (signal or uncorrelated lepton) but not the absolute efficiencies or the relative efficiencies between the two classes of events.

This calculation of uncorrelated background fractions is done with the OS and SS yields extracted from each of the 36 separate subdivisions of the data: `onOffRes` (2)

Category	peak yield	f_{fake}	$f_{\mathrm{continuum}}$	$f_{\mathrm{uncorr.}}$	$f_{\mathrm{comb.}}$
elec;OS;OnRes	17621 ± 146	$0.2849 \pm 0.0031\%$	$4.13 \pm 0.51\%$	$6.02 \pm 0.40\%$	$18.33 \pm 0.26\%$
muon;OS;OnRes	13530 ± 128	$3.675 \pm 0.044\%$	$3.79 \pm 0.63\%$	$5.63 \pm 0.53\%$	$19.06 \pm 0.31\%$
fake;OS;OnRes	33630 ± 211	$88.42 \pm 0.84\%$	$0.5 \pm 1.4\%$	$0.7 \pm 1.2\%$	$28.26 \pm 0.30\%$
elec;SS;OnRes	1364 ± 43	$0.875 \pm 0.030\%$	$1.8 \pm 1.3\%$	$73.4 \pm 5.5\%$	$26.9 \pm 1.4\%$
muon;SS;OnRes	1125 ± 38	$10.50 \pm 0.39\%$	$3.9 \pm 2.4\%$	$63.9 \pm 6.5\%$	$25.7 \pm 1.4\%$
fake;SS;OnRes	7366 ± 100	$95.9 \pm 1.9\%$	$0.2 \pm 2.9\%$	$2.9 \pm 5.2\%$	$30.48 \pm 0.69\%$
elec;OS;OffRes	81.4 ± 9.9	$2.02 \pm 0.26\%$	$98 \pm 12\%$	$0 \pm 0\%$	$16.4 \pm 3.2\%$
muon;OS;OffRes	72.5 ± 9.3	$22.4 \pm 3.0\%$	$77.6 \pm 9.9\%$	$0 \pm 0\%$	$20.8 \pm 4.2\%$
fake;OS;OffRes	988 ± 36	$98.3 \pm 5.2\%$	$1.7 \pm 3.7\%$	$0 \pm 0\%$	$30.0 \pm 1.8\%$
elec;SS;OffRes	2.8 ± 2.0	$4.1 \pm 3.0\%$	$96 \pm 69\%$	$0 \pm 0\%$	$48 \pm 70\%$
muon;SS;OffRes	7.5 ± 3.0	$36 \pm 15\%$	$64 \pm 26\%$	$0 \pm 0\%$	$24 \pm 15\%$
fake;SS;OffRes	194 ± 17	$99 \pm 12\%$	$0.7 \pm 8.7\%$	$0 \pm 0\%$	$38.1 \pm 5.9\%$

Table 5.5: **Peak** *yields (i.e. with combinatoric background subtracted. The successive fractions for fake-lepton, continuum and uncorrelated background are also indicated. The last column shows the fraction of combinatoric background in the original distributions.) in the 12 control and signal samples described in this chapter. They include only events with a real D^* and two backgrounds which cannot be determined with data control samples. After subtracting the fake-lepton, continuum and uncorrelated fraction of events to the peak yields we are left with a combination of real signal plus correlated background and $B \to D^* \ell \nu X$ background.*

$\otimes$ `leptID` (3) $\otimes$ `svtDch` (2) $\otimes$ `Dmode` (3). Due to the different selection efficiencies for the two `svtDch` categories, whose origin has been explained above, we assume that one efficiency matrix is valid for all SVT-only samples and one for all SVT+DCH samples.

5.5 Summary of Fit Extracted Yields

Once we evaluated combinatoric, fake-lepton, continuum and uncorrelated background fractions with the methods described above we can extract δm peak yields. By saying "peak yields" we mean they are not only signal yields because they still include correlated and $B \to D^{*-}\ell^+\nu_\ell X$ backgrounds. In table 5.5 and table 5.6 there is a summary with all these partial yields and total yields with peaking background fractions and combinatoric background fractions. In figures 5.10, 5.11 and 5.12 we show the δm plots with the combinatoric and peaking backgrounds evaluated so far.

What we call "signal" still includes the backgrounds from correlated and $B \to D^{*-}\ell^+\nu_\ell X$ events.

Category	total yield	f_{fake}	$f_{\text{continuum}}$	$f_{\text{uncorr.}}$	$f_{\text{comb.}}$
elec Kpi	7861 ± 90	$0.2879 \pm 0.0044\%$	$2.73 \pm 0.57\%$	$6.57 \pm 0.43\%$	$7.72 \pm 0.16\%$
elec K3pi	6530 ± 81	$0.2787 \pm 0.0050\%$	$4.60 \pm 0.82\%$	$5.04 \pm 0.45\%$	$30.69 \pm 0.34\%$
elec Kpipi0	7185 ± 88	$0.2769 \pm 0.0046\%$	$5.48 \pm 0.88\%$	$6.38 \pm 0.41\%$	$21.83 \pm 0.28\%$
muon Kpi	6183 ± 79	$3.606 \pm 0.059\%$	$4.37 \pm 0.91\%$	$5.03 \pm 0.53\%$	$7.18 \pm 0.17\%$
muon K3pi	4975 ± 71	$3.750 \pm 0.073\%$	$3.03 \pm 0.92\%$	$5.21 \pm 0.63\%$	$33.44 \pm 0.40\%$
muon Kpipi0	5557 ± 77	$3.579 \pm 0.064\%$	$3.45 \pm 0.90\%$	$7.01 \pm 0.57\%$	$22.60 \pm 0.33\%$

Table 5.6: **Total** *yields (signal plus all backgrounds) divided in the 3* `Dmode` *(we recall that here $D^0 \to K\pi\pi\pi$ and $D^0 \to K_S\pi\pi$ are merged together) signal electron and muon samples estimated with the fit method described in this chapter. All background fractions but $f_{\text{comb.}}$ refers to the amount of peaking background while $f_{\text{comb.}}$ represents the fraction of combinatoric background in the total yield.*

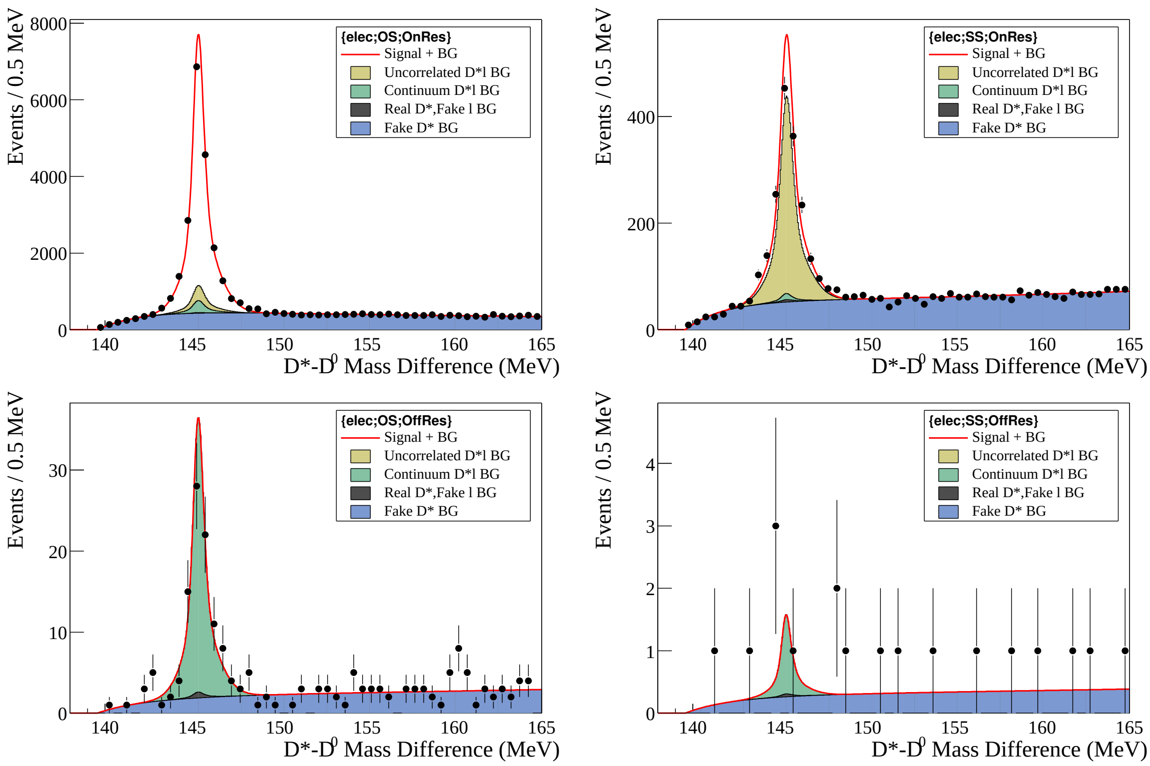


Figure 5.10: Uncorrelated, continuum, fake-lepton and combinatoric contributions for electron control samples. From top to bottom: on-resonance and off-resonance. From left to right: Opposite-Side and Same-Side.

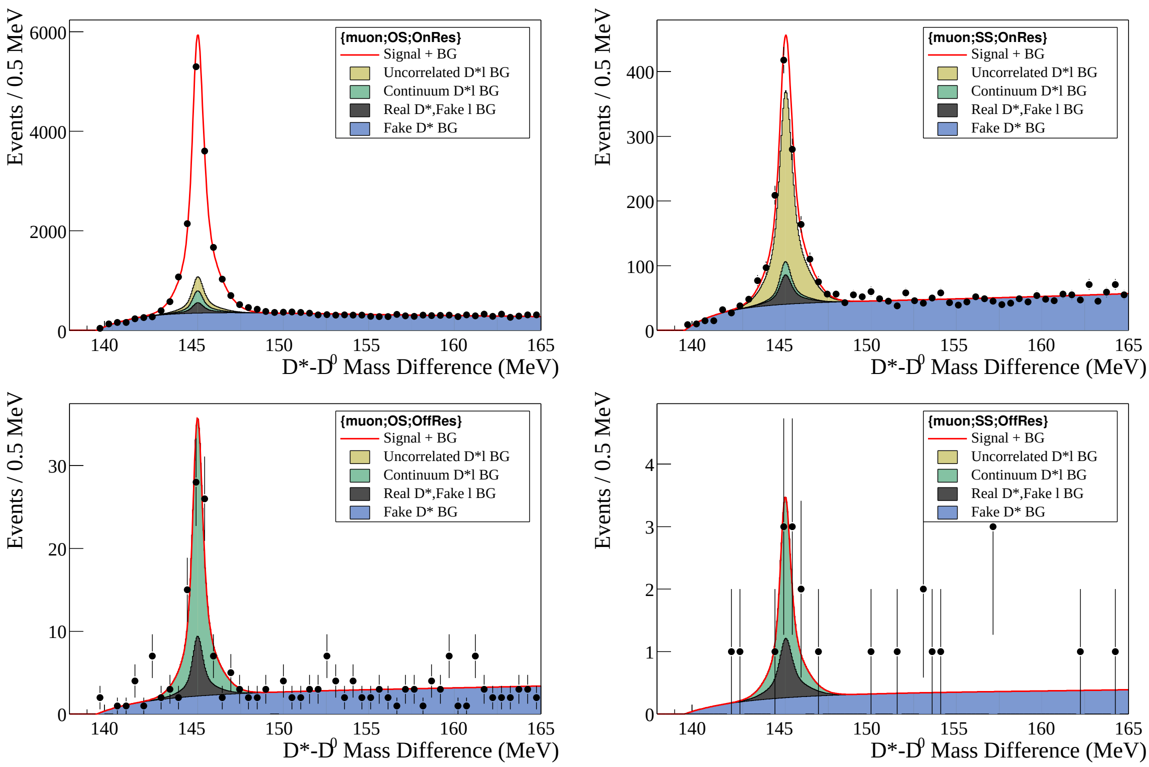


Figure 5.11: Uncorrelated, continuum, fake-lepton and combinatoric contributions for muon control samples. From top to bottom: on-resonance and off-resonance. From left to right: Opposite-Side and Same-Side.

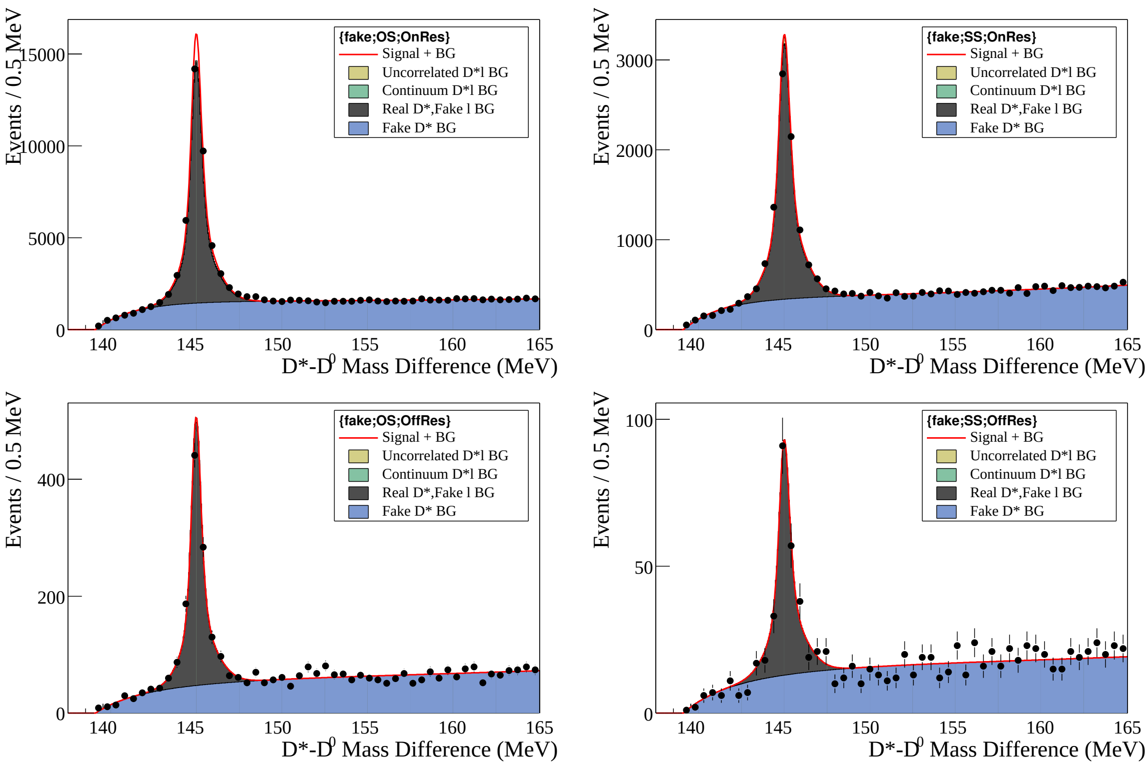


Figure 5.12: Uncorrelated, continuum, fake-lepton and combinatoric contributions for fake control samples. From top to bottom: on-resonance and off-resonance. From left to right: Opposite-Side and Same-Side.

5.6 Correlated lepton background

The correlated lepton background is characterized by a correctly identified D^* and ℓ coming from the same B meson decay but the lepton is not a primary daughter of the B [2]. Roughly 2/3 of these events are due to B meson decay to a pair of charmed mesons, one of which is the reconstructed D^* and the other produces a cascade lepton. A possible quark view of this decay is, for example, $\overline{b} \to \overline{c}cs$ where $c \to s\ell^+\nu_\ell$. Another contribution to correlated background comes from the decay $B \to D^*\tau\nu_\tau(X)$ which roughly amounts for 1/4 of it. The remaining contribution is due to the decay of a B to a D^* plus one or several lighter particles through either photon conversion ($\gamma \to e^+e^-$) or through Dalitz decays of π or η mesons (e.g. $\pi^0 \to \gamma e^+e^-$ or $\eta \to \gamma e^+e^-$). Not all of these decays are well understood and measured, as we can see from table 5.7, but we showed in Figure 4.18 that the cut on $p^*_\ell > 1.2\,\text{GeV}$ dramatically reduces this background so we expect a small contribution from these decays. Since they arise from the same single B decays, correlated background events must lie either within or below the physical region of $\cos\theta_{B^0,D^*\ell}$ by the same argument that applies to $B \to D^*\ell\nu X$ decays.

Decay	Branching fraction
$D^*D^*K^*$	$(7.4\%)^\star$
$D_s^*D^*$	$(2.0 \pm 0.7)\%$
D^*D	$(2.09 \times 10^{-3})^\star$
D_sD^*	$(9.6 \pm 3.4) \times 10^{-3}$
D^*D^*	$(6.2^{+4.1}_{-3.1}) \times 10^{-4}$
$D^*\tau\nu_\tau$	$(1.6\%)^\star$
$D^*\rho^\pm$	$(6.8 \pm 3.4) \times 10^{-3}$
$D^*\pi^\pm(\pi^0, \eta)$	$(1.5 \pm 0.5)\%$
$D^*a_1^\pm$	$(1.3 \pm 0.27)\%$

Table 5.7: *B meson decay modes list of events contributing to the correlated background. Most of the BaBar Monte Carlo simulation Branching Fractions for these modes are taken from PDG 2000. The ones with a $\star$ come from theoretical estimates.*

In order to estimate the correlated background fractions in each of our data sample we used Monte Carlo simulations of generic $B^0\overline{B}^0$ and B^+B^- events. Once we obtain these fractions we scale them properly by the ratio of data luminosity to the equivalent luminosity of the Monte Carlo sample to get the correct amount

[2] we mean events like $B \to D^*\ X$ where $X \to \ell Y$.

of correlated background. The fraction of this background will be shown with all PID and tracking efficiency corrections in the Branching fraction plots in the next chapter. We estimate that in all 4 `Dmode` and e and μ channels the correlated background fractions are about 1% of the total yield.

Chapter 6

The measurement of the branching fraction

In this chapter we describe the method adopted to perform the subtraction of the residual background and to extract the $B^0 \to D^{*-}\ell^+\nu_\ell$ signal yields and subsequently the branching fraction. We also show the evaluation of systematic uncertainties for each of the 8 signal categories (4 D^0 decay modes and 2 lepton flavors).

The fractions of combinatoric, fake-lepton, continuum and uncorrelated background are computed in each of the 8 signal categories by using δm fits shown in the previous chapter. The amount of correlated background fraction has been evaluated using Monte Carlo information.

There is still another source of δm peaking background that has to be measured, which is due to $B \to D^{**}\ell\nu_\ell$ and 4-body $B \to D^*\pi\ell\nu_\ell$ decays. In both channels of such background the final state is the same as signal **plus** another particle, typically a pion. To subtract this background, we need to use the Monte Carlo, in which both processes are simulated. In section 6.4. we explain how we disentangle it from the true signal in order to measure $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$. Throughout the chapter we will use D^{**} as a shorthand for both the resonant $B \to D^{**}\ell\nu_\ell$ and non resonant $B \to D^*\pi\ell\nu_\ell$ contribution to the $B \to D^*\ell\nu_\ell X$ background.

We use Monte Carlo simulation to calculate selection efficiencies for each of the 8 signal categories. These selection efficiencies include already tracking and particle identification efficiencies (evaluated on Monte Carlo). Therefore, we have to correct for any data-Monte Carlo differences both for tracking and for particle identification. We show in the sections below how to apply these corrections.

6.1 PID corrections

We show in Appendix F the PID efficiencies evaluated with Monte Carlo, weighted according to angle and momentum. Applying the same weighted average as described in section 5.2 for the data PID efficiencies (weighted according to the luminosity of each subsample of data collected) we can calculate the equivalent Monte Carlo efficiency and then correct for data-Monte Carlo differences in the calculation of selection efficiencies used to extract $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$. The leptons are selected according to the *electronVeryTight* and *muonVeryTight* selectors. The charged Ks are selected with the *notAPion* selector for the $D^0 \to K\pi$ mode and with the *KaonTight* selector in the $D^0 \to K\pi\pi\pi$ and $D^0 \to K\pi\pi^0$ modes as decscribed in section 4.5.

The overall correction factor to be applied to selection efficiencies for each of the 8 signal categories is:

$$C_{PID} = \frac{\epsilon^{data}(\ell \to \ell^{VT})}{\epsilon^{MC}(\ell \to \ell^{VT})} \cdot \frac{\epsilon^{data}(K^\pm \to K^\pm_{selector})}{\epsilon^{MC}(K^\pm \to K^\pm_{selector})} \tag{6.1}$$

where $\ell = e^\pm, \mu^\pm$ and *selector* depends on the `Dmode` decay for the charged K as decribed above. Table 6.1 shows the efficiencies of the selectors adopted in this

selector	Data efficiency	MC efficiency
$\epsilon(e \to e^{VT})$	0.91121 ± 0.00100	0.91046 ± 0.00048
$\epsilon(\mu \to \mu^{VT})$	0.60298 ± 0.00284	0.72194 ± 0.00055
$\epsilon(K \to K^{T})$	0.8081 ± 0.0028	0.7725 ± 0.0089
$\epsilon(K \to K^{notAPion})$	0.97265 ± 0.00093	0.97900 ± 0.00100

Table 6.1: *PID efficiencies used to extract C_{PID} for the selectors used both in data and in Monte Carlo.*

analysis both in data and in Monte Carlo and their uncertainties. The corresponding systematic uncertainty on the branching fraction is obtained by propagating these PID uncertainties accordingly.

6.2 Tracking efficiencies, K^0_S and π^0 identification corrections

The tracking efficiency corrections, C_{trk}, are applied with the same criteria as in the previuos section: the discrepancies of the tracking efficiencies between data and

Monte Carlo are quantified for each tracking category as defined in section 4.4. We include in this section the data and Monte Carlo discrepancy in the K_S^0 and π^0 identification. We then apply tracking efficiencies, K_S^0 and π^0 corrections in the calculation of selection efficiencies used to extract $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ for each of the 8 signal categories. In the next sections when referring to tracking systematic errors we mean that they include both systematic uncertainties in the tracking and in the detection of K_S^0 and π^0.

6.2.1 GoodTrackLoose efficiency

BABAR is in the fortunate position of having two devices (the SVT and DCH) that are capable of independently reconstructing tracks. This allows the possibility of measuring the efficiency of each device by using the presence of a charged track in the other as a trigger. In principle this allows one to measure both the DCH and SVT efficiencies. In practice, measuring the DCH efficiency turns out to be simpler and more useful, for the following reasons:

- Long-lived particles (e.g. K_S^0) may decay outside of the SVT volume, thereby producing tracks only in the DCH. The rate of such decays would need to be carefully measured to calculate the SVT efficiency.
- One of the track definitions used in *BABAR* is GoodTrackLoose (GTL), which places a requirement that there be at least 12 DCH hits on a track. The only other requirements are that the track point back to the beamspot within a area much larger than the DCH resolution.

The method adopted consists in starting with an SVT track and measuring how often a corresponding drift chamber track is found. Since the only difference between the definition of GTL and GoodTrackVeryLoose (GTVL) for tracks with $p_T > 100$ MeV is that the former requires 12 DCH hits, one can define the efficiency as:

$$\epsilon_{\text{raw}} = \frac{\text{Number of GTL with } N_{SVT} \text{ SVT hits}}{\text{Number of GTVL with } N_{SVT} \text{ SVT hits}} \tag{6.2}$$

The parameter N_{SVT} is set to 10. One would expect that ϵ_{raw} should vary as a function of track p_T, θ, and ϕ. In addition, previous studies have shown that the DCH efficiency decreases as the occupancy of the chamber increases [66]. This is due to the confusion created by the additional hits. Since the distribution of tracks in different decays may differ significantly in these quantities, the efficiency is calculated in bins. The bin definitions are:

- p_T: 6 bins, covering the following ranges (in GeV): 0.10-0.18; 0.18-0.31; 0.31-0.45; 0.45-0.65; 0.65-1.1; > 1.1. Binning is finer at low p_T since the efficiency rises with p_T up to about 0.5 GeV.

- θ: 6 bins, covering the following ranges (in radians) : < 0.41; 0.41-0.83; 0.83-1.15; 1.15-1.53; 1.53-2.41; > 2.41. These bins are chosen to have roughly equal population in each bin.

- ϕ: 3 bins, covering the following ranges (in radians) : $<$ -1.575; -1.575 - -0.3937; >-0.3937. This assigns one bin to a region in phi with two dead DCH superlayers, and gives two large bins for everything else.

- Track multiplicity (defined as number of GTVL): 5 bins: < 6; 7-8; 9-11; 12-14; > 14

The efficiencies determined by this method are tabulated in over 500 bins. The plots shown in figure 6.1 and 6.2 are intended to demonstrate the efficiency for tracks within the acceptance of the DCH, and therefore only the bins represented $p_T > 180$ MeV/c and $0.41 < \theta < 2.41$ are included in the plot . The data and Monte Carlo GTL efficiencies are shown in figure 6.1.

One notes that the GoodTracksLoose efficiency is quite high, above 95% except at the lowest p_T and highest multiplicity. In all cases one observes that the Monte Carlo is generally too optimistic in its assessment of the GoodTracksLoose efficiency, but that the absolute magnitude of the difference is modest. Profile plots of the data/MC efficiency ratios are presented in figure 6.2. One sees that not only are the differences between data and Monte Carlo efficiencies small, they are also relatively uniform.

When calculating selection efficiencies used to extract $\mathcal{B}(B^0 \rightarrow D^{*-}\ell^+\nu_\ell)$ we applied:

- average tracking efficiency corrections and systematic errors for charged kaons and charged pions from D^0 decay which are GoodTrackLoose;

- average tracking efficiency correction and systematic error for electron and muon tracks (our selection criterium requires the lepton to be a GoodTrack-Tight). To perform tracking efficiency corrections we treat them as GoodTrackLoose: the differences between the two categories are in the requirement

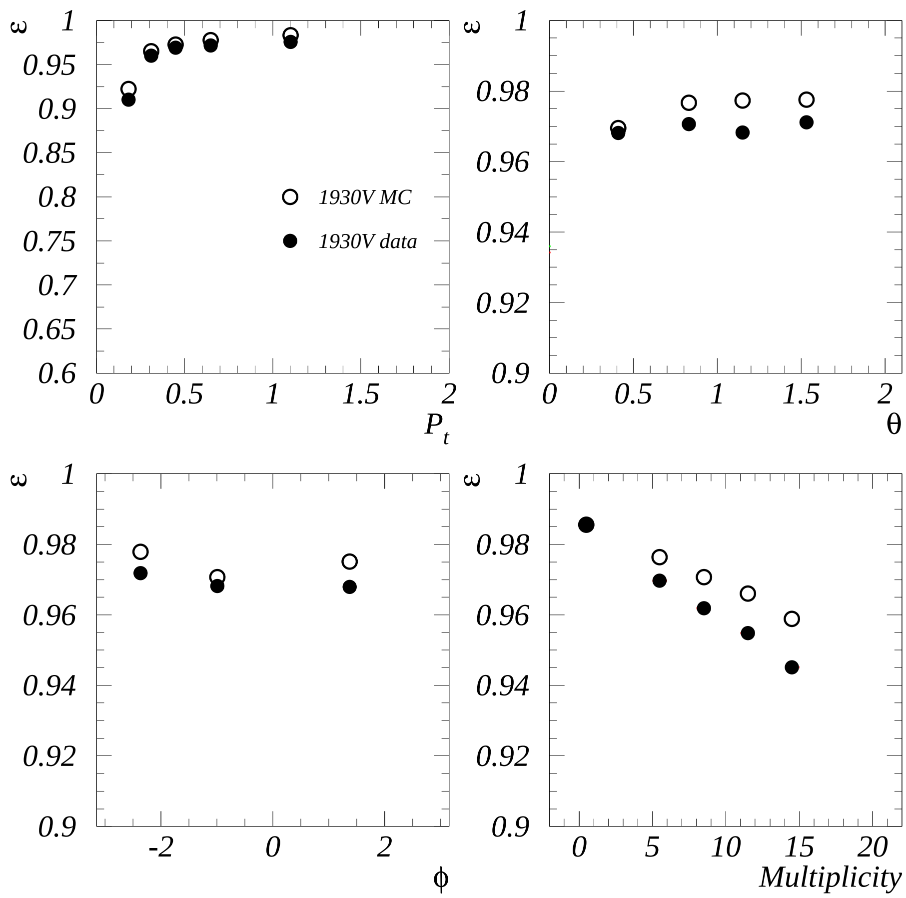


Figure 6.1: Data and Monte Carlo GoodTracksLoose efficiencies for the data sample adopted in this analysis.

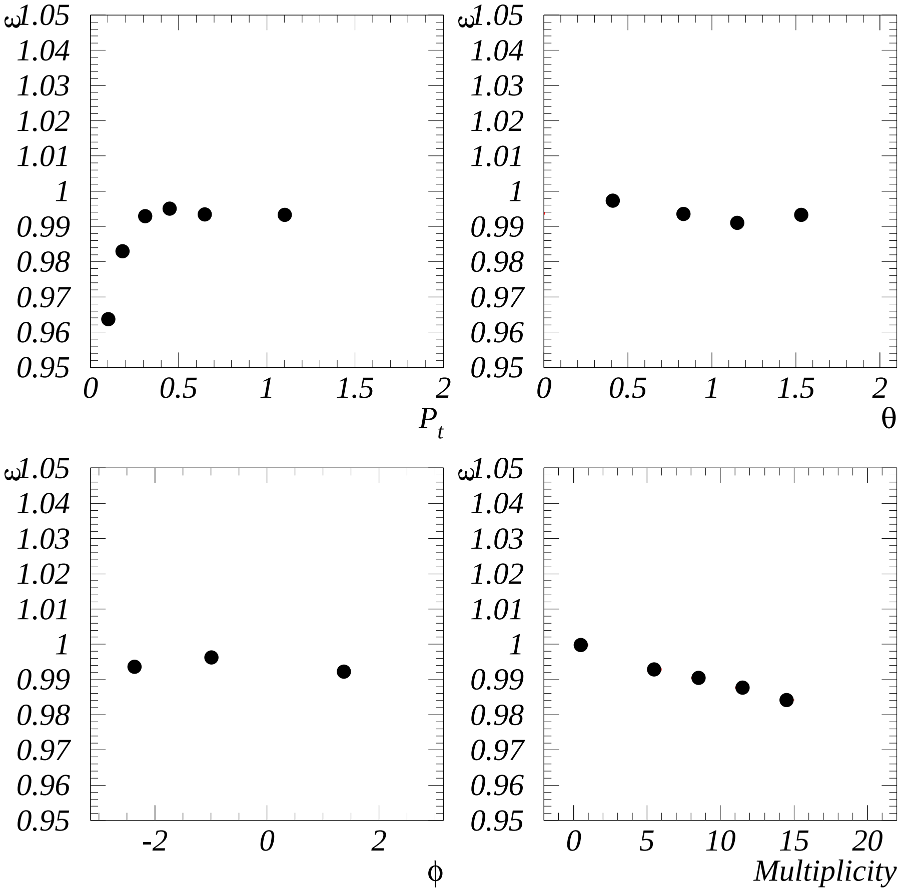


Figure 6.2: Ratio of data and Monte Carlo GoodTracksLoose efficiencies for the data sample adopted in this analysis.

on the impact parameters and in the minumum number of DCH hits of a track[1].

6.2.2 K_S^0 efficiency corrections

To correct for data and Monte Carlo discrepancies in K_S^0 selection efficiency we follow the recipe developed for *BABAR* in [66]. We adopt no tracking efficiency correction and a 1.6% systematic error on this assumption, using the results obtained in [66], for all charged pions coming from K_S^0 decay which are selected from the ChargedTracks list. Another data-Monte Carlo correction applied is related to average flight length of the K_S^0 coming from a D^0 deacy and its relative systematic error. The data-Monte Carlo correction to be applied is evaluated to be 0.934 and the relative systematic uncertainty, as calculated in [66], is 1.6%. This gives a total 2.3% relative systematic uncertainty on K_S^0 data-Monte Carlo detection efficiency difference.

6.2.3 π^0 efficiency corrections

The differences in π^0 data and Monte Carlo detection efficiencies have been studied by a dedicated analysis group in *BABAR* [71]. Applying the suggested recipe we expect the systematic uncertainties on this correction to be the largest among all other known uncertainties in this analysis. The 0.963 correction factor to be applied to our selected π^0s from the $D^0 \rightarrow K\pi\pi^0$ decays is evaluated to give a total 5.7% relative systematic uncertainties on the branching fraction measurement.

6.2.4 Soft pion efficiency

In order to assess a data-Monte Carlo correction for the soft pion tracking efficiency we use the results obtained in [66] on relative efficiencies with the helicity tecnique (used also by CLEO [67, 68]).

This method uses the angle between the soft pion (π_s) direction of motion in the D^* frame and the direction of motion of the D^* in the laboratory frame. We call it the helicity angle θ^* (figure 6.3). For a given D^* momentum in the laboratory frame the relation between the energy of the π_s (in the laboratory frame) and $cos\theta^*$ is linear. The $p(D^*)$ distribution of the D^* momentum in the laboratory frame is

[1] a *BABAR* study shows no evident effects and shifts in detection efficiencies when using GTT tracking efficiencies tables [73]

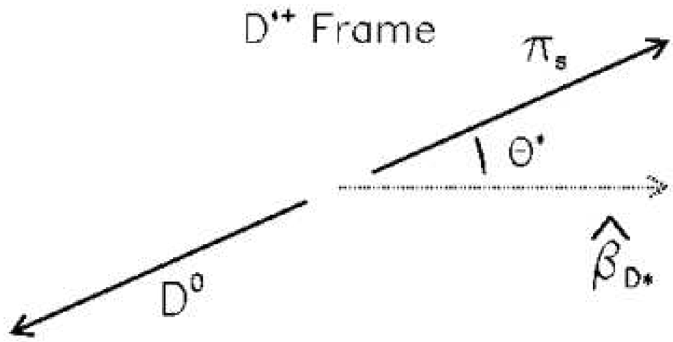


Figure 6.3: Definition of the slow pion helicity angle θ^*. As can be seen it is the angle between the slow pion direction of motion in the D^* frame and the D^* momentum direction in the laboratory frame.

divided in 8 bins of different width and the distribution of events for each bin as a function of $cos\theta^*$ is expected to be symmetric:

$$\frac{dN}{dcos\theta^*} = n(1 + \alpha cos^2\theta^*) \ , \tag{6.3}$$

where n is a normalization factor and α takes into account the helicity of the D^* decay products, i.e. the D^0 and the π_s. We can assume that any asymmetry in this distribution, after reconstruction, is due to a relative efficiency difference that can be mapped on to a specific part of the π_s momentum spectrum. For slower D^*, also the π_s is slower and so we expect large asymmetries in $cos\theta^*$ distribution corresponding to the first few bins. The distribution in $cos\theta^*$ is fitted bin per bin in $p(D^*)$ with the function 6.3 convoluted with the efficency function 6.4. The efficiency curves used both for data and Monte Carlo in the fit have the form:

$$\epsilon(p) = \begin{cases} 1 - \frac{1}{\beta(p-p_0)+1} & \text{if } p > p_0 \\ 0 & \text{if } p \leq p_0 \end{cases} \tag{6.4}$$

where p_0 is the starting point and β describes how fast the efficiency function rises to the plateau value. The results are presented in table 6.2.

sample	β (c/ MeV)	p_0 (MeV/c)	$\chi^2/d.o.f.$
data 2001 DCH 1930	170 ± 11	65.2 ± 0.2	2.81
generic MC	148 ± 16	65.0 ± 0.4	1.40

Table 6.2: *Results of the fit on our data and Monte Carlo sample.*

From these efficiency curves convoluted with the momentum distributions of our selected events, as can be seen in figure 6.4, we extract the momentum weighted integrated efficiencies:

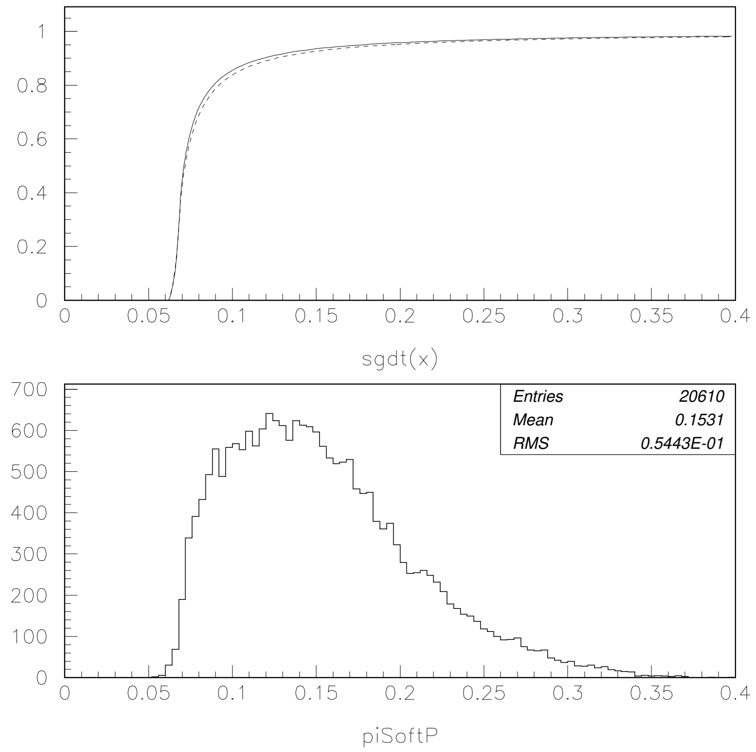


Figure 6.4: Top plot: efficiency curves as a function of π_{soft} laboratory frame momentum (in GeV/c) used for soft pion relative efficiencies studies. The data tracking efficienciy function is represented with a continuous line, while the Monte Carlo one is shown with a dashed line. The bottom plot shows the soft pion momentum distribution in the laboratory frame (in GeV/c) of our selected events in Monte Carlo.

- data efficiency $= 0.899 \pm 0.006$
- Monte Carlo efficiency $= 0.888 \pm 0.011$

and then calculate the data-Monte Carlo correction on the relative efficiencies:

- data-Monte Carlo relative efficiency correction $= 1.012 \pm 0.014$

where the uncertainties come from the propagation of the errors on the parameters used in the efficiency curves.

We recall, again, that the definition of our selection efficiency, calculated from Monte Carlo, also includes tracking efficiencies. For this reason we have to correct for data-Monte Carlo differences in tracking efficiencies. The relative correction showed above in the case of the slow pion does not take into account that the asymptotic tracking efficiency used in equation 6.4 does not necessarily be equal to one. This method suffers from a systematic uncertainty since we are interested in absolute slow pion tracking efficiencies: a potential difference in the high soft pion momentum ($p_{\pi_s} > 200\,\mathrm{MeV}/c$) tracking efficiency $\epsilon_{asymptotic}$ between data and Monte Carlo. We have to evaluate and correct for such discrepancy, and calculate a systematic error on this correction. The soft pion is selected in the GTVL list of charged tracks. The difference among GTVL absolute efficiency in data and Monte Carlo at high soft-pion momenta has been evaluated by a dedicated study [66]. The result is that for GTVL charged tracks at high momentum (more than about 200 MeV/c) no correction has to be applied to Monte Carlo absolute efficiency and a 1.3% systematic error is evaluated. This is a conservative approach because, as can be seen in figure 6.5, we discard the information that many of them are actually GTL (88.3% in our sample if the soft pion momentum is greater than 180 MeV).

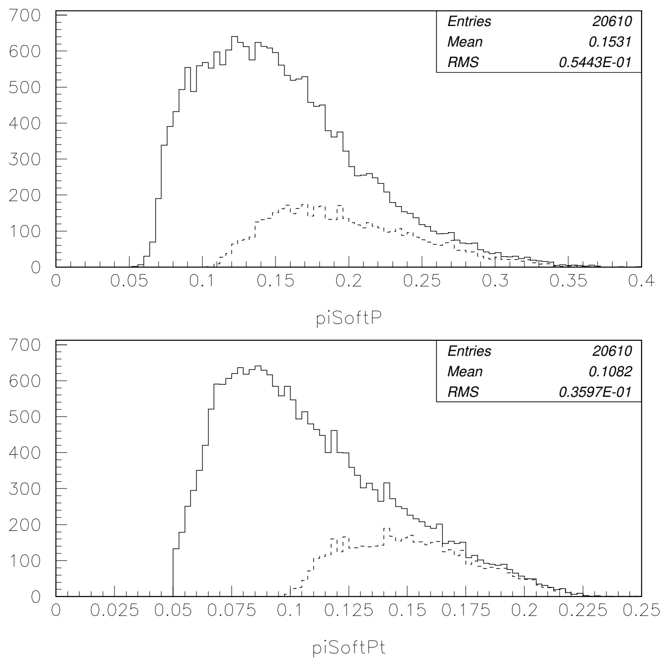


Figure 6.5: Distribution of GTVL (continuos line) and GTL (dashed line) of Monte Carlo selected events as a function of soft pion momentum (top plot) and transverse momentum (bottom plot) both evaluated in the laboratory frame and expressed in GeV/c.

There are preliminary studies [69] which confirm that the fitted β and p_0 parameters are consistent inside their error bars if the analysis is redone in bins of polar angle of the soft pion in the laboratory frame. We assume no further systematic uncertainty from this effect so we apply no correction and add 1.3% systematic uncertainty as suggested in [66]. Another check we performed is how well Monte Carlo describes the distributions of the relevant soft pion variables. As can be inferred from figure 6.6 there is no need to apply a systematic uncertainty on Monte Carlo modeling of the soft pion transverse momentum shape. As a final result the soft

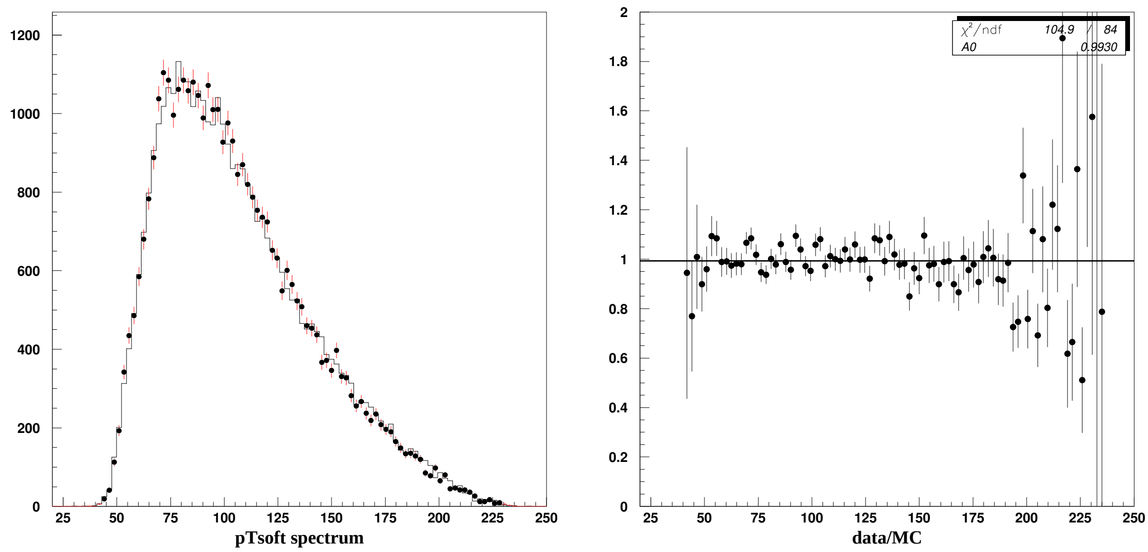


Figure 6.6: On the left we show data and $B^0\overline{B}^0 + B^+B^- + c\overline{c}$ luminosity weighted Monte Carlo agreement of the selected soft pion p_T shapes in the laboratory frame (in MeV/c). Monte Carlo has been scaled to have the equivalent luminosity of data. On the left we show the result of the fit of data over Monte Carlo content in each bin of the histogram. The result of the fit is consistent with no correction and the χ^2 is good.

pion correction to apply is shown in table 6.3.

type of correction	value	error (%)
relative efficiency	1.012	1.4
GTVL correction	1.000	1.3
total correction	1.012	1.9

Table 6.3: *Total soft pion tracking efficiency correction and systematic uncertainty on it.*

In table 6.4 we show the C_{trk} to be applied[2] to each D^0 decay mode.

[2]here we show and include in C_{trk} also the π^0 and K_S^0 data-Monte Carlo corrections to be applied

	$K\ \pi$ `Dmode`	$K3\pi$ `Dmode`	$K\pi\pi^0$ `Dmode`	$K^0_S\pi\pi$ `Dmode`
D^0 tracks corr.	0.981	0.940	0.976	0.934
syst. (%)	1.1	1.6	1.1	1.1
ℓ track corr.	0.988	0.988	0.988	0.988
syst. (%)	0.8	0.8	0.8	0.8
π_{soft} track corr.	1.012	1.012	1.012	1.012
syst. (%)	1.9	1.9	1.9	1.9
π^0 eff.	1	1	0.963	1
syst. (%)	0	0	5.7	0
K^0_S fl. lenght. corr. π CH. TR.	1	1	1	0.934
syst. (%)	0	0	0	2.3

Table 6.4: *Tracking efficiencies corrections. To get the total C_{trk} for each* `Dmode` *we need to multiply all the entries of a column.*

We obtain the systematic unceratinties due to tracking by propagating these uncertainties in the branching ratio measurement.

6.3 PDG Branching Fractions corrections

The branching ratios of the $D^* \to D^0\pi$ decay and of the four D^0 decay modes used in this analysis are different in the *BABAR* Monte Carlo production with respect to the ones quoted in PDG2002 [5]. This gives a correction factor on the true yields of D^* and D^0 when determining $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$. We have to correct for this discrepancy and assign the PDG2002 quoted uncertainties as a systematic error on this correction. We correct for BR(PDG)/BR(MC) with the values shown in table 6.5. The relative PDG2002 errors on D^* and D^0 branching ratios enter as relative errors on the branching fraction measurement. These uncertainties give a relative 2.4% systematic error on the $D^0 \to K\pi$ mode (the smallest one among the 4 `Dmode`) and contribute to a 4.2%, 6.9% and 6% relative error on the $D^0 \to K\pi\pi\pi$, $D^0 \to K\pi\pi^0$ and $D^0 \to K_S\pi\pi$ modes, respectively. This type of systematic uncertainties are irreducible until more accurate measurements of these D^0 branching fractions are available. This will be possible with the huge amount of data, more than 90 fb^{-1}, collected both on and off the $\Upsilon(4S)$ resonance in *BABAR*.

	$K\ \pi$ `Dmode`	$K3\pi$ `Dmode`	$K\pi\pi^0$ `Dmode`	$K_S^0\pi\pi$ `Dmode`	$D^* \rightarrow D^0\pi$
PDG2002 (%)	3.80	7.46	13.1	2.96	67.7
error (%)	0.09	0.31	0.9	0.18	0.5
BABAR MC (%)	3.83	8.62	13.9	2.7	68.3

Table 6.5: *PDG values of D^* and D^0 branching fractions with quoted errors. We show also the branching fractions adopted in BABAR Monte Carlo.*

6.4 Background subtraction

The variable $\cos\theta_{B^0,D^*\ell}$ is computed assuming that the B^0 decayed to a $D^{*-}\ell^+\nu_\ell$, i.e. that the missing mass of the D^* ℓ to a B^0 is zero. Such variable can tell us how true this statement is. We use the $\cos\theta_{B^0,D^*\ell}$ variable to disentagle the D^{**} background from the signal. The Monte Carlo distributions of $\cos\theta_{B^0,D^*\ell}$ for signal and D^{**} are known (see section 4.13.2 and figures 4.20 and 4.21) and different from each other. Separate fits to the distributions of the 8 signal channels are preformed to determine the fraction of D^{**} events in each of them.

The $\cos\theta_{B^0,D^*\ell}$ distribution of all selected events spans the range [-10,5]. We divide this interval into 150 bins and, with the fitted fraction of D^{**} events, perform a bin-by-bin subtraction of the backgrounds.

We checked that the $\cos\theta_{B^0,D^*\ell}$ shapes of D^{**} events coming separately from $B^0\overline{B}^0$ and B^+B^- are similar, as can be seen in figure 6.7. This is the reason why we adopt a unique distribution for the D^{**} background in the fit. The distribution is extracted adding appropriatly all non resonant and resonant D^{**} exclusive modes in B^0 and $B^\pm$ decays shown in table 4.8.

Since many of the exclusive D^{**} modes have not been experimentally discovered so far, the branching fractions and form factors adopted for this modes in *BABAR* Monte Carlo come from theoretical estimates which have large uncertainties. This introduces a systematic effect in the estimate of signal events and has been taken into account as we will show below.

6.4.1 Fit to the D^{**} contribution

We already said that we divide the $\cos\theta_{B^0,D^*\ell}$ distribution in 150 bins. Each bin j contains $N_{dt}(j)$ events in data and $N_{mc}(j)$ in Monte Carlo simulations. The latter is defined as:

$$N_{mc}(j) = f_{D^*}N_{D^*}(j) + f_{D^{**}}N_{D^{**}}(j) + N_{bkg}(j)$$

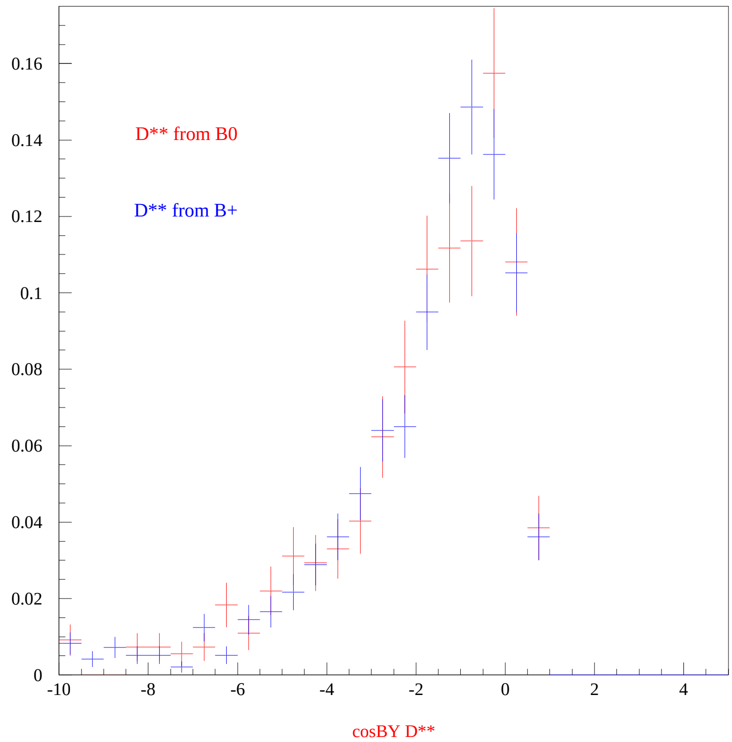


Figure 6.7: A comparison of the $\cos\theta_{B^0,D^*\ell}$ distributions, extracted from Monte Carlo, adding all the $B \to D^{**}\ell\nu_\ell$ and non resonant $B \to D^*\pi\ell\nu_\ell$ modes (shown in table 4.8) separately from B^0 or B^+ decays. As can be seen from the figure there is no evident benefit in keeping the shapes separated when fitting for the D^{**} amount of background because they are very similar (in figure the shapes are normalized to have the same area).

where N_{bkg} are the events of all the backgrounds (combinatoric included) but the D^{**}, and the parameters f_{D^*} and $f_{D^{**}}$ indicates the scaling factors that has to be applied to the *BABAR* Monte Carlo predictions for $N_{D^*}(j)$ and $N_{D^{**}}(j)$ in order to best match the $\cos\theta_{B^0,D^*\ell}$ distribution in data. These parameters are obtained from a fit which minimizes the following χ^2 distribution

$$\chi^2 = \sum_{j=1}^{150} \frac{[N_{dt}(j) - N_{mc}(j)]^2}{\sigma^2(j)} \tag{6.5}$$

where

$$\sigma^2(j) = \sigma^2_{dt}(j) + (f_{D^*}\sigma_{D^*}(j))^2 + (f_{D^{**}}\sigma_{D^{**}}(j))^2 + \sigma_{bkg}(j))^2.$$

If $N_{dt}(j)$ is less than 10, its content is merged with $N_{dt}(j+1)$, and the procedure is repeated. The fit is performed with a fortran program interfaced to MINUIT package [78] for the χ^2 minimization.

The χ^2 distributions are obtained separately for each of the 4 `Dmode` and 2 lepton signal categories for a total of 8 χ^2 minimization fits.

Some remarks on different fit approaches

We performed each fit to the D^{**} contribution with two different approaches:

1. we kept the value $f^i_{D^{**}}/f^i_{D^*}$ common to all 8 channels and fit for 8 different $f^i_{D^*}$ ($i = 1, \ldots, 8$) in each signal category. We refer to this method as the 9 parameters fit;

2. we performed the fit with $f^i_{D^*}$ and $f^i_{D^{**}}$ floating in each of the 8 channels, for a total of 16 parameters in the fit, with the constraint coming from a *BABAR* internal analysis result [76] that measure approximately $\mathcal{B}(B \to D^{**}\ell\nu) = 1.75 \pm 0.3\%$ for resonant D^{**} decays and $\mathcal{B}(B \to D^*\pi\ell\nu) = 1.2 \pm 0.4\%$ for the non-resonant modes contributing to our D^{**} background. This gives a sum of the resonant and non-resonant contributions equal to 2.95±0.50% for each type of lepton (e or μ).

The results of the fitted parameters and their statistical errors are shown in table 6.6 for the case of the 9 parameters fit.

The final state radiation of the electron, when the photon is not recovered in the event, creates a tail of signal events outside the $\cos\theta_{B^0,D^*\ell}$ signal region in the negative value of $\cos\theta_{B^0,D^*\ell}$. This tail falls exactly in the same place of the D^{**}

parameter number	parameter	fitted value	statistical error
1	$f_{D^{**}}/f_{D^*}$	0.53467	0.03835
2	f_{D^*} Kpi e	0.87677	0.01399
3	f_{D^*} Kpi mu	0.87807	0.01505
4	f_{D^*} K3pi e	0.84016	0.01808
5	f_{D^*} K3pi mu	0.82123	0.01925
6	f_{D^*} Kpipi0 e	0.83931	0.01558
7	f_{D^*} Kpipi0 mu	0.86458	0.01716
8	f_{D^*} Kspipi e	0.8167	0.0517
9	f_{D^*} Kspipi mu	0.7317	0.0521

Table 6.6: *Fitted values of f_{D^*} and $f_{D^{**}}/f_{D^*}$ with the 9 parametrs fit described above. The number shown says which scale factor one has to apply on the Monte Carlo generated $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ (for f_{D^*}) and on $\mathcal{B}(B \to D^{**}\ell\nu_\ell) + \mathcal{B}(B \to D^*\pi\ell\nu)$ (for $f_{D^{**}}$) to best match the data $\cos\theta_{B^0,D^*\ell}$ distribution.*

background as we discussed in chapter 4. In this situation the χ^2 fit to the $\cos\theta_{B^0,D^*\ell}$ distribution looses the power to discriminate among signal and D^{**} background. Therefore the radiative tail of the $\cos\theta_{B^0,D^*\ell}$ distribution in the case of the electron signal channels makes quite difficult to disentangle D^{**} background from signal. The situation is different in the muon case. The final state radiation is almost negligible and the $\cos\theta_{B^0,D^*\ell}$ tail at low values is dominated by the D^{**} background. The fit in this case has only one shape to fit in that region. For this reason we choose the 9 parameters fit since expect that the results in the muon channels descibe better the amount of D^{**} background.

The constraint applied in the 9 parameters fit gives much better agreement between the results for the branching ratio evaluated in the muon and in the electron channels since we can benefit of the clear signal to D^{**} background separation in the 4 muon channels. In table 6.7 we show all the correlations in the 9 parameters fit.

In appendix I we show the plots and the branching fraction results with the method of the 16 parameters constrained fit used as a cross check.

parameter number	GLOBAL	1	2	3	4	5	6	7	8	9
1	0.547	1.000	-0.334	-0.250	-0.237	-0.192	-0.257	-0.222	-0.089	-0.078
2	0.334	-0.334	1.000	0.084	0.079	0.064	0.086	0.074	0.030	0.026
3	0.250	-0.250	0.084	1.000	0.059	0.048	0.064	0.056	0.022	0.020
4	0.237	-0.237	0.079	0.059	1.000	0.045	0.061	0.053	0.021	0.018
5	0.192	-0.192	0.064	0.048	0.045	1.000	0.049	0.043	0.017	0.015
6	0.257	-0.257	0.086	0.064	0.061	0.049	1.000	0.057	0.023	0.020
7	0.222	-0.222	0.074	0.056	0.053	0.043	0.057	1.000	0.020	0.017
8	0.089	-0.089	0.030	0.022	0.021	0.017	0.023	0.020	1.000	0.007
9	0.078	-0.078	0.026	0.020	0.018	0.015	0.020	0.017	0.007	1.000

(Column header over GLOBAL–9: CORRELATION COEFFICIENTS)

Table 6.7: *Correlations of all fitted parameters numbered as shown in table 6.6.*

6.5 Branching fraction evaluation

We can extract the $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ in all the 8 signal channels after obtaining the yields, evaluating the fraction of all backgrounds (D^{**} included) and after applying all the data-Monte Carlo correction factors to selection efficiencies.

The $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)_{MC}$ in MC is 5.60%. After scaling the background events (N_{bkg}) and the Monte Carlo selected signal events (N_s) with the data-Monte Carlo correction factors calculated in the previous sections, the branching fraction is:

$$\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell) = \mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)_{MC}\frac{N_{data} - N_{bkg}}{N_s} \tag{6.6}$$

where N_{data} are the selected events in data in each of the 8 signal channels. While the fit for the D^{**} fraction is performed on the range [-10,5] in $\cos\theta_{B^0,D^*\ell}$, the branching fraction is evaluated in the smaller range [-1.2,1.2]. The decision of using a larger range than the expected physical value for $\cos\theta_{B^0,D^*\ell}$ is motivated by resolution studies on $\cos\theta_{B^0,D^*\ell}$ variable [74].

In appendix H we show the relationship between equation 6.6 and equation H.1 that contain selection efficiencies in it. We illustrate also the selection efficiencies calculated in each of the 8 channels adopted for the $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ measurement.

6.6 Systematic uncertainties studies

We show below all the studies performed to quantify the systematic uncertainties of the branching fraction measurement in all 8 signal channels.

6.6.1 D^{**} model systematic uncertainty

We already mentioned that the form factors and the branching fractions of all the D^{**} exclusive modes are not well known. The relative amount of each exclusive mode to the total contribution is another unknown physical quantity. In order to estimate the sytematic effects coming from this lack of knwoledge we used a method already used by CLEO [75]. We assume that we have 100% uncertainty on the relative composition of the D^{**} background and we perform the fit many times assuming only 1 exclusive mode shape as the global D^{**} shape in the fit [3]. The exclusive D^{**} channels used for this study are shown in table 4.8. The specific D^{**} Monte Carlo available in *BABAR* contains events where the D^0 is always forced to decay as $D^0 \to K\pi$. We are obliged to use only this `Dmode` and then assume the same systematic uncertainty for the other D^0 modes.

In table 6.8 we can see all the deviations from the fitted value and the maximum and minimum one. The systematic error on the branching ratio is defined as half of the difference among the maximum and minimum deviation from the originally extracted values. Performing the errors average among the electron and muon channels we get a relative 2% difference from the fitted $B^0 \to D^{*-}\ell^+\nu_\ell$ branching ratio (see section 6.7 for the fitted branching ratio results). This systematic uncertainty is applied in all the 4 `Dmode` for both lepton species.

The approach adopted to assess D^{**} model uncertainties can be considered somehow conservative. We adopt a 2% relative error on final branching fraction results because we studied the effect only in the $D^0 \to K\pi$ channel and the effect can be different from one `Dmode` to another. Adopting 100% ignorance on the relative composition of the D^{**} background when using the shapes for the χ^2 fit is very conservative. Further measurements of $B \to D^{**}\ell\nu_\ell$ exclusive channels will decrease a lot this systematic effect. This error can be reduced also using specific $B \to D^{**}\ell\nu_\ell$ Monte Carlo for all D^0 decays adopted in the analysis and not only for the $D^0 \to K\pi$ mode as it is now available in *BABAR* Monte Carlo. This uncertainty is one of the dominant, but not the largest, sources of systematic error in this measurement.

[3] we recall that for the basic fit we use the shape coming from the generic $B^0\overline{B}^0$ and B^+B^- Monte Carlo properly weighted to match data luminosity

D^{**} mode	$D^0 \to K\pi\ e$	$D^0 \to K\pi\ \mu$
$B^- \to D_1^0 l^- \bar{\nu}$	-0.063	-0.196
$\bar{B}^0 \to D_1^+ l^- \bar{\nu}$	-0.064	-0.179
$B^- \to D'^0 l^- \bar{\nu}$	0.025	-0.100
$\bar{B}^0 \to D'^+ l^- \bar{\nu}$	0.007	-0.057
$B^- \to D_2^{*0} l^- \bar{\nu}$	-0.073	-0.175
$\bar{B}^0 \to D_2^{*+} l^- \bar{\nu}$	-0.051	-0.161
$B^- \to D_1^{*0} l^- \bar{\nu}$	-0.060	-0.221
$\bar{B}^0 \to D_1^{*+} l^- \bar{\nu}$	-0.098	-0.183
$B^- \to D^{*\prime 0} l^- \bar{\nu}$	-0.010	-0.111
$\bar{B}^0 \to D^{*\prime +} l^- \bar{\nu}$	-0.059	-0.119
$\bar{B}^0 \to D^{*+} \pi^0 l^- \bar{\nu}$	0.033	-0.133
$B^- \to D^{*+} \pi^- l^- \bar{\nu}$	0.053	-0.130
max	0.053	-0.057
min	-0.098	-0.221
1/2 difference	0.076	0.082

Table 6.8: *Absolute difference, measured in %, from fitted value of the branching fraction for different D^{**} shapes hypotheses. The difference are shown separately for the electron and the muon channel and only for the $D^0 \to K\pi$ decay for the reasons explained above.*

6.6.2 Uncertainties from D^{**} fit

We already discussed above the 2% systematics effect from D^{**} modeling. There is also the statistical error coming from $f_{D^{**}}$ extracted from the fit. We propagate it in the branching fraction measurement.

6.6.3 Uncertainties in the number of generated B^0

The method to calculate the number of $\Upsilon(4S)$ and collected in *BABAR* [79] is based on a selection of hadronic events and is used with data collected both on and off resonance. The number of $\Upsilon(4S)$ mesons produced is obtained assuming that the increase in the hadronic to di-muons ratio in on resonance data with respect to off resonance data is due only by the $\Upsilon(4S)$ production. The number of events where a $\Upsilon(4S)$ has been produced satisfying selection criteria is:

$$N_{\Upsilon} = N_{MH} - N_{\mu\mu} \cdot R_{off} \cdot k$$

where:

- N_{MH} is the number of hadronic events in the on resonance data sample passing selection criteria;
- $N_{\mu\mu}$ is the number of di-muons events in the on resonance sample;
- $R_{off} = N_{MH}^{off}/N_{\mu\mu}^{off}$ is the number of hadrons to di-muons events ratio in the off resonance sample;
- $k \sim 1$ is a correction factor that takes into account the cross section and efficiency differences among on and off resonance data taking.

The total number of $\Upsilon(4S)$ mesons produced is given by:

$$N_{\Upsilon}^{tot} = \frac{N_{\Upsilon}}{\epsilon_{B\bar{B}}}$$

where $\epsilon_{B\bar{B}}$ is the selection efficiency of $B\overline{B}$ events extracted from Monte Carlo simulations.

The B counting algorithm applied to our data sample estimates $(38.040 \pm 0.015(\text{stat}) \pm 0.418(\text{syst})) \cdot 10^6 B\overline{B}$ events. Propagating these errors we get the systematic error from luminosity uncertainties in the total number of $B\overline{B}$ pair produced. The uncertainties will be shown in the tables below when summarizing all the systematic uncertainties. They amount for about 1.1% as relative error on the measured $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$.

6.6.4 Uncertainties from $\mathcal{B}(\Upsilon(4S) \to B^0\overline{B}^0)$

From the B counting algorithm we know the number of $\Upsilon(4S)$ produced in the data sample we analyze. The next step consists in extracting the number of $B^0\overline{B}^0$ produced in the on resonance data sample adopted for this measurement. There are two possible approaches:

1. assume that $f_{00} = \mathcal{B}(\Upsilon(4S) \to B^0\overline{B}^0) = f_{+-} = \mathcal{B}(\Upsilon(4S) \to B^+B^-) = 0.5$ and assess no systematic uncertainties on this. This method has been adopted by Belle [19].
2. use the world average measurement on $f_{+-}/f_{00} = 1.072 \pm 0.058$ [5] and extract $f_{00} = 0.483 \pm 0.014$. This gives a relative 2.9% systematic uncertainties on the $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$. CLEO [13] quotes his $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ and $|V_{cb}|$ measurements including this systematic error.

We agree with the second approach so we assume a smaller $\mathcal{B}(\Upsilon(4S) \to B^0\overline{B}^0)$. This implies a larger $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ value and an additional systematic uncertainty if compared with the Belle result.

The correct fraction of $\Upsilon(4S) \to B^0\overline{B}^0$ decays, f_{00}, is an important measurement since it affects all B branching fraction measurements exploited at the $\Upsilon(4S)$. The huge dataset available both in *BABAR* and Belle has to be used to measure with high accuracy f_{+-}/f_{00} since theoretical models generally prefer $f_{00} < f_{+-}$ [5].

We propagate the systematic uncertainty, a relative 2.9%, in the branching fraction we measure as can be seen later.

6.6.5 Uncertainties from the amount of background fractions extracted from δm fit

The background fractions extracted with the fit to the δm distributions, as shown in chapter 5, are quoted with their uncertainties as can be seen in table 5.6. We propagate the uncertainties on these fractions in order to evaluate the systematic uncertainty contribution to the branching ratio measured in all 8 channels (4 D^0 modes and e or μ).

6.6.6 Vertexing algorithm systematic uncertainty studies

The vertexing algorithm adopted to reconstruct the $D^*\ell$ candidate has already been explained in section 4.13.1. Here we want to assess the systematic uncertainty due to the method adopted. Because the vertexing algorithms, when performing the refitting, adjust the D^* four momentum and position, we might also expect to see a dependence of the δm mass difference resolution on the choice of the vertexing algorithm. Following *BABAR* studies made on vertexing control samples [61], we adopt the strategy of looking at different approaches from the one used in this analysis:

- the standard approach, used in this analysis, which treats the lepton, the beamspot, the D^* and its daughters simultaneously to find the best vertex and is known to perform the best on average [62] (D^*, ℓ, beamspot);
- the same approach as before apart from the absence of the beamspot constraint request in the fit (D^*,ℓ);

- the same approach as the standard one apart from the asbsence of the lepton constraint in the fit (D^*, beamspot).

We assess systematic errors redoing from the beginning all the analysis with the last two vertexing methods[4] and evaluating the shift from the reference value of the standard vertexing method. Results are reported in table 6.9. Taking the largest discrepancy, a conservative relative 1% error on the branching ratio is applied to all 8 channels.

channel	$\Delta\mathcal{B}$ no beamspot	$\Delta\mathcal{B}$ no lepton
Kpi e	-0.005	-0.044
Kpi m	-0.003	-0.031
K3pi e	0.000	-0.040
K3pi m	0.000	-0.033
Kpipi0 e	-0.002	-0.016
Kpipi0 m	-0.008	-0.037
Kspipi e	0.008	-0.052
Kspipi m	-0.003	-0.090

Table 6.9: *Summaries of vertexing systematic checks with 9 parameters fit. All the absolute deviations, $\Delta\mathcal{B} = \mathcal{B}_{standard\ vtx} - \mathcal{B}_{new\ vtx}$, from the branching fraction measured with the standard vertexing technique are expressed in %. To assess the systematic effect we choose the largest difference among the reference value and the value obtained with no beamspot and no lepton constraint in the vertexing fit, respectively.*

6.6.7 Lepton momentum cut systematic uncertainty

Another source of systematic error is related to the cut on the lepton momentum in the CM frame. In this analysis we select $p^*_\ell > 1.2\,\mathrm{GeV}/c$. We know, from the discussion on the kinematic cuts, that a different cut on p^*_ℓ implies a different amount of some type of backgrounds like the correlated background and also a different amount of $D^{*-}\ell^+\nu_\ell\ X$ background. A tighter lepton cut can slightly affect the fake-lepton background fraction too.

In order to assess a systematic error on this selection cut we perform the $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ measurement in all the 8 channels with two tighter p^*_ℓ cuts, one at 1.3 GeV/c and one at 1.4 GeV/c. We can see the results of this study in table 6.10. Taking

[4]we recall the methods have different efficiencies, so we were forced to restart from the beginning the anlysis for each of the two novel vertexing techniques

the largest discrepancy a conservative relative 1% error from this systematic effect is applied to the branching ratio in all 8 channels.

channel	$\Delta\mathcal{B}\ p_\ell^* > 1.3\,\mathrm{GeV}/c$	$\Delta\mathcal{B}\ p_\ell^* > 1.4\,\mathrm{GeV}/c$
Kpi e	0.026	-0.041
Kpi m	0.008	-0.061
K3pi e	-0.048	-0.075
K3pi m	0.001	-0.038
Kpipi0 e	-0.037	-0.015
Kpipi0 m	0.025	-0.081
Kspipi e	-0.114	-0.078
Kspipi m	-0.063	-0.216

Table 6.10: *Summaries of p_ℓ^* systematic checks with 9 parameters fit. All the absolute deviations, $\Delta\mathcal{B} = \mathcal{B}_{p_\ell^*>1.2\,\mathrm{GeV}/c} - \mathcal{B}_{other\ p_\ell^*\ cut}$, from the measured branching ratio are expressed in %. To assess the systematic effect we choose the largest difference among the reference value evaluated with the cut $p_\ell^* > 1.2\,\mathrm{GeV}/c$ and the one obtained with $p_\ell^* > 1.3\,\mathrm{GeV}/c$ and $p_\ell^* > 1.4\,\mathrm{GeV}/c$, respectively.*

6.6.8 χ^2 fit binning effects

We perform the fit with different requests on the minimal amount of entries in one bin and assess a systematic effect which gives on average a relative 1% error on the branching fraction measurement. All the variations from the default value (10 entries minimum per bin) are shown in table 6.11 for the 9 parameters fit.

6.7 Branching fraction measurement results

In table 6.12 we show the number of selected events in each of the 8 channels together with the fraction of each background and the fraction of $B^0 \to D^{*-}\ell^+\nu_\ell$ signal, after the fit, in the $\cos\theta_{B^0,D^*\ell}$ range [-1.2,1.2] adopted for the branching fraction measurement.

In table 6.13 we show the list of the $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ evaluated on 8 different channels, together with all the statistical and systematic errors.

In figures 6.8 and 6.9 we show all the plots of the $\cos\theta_{B^0,D^*\ell}$ distributions in each of the 8 channles where the $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ is evaluated. The figures show also the

minimum entries/bin	100	50	40	20	6	0
Kpi e	0.035	0.025	0.025	0.012	-0.011	-0.083
Kpi m	0.036	0.025	0.026	0.012	-0.010	-0.081
K3pi e	0.041	0.028	0.029	0.014	-0.011	-0.091
K3pi m	0.043	0.029	0.030	0.014	-0.013	-0.098
Kpipi0 e	0.035	0.024	0.025	0.012	-0.010	-0.077
Kpipi0 m	0.036	0.025	0.025	0.012	-0.011	-0.082
Kspipi e	0.041	0.028	0.028	0.013	-0.012	-0.090
Kspipi m	0.038	0.028	0.029	0.012	-0.010	-0.074

Table 6.11: *Summary of binning systematic checks with 9 parametrs fit. To asses the systematic effect we choose the largest absolute difference* $\Delta\mathcal{B} = \mathcal{B}_{10\ entries\ per\ bin} - \mathcal{B}_{n\ entries\ per\ bin}$ *for each of the 8 channels. Al the absolute deviations are expressed in %.*

total number of events selected in each channel and the number of $B^0 \to D^{*-}\ell^+\nu_\ell$ signal events estimated from the fit together with the number of events in each background type. These numbers refer to the fit range $-10 < \cos\theta_{B^0,D^*\ell} < 5$.

Channel	N_{data}	f_{D^*}	$f_{D^{**}}$	$f_{\rm fake}$	$f_{\rm uncorr.}$	$f_{\rm corr.}$	$f_{\rm continuum}$	$f_{\rm comb.}$
e-Kpi	6444	85.17%	4.05%	0.16%	2.84%	0.35%	1.32%	6.11%
mu-Kpi	5286	85.58%	3.99%	1.89%	1.78%	0.55%	1.73%	4.49%
e-K3pi	4429	72.96%	3.98%	0.09%	1.59%	0.40%	1.51%	19.48%
mu-K3pi	3457	74.35%	4.43%	1.42%	1.21%	0.48%	0.88%	17.24%
e-Kpipi0	5676	76.10%	3.56%	0.14%	2.36%	0.28%	2.27%	15.28%
mu-Kpipi0	4594	78.21%	3.75%	1.69%	2.11%	0.50%	1.14%	12.61%
e-Kspipi	670	70.48%	3.85%	0.00%	1.35%	0.18%	1.52%	22.61%
mu-Kspipi	485	71.98%	3.58%	1.66%	1.40%	0.16%	0.69%	20.52%

Table 6.12: *Number of data selected events (N_{data}) in each of the 8 signal channles divided by D^0 mode and lepton flavor. We show also the fractions of each background type in all 8 channels and the fraction of signal events selected after the fit, f_{D^*}. All these numbers are used to extract N_s and N_{bkg} used in equation 6.6.*

In table 6.14 we show the $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ calculated for different division of the 8 channles analyzed: by D^0 mode, by ℓ flavor and the average of all the 8 channels. In figure 6.10 we show the results on $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ we obtain in this analysis. In figure 6.11 we show some of the results we obtain compared with the measurements performed by other experiments: we show only the results of the $D^0 \to K\pi$ mode both for electrons and muons since exclusive $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ measurements performed by CLEO and Belle use only the $D^0 \to K\pi$ mode. In this way we have a better comparison between this measurement and the results of

	$\mathcal{B}$	stat DATA	stat MC	syst δm FIT	syst PID eff.	syst TRK eff.	syst VTX	syst p^*_ℓ	syst BIN	syst D^{**} fit	syst D^{**} model	syst L	syst BR D^0	syst BR D^*	syst f_{00}
e-Kpi	4.945	0.062	0.032	0.039	0.009	0.115	0.049	0.049	0.049	0.021	0.099	0.054	0.117	0.037	0.138
mu-Kpi	4.895	0.067	0.032	0.057	0.024	0.114	0.049	0.049	0.049	0.021	0.098	0.054	0.116	0.036	0.137
e-K3pi	4.743	0.071	0.034	0.048	0.057	0.126	0.047	0.047	0.047	0.025	0.095	0.052	0.197	0.035	0.133
mu-K3pi	4.583	0.078	0.033	0.052	0.059	0.122	0.046	0.046	0.046	0.027	0.092	0.050	0.190	0.034	0.128
e-Kpipi0	4.728	0.063	0.032	0.050	0.057	0.299	0.047	0.047	0.047	0.020	0.095	0.052	0.325	0.035	0.132
mu-Kpipi0	4.791	0.071	0.032	0.054	0.062	0.303	0.048	0.048	0.048	0.021	0.096	0.053	0.329	0.035	0.134
e-Kspipi	4.762	0.184	0.129	0.049	0.006	0.162	0.048	0.048	0.048	0.035	0.095	0.052	0.282	0.035	0.133
mu-Kspipi	4.306	0.196	0.118	0.050	0.021	0.147	0.043	0.043	0.043	0.030	0.086	0.047	0.255	0.032	0.121

Table 6.13: *Measured $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$, statistical and systematics uncertainties for all the 8 signal channels in the 9 free parameters fit.*

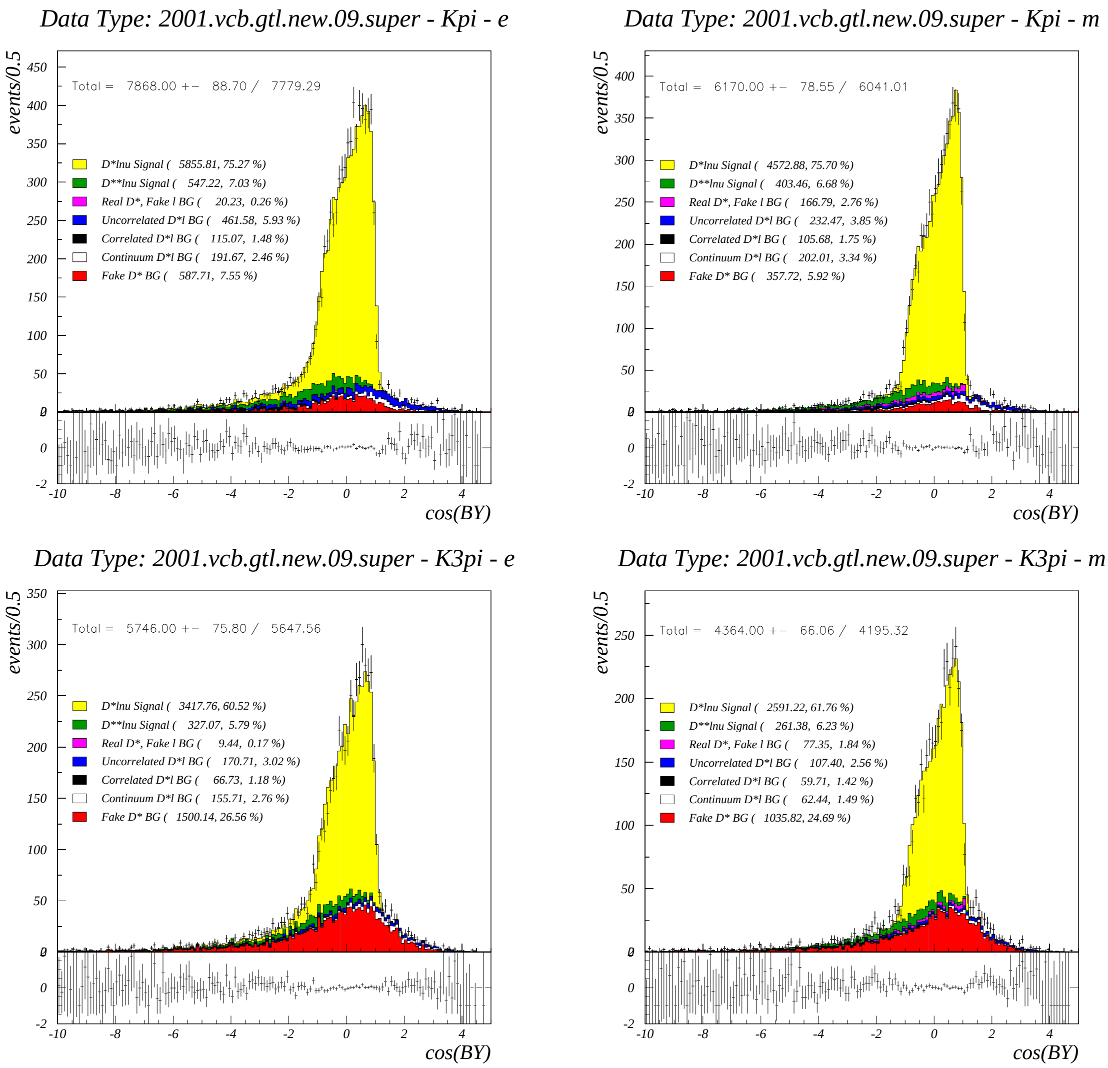


Figure 6.8: Results of the χ^2 fits to extract amount of D^{**} background and amount of the total sample selected with corrected backgrounds in $D^0 \to K\pi$ and $D^0 \to K\pi\pi\pi\ \ell$ channels in the 9 free parameters fit. Each plot shows also the data/Monte Carlo ratio for each of the 150 bins.

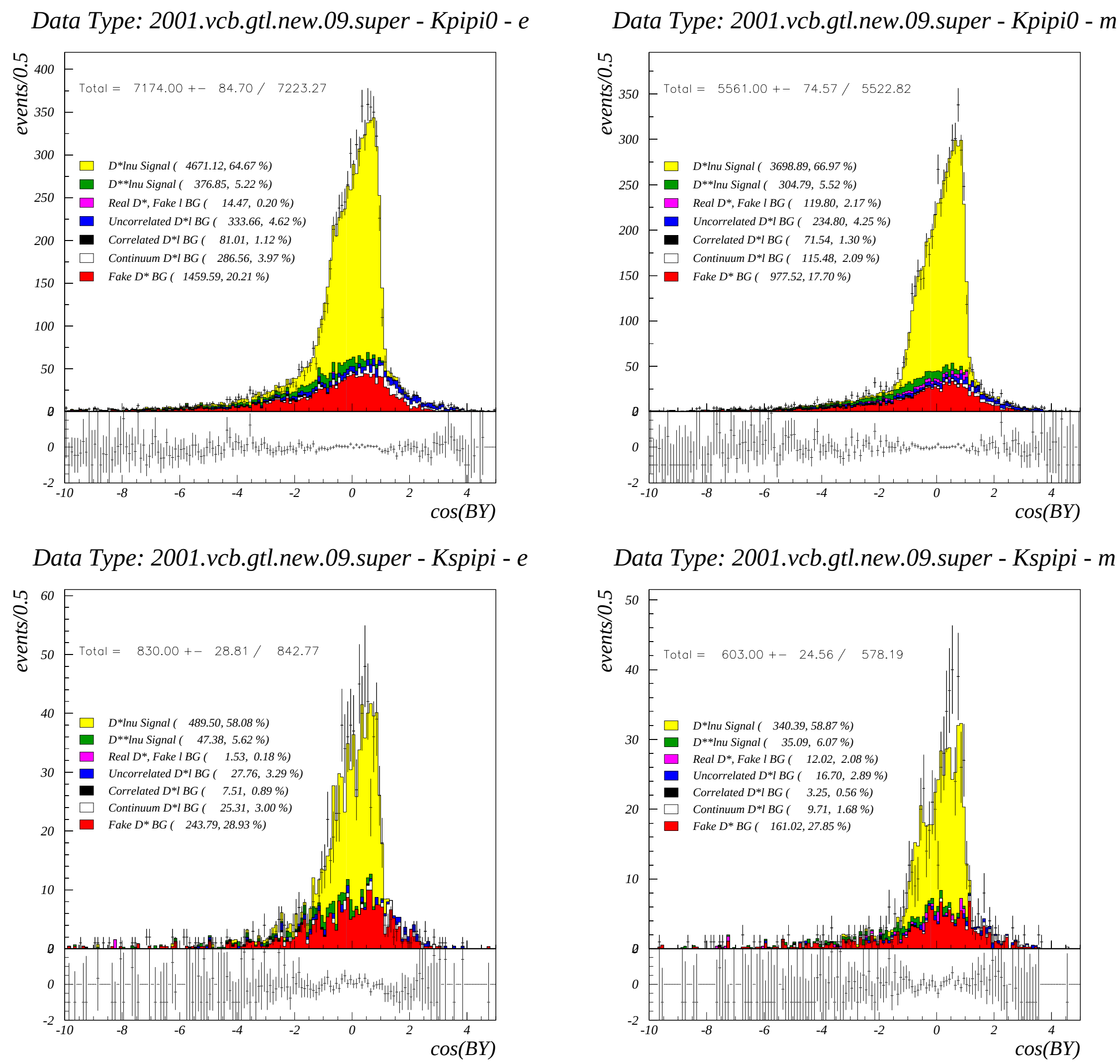


Figure 6.9: Results of the χ^2 fits to extract amount of D^{**} background and amount of the total sample selected with corrected backgrounds in $D^0 \to K\pi\pi^0$ and $D^0 \to K_S\pi\pi\ \ell$ channels in the 9 free parametrs fit. Each plot shows also the data/Monte Carlo ratio for each of the 150 bins.

	$\mathcal{B}$	stat DATA	stat MC	syst δm FIT	syst PID eff.	syst TRK eff.	syst VTX	syst p^*_ℓ	syst BIN	syst D^{**} fit	syst D^{**} model	syst L	syst BR D^0	syst BR D^*	syst f_{00}
$D^0 \to K\pi$	4.898	0.047	0.024	0.034	0.012	0.115	0.049	0.049	0.049	0.015	0.098	0.054	0.117	0.036	0.137
$D^0 \to K\pi\pi\pi$	4.632	0.055	0.025	0.037	0.041	0.124	0.046	0.046	0.046	0.019	0.093	0.051	0.193	0.034	0.130
$D^0 \to K\pi\pi^0$	4.740	0.050	0.025	0.039	0.042	0.300	0.047	0.047	0.047	0.016	0.095	0.052	0.326	0.035	0.132
$D^0 \to K_S\pi\pi$	4.476	0.139	0.088	0.036	0.012	0.153	0.045	0.045	0.045	0.023	0.089	0.049	0.266	0.033	0.125
all e	4.785	0.044	0.024	0.027	0.029	0.074	0.047	0.047	0.047	0.014	0.096	0.052	0.092	0.036	0.134
all μ	4.707	0.049	0.025	0.031	0.038	0.073	0.047	0.047	0.047	0.014	0.094	0.052	0.091	0.035	0.132
all channels	4.746	0.037	0.020	0.021	0.022	0.074	0.047	0.047	0.047	0.010	0.095	0.052	0.091	0.035	0.133

Table 6.14: *Measured $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$, statistical and systematics uncertainties obtained grouping the 8 channles in: different D^0 modes, all e or all μ modes and considering all channels together. All the errors and branching fraction values are quoted weigthing properly all the systematic correlated and uncorrelated errors in each of the 8 channles adopted for this measurement. All these numbers refers to the D^{**} background amount evaluated with the 9 free parameters fit.*

these experiments. We show also the average of our measurement performed on all 8 channels.

As can be seen in figure 6.11 the latest CLEO measurement of $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ with exclusive reconstruction [13] is out more than three standard deviations with respect to the PDG2002 world average. CLEO claims that this measurement supercedes the old ones they published. For this reason we decided not to show previous CLEO published measurements [14, 17] in figure 6.11. It is important to know that the PDG2002 world average is evaluated with the old published CLEO measurements and not with the latest one. We decided to compare the results of our $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ measurement with PDG2002 world average [5] and wait for the new offcial world average that will be done after the 2003 winter conferences preliminary results will be available.

Our measurement agrees very well with PDG2002 world average and the uncertainties we quote are almost of the same size of the world average uncertainties.

6.8 Future perspectives

This measurement is performed on a subset of the available *BABAR* datasample. It is evident from table 6.14 and table 6.13 that this measurement is dominated by systematic uncertainties hence not using the whole *BABAR* dataset is not a limiting factor on the precision of the measurement. For the 2003 winter conferences we will release this mesurement with the entire *BABAR* dataset. We already analyzed about 80 fb^{-1} of *BABAR* on-resonace data and we are finalizing the systematic

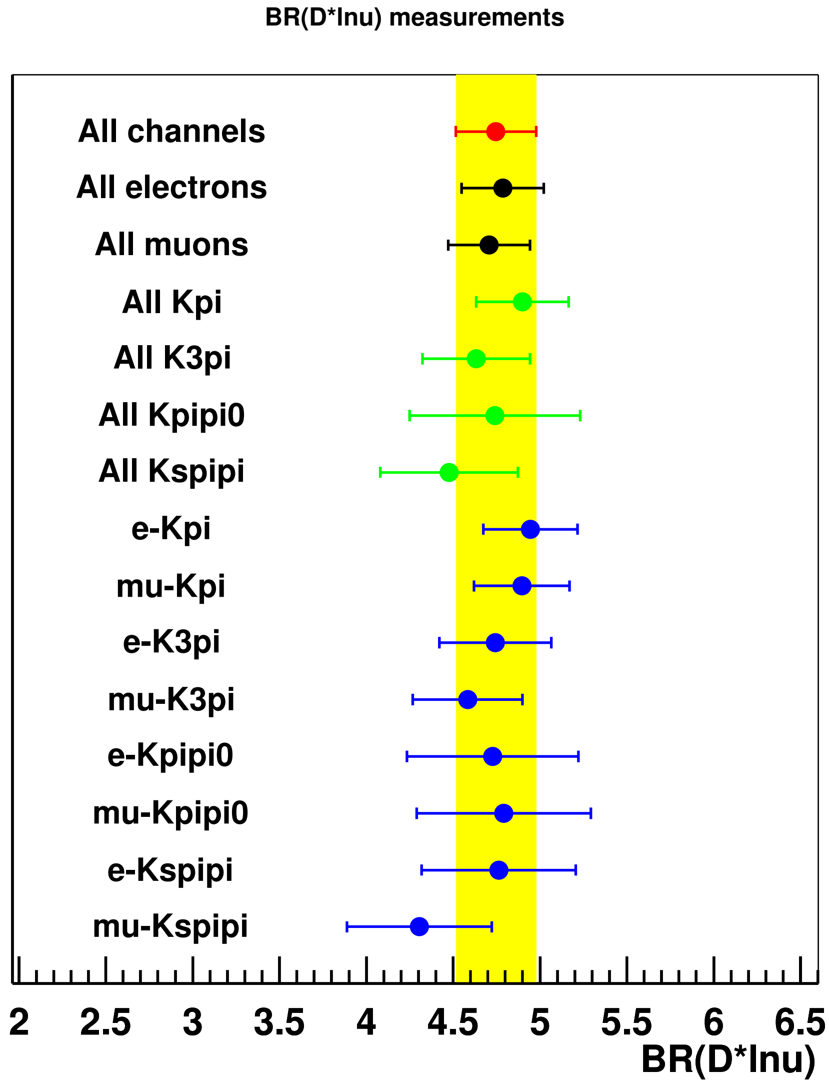


Figure 6.10: Summary of all $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ measurements performed in this analysis. We show, from top to bottom, the branching fraction measurement of: the average of all 8 signal channles (in red), the average of all 4 electron channels (in black), the average of all 4 muon channels (in black), the average of each `Dmode` independently of the the lepton flavor (in green) and all the 8 channels measurements (in blue). They are all consistent among each other within one standard deviation.

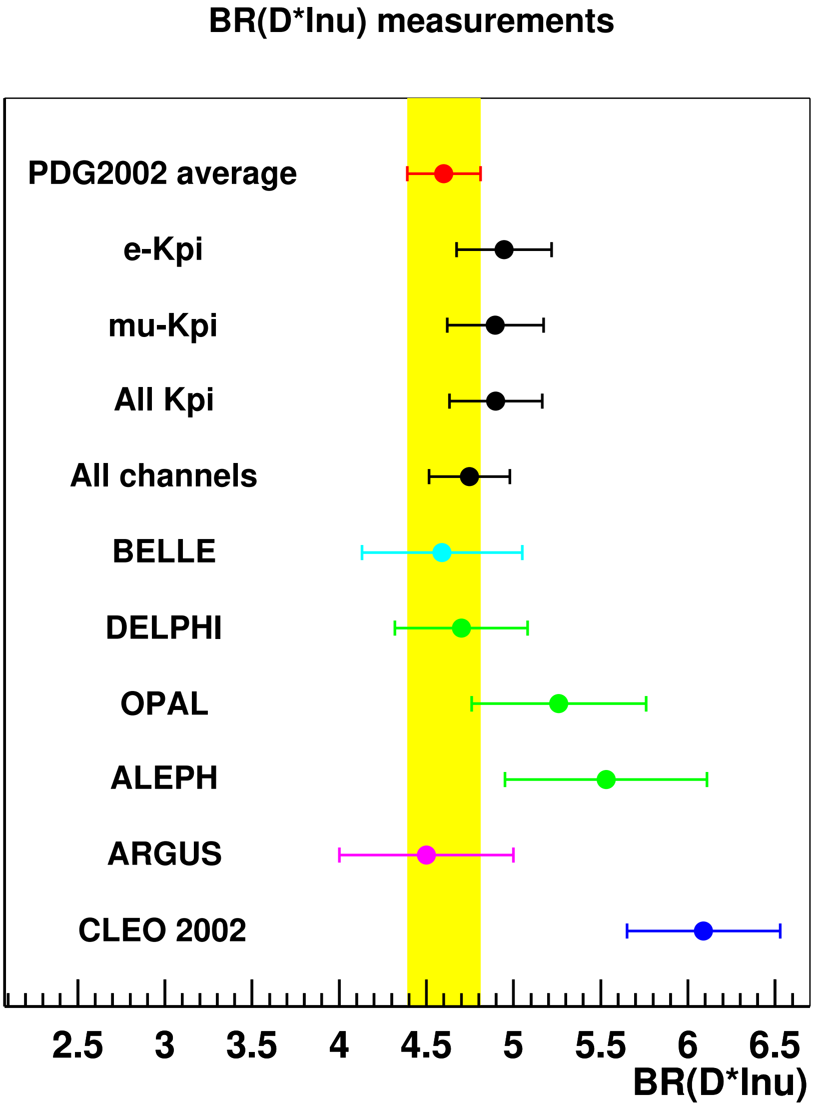


Figure 6.11: Summary of all $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ available nowadays compared with the PDG2002 world average value. We show also the results of our measurement for the electron $D^0 \to K\pi$, for the muon $D^0 \to K\pi$ channels and for the average of the $D^0 \to K\pi$ independently of the lepton flavor. We show also our average on all channels. As can be seen they are all consistent with the world average within one standard deviation. We decided to show the $D^0 \to K\pi$ channles since Belle and CLEO experiments perform exclusive measurements of $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ only in the $D^0 \to K\pi$ mode. Belle experiment uses only the electron and measure $\mathcal{B}(B^0 \to D^{*-}e^+\nu_e)$ only. As can be seen we show only the latest CLEO result since they claim that this result superceed all their previous $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ measurements. The PDG2002 world average is not performed with the latest CLEO published measurement.

checks [80, 81]. There are two main reasons for publishing this result with all the data available so far in *BABAR*:

1. the new PDG world average will be dominated by our result: the 80 fb^{-1} measurement of $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ uses 4 times the amount of Belle result and 7 channels more than the only electron lepton flavor and $D^0 \to K\pi$ mode they use;

2. the next step of this analysis is the $|V_{cb}|$ extraction. This measurement is very sensitive to the amount of events in the low $w = v_{D^*} \cdot v_B$ region. Using the whole *BABAR* data set we can decrease the statistical error on the amount of selected events in the low w region and consequently measure $|V_{cb}|$ with a smaller uncertainty.

The results on $|V_{cb}|$ and $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ we will provide for the 2003 winter conferences will be the most precise measurements of these quantities performed so far.

Our results are also important from the theoretical point of view. After publishing their latest results both on $|V_{cb}|$ and $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$, CLEO claims that their measured values (more than three standard deviations far from the world average in both measurements) are not in agreement with HQET and that HQET is not anymore a proper description of semileptonic B decays.

As can be inferred from our measurements we are in perfect agreement with HQET predictions and with almost all the other published results, CLEO excluded.

Our results confirm that HQET still gives a theoratically consistent picture of semileptonic B decays.

Conclusions

We present the measurement of the $B^0 \to D^{*-}\ell^+\nu_\ell$ branching fraction performed with about 38×10^6 B^0 mesons produced in $\Upsilon(4S)$ decays and collected at the *BABAR* experiment at the SLAC PEP-II asymmetric B Factory.

The $B^0 \to D^{*-}\ell^+\nu_\ell$ decay (and its charge-conjugate mode) is reconstructed in the following decay channels:

$$B^0 \to D^{*-}\ell^+\nu_\ell \; (D^* \to D^0\pi \; , \; D^0 \to K\pi \; , \; D^0 \to K\pi\pi\pi \; , \; D^0 \to K\pi\pi^0 \; , \; D^0 \to K_S\pi\pi \;).$$

We reconstruct both electrons and muons for a total of 8 different exclusive decay channels (4 D^0 modes and both e and μ).

The results of the measurement performed using only the $D^0 \to K\pi$ mode are:

$$\begin{aligned}
\mathcal{B}(B^0 \to D^{*-}e^+\nu_e) &= (4.945 \pm 0.070(\text{stat}) \pm 0.263(\text{syst}))\% \\
\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu) &= (4.895 \pm 0.074(\text{stat}) \pm 0.266(\text{syst}))\%
\end{aligned}$$

The results of the measurement in the electron channel and the muon channel, properly averaged over all D^0 decay modes, are:

$$\begin{aligned}
\mathcal{B}(B^0 \to D^{*-}e^+\nu_e) &= (4.785 \pm 0.050(\text{stat}) \pm 0.231(\text{syst}))\% \\
\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu) &= (4.707 \pm 0.055(\text{stat}) \pm 0.230(\text{syst}))\%
\end{aligned}$$

Averaging over all 8 channels we get:

$$\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell) = (4.746 \pm 0.232)\% \; ,$$

to be compared with the PDG2002 world average value [5]:

$$\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)_{PDG} = (4.60 \pm 0.21)\% \; .$$

Our result agrees with the world average within one standard deviation. This measurement has been performed with the largest $B^0 \to D^{*-}\ell^+\nu_\ell$ data sample reconstructed ever. The result is of great intereset because, in 2002, a new result on $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ published by the CLEO collaboration [13] showed a more than 3 standard deviation effect from the $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ world average value, they measure:

$$\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)_{CLEO} = (6.09 \pm 0.44)\% \ .$$

As a consequence CLEO claimed that Heavy Quark Effective Theory (HQET) is no longer adequate to describe semileptonic B decays. Our result goes in the opposite direction confirming HQET predictions; this measurement has the smallest uncertainty obtained by a single measurement so far. Our result is supported by all other published measurements and in particular is consistent within one standard deviation with the $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ measured by the Belle collaboration (an asymmetric B Factory as *BABAR*). The selected number of events in the Belle analysis is the second largest sample ever analyzed. We conclude that both asymmetric B Factories results on semileptonic B decays do not show any deviation from HQET predictions.

The next step of this analysis is the measurement of the CKM matrix element $|V_{cb}|$ with a fit to the differential $B^0 \to D^{*-}\ell^+\nu_\ell$ decay distribution based on HQET predictions of the from factors of this decay.

Our measurement is under review inside the *BABAR* collaboration in order to be published as a preliminary result for the 2003 winter conferences.

Appendix A

Form Factors parametrization in HQET framework

HQET is written in terms of the four-velocities v_B and v_{D^*} instead of the four-momenta p_B and p_{D^*}. The $B^0 \to D^{*-}\ell^+\nu_\ell$ matrix element can be written as:

$$\begin{aligned}\langle D^*(v_{D^*},\epsilon)|(V^\mu - A^\mu)|B(v_B)\rangle &= \sqrt{m_B m_{D^*}}[ih_V(w)\epsilon^{\mu\nu\alpha\beta}\epsilon^*_\nu(v_{D^*})_\alpha(v_B)_\beta \\ &- h_{A_1}(w)\epsilon^{*\mu}(w+1) + h_{A_2}(w)(v_B)^\mu\epsilon^*\cdot v_B \\ &+ h_{A_3}(w)(v_{D^*})^\mu\epsilon^*\cdot v_B]\ . \end{aligned} \tag{A.1}$$

This equation contains four HQET form factors, $h_V(w)$, $h_{A_1}(w)$, $h_{A_2}(w)$ and $h_{A_3}(w)$. However, there are only three independent form factors, as $h_{A_2}(w)$ and $h_{A_3}(w)$ always occur in observables in the combination

$$h_{A_3}(w) + rh_{A_2}(w)\ , \tag{A.2}$$

where $r = m_{D^*}/m_B$ again. These HQET form factors are related to $V(q^2)$, $A_1(q^2)$ and $A_2(q^2)$, defined in section 1.3, by [8]:

$$\begin{aligned}\frac{2\sqrt{m_B m_{D^*}}}{m_B + m_{D^*}}V(q^2) &= h_V(w)\ , \\ \frac{m_B + m_{D^*}}{2\sqrt{m_B m_{D^*}}}A_1(q^2) &= \frac{w+1}{2}h_{A_1}(w)\ , \\ \frac{2\sqrt{m_B m_{D^*}}}{m_B + m_{D^*}}A_2(q^2) &= h_{A_3}(w) + rh_{A_2}(w)\ . \end{aligned} \tag{A.3}$$

Following Neubert [16], we write two ratio of form factors:

$$R_1(w) \equiv \frac{h_V(w)}{h_{A_1}(w)} = \left[1 - \frac{q^2}{(m_B + m_{D^*})^2}\right] \frac{V(q^2)}{A_1(q^2)} \,, \tag{A.4}$$

$$R_2(w) \equiv \frac{h_{A_3}(w) + r h_{A_2}(w)}{h_{A_1}(w)} = \left[1 - \frac{q^2}{(m_B + m_{D^*})^2}\right] \frac{A_2(q^2)}{A_1(q^2)} \,. \tag{A.5}$$

These ratios are motivated by heavy-quark symmetry considerations. In the heavy-quark symmetry limit, the HQET form factors have the relations:

$$h_V(w) = h_{A_1}(w) = h_{A_3}(w) = \xi(w) \,, \tag{A.6}$$

$$h_{A_2}(w) = 0 \,, \tag{A.7}$$

where $\xi(w)$ is the Isgure-Wise function. In this limit, the form factor ratios are simply unity for all w. Outside the heavy-quark symmetry limit, $R_1(1)$ and $R_2(1)$ are corrected to values away from unity, just as $\mathcal{F}(1)$ differs from unity. However, the expected w dependence of $R_1(w)$ and $R_2(w)$ is very weak.

The helicity amplitudes $H_0(w)$, $H_\pm(w)$, defined in section 1.3, can be expressed in terms of $h_{A_1}(w)$, $R_1(w)$ and $R_2(w)$. These relations are [16]:

$$H_0(w) = (m_B - m_{D^*})\sqrt{\frac{m_B m_{D^*}}{q^2(w)}}(w+1)h_{A_1}(w)\left[1 + \left(\frac{w-1}{1-r}\right)(1 - R_2(w))\right] \tag{A.8}$$

and

$$\begin{aligned} H_\pm(w) &= (m_B - m_{D^*})\sqrt{\frac{m_B m_{D^*}}{q^2(w)}}(w+1)h_{A_1}(w) \\ &\times \frac{\sqrt{1 - 2wr + r^2}}{1-r}\left[1 \mp \sqrt{\frac{w-1}{w+1}}R_1(w)\right] \,. \end{aligned} \tag{A.9}$$

Appendix B

Vertex reconstruction: formalism and strategy

The general vertexing algorithm is based on the generalized least squares method using the well-known Lagrange Multiplier technique [58]. In this method it is assumed that the constraint equations can be linearized and summarized in three matrices, labeled as $\tilde{A}$, $\tilde{B}$ and $\vec{c}$.

Let $\vec{\eta}$ represent the measurable quantities put in an vector of dimension n. The actually measured quantities $\vec{y}$ (for instance the measure track parameters) deviate from $\vec{\eta}$ by errors $\vec{\delta}$. We assume that the errors δ_i, $i = 1, ..., n$, are normally distributed. There are r unknowns grouped in a vector $\vec{x}$. The $\vec{x}$ and $\vec{\eta}$ are related by m functions (constraints),

$$f_k(\vec{x}, \vec{\eta}) = f_k(\vec{x}, \vec{y} + \vec{\delta}) = 0, \quad k = 1, ..., m \tag{B.1}$$

Assuming that functions f_k can be linearized at $\vec{x} = \vec{x}_0$ and $\vec{\eta} = \vec{\eta}_0$, the χ^2 anszat can be written as [58]:

$$\chi^2 = \vec{\delta}^T \tilde{W}_y \vec{\delta} + 2\vec{\mu}^T \left(\tilde{A}\vec{\xi} + \tilde{B}\vec{\delta} + \vec{c} \right) \tag{B.2}$$

where $\vec{\mu}$ is the vector of Lagrange multipliers, and

$$\vec{\xi} = \vec{x} - \vec{x_0} \tag{B.3}$$

$$\vec{\delta} = \vec{\eta} - \vec{\eta_0} \tag{B.4}$$

$\tilde{W}_y$ is the weight matrix of parameters. Matrices $\tilde{A}$ and $\tilde{B}$ are, respectively, the derivatives (estimated at point $\vec{x}_0$, $\vec{\eta}_0$) of the constraints with respect to the unknowns and parameters:

$$a_{kl} = \left(\frac{\partial f_k}{\partial x_l}\right)_{\vec{x}_0,\vec{\eta}_0} \quad , \qquad \tilde{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1r} \\ a_{21} & a_{22} & \dots & a_{2r} \\ \dots & & & \\ a_{m1} & a_{m2} & \dots & a_{mr} \end{pmatrix} \tag{B.5}$$

$$b_{kl} = \left(\frac{\partial f_k}{\partial \eta_l}\right)_{\vec{x}_0,\vec{\eta}_0} \quad , \qquad \tilde{B} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & & & \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix} \tag{B.6}$$

$\vec{c}$ is the vector of values of the constraints at point $\vec{x}_0$, $\vec{\eta}_0$,

$$c_k = f_k(\vec{x}_0, \vec{\eta}_0) \ , \qquad \vec{c} = \begin{pmatrix} c_1 \\ c_2 \\ \dots \\ c_m \end{pmatrix} \tag{B.7}$$

As a first approximation, we take $\vec{\eta}_0 = \vec{y}$. The expansion point requires also $\vec{x}_0$. Is here where the first guess enters in: a common issue in vertexing is that an initial guess of the location of the vertex is required. Usually the primary interaction point can be used, but the convergence radius is of the order of a few centimeters. This is good enough for decays of short-lived particles (B and D) but not for long-lived particles (V^0's). This problem is handled generally for all cases by solving analytically (with an iterative procedure) the point of closest approach of the two tracks or candidates (both with defined spatial information), using a second-order approximation for the track at each point. When more than two tracks appear, the point of closest approach of the closest pair of tracks is used. No cut on the minimal distance of closest approach is applied. When the search for the poca point fails, the *BABAR* origin is assumed and the vertex is still attempted. In the case of neutrals, the *BABAR* origin is assumed by default for estimating its momentum.

Resolving (B.2) for unknowns and parameters [58], we get:

$$\vec{\xi} = -\left(\tilde{A}^T \tilde{G}_B \tilde{A}\right)^{-1} \tilde{A}^T \tilde{G}_B \vec{c} \tag{B.8}$$

$$\vec{\delta} = -\tilde{W}_y^{-1} \tilde{B}^T \tilde{G}_B \left(\vec{c} + \tilde{A}\vec{\xi}\right) \tag{B.9}$$

where

$$\tilde{G}_B = \left(\tilde{B}\tilde{W}_y^{-1}\tilde{B}^T\right)^{-1} \tag{B.10}$$

The covariance matrices for unknowns and parameters are:

$$\tilde{G}_x^{-1} = \tilde{G}_\xi^{-1} = \left(\tilde{A}^T \tilde{G}_B \tilde{A}\right)^{-1} \tag{B.11}$$

$$\tilde{G}_\eta^{-1} = \tilde{W}_y^{-1} - \tilde{W}_y^{-1}\tilde{B}^T\tilde{G}_B\tilde{B}\tilde{W}_y^{-1} + \tilde{W}_y^{-1}\tilde{B}^T\tilde{G}_B\tilde{A}\tilde{G}_\xi^{-1}\tilde{A}^T\tilde{G}_B\tilde{B}\tilde{W}_y^{-1} \tag{B.12}$$

and the correlation matrix between unknowns and parameters is:

$$\tilde{G}_{x\eta}^{-1} = -\tilde{W}_y^{-1}\tilde{B}^T\tilde{G}_B\tilde{A}\tilde{G}_\xi^{-1} \tag{B.13}$$

The minimum of (B.2) can be written as:

$$\chi^2_{min} = \vec{\delta}^T \left(\tilde{B}^T\tilde{G}_B\tilde{B}\right)\vec{\delta} \tag{B.14}$$

which follows a χ^2-distribution with $m-r$ degrees of freedom. Equation (B.14), projected over a "group" of parameters allow us to estimate the contribution of it to the total χ^2. By "group" of parameters we mean any set of parameters for which the covariance matrix $\tilde{W}_y^{-1}$ is block diagonal.

When (B.1) are already linear, equations (B.3), (B.4), (B.11) and (B.12) give the final solution. In general constraints are not linear, and then we need a better approximation, which can be obtained replacing $\vec{x}_0$, $\vec{\eta}_0 = \vec{y}$ by $\vec{x}$, $\vec{\eta}$ into (B.5), (B.6) and (B.7). This iteration process is repeated until a satisfactory solution is found. The default convergence criteria consist in requiring a change in χ^2, as given by equation (B.14), between two successive iterations less than 0.005, with a maximum of 6 iterations.

The equations and iterative procedure described above are implemented in the class `VtxB0_DstarlnuAlgorithm` in the VtxFitter package of *BABAR* analysis software.

In the particular case that parameters are candidates expressed in the coordinate-momentum representation, the covariance matrix $\tilde{W}_y^{-1}$ will be block diagonal, and each block can be written as

$$\tilde{W}_{y,i}^{-1} = \begin{pmatrix} \tilde{W}_{xx,i}^{-1} & \tilde{W}_{xp,i}^{-1} \\ \tilde{W}_{xp,i}^{-1} & \tilde{W}_{pp,i}^{-1} \end{pmatrix} \tag{B.15}$$

where $\tilde{W}_{xx,i}^{-1}$, $\tilde{W}_{xp,i}^{-1}$ and $\tilde{W}_{pp,i}^{-1}$ are the 3×3 covariance matrices for position and momentum. It should be noted that with this approach the position-momentum correlations are accounted for, which is important for kinematic fitting[1] but less relevant for pure vertex fitting.

[1] e.g. the dramatic improvement in δm resolution in D^* decays is possible thanks to these correlations.

The iterative procedure for track selection removes the candidate with the largest contribution to the total χ^2. The procedure is iterated until there are no tracks contributing more than 6 units to the χ^2 or only n tracks remain, where n (which can be 0) is a parameter to be tuned. The best value of parameter n has been determined studying a sample of $B^- \rightarrow D^0(K^-\pi^+)\pi^-$ events. In order to keep as high as possible the reconstruction efficiency, $n = 0$ was adopted as default configuration. In such a configuration it has been verified that applying a cut on χ^2 probability at 0.1% has basically no effect (the cut is somehow implicit in the track removal strategy), beyond the small drop in efficiency which is acceptable for the purposes and large data sample used for this analysis. In some cases we tightened the cut on vertex χ^2 probability to 1%.

Appendix C

Dalitz Plot Calculation

For a three-body decay $M \to a + b + c$, the full kinematics can be described by two parameters; e.g., the invariant mass of $a + b$ and $a + c$ systems (m_{ab}, m_{ac})[1]. The phase space factor is uniform on (m_{ij}, m_{ik}) plane within the kinematically allowed range, where $i \neq j \neq k$, $i \neq k$. The event distribution in this plane is called the Dalitz plot. If a fraction of this three-body decay amplitude comes from a resonance that decays to particles 1 and 2; i.e., $M \to (ab) + c \to a + b + c$, a peak will show up in $a + b$ invariant mass distribution and a higher-density band perpendicular to m_{ab} axis will show up in the Dalitz plot. By studying the Dalitz plot, one can extract the fractions of decay amplitudes and phases through resonances.

For the analysis described in this thesis, the known resonance structure on the Dalitz plot is used to select events within the high density regions in order to increase signal to background ratios (see Sec. 4.10). In this appendix, we document the formulas and parameters (based on the E687 experiment at the Fermilab [56]) used to calculate the expected Dalitz plot density.

The total amplitude of a three-body decay can be expressed as

$$A_{\text{total}} = A_0 e^{i\phi_0} + \sum_i A_i e^{i\phi_i} B_{\text{bw}} S \,,$$

where the first term represents the non-resonant component, A_i and ϕ_i are the amplitude and phase for resonance i, respectively, B_{bw} is the Breit-Wigner function, and S is an angular correlation function due to the spin of the resonance. In particular, for a resonance $r \to a + b$, B_{bw} can be written as

$$B_{\text{bw}}(a, b|r) = \frac{F_D F_r}{m_r^2 - m_{ab}^2 - i\Gamma m_r} \,,$$

[1] see, for example, a review in [5]

where m_r is the peak of the resonance r, Γ an energy-dependent width, and F_D and F_r are Blatt-Weisskopf peneration factors [70]. The angular functions S are different for different resonance spin:

$$\begin{aligned} &S = 1 && \text{for spin} = 0\,, \\ &S = 2p_b p_c \cos\theta_{bc} && \text{for spin} = 1\,, \\ &S = 2p_b^2 p_c^2 (3\cos^2\theta_{bc} - 1) && \text{for spin} = 2\,, \end{aligned}$$

where p_b and p_c are the magnitudes of three-momentum of b and c and θ_{bc} is the angle between $\vec{p}_b$ and $\vec{p}_c$, all measured in the $a+b$ rest frame.

The factors F_D and F_r are chosen to be

$$\begin{aligned} &F \sim 1 && \text{for spin} = 0\,, \\ &F \sim \frac{1}{\sqrt{R^2 p_b^2}} && \text{for spin} = 1\,, \\ &F \sim \frac{1}{\sqrt{9 + 3R^2 p_b^2 + R^4 p_b^4}} && \text{for spin} = 2\,, \end{aligned}$$

where $R = 1.5\ \text{GeV}^{-1}$ is the Blatt and Weiskopf meson radius[2].

The energy-dependent width for spin=s is

$$\Gamma = \Gamma_0 \frac{m_r}{m_{ab}} \left(\frac{p_b}{p_b^R} \right)^{2s+1} \frac{F^2}{F_R^2}\,,$$

where p_b^R is the b momentum in the $a+b$ rest frame when $m_{ab} = m_r$, and F_R is the F factor when p_b is substituted by p_b^R.

The parameters used to calculate the Dalitz plot density are taken from the E687 result [56], and are shown in Table C.1. The results presented [56] are in terms of decay fractions, not amplitude A. Here we have normalized the amplitudes so that the ratio of the integration of the signal intensity for a certain mode to the integration with all modes combined matches the decay fractions in [56].

[2] In principle this value should be different for different resonances, but the final result is not sensitive to the choice of R.

$D^0 \to K^- \pi^+ \pi^0$					
resonance	A	ϕ	Γ_0 (GeV)	m_r (GeV)	spin
K^{*0}	0.0862	$-2°$	0.0507	0.8961	1
ρ^+	0.1942	$0°$	0.1502	0.7693	1
K^{*-}	0.0813	$162°$	0.0508	0.8917	1
non-res	0.1784	$-122°$	-	-	-
$D^0 \to \overline{K}^0 \pi^+ \pi^-$					
resonance	A	ϕ	Γ_0 (GeV)	m_r (GeV)	spin
K^{*-}	0.1703	$0°$	0.0508	0.8917	1
K^{*0} (1430)	0.1222	$-166°$	0.294	1.412	0
ρ^0	0.1324	$-136°$	0.1502	0.7693	1
$f_0(975)$	0.0283	$38°$	0.07	0.98	0
$f_2(1270)$	0.6032	$-174°$	0.1851	1.2754	2
$f_0(1400)$	0.1622	$-45°$	0.35	1.4	0

Table C.1: *Parameters used for Dalitz plot calculations.*

Appendix D

Summary of final selection cuts

In this appendix we summarize our final selection cuts.

Our final event selection criteria are:

- $|\cos\theta^*_{\rm thrust}| < 0.85$ where $\theta^*_{\rm thrust}$ is the angle between the thrust of the $D^*\ell$ candidate and that of the rest of the event.
- The lepton candidate is selected from GoodTracksTight list.
- The lepton momentum in the CM frame satisfies $p^*_\ell > 1.2$ GeV/c.
- For e and μ samples (Table 4.2), the lepton candidate passes the veryTight e or μ selector. For fake samples, the lepton candidate fails both e and μ Loose selectors.
- π^0 is reconstructed from two photons with raw invariant mass within ±15.75 MeV of the PDG π^0 mass. The π^0 is fit with a mass contraint and has χ^2 probability greater than 1%.
- K^0_S is reconstructed from a pair of charged π's (from ChargedTracks list) with raw invariant mass within ±15 MeV of the PDG K^0_S mass. The vertexing χ^2 probability is greater than 1%.
- For the charged daughters of D^0 candidates, π's are selected from GoodTracksLoose, and K is selected from GoodTracksLoose list for $K\pi$, $K\pi\pi\pi$, $K\pi\pi^0$ mode.
- Raw D^0 mass is within ±17 MeV for $K\pi$, $K\pi\pi\pi$ and $K_S\pi\pi$ and ±34 MeV for $K\pi\pi^0$ of the PDG D^0 mass. D^0 is vertexed with mass contraint and the χ^2 probability is greater than 0.1%.

- For $K\pi\pi^0$ and $K_S\pi\pi$ modes, the Dalitz probability density[1] is greater than 0.1.

- Charged kaon passes Tight criterion for $K\pi\pi\pi$ and $K\pi\pi^0$, or notPion for $K\pi$ mode using KaonSMSSelector.

- Soft π is selected from GoodTracksVeryLoose list.

- Soft π satisfies $p_{\pi_{\rm soft}} < 450$ MeV/c, and $p^t_{\pi_{\rm soft}} > 50$ MeV/c.

- Momentum of the D^* in the CM frame satisfies 0.5 GeV/$c < p^*_{D^*} < 2.5$ GeV/c.

- D^*-D^0 mass difference δm before $D^*\ell$ refitting (but using mass-constrained D^0) is less than 170 MeV.

- $D^*\ell$ vertex is fitted with beamspot contraint. The χ^2 probability is required to be greater than 1%.

- For Opposite-Side sample, the $D^*\ell$ candidate satisfies:
  $\cos\theta^*_{D^*\ell} < 0$ and $|\cos\theta^*_{B-D^*\ell}| < 10$.
  For Same-Side sample, it satisfies $\cos\theta^*_{D^*\ell} > 0$ and in this case we evaluate $\cos\theta*_{B-D^*(-\ell)}$, where $-\ell$ represents the lepton four-momentum after it is flipped by 180° and it has to satisfy $|\cos\theta *_{B-D^*(-\ell)}| < 10$. (See Sec. 4.13.2 for definitions of these angles.)

- Refit for lepton, $\pi_{\rm soft}$ and D^0 daughters converges.

- We select a candidate in the refitted $143\,{\rm MeV} < \delta m < 148\,{\rm MeV}$ window when the soft pion has hits only in the SVT and $144\,{\rm MeV} < \delta m < 147\,{\rm MeV}$ when it has hits also in the DCH (this allows to take into account better δm resolution when the the soft pion is reconstructed also in the DCH)

[1] We calculate the decay amplitude squared based on measurements of amplitudes and phases by E687[56] and the four-momenta of the D^0 decay products in our data, assuming perfect resolutions. The maximum is normalized to unity. See Appendix A for details.

Appendix E

δm fits

Here we show all the δm plots of the 72 yield fits described in section 5.1.

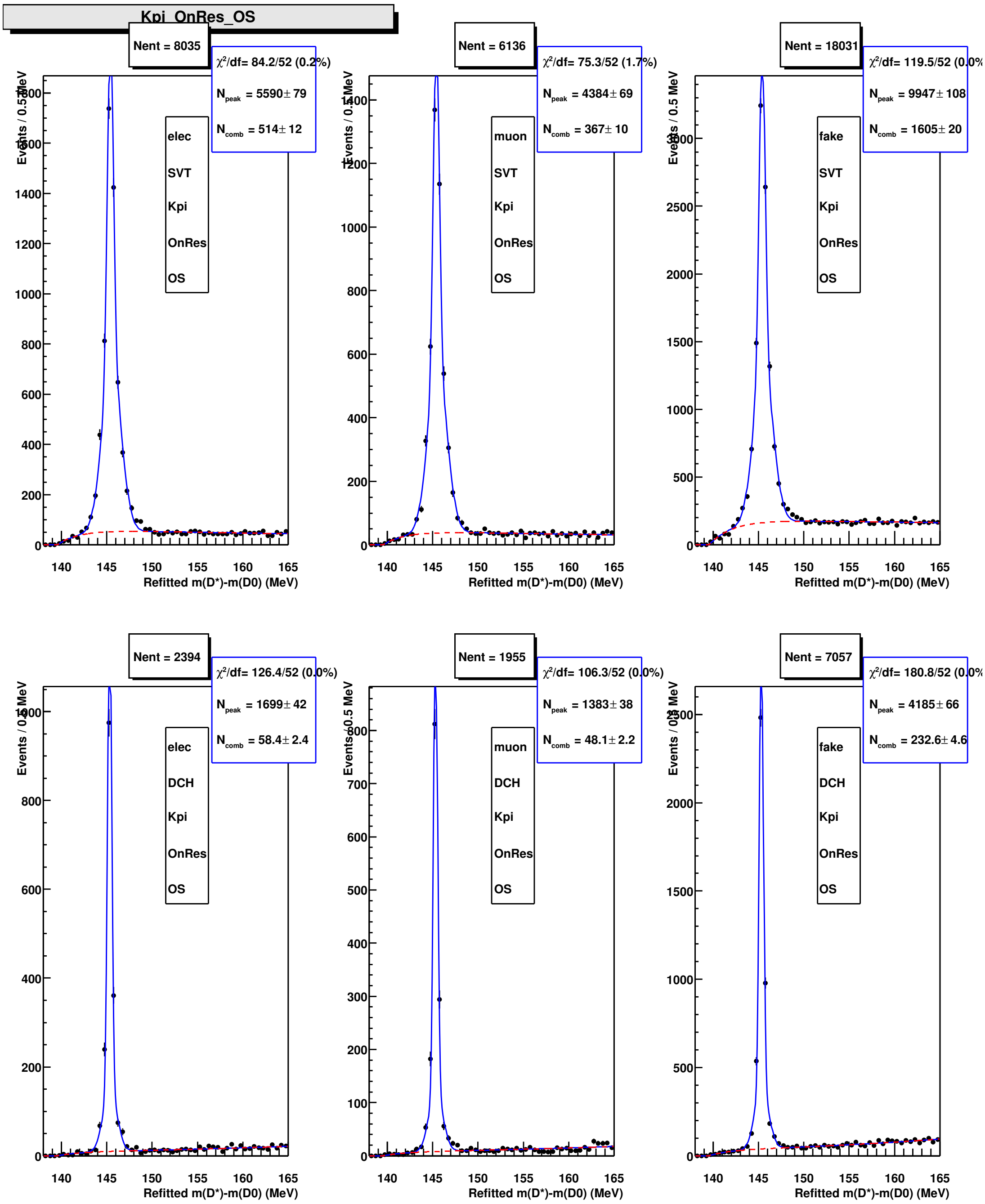


Figure E.1: δm fits used to determine the number of signal and background events in each subsample. We show the fits to the $D^0 \to K\pi$ mode, On-Resonance in the OppositeSide categories for events with soft pion reconstructed only in the SVT (top) or both in the SVT and the DCH (bottom).

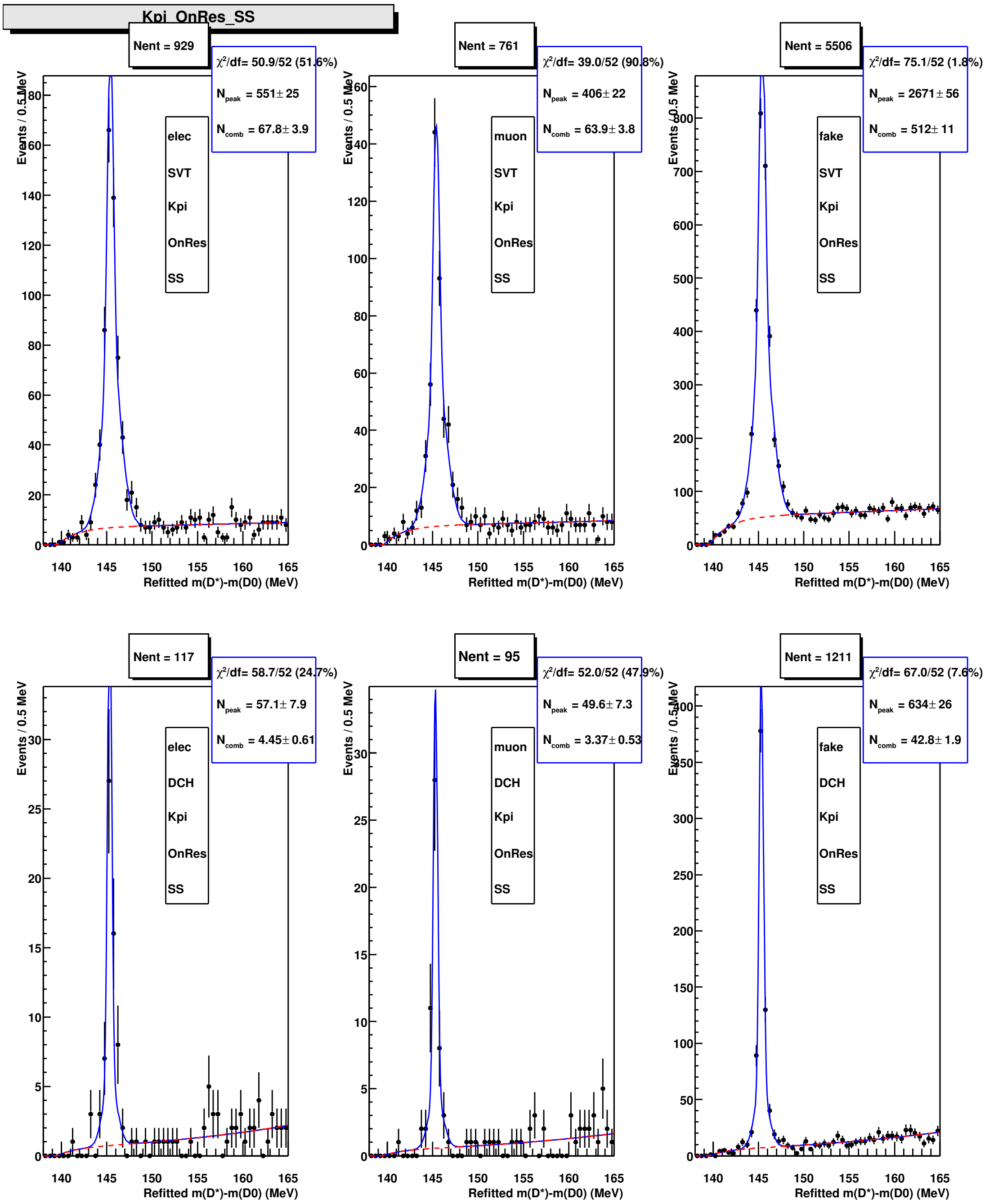


Figure E.2: δm fits used to determine the number of signal and background events in each subsample. We show the fits to the $D^0 \to K\pi$ mode, On-Resonance in the SameSide categories for events with soft pion reconstructed only in the SVT (top) or both in the SVT and the DCH (bottom).

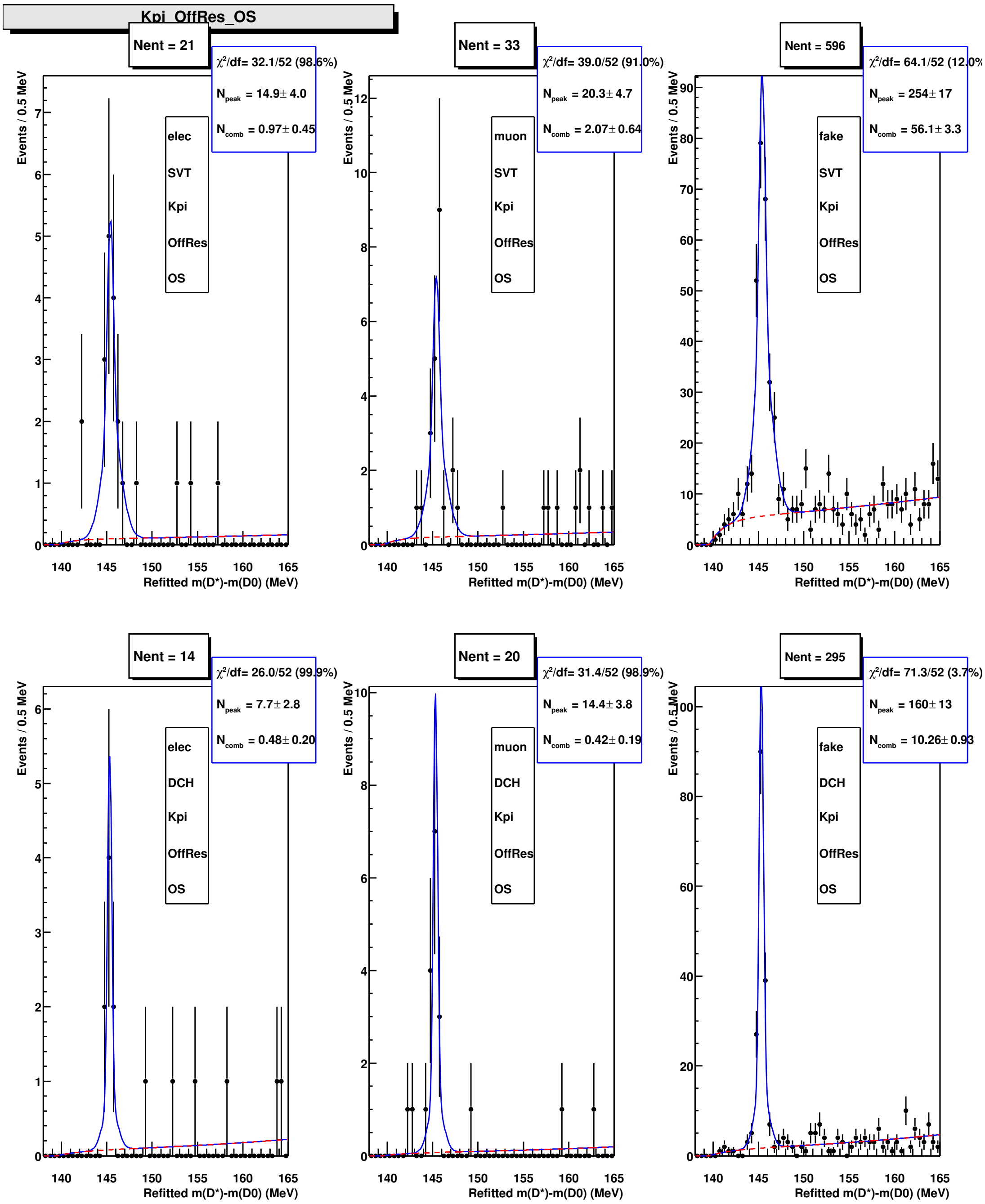


Figure E.3: δm fits used to determine the number of signal and background events in each subsample. We show the fits to the $D^0 \to K\pi$ mode, Off-Resonance in the OppositeSide categories for events with soft pion reconstructed only in the SVT (top) or both in the SVT and the DCH (bottom).

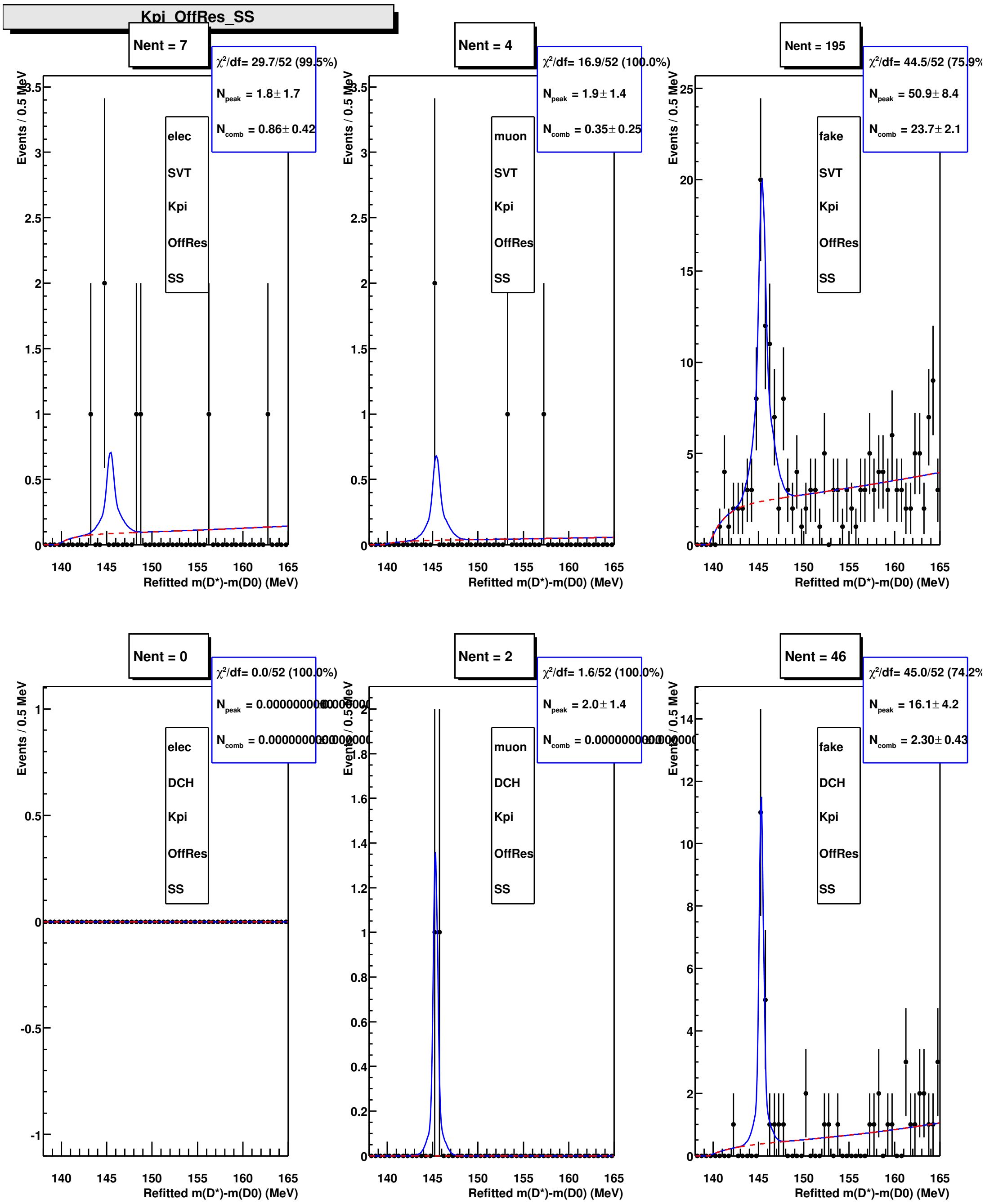


Figure E.4: δm fits used to determine the number of signal and background events in each subsample. We show the fits to the $D^0 \to K\pi$ mode, Off-Resonance in the SameSide categories for events with soft pion reconstructed only in the SVT (top) or both in the SVT and the DCH (bottom).

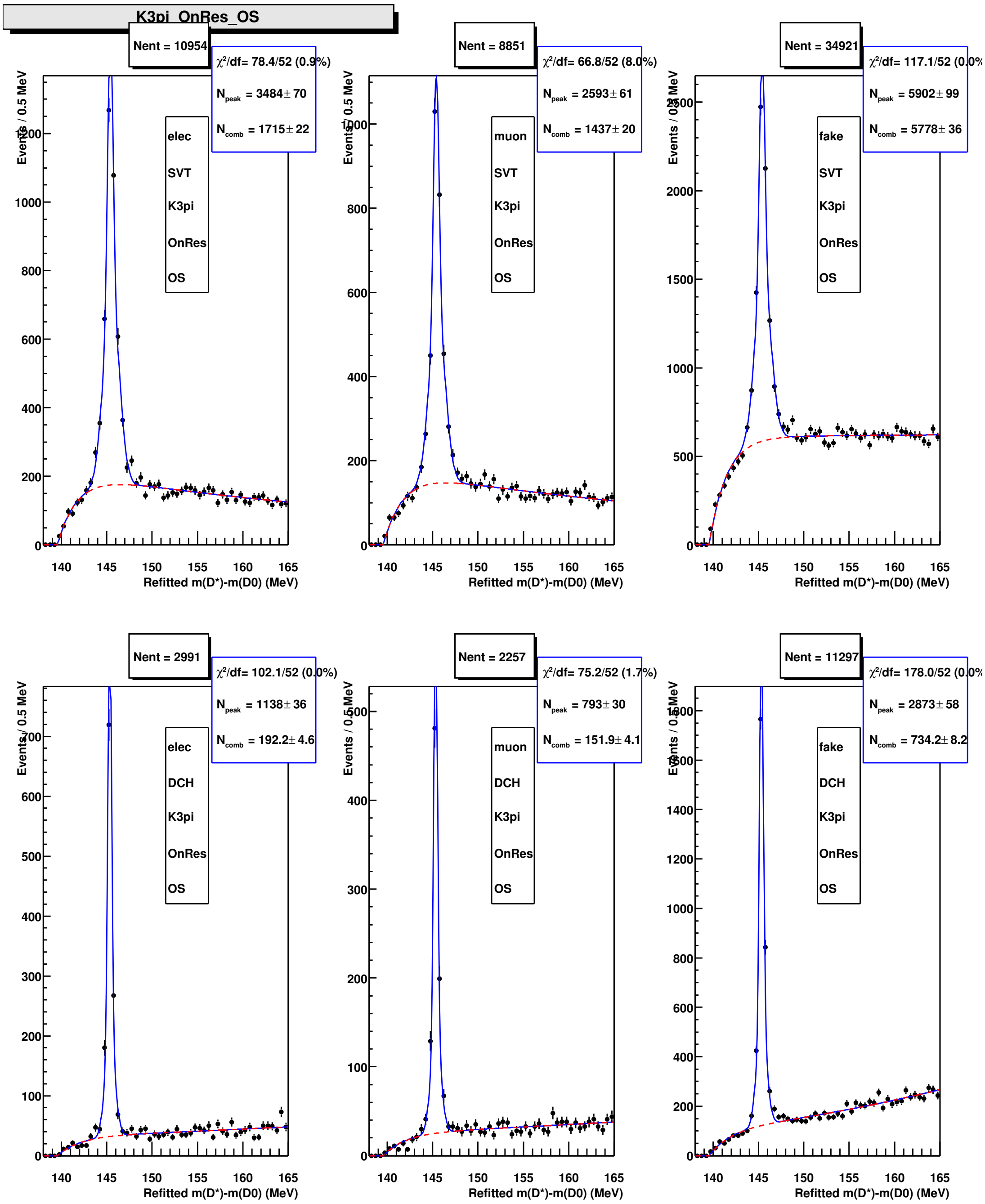


Figure E.5: δm fits used to determine the number of signal and background events in each subsample. We show the fits to the $D^0 \to K\pi\pi\pi$ mode, On-Resonance in the OppositeSide categories for events with soft pion reconstructed only in the SVT (top) or both in the SVT and the DCH (bottom). We recall that for this study we merged together the mode $D^0 \to K\pi\pi\pi$ and $D^0 \to K_S\pi\pi$ in a single `Dmode`.

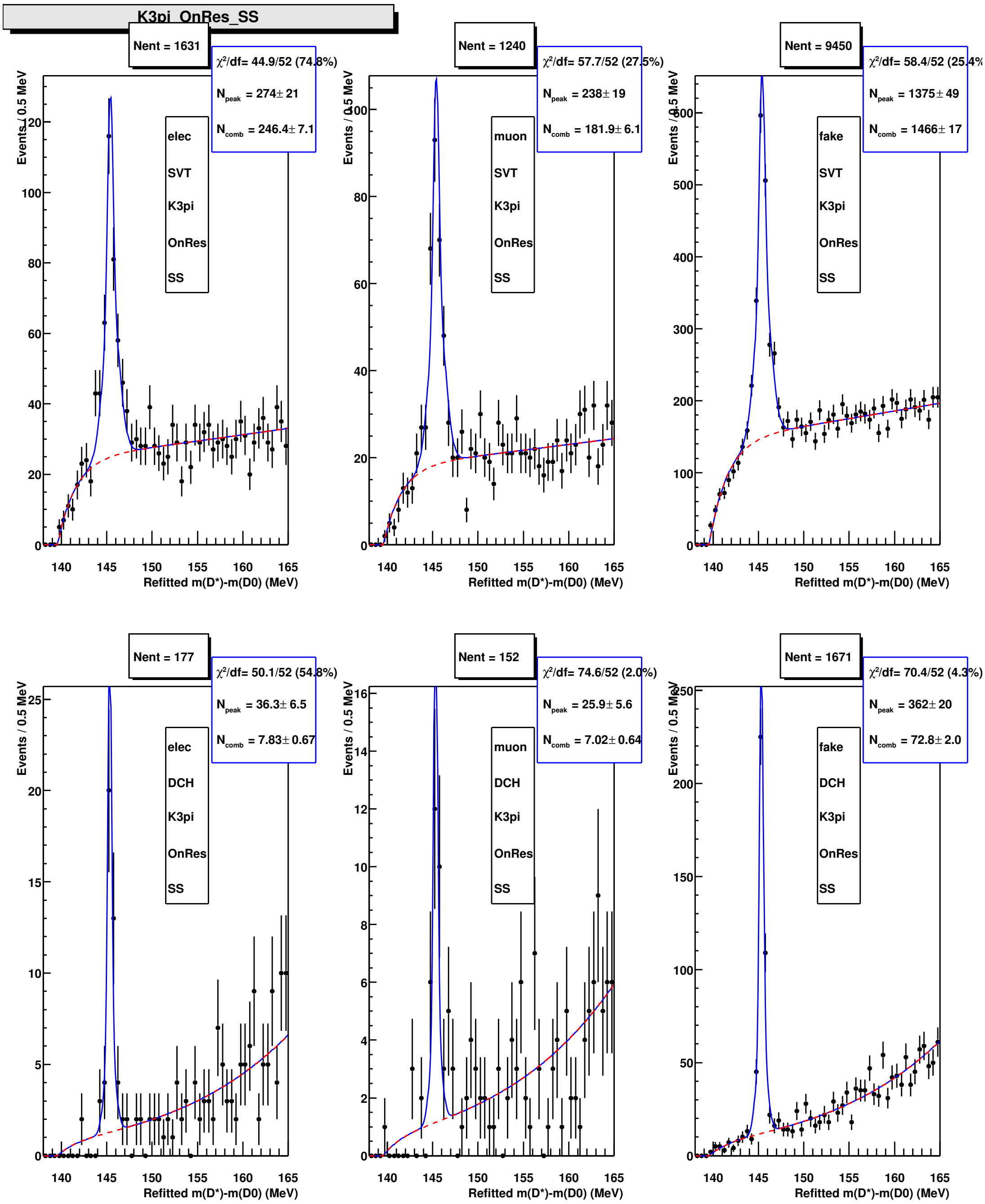


Figure E.6: δm fits used to determine the number of signal and background events in each subsample. We show the fits to the $D^0 \to K\pi\pi\pi$ mode, On-Resonance in the SameSide categories for events with soft pion reconstructed only in the SVT (top) or both in the SVT and the DCH (bottom). We recall that for this study we merged together the mode $D^0 \to K\pi\pi\pi$ and $D^0 \to K_S\pi\pi$ in a single `Dmode`.

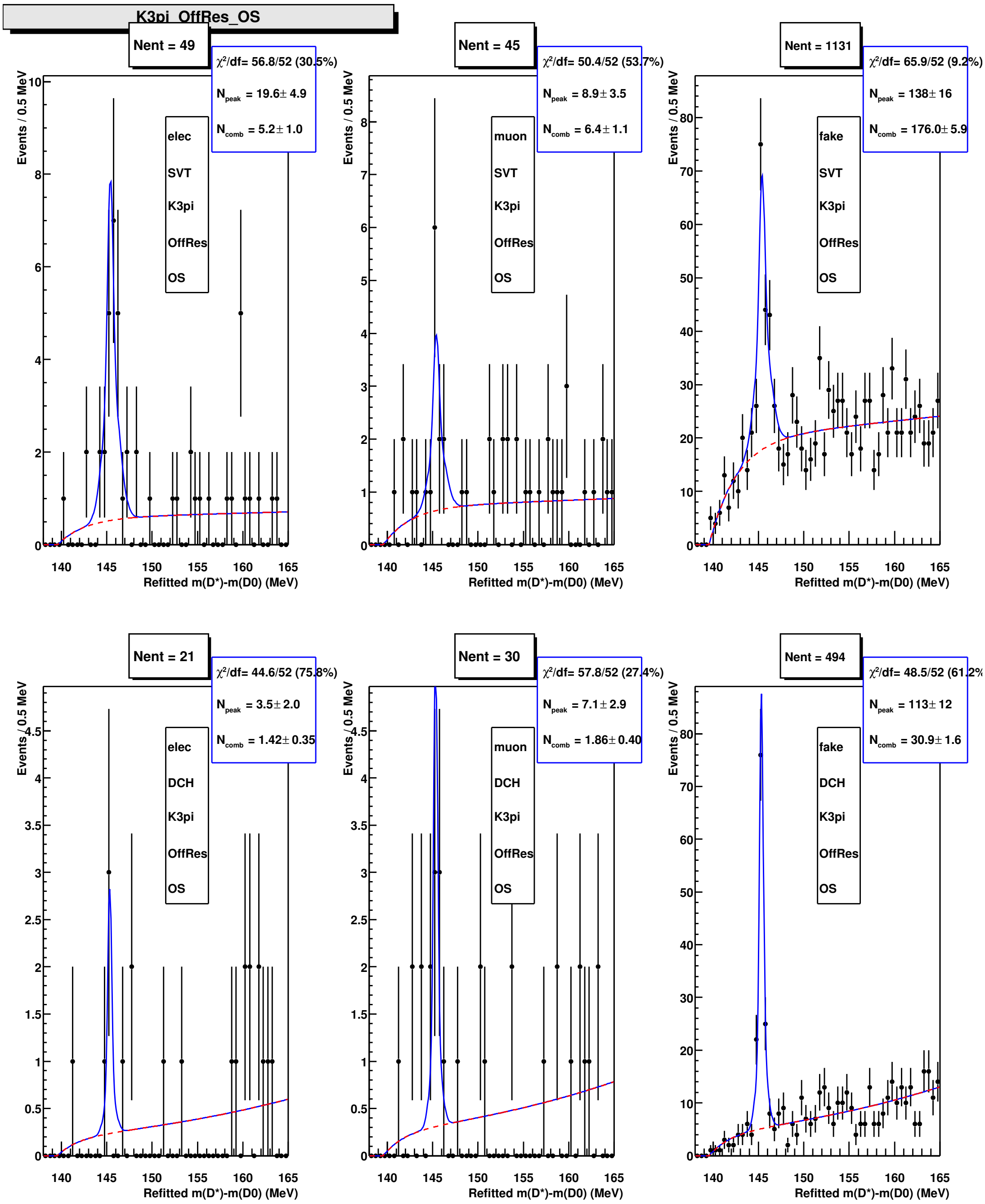


Figure E.7: δm fits used to determine the number of signal and background events in each subsample. We show the fits to the $D^0 \to K\pi\pi\pi$ mode, Off-Resonance in the OppositeSide categories for events with soft pion reconstructed only in the SVT (top) or both in the SVT and the DCH (bottom). We recall that for this study we merged together the mode $D^0 \to K\pi\pi\pi$ and $D^0 \to K_S\pi\pi$ in a single `Dmode`.

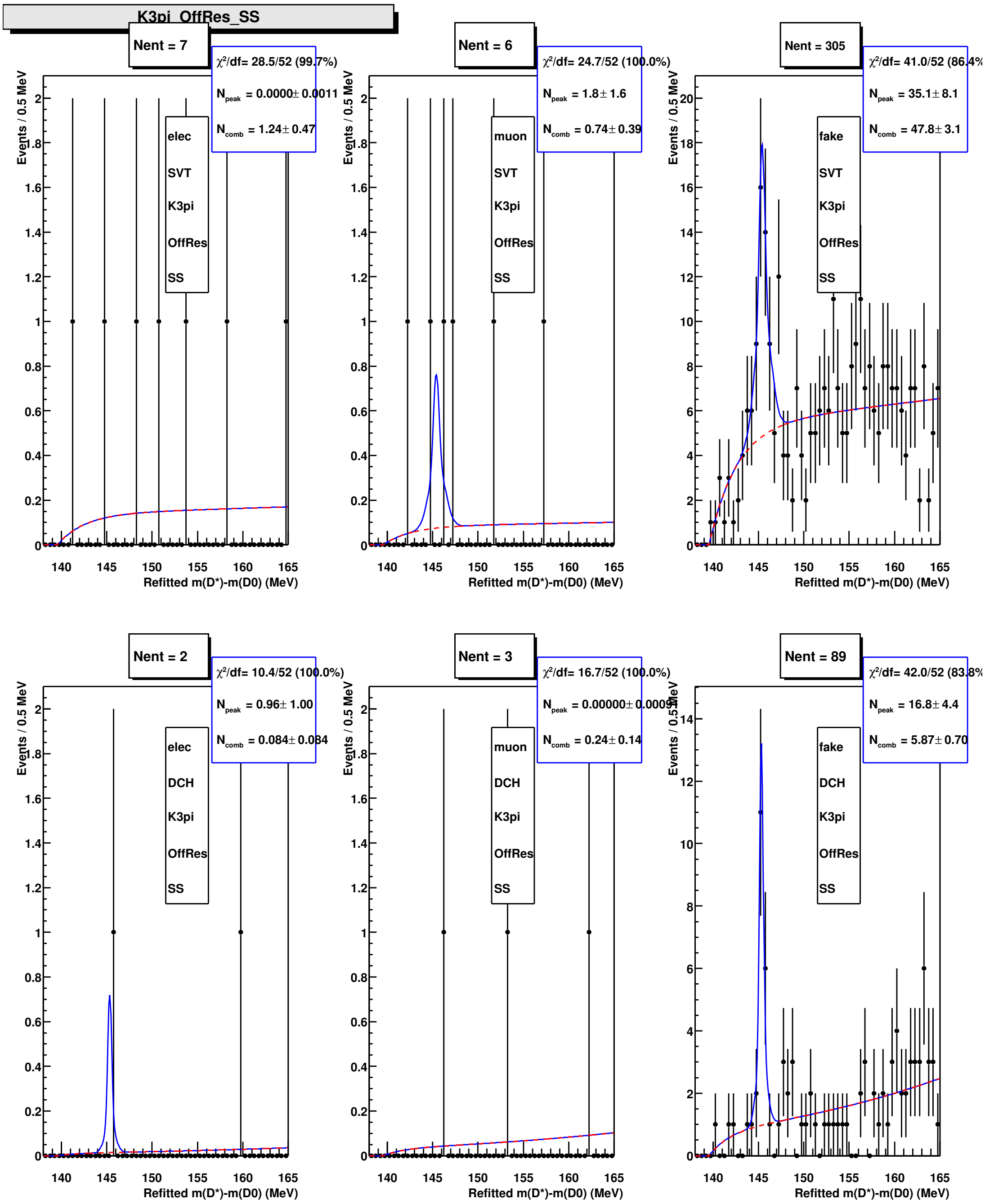


Figure E.8: δm fits used to determine the number of signal and background events in each subsample. We show the fits to the $D^0 \to K\pi\pi\pi$ mode, Off-Resonance in the SameSide categories for events with soft pion reconstructed only in the SVT (top) or both in the SVT and the DCH (bottom). We recall that for this study we merged together the mode $D^0 \to K\pi\pi\pi$ and $D^0 \to K_S\pi\pi$ in a single `Dmode`.

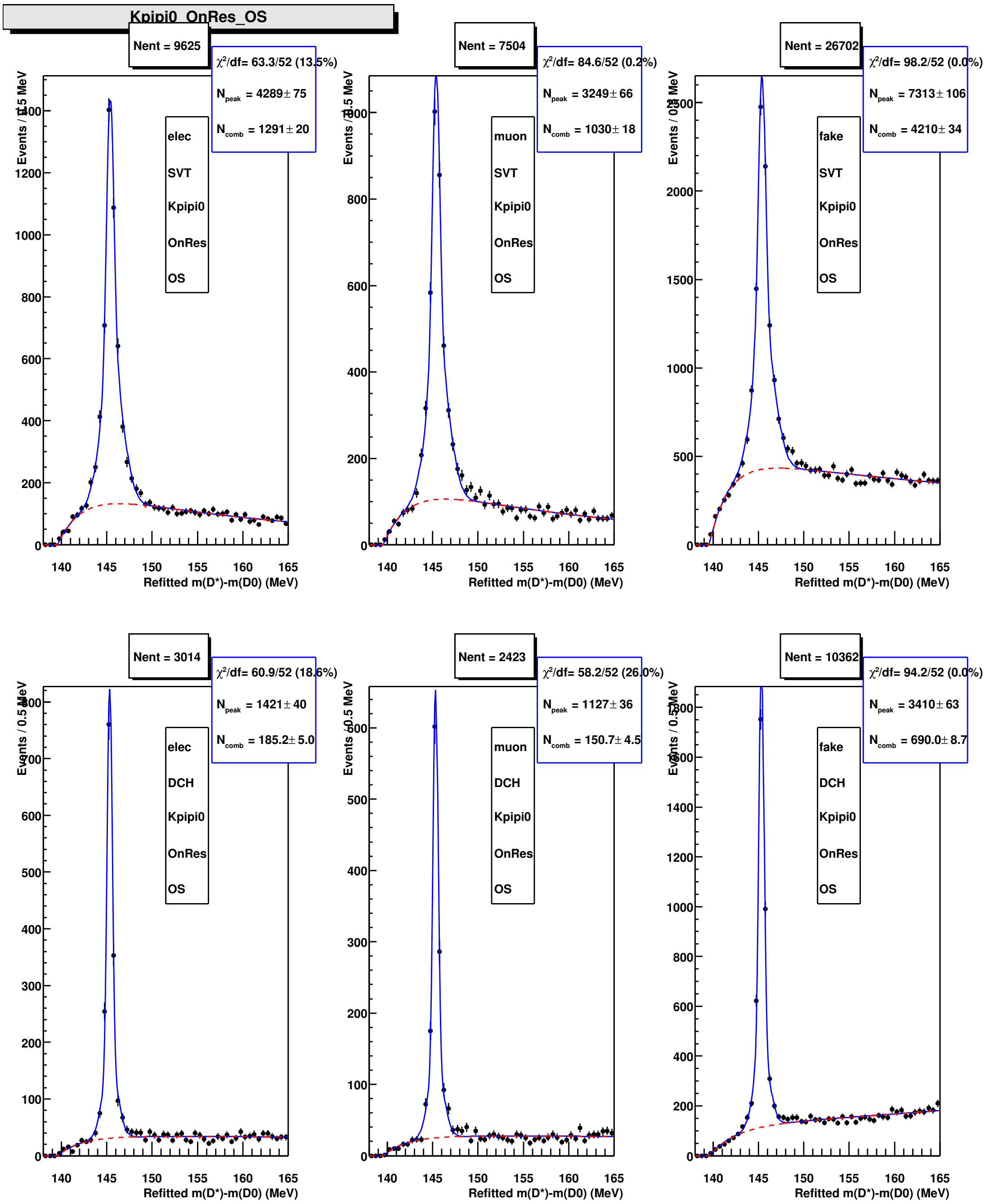


Figure E.9: δm fits used to determine the number of signal and background events in each subsample. We show the fits to the $D^0 \rightarrow K\pi\pi^0$ mode, On-Resonance in the OppositeSide categories for events with soft pion reconstructed only in the SVT (top) or both in the SVT and the DCH (bottom).

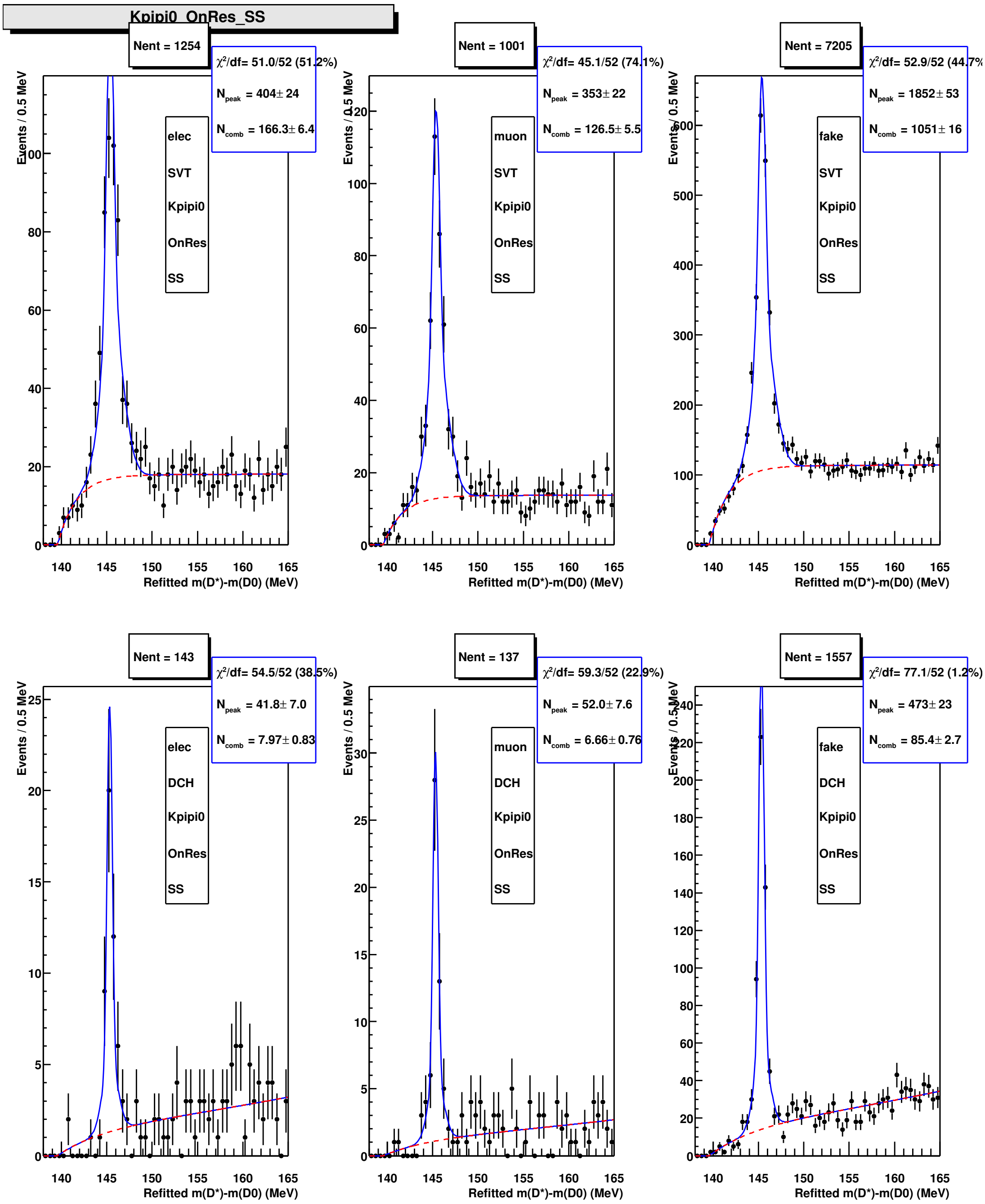


Figure E.10: δm fits used to determine the number of signal and background events in each subsample. We show the fits to the $D^0 \to K\pi\pi^0$ mode, On-Resonance in the SameSide categories for events with soft pion reconstructed only in the SVT (top) or both in the SVT and the DCH (bottom).

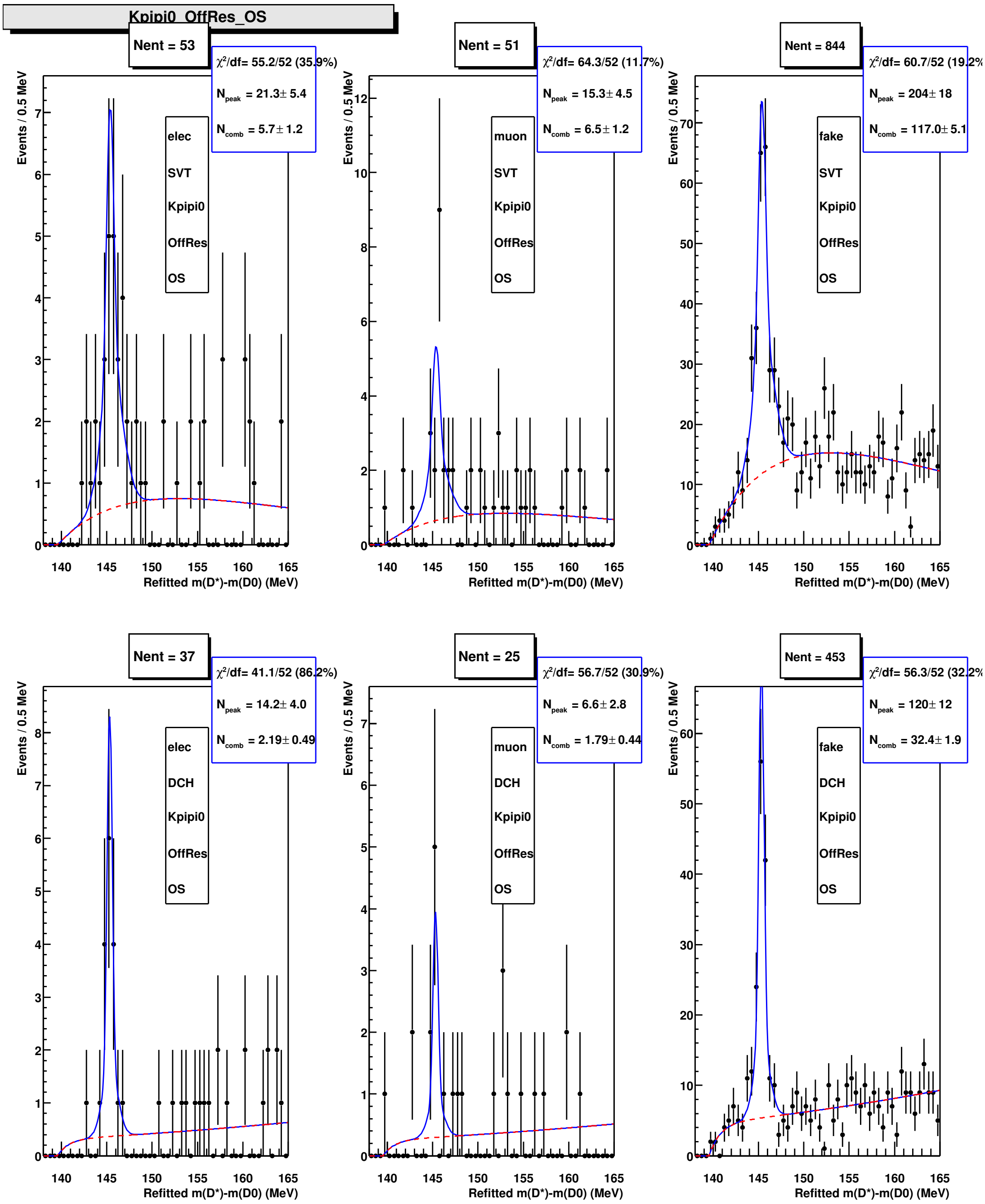


Figure E.11: δm fits used to determine the number of signal and background events in each subsample. We show the fits to the $D^0 \to K\pi\pi^0$ mode, Off-Resonance in the OppositeSide categories for events with soft pion reconstructed only in the SVT (top) or both in the SVT and the DCH (bottom).

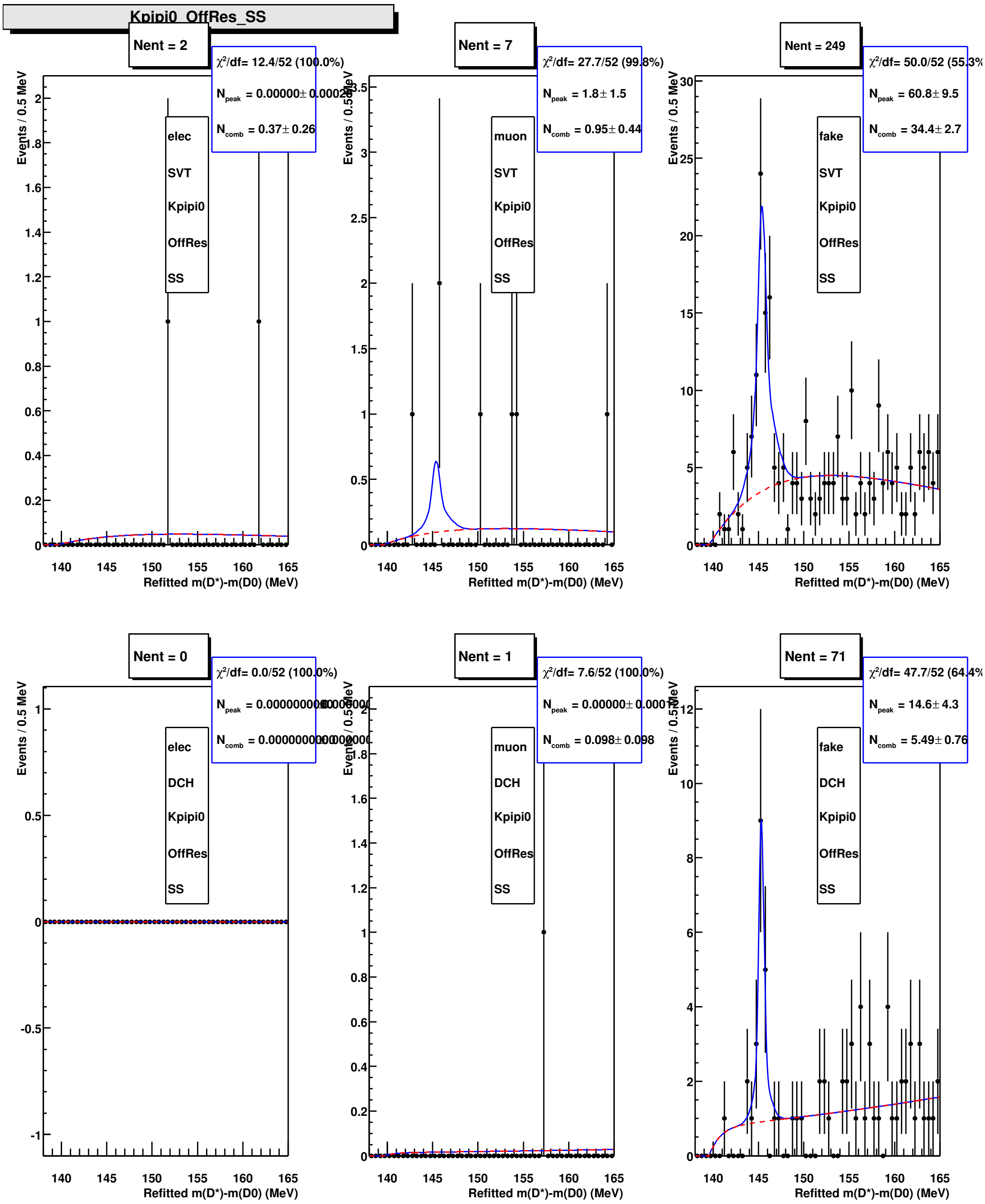


Figure E.12: δm fits used to determine the number of signal and background events in each subsample. We show the fits to the $D^0 \to K\pi\pi^0$ mode, Off-Resonance in the SameSide categories for events with soft pion reconstructed only in the SVT (top) or both in the SVT and the DCH (bottom).

Appendix F

Particle Identification selector plots and tables

F.0.1 PID tables efficiencies from Data Control Samples

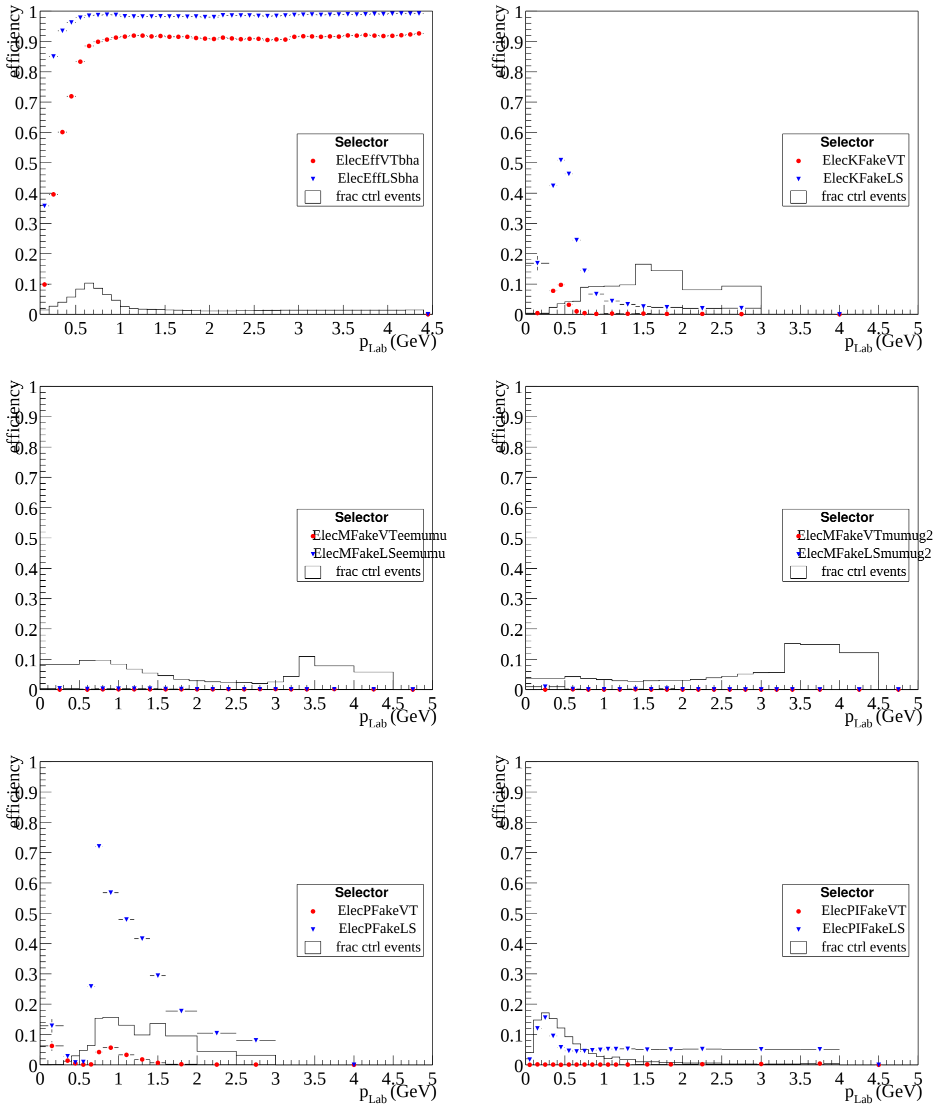


Figure F.1: PID efficiencies for VeryTight (red dots) and Loose (blue triangles) selectors. The plots from top to bottom represent different control samples: Bhabha electron control sample; K, μ, π and p control samples for e efficiency. All efficiencies are retrieved from PID tables created on data control samples. The histogram at the bottom of each plot shows the distribution of events in the control sample, normalized to unit area.

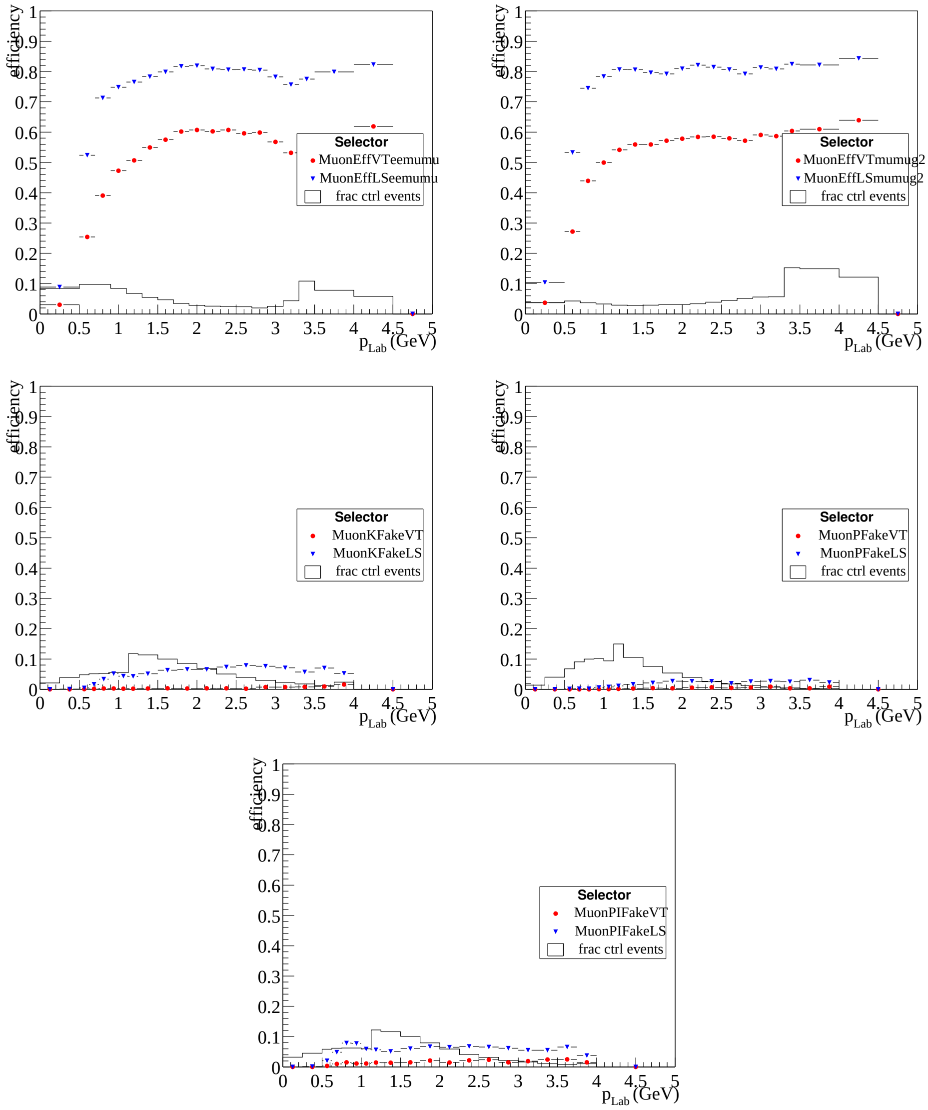


Figure F.2: PID efficiencies for VeryTight (red dots) and Loose (blue triangles) selectors. The plots from top to bottom represent different control samples: $ee\mu\mu$ and $\mu\mu\gamma 2$ μ control sample; K, π and p control samples for μ efficiency. All efficiencies are retrieved from PID tables created on data control samples. The histogram at the bottom of each plot shows the distribution of events in the control sample, normalized to unit area.

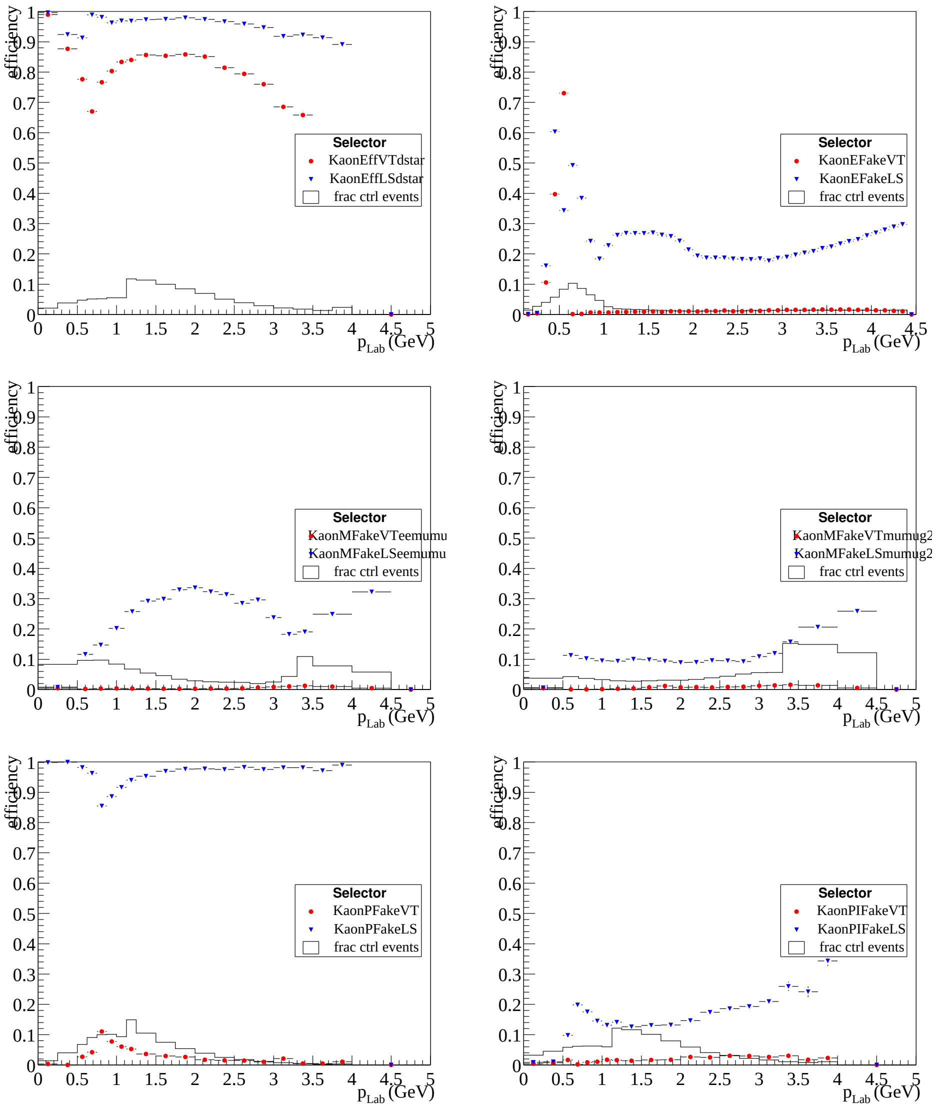


Figure F.3: PID efficiencies for VeryTight (red dots) and Loose (blue triangles) selectors. The plots from top to bottom represent different control samples: here KaonEffVTdstar represent the KaonTight selector (the red dots) and NotAPion selector (blue triangles) K control sample; e, μ, π and p control samples for K efficiency. All efficiencies are retrieved from PID tables created on data control samples. The histogram at the bottom of each plot shows the distribution of events in the control sample, normalized to unit area.

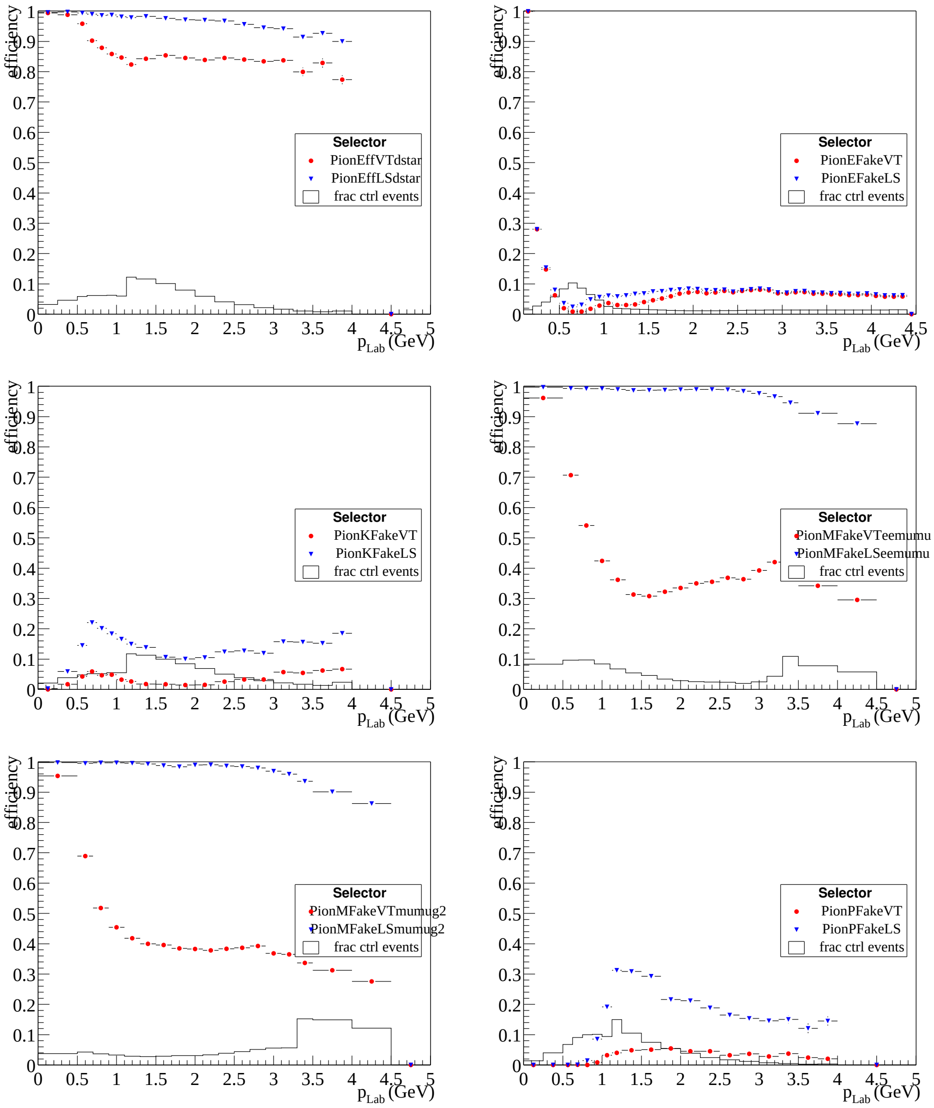


Figure F.4: PID efficiencies for VeryTight (red dots) and Loose (blue triangles) selectors. The plots from top to bottom represent different control samples: we show only for reference pion efficiencies for the π control sample. We show also e, μ K and p control samples for π efficiency. All efficiencies are retrieved from PID tables created on data control samples. The histogram at the bottom of each plot shows the distribution of events in the control sample, normalized to unit area.

F.0.2 Particle Identification efficiency Tables from Monte Carlo Control Samples

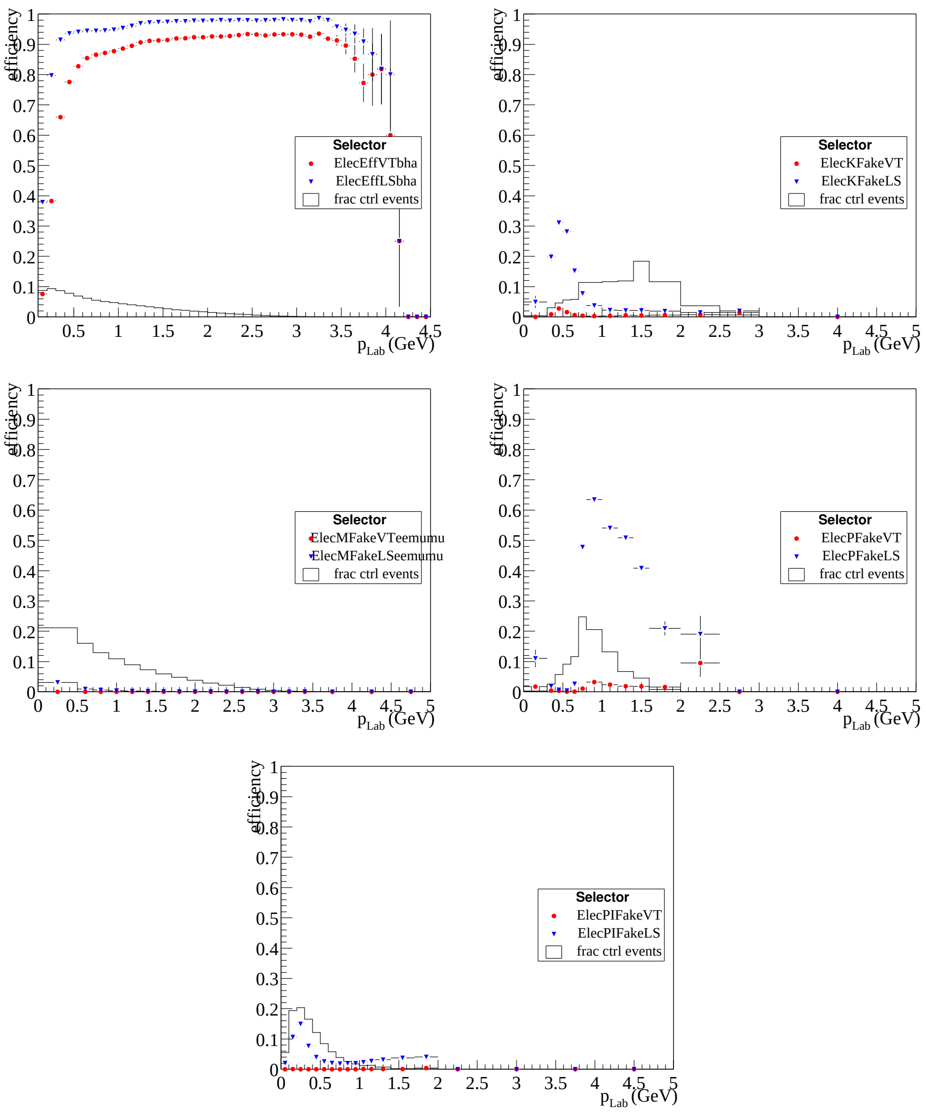


Figure F.5: PID efficiencies for VeryTight (red dots) and Loose (blue triangles) selectors. The plots from top to bottom represent different control samples: Bhabha electron control sample; K, μ, π and p control samples for e efficiency. All efficiencies are retrieved from PID tables created on MC control samples. The histogram at the bottom of each plot shows the distribution of events in the control sample, normalized to unit area.

For the purpose of PID systematics checks we applied the same technique described when evaluating fake-lepton background with PID tables extracted from MC control samples. Here is the equivalent of Table 5.2 with MC PID tables.

sample	veryTight	Loose
ElecEffVTbha,ElecEffLSbha	0.91046 ± 0.00048	0.96541 ± 0.00040
ElecMFakeVTeemumu,ElecMFakeLSeemumu	0.000029 ± 0.000019	0.000976 ± 0.000028
ElecPIFakeVT,ElecPIFakeLS	0.0012 ± 0.0052	0.0272 ± 0.0054
ElecKFakeVT,ElecKFakeLS	0.00563 ± 0.00092	0.0229 ± 0.0014
ElecPFakeVT,ElecPFakeLS	0.0296 ± 0.0095	0.336 ± 0.014
MuonEffVTeemumu,MuonEffLSeemumu	0.72194 ± 0.00055	0.88864 ± 0.00038
MuonPIFakeVT,MuonPIFakeLS	0.01405 ± 0.00068	0.0612 ± 0.0012
MuonKFakeVT,MuonKFakeLS	0.00451 ± 0.00050	0.0712 ± 0.0013
MuonPFakeVT,MuonPFakeLS	0.00190 ± 0.00015	0.01449 ± 0.00042
PionEffVTdstar,PionEffLSdstar	0.8414 ± 0.0045	0.9814 ± 0.0011
PionMFakeVTeemumu,PionMFakeLSeemumu	0.2413 ± 0.0027	0.99048 ± 0.00049
PionKFakeVT,PionKFakeLS	0.02235 ± 0.00099	0.1341 ± 0.0022
PionPFakeVT,PionPFakeLS	0.0594 ± 0.0021	0.2512 ± 0.0060
KaonEffVTdstar,KaonEffLSdstar	0.7725 ± 0.0089	0.9790 ± 0.0010
KaonMFakeVTeemumu,KaonMFakeLSeemumu	0.00720 ± 0.00048	0.2190 ± 0.0091
KaonPIFakeVT,KaonPIFakeLS	0.01505 ± 0.00078	0.1511 ± 0.0016
KaonPFakeVT,KaonPFakeLS	0.03440 ± 0.00076	0.9587 ± 0.0012

Table F.1: *Summary of angle- and momentum-weighted PID efficiencies for all selectors used to evaluate ℓ and K ID efficiencies. The lepton ID efficiencies are evaluated from a weighted average according to each luminosity data taking period the 2001 sample has been divided.*

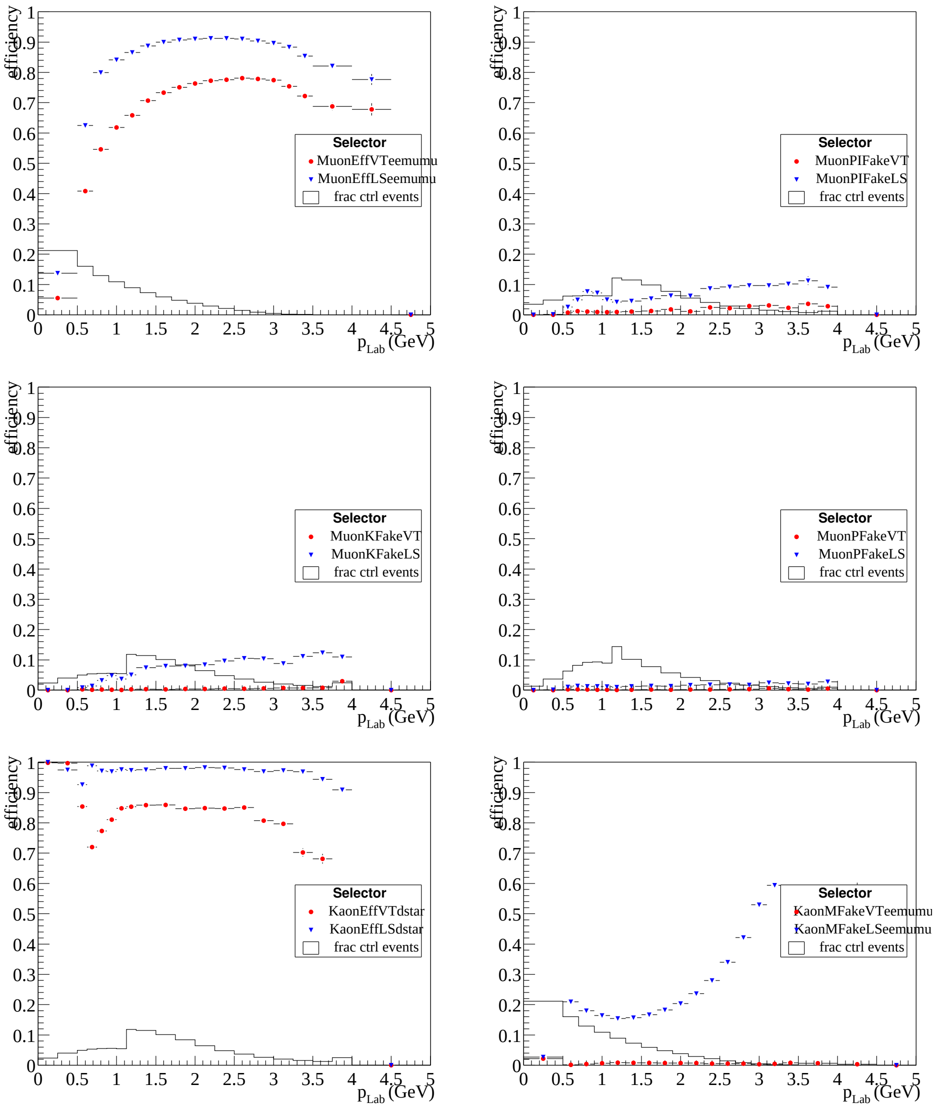


Figure F.6: PID efficiencies for VeryTight (red dots) and Loose (blue triangles) selectors. The plots from top to bottom represent different control samples: $ee\mu\mu$ and $\mu\mu\gamma 2$ μ control sample; K, π and p control samples for μ efficiency. Here KaonEffVTdstar represent the KaonTight selector (the red dots) and NotAPion selector (blue triangles) K control sample; we show also μ control sample for K efficiency. All efficiencies are retrieved from PID tables created on MC control samples. The histogram at the bottom of each plot shows the distribution of events in the control sample, normalized to unit area.

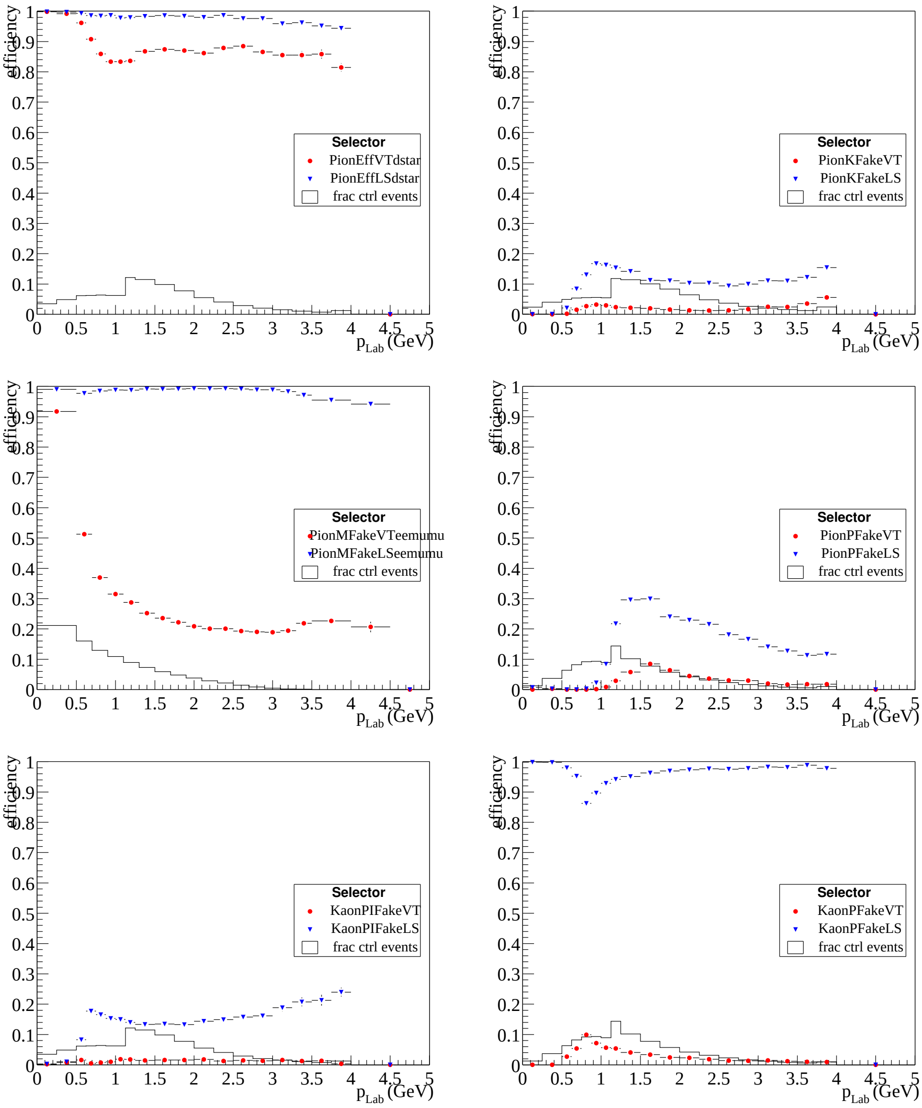


Figure F.7: The plots from top to bottom represent different control samples: we show only for reference pion efficiencies for the π control sample. We show also e, μ K and p control samples for π efficiency. We show also π and p control sample for K efficiency. All efficiencies are retrieved from PID tables created on MC control samples. The histogram at the bottom of each plot shows the distribution of events in the control sample, normalized to unit area.

Appendix G

Fit Yields in Monte Carlo samples

Here we show the yields of events evaluated with the δm fit method described in chapter 5 divided by control samples as defined in section 4.3.

We recall that

$$N_{\text{comb}} = \text{yield} \cdot \frac{f_{\text{comb}}}{1 - f_{\text{comb}}}$$

and

$$\text{totalyield} = \text{yield} + N_{\text{comb}}$$

and that all the fake-lepton, continuum and uncorrelated fractions refers to the peak yield indicate as yield in these tables. Another important thing to say is that we still do not calculate the amount of correlated and D^{**} background for each kind of Monte Carlo sample we show below.

Category	yield	f_{fake}	$f_{\text{continuum}}$	$f_{\text{uncorr.}}$	$f_{\text{comb.}}$
elec;OS;OnRes	59804 ± 263	$0.3048 \pm 0.0018\%$	$0 \pm 0\%$	$6.05 \pm 0.19\%$	$13.73 \pm 0.11\%$
muon;OS;OnRes	57670 ± 258	$2.070 \pm 0.013\%$	$0 \pm 0\%$	$6.08 \pm 0.20\%$	$13.60 \pm 0.11\%$
fake;OS;OnRes	94408 ± 338	$88.37 \pm 0.48\%$	$0 \pm 0\%$	$0.72 \pm 0.35\%$	$18.67 \pm 0.11\%$
elec;SS;OnRes	4577 ± 76	$0.877 \pm 0.016\%$	$0 \pm 0\%$	$74.4 \pm 2.7\%$	$20.61 \pm 0.57\%$
muon;SS;OnRes	4640 ± 76	$5.66 \pm 0.10\%$	$0 \pm 0\%$	$71.1 \pm 2.7\%$	$20.31 \pm 0.55\%$
fake;SS;OnRes	19218 ± 154	$95.6 \pm 1.1\%$	$0 \pm 0\%$	$3.3 \pm 1.7\%$	$20.98 \pm 0.28\%$

Table G.1: *Results of the partial yields (peaking yields) obtained from the fit method described on chapter 5 on the $B^0\overline{B}^0$ Monte Carlo sample used in this analysis.*

Category	yield	f_{fake}	$f_{\text{continuum}}$	$f_{\text{uncorr.}}$	$f_{\text{comb.}}$
elec;OS;OnRes	1172 ± 50	$6.63 \pm 0.29\%$	$0 \pm 0\%$	$81.8 \pm 6.0\%$	$76 \pm 14\%$
muon;OS;OnRes	1878 ± 66	$27.09 \pm 0.96\%$	$0 \pm 0\%$	$49.6 \pm 3.5\%$	$66.5 \pm 7.1\%$
fake;OS;OnRes	35814 ± 220	$99.29 \pm 0.87\%$	$0 \pm 0\%$	$0.51 \pm 0.55\%$	$31.20 \pm 0.33\%$
elec;SS;OnRes	963 ± 38	$1.257 \pm 0.054\%$	$0 \pm 0\%$	$98.5 \pm 7.0\%$	$39.1 \pm 2.8\%$
muon;SS;OnRes	971 ± 38	$8.16 \pm 0.35\%$	$0 \pm 0\%$	$90.9 \pm 6.8\%$	$39.4 \pm 2.8\%$
fake;SS;OnRes	5717 ± 91	$97.0 \pm 2.2\%$	$0 \pm 0\%$	$3.0 \pm 3.3\%$	$36.9 \pm 1.0\%$

Table G.2: *Results of the partial yields (peaking yields) obtained from the fit method described on chapter 5 on the B^+B^- Monte Carlo sample used in this analysis.*

Category	yield	f_{fake}	$f_{\text{continuum}}$	$f_{\text{uncorr.}}$	$f_{\text{comb.}}$
elec;OS;OnRes	1525 ± 42	$3.49 \pm 0.10\%$	$0 \pm 0\%$	$1.25 \pm 0.89\%$	$16.67 \pm 0.78\%$
muon;OS;OnRes	1808 ± 46	$19.26 \pm 0.51\%$	$0 \pm 0\%$	$1.7 \pm 1.2\%$	$17.47 \pm 0.75\%$
fake;OS;OnRes	24618 ± 174	$98.86 \pm 0.99\%$	$0 \pm 0\%$	$0.02 \pm 0.86\%$	$21.45 \pm 0.25\%$
elec;SS;OnRes	59.7 ± 8.9	$29.3 \pm 4.4\%$	$0 \pm 0\%$	$30 \pm 22\%$	$26.7 \pm 6.5\%$
muon;SS;OnRes	164 ± 14	$69.0 \pm 6.1\%$	$0 \pm 0\%$	$18 \pm 12\%$	$19.5 \pm 2.8\%$
fake;SS;OnRes	8015 ± 98	$99.9 \pm 1.7\%$	$0 \pm 0\%$	$0.1 \pm 2.4\%$	$18.60 \pm 0.37\%$

Table G.3: *Results of the partial yields (peaking yields) obtained from the fit method described on chapter 5 on the $c\overline{c}$ Monte Carlo sample used in this analysis.*

Category	yield	f_{fake}	$f_{\text{continuum}}$	$f_{\text{uncorr.}}$	$f_{\text{comb.}}$
elec;OS;OnRes	5500 ± 76	$0.0149 \pm 0.0016\%$	$0 \pm 0\%$	$2.88 \pm 0.46\%$	$2.92 \pm 0.16\%$
muon;OS;OnRes	4896 ± 71	$0.110 \pm 0.012\%$	$0 \pm 0\%$	$2.65 \pm 0.48\%$	$2.95 \pm 0.17\%$
fake;OS;OnRes	1339 ± 37	$28.1 \pm 3.0\%$	$0 \pm 0\%$	$1.9 \pm 3.8\%$	$2.81 \pm 0.29\%$
elec;SS;OnRes	260 ± 17	$0.402 \pm 0.032\%$	$0 \pm 0\%$	$58 \pm 10\%$	$4.50 \pm 0.85\%$
muon;SS;OnRes	226 ± 15	$3.03 \pm 0.26\%$	$0 \pm 0\%$	$54 \pm 11\%$	$3.94 \pm 0.86\%$
fake;SS;OnRes	522 ± 23	$91.6 \pm 6.1\%$	$0 \pm 0\%$	$4.7 \pm 9.4\%$	$4.55 \pm 0.59\%$

Table G.4: *Results of the partial yields (peaking yields) obtained from the fit method described on chapter 5 on the signal $D^0 \to K\pi$ Monte Carlo sample used in this analysis.*

Category	yield	f_{fake}	$f_{\text{continuum}}$	$f_{\text{uncorr.}}$	$f_{\text{comb.}}$
elec;OS;OnRes	1440 ± 40	$0.0093 \pm 0.0031\%$	$0 \pm 0\%$	$2.43 \pm 0.91\%$	$7.15 \pm 0.49\%$
muon;OS;OnRes	1463 ± 40	$0.060 \pm 0.020\%$	$0 \pm 0\%$	$1.88 \pm 0.84\%$	$6.90 \pm 0.48\%$
fake;OS;OnRes	342 ± 19	$17.9 \pm 6.1\%$	$0 \pm 0\%$	$1.6 \pm 7.4\%$	$8.9 \pm 1.1\%$
elec;SS;OnRes	61.7 ± 8.6	$0.407 \pm 0.070\%$	$0 \pm 0\%$	$54 \pm 22\%$	$12.6 \pm 3.1\%$
muon;SS;OnRes	57.0 ± 8.0	$2.89 \pm 0.50\%$	$0 \pm 0\%$	$46 \pm 22\%$	$11.4 \pm 2.9\%$
fake;SS;OnRes	126 ± 12	$91 \pm 13\%$	$0 \pm 0\%$	$4 \pm 19\%$	$7.6 \pm 1.4\%$

Table G.5: *Results of the partial yields (peaking yields) obtained from the fit method described on chapter 5 on the signal $D^0 \to K\pi\pi\pi$ Monte Carlo sample used in this analysis.*

Category	yield	f_{fake}	$f_{\text{continuum}}$	$f_{\text{uncorr.}}$	$f_{\text{comb.}}$
elec;OS;OnRes	1945 ± 47	$0.0175 \pm 0.0028\%$	$0 \pm 0\%$	$2.68 \pm 0.83\%$	$9.06 \pm 0.54\%$
muon;OS;OnRes	1741 ± 44	$0.128 \pm 0.021\%$	$0 \pm 0\%$	$2.20 \pm 0.83\%$	$9.05 \pm 0.57\%$
fake;OS;OnRes	498 ± 24	$31.3 \pm 5.2\%$	$0 \pm 0\%$	$1.6 \pm 6.7\%$	$9.1 \pm 1.0\%$
elec;SS;OnRes	87 ± 11	$0.485 \pm 0.070\%$	$0 \pm 0\%$	$56 \pm 19\%$	$17.9 \pm 4.2\%$
muon;SS;OnRes	72.6 ± 9.5	$3.81 \pm 0.58\%$	$0 \pm 0\%$	$50 \pm 20\%$	$15.8 \pm 4.2\%$
fake;SS;OnRes	207 ± 15	$93 \pm 10\%$	$0 \pm 0\%$	$4 \pm 15\%$	$9.5 \pm 1.7\%$

Table G.6: *Results of the partial yields (peaking yields) obtained from the fit method described on chapter 5 on the signal $D^0 \to K\pi\pi^0$ Monte Carlo sample used in this analysis.*

Category	yield	f_{fake}	$f_{\text{continuum}}$	$f_{\text{uncorr.}}$	$f_{\text{comb.}}$
elec;OS;OnRes	252 ± 16	$0.0148 \pm 0.0068\%$	$0 \pm 0\%$	$4.6 \pm 2.4\%$	$3.54 \pm 0.97\%$
muon;OS;OnRes	216 ± 15	$0.113 \pm 0.052\%$	$0 \pm 0\%$	$3.5 \pm 2.4\%$	$4.7 \pm 1.2\%$
fake;OS;OnRes	54.6 ± 7.3	$31 \pm 15\%$	$0 \pm 0\%$	$3 \pm 19\%$	$0.181 \pm 0.086\%$
elec;SS;OnRes	15.8 ± 4.0	$0.214 \pm 0.080\%$	$0 \pm 0\%$	$70 \pm 42\%$	$0 \pm 0\%$
muon;SS;OnRes	11.4 ± 3.4	$1.94 \pm 0.79\%$	$0 \pm 0\%$	$63 \pm 48\%$	$8.0 \pm 4.7\%$
fake;SS;OnRes	24.1 ± 4.8	$90 \pm 27\%$	$0 \pm 0\%$	$6 \pm 36\%$	$0.058 \pm 0.059\%$

Table G.7: *Results of the partial yields (peaking yields) obtained from the fit method described on chapter 5 on the signal $D^0 \to K_S\pi\pi$ Monte Carlo sample used in this analysis.*

Appendix H

Selection efficiencies for branching fraction evaluation

The equation 6.6 adopted to calculate $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ is equivalent to

$$\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell) = \frac{N_{data} - N_{bkg}}{\epsilon N_{B^0}\mathcal{B}(D^* \to D^0\pi)_{PDG}\mathcal{B}(D^0)_{PDG}} \ , \tag{H.1}$$

where ϵ is the selection efficiency (corrected for $C_{PID} \cdot C_{trk}$), evaluated on generic $B^0\overline{B}^0$ Monte Carlo events, N_{B^0} is the number of B^0 or $\overline{B}^0$ generated events, $\mathcal{B}(D^* \to D^0\pi)_{PDG}$ and $\mathcal{B}(D^0)_{PDG}$ are the D^* and D^0 branching ratio taken from PDG2002, respectively. N_{data} and N_{bkg} are defined in section 6.5. In table H.1 and H.2 we show the selection efficiencies for all 8 channels both in the fit range $-10 < \cos\theta_{B^0,D^*\ell} < 5$ and in the range $-1.2 < \cos\theta_{B^0,D^*\ell} < 1.2$ used for branching ratio evaluation.

		raw	after $C_{PID} \cdot C_{trk}$
$\epsilon(B^0 \to D^*e\nu_e, D^0 \to K\pi)$	=	12.896 ± 0.087	12.578 ± 0.086
$\epsilon(B^0 \to D^*e\nu_e, D^0 \to K\pi\pi\pi)$	=	3.972 ± 0.032	3.905 ± 0.032
$\epsilon(B^0 \to D^*e\nu_e, D^0 \to K\pi\pi^0)$	=	3.110 ± 0.022	3.076 ± 0.022
$\epsilon(B^0 \to D^*e\nu_e, D^0 \to K_S\pi\pi)$	=	0.413 ± 0.013	0.360 ± 0.012
$\epsilon(B^0 \to D^*\mu\nu_\mu, D^0 \to K\pi)$	=	12.120 ± 0.084	9.865 ± 0.076
$\epsilon(B^0 \to D^*\mu\nu_\mu, D^0 \to K\pi\pi\pi)$	=	3.710 ± 0.031	3.043 ± 0.028
$\epsilon(B^0 \to D^*\mu\nu_\mu, D^0 \to K\pi\pi^0)$	=	2.888 ± 0.022	2.384 ± 0.020
$\epsilon(B^0 \to D^*\mu\nu_\mu, D^0 \to K_S\pi\pi)$	=	0.387 ± 0.013	0.282 ± 0.011

Table H.1: *Summary of selection efficiencies for each of the 8 channels before (raw) and after $C_{PID} \cdot C_{trk}$ corrections. These efficiencies are evaluated in the fit range $-10 < \cos\theta_{B^0,D^*\ell} < 5$.*

		raw	after $C_{PID} \cdot C_{trk}$
$\epsilon(B^0 \to D^* e\nu_e, D^0 \to K\pi)$	$=$	12.052 ± 0.084	11.755 ± 0.083
$\epsilon(B^0 \to D^* e\nu_e, D^0 \to K\pi\pi\pi)$	$=$	3.743 ± 0.031	3.679 ± 0.031
$\epsilon(B^0 \to D^* e\nu_e, D^0 \to K\pi\pi^0)$	$=$	2.846 ± 0.021	2.815 ± 0.021
$\epsilon(B^0 \to D^* e\nu_e, D^0 \to K_S\pi\pi)$	$=$	0.389 ± 0.013	0.340 ± 0.012
$\epsilon(B^0 \to D^* \mu\nu_\mu, D^0 \to K\pi)$	$=$	12.016 ± 0.084	9.780 ± 0.076
$\epsilon(B^0 \to D^* \mu\nu_\mu, D^0 \to K\pi\pi\pi)$	$=$	3.685 ± 0.031	3.022 ± 0.028
$\epsilon(B^0 \to D^* \mu\nu_\mu, D^0 \to K\pi\pi^0)$	$=$	2.793 ± 0.021	2.305 ± 0.019
$\epsilon(B^0 \to D^* \mu\nu_\mu, D^0 \to K_S\pi\pi)$	$=$	0.383 ± 0.013	0.279 ± 0.011

Table H.2: *Summary of selection efficiencies for each of the 8 channels before (raw) and after $C_{PID} \cdot C_{trk}$ corrections. These efficiencies are calculated in the range $-1.2 < \cos\theta_{B^0,D^*\ell} < 1.2$ used for branching ratio evaluation.*

Appendix I

Cross checks results on $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ with different χ^2 fit

Here we summarize all the plots and the results of branching fraction measurement in all the 8 channels (4 D^0 mode and e and μ) extracted with a different χ^2 fit for D^{**} background estimation.

All the results are well consistent with the method shown in chapter 6. This method has been used as a cross check on possible bias of the fitting method adopted for the branching fraction measurement. We recall that this method does not constrain the D^{**} to D^* relative amount to be the same in all 8 signal channels but imposes only a constraint in the χ^2, defined in formula 6.5. The constraint on the χ^2 consists in adopting a measurement obtained in *BABAR*, but still not published, on the indirect estimation of $\mathcal{B}(B \to D^{**}\ell\nu)$ and $\mathcal{B}(B \to D^*\pi\ell\nu)$. We consider those results only preliminary numbers used as a cross check on our official fit method adopted in this thesis.

In the case of the 16 parameters fit with the branching ratio constraint for the D^{**} amount [76] we found:

- there is no need to force the $\mathcal{B}(D^{**})/\mathcal{B}(D^*)$ ratio to be the same for all channels;
- the correlations between the D^* and D^{**} parameters are at the same level as they are in the 9 parameters fit;
- there is a good agreement between the results with the muon and the electron channel.

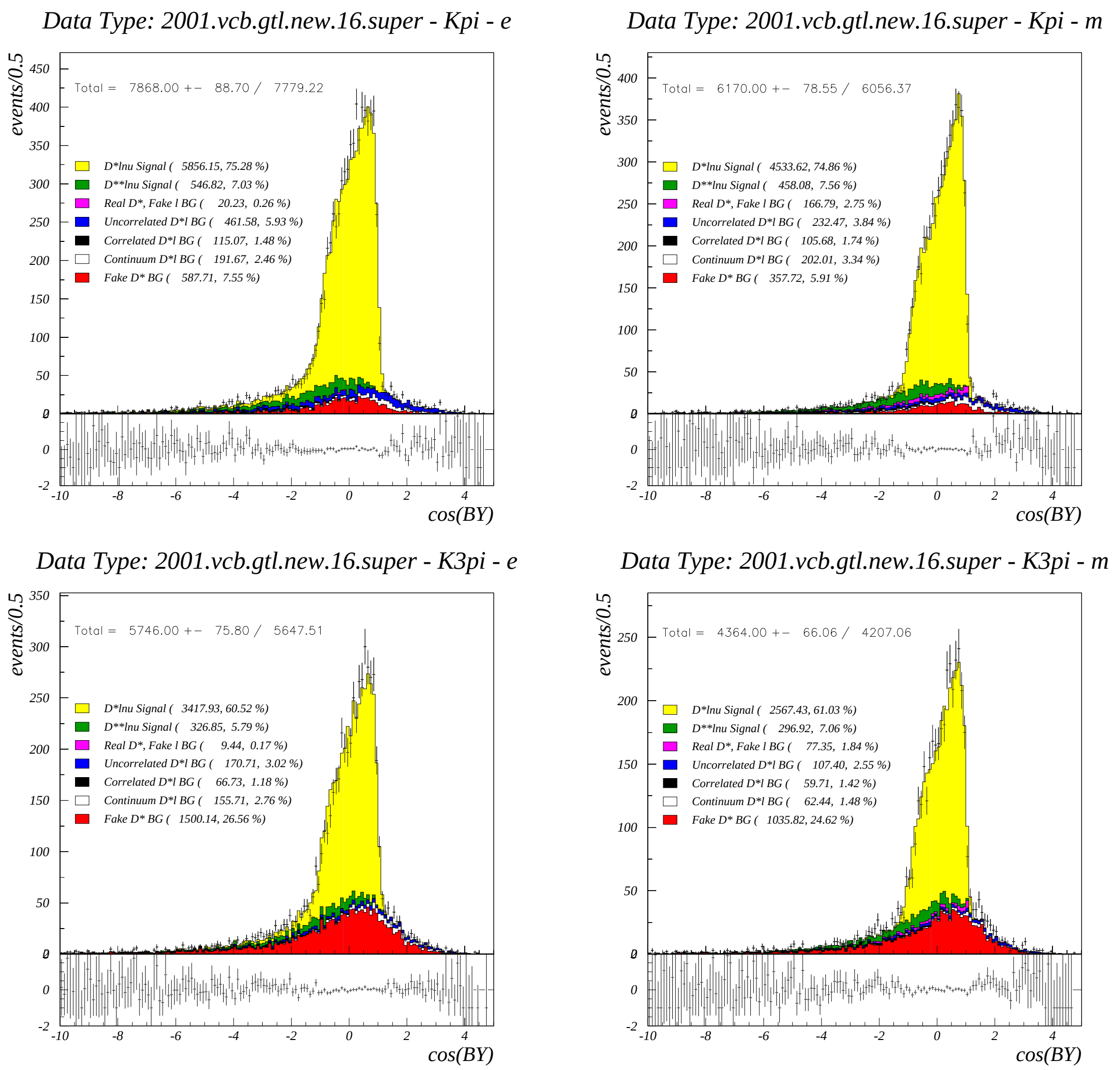


Figure I.1: Results of the fits and total sample with corrected backgrounds in $D^0 \to K\pi$ and $D^0 \to K\pi\pi\pi\ \ell$ channels for the D^{**} constrained 16 parameters fit. Each plot shows also the data/Monte Carlo ratio for each of the 150 bins.

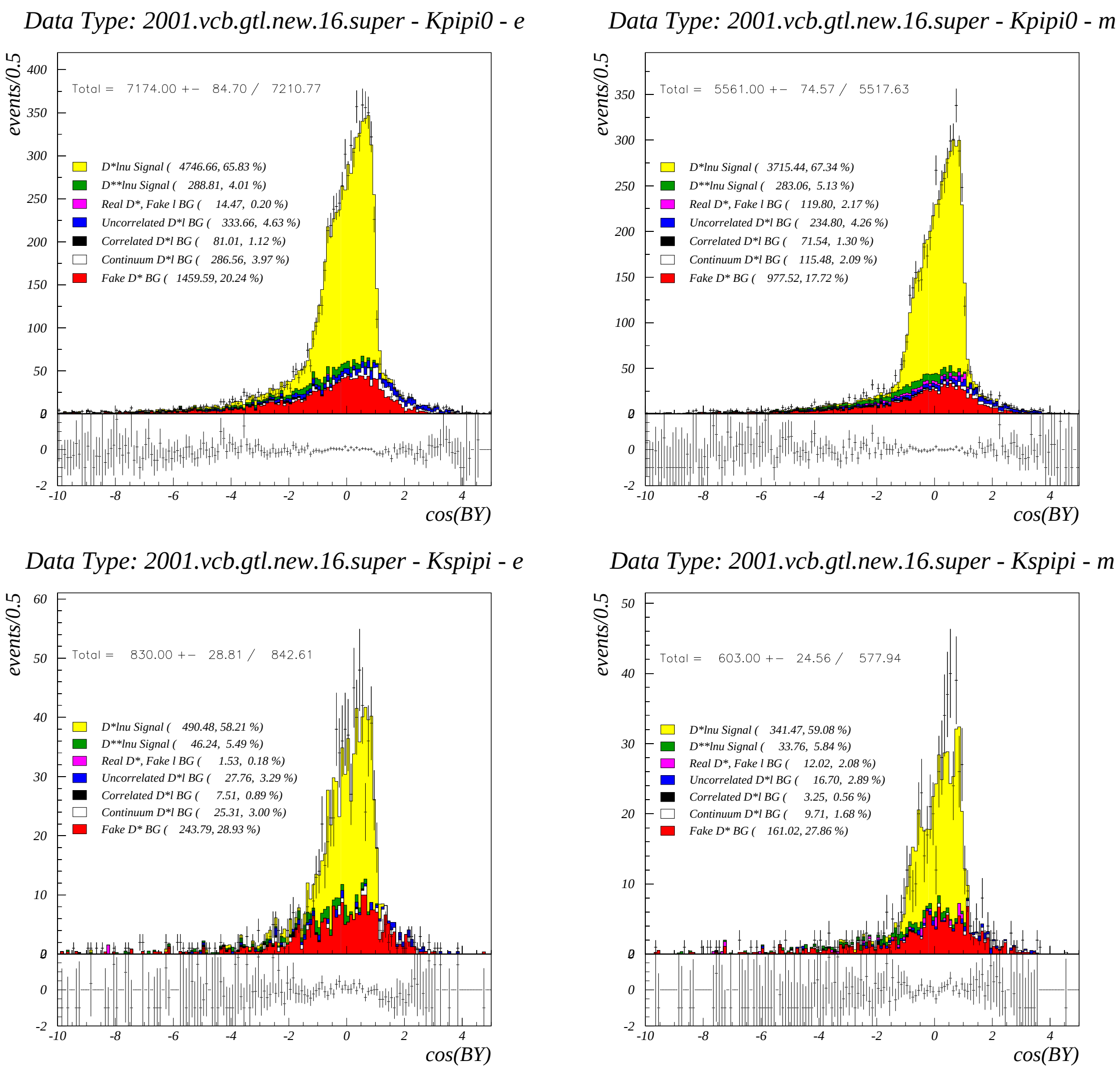


Figure I.2: Results of the fits and total sample with corrected backgrounds in $D^0 \to K\pi\pi^0$ and $D^0 \to K_S\pi\pi\ \ell$ channels the D^{**} constrained 16 parameters fit. Each plot shows also the data/Monte Carlo ratio for each of the 150 bins.

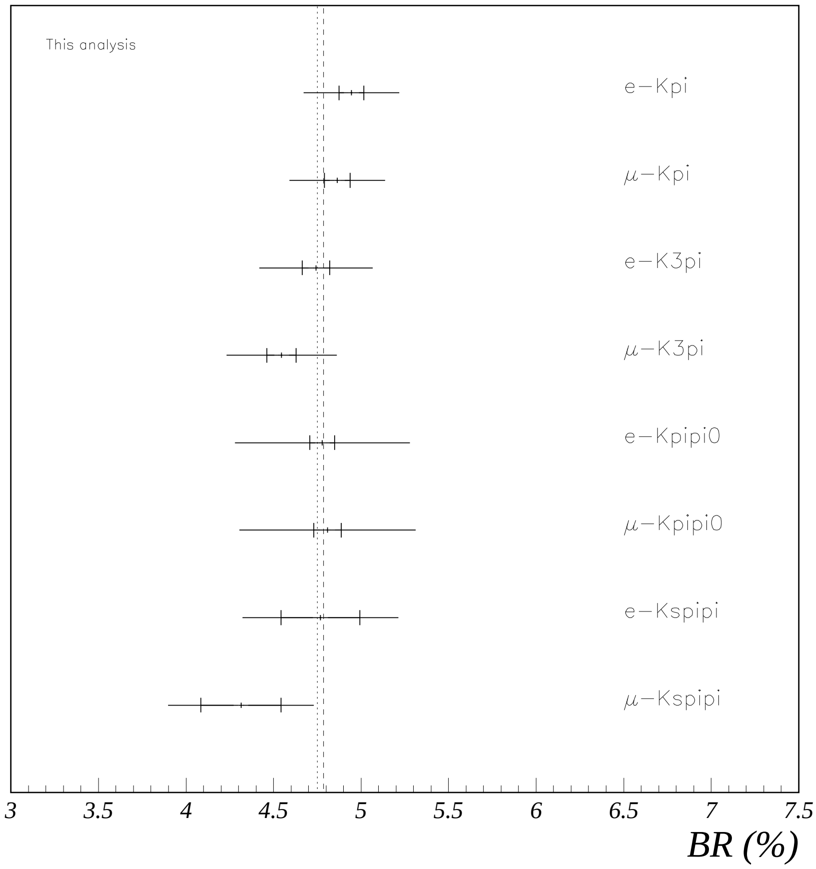


Figure I.3: Summary of all BR measurements with the D^{**} constrained 16 parameters fit. Plotted are statistical (inside bar) and statistical plus sytematic errors. The dashed vertical line corresponds to the average central value calculated assuming all the systematic errors correlated, while the dotted vertical line corresponds to the average central value calculated assuming all systematic errors to be uncorrelated. They are plotted only for reference.

parameter number	parameter	fitted value	statistical error
1	f_{D^*} Kpi e	0.87651	0.01489
2	$f_{D^{**}}$ Kpi e	0.47055	0.04671
3	f_{D^*} Kpi mu	0.87146	0.01553
4	$f_{D^{**}}$ Kpi mu	0.52342	0.04605
5	f_{D^*} K3pi e	0.84251	0.01939
6	$f_{D^{**}}$ K3pi e	0.43268	0.05771
7	f_{D^*} K3pi m	0.81451	0.02001
8	$f_{D^{**}}$ K3pi m	0.49171	0.05538
9	f_{D^*} Kpipi0	0.85298	0.01707
10	$f_{D^{**}}$ Kpipi0	0.34445	0.05906
11	f_{D^*} Kpipi0	0.86864	0.01817
12	$f_{D^{**}}$ Kpipi0	0.42878	0.05783
13	f_{D^*} Kspipi	0.8196	0.0528
14	$f_{D^{**}}$ Kspipi	0.8367	0.1417
15	f_{D^*} Kspipi	0.7346	0.0530
16	$f_{D^{**}}$ Kspipi	0.7453	0.1282

Table I.1: *Fitted values of f_{D^*} and $f_{D^{**}}/f_{D^*}$ with the 16 parametrs constrained fit bescribed above. The number shown says which scale factor one has to apply on the Monte Carlo generated $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ (for f_{D^*}) and on $\mathcal{B}(B \to D^{**}\ell\nu_\ell) + \mathcal{B}(B \to D^*\pi\ell\nu)$ (for $f_{D^{**}}$) to best match the data $\cos\theta_{B^0,D^*\ell}$ distribution in each of the 8 signal channels.*

	1-2	3-4	5-6	7-8	9-10	11-12	13-14	15-16
correlation	0.326	0.185	0.269	0.155	0.346	0.241	0.183	0.253

Table I.2: *Correlations of all fitted parameters numbered as is shown in table I.1. We show only the correlation of the pairs of f_{D^*} and $f_{D^{**}}$ of each of the 8 channels. There are no correlation among parameters outside the same signal channel by definition.*

	$\mathcal{B}$	stat DATA	stat MC	syst δm FIT	syst PID eff.	syst TRK eff.	syst VTX	syst p_ℓ^*	syst BIN	syst D^{**} fit	syst D^{**} model	syst L	syst BR D^0	syst BR D^*	syst f_{00}
e-Kpi	4.945	0.062	0.032	0.039	0.009	0.115	0.049	0.049	0.049	0.024	0.099	0.054	0.117	0.037	0.138
mu-Kpi	4.864	0.067	0.031	0.057	0.024	0.113	0.049	0.049	0.049	0.023	0.097	0.054	0.115	0.036	0.136
e-K3pi	4.743	0.071	0.034	0.048	0.057	0.126	0.047	0.047	0.047	0.034	0.095	0.052	0.197	0.035	0.133
mu-K3pi	4.546	0.077	0.033	0.052	0.059	0.121	0.045	0.045	0.045	0.035	0.091	0.050	0.189	0.034	0.127
e-Kpipi0	4.779	0.063	0.033	0.050	0.058	0.303	0.048	0.048	0.048	0.030	0.096	0.053	0.328	0.035	0.134
mu-Kpipi0	4.808	0.071	0.032	0.054	0.062	0.305	0.048	0.048	0.048	0.029	0.096	0.053	0.330	0.036	0.135
e-Kspipi	4.768	0.184	0.129	0.049	0.006	0.163	0.048	0.048	0.048	0.042	0.095	0.052	0.282	0.035	0.133
mu-Kspipi	4.314	0.196	0.118	0.050	0.021	0.147	0.043	0.043	0.043	0.034	0.086	0.047	0.255	0.032	0.121

Table I.3: *Values of $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ and statistical and systematics uncertainties in all the 8 signal channels for the 16 parameters fit with the constraint of D^{**} from [76].*

Appendix I : Cross checks results on $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ with different χ^2 fit

Acknowledgments

Obtaining a PhD in Particle Particle Physics requires sacrifice, hard work and suggestions from more experienced researchers: *no pain no gain*.

The first person I would like to thank wishes to remain anonymous: we can go skiing now that I submitted the thesis!!

Then I would like to thank Dr. Giuseppe Della Ricca: we started this analysis together and since the very first day he pushed hard on me to put all my efforts to reach the goal of completing this measurement by the end of my graduate studies. Thanks for all the enlightening suggestions and the help he gave me, especially for taking care of the bookkeeping and for providing his "magic scripts" that simplified a lot the treatment of such a huge amount of data. Thanks also for finding typos in this thesis.

Our analysis group is composed also by Dr. Diego Monorchio, Dr. Guglielmo De Nardo and Dr. Marcello Rotondo. Many thanks to Diego for his good job but I am sorry to remind him Triestina is still ahead of Napoli...
A *grazie* to Marcello: we spent many nights in his office talking about analysis and physics and I always enjoyed our discussions. We have a very similar approach to physics so it is always very easy to understand each other.
The next step of this analysis (the $|V_{cb}|$ measurement) will benefit from the help of Prof. Franco Simonetto. Many thanks to Franco for all the enlightening suggestions he gave me.

I want to thank the Stanford University group, Dr. Tim Meyer and Dr. Chih Cheng, and Prof. David Kirkby for all the support and the tools they provided to analyze $B^0 \to D^{*-}\ell^+\nu_\ell$ events.

I would like to thank also Dr. Concezio Bozzi for carefully reading the last chapter of this thesis.

During these three years I also experienced the thrill of being involved in data taking detector operations. When I was SVT commissioner I received a lot of help and friendship by Dr. Giovanni Calderini: *grazie di cuore Calde*. While acting as *BABAR* Deputy Run Coordinator I received a lot of help from Dr. Yannis Karyotakis.

Three years at SLAC means not only work but also many new friends. Thanks to everybody: I spent a wonderful time with you. We discovered a new particle, played very well the piano, enjoyed night-life in San Francisco (Rubyskye), surfed in Santa Cruz, skied in the wonderful Sierra Mountains and more much more than this. I promise you I will always behave as required by my nickname.
While at Stanford I lived in a wonderful place with a wonderful American grandmother: thanks Violetta!

I want to thank also all the friends I have in Ferrara and especially all the *Via Gavioli* gang. A special thank to Luca Boatto, he knows I owe him a lot, really a lot.

A big thank to my best friend: *grazie Raffi* for showing me that, if there is true friendship, 6000 miles are nothing; Pranz and you are really special people.

My parents always loved me and helped me in the good and in the bad times: I am lucky to have you as mom and dad.

Last and not least: Sarah (alias Conconi), my *Penelope*. You waited so long for me while I was far away. I spent the last three years like a globetrotter: now I am back and we have all the time on for us *ciccia*!

Bibliography

[1] S.L. Glashow, Nucl. Phys. **B 22** 579 (1961);
S. Weinberg, Phys. Rev. Lett. **19** 1264 (1967);
A. Salam, in Proc. of the 8th Nobel Symp., p. 367, ed N. Svartholm, Almqvist and Wiksell, Stockholm, 1968.

[2] J. B. Richman and P. R. Burchat, Rev. Mod. Phys. **67**, 893 (1995).

[3] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963);
M. Kobayashi, K. Maskawa, Prog. Theor. Phys. **49** 652 (1973).

[4] L.L. Chau and W.Y. Keung, Phys. Rev. Lett. **53**, 1802 (1984);
H. Harari and M. Leurer, Phys. Lett. B **181**, 123 (1986);
H. Fritzsch and J. Plankl, Phys. Rev. D **35**, 1732 (1987);
F. Botella and L.L. Chau, Phys. Lett. B **168**, 97 (1986).

[5] Particle Data Group, K. Hagiwara *et al.*, Phys. Rev. D **66**, 010001 (2002).

[6] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).

[7] *BABAR* Collaboration, B. Aubert *et al.*, SLAC-PUB-9293, hep-ex/0207042;
BELLE Collaboration, K. Abe *et al.*, BELLE-CONF-0201, hep-ex/0207098.

[8] For a pedagogical review see:
J. B. Richman, Les-Houches School in Particle Physics (1997) and references therein.

[9] N. Isgur and M.B. Wise, Phys. Lett. B **232**, 113 (1989);
N. Isgur and M.B. Wise, Phys. Lett. B **237**, 527 (1990).

[10] M.E. Luke, Phys. Lett. B **252**, 447 (1990).

[11] The *BABAR* Physics Book, P.F. Harrison and H.R. Quinn, SLAC-R-504 (1998).

[12] S. Hashimoto *et al.*, FERMILAB-PUB-01/317-T, hep-ph/0110253.

[13] CLEO Collaboration, R.A. Briere *et al.*, LNS Preprint CLNS 01-1776 (2002); hep-ex/0203032.

[14] CLEO Collaboration, J. Duboscq *et al.*, Phys. Rev. Lett. **76**, 3898 (1996).

[15] I. Caprini, L. Lellouch and M. Neubert, Nucl. Phys. B **530**, 153 (1998).

[16] M. Neubert, Phys. Rep. **245**, 259 (1994).

[17] CLEO Collaboration, B. Barish *et al.*, Phys. Rev. D **51**, 1014 (1995).

[18] S. Stone, in *B decays, 2nd Edition*, S. Stone editor, 283 (1994).

[19] BELLE Collaboration, K. Abe *et al.*, Phys. Lett. B **526**, 247 (2002) ; hep-ex/0111060.

[20] ALEPH Collaboration, D. Buskulic *et al.*, Phys. Lett. B **935**, 373 (1997).

[21] DELPHI Collaboration, P. Abreu *et al.*, Phys. Lett. B **510**, 55 (2001).

[22] OPAL Collaboration, G. Abbiendi *et al.*, Phys. Lett. B **482**, 15 (2000).

[23] LEP V_{cb} Working Group, Internal Note,
http://lepvcb.web.cern.ch/LEPVCB/

[24] PEP-II upgrade plans for the 2002 shutdown, Internal Note,
http://www.slac.stanford.edu/˜ sullivan/PEPII_upgrades.ppt

[25] Private communication with Dr. Michael Sullivan (PEP-II Run Coordinator).

[26] Private communication with Dr. Rainer Bartoldus.

[27] CLEO Collaboration, S.E. Csorna *et al.*, Phys. Rev. D **61**, 111101 (2000).

[28] B.N. Ratcliff, SLAC-PUB-5496 (1992);
B.N. Ratcliff, SLAC-PUB-6067 (1993);
P. Coyle *et al.*, Nucl. Instr. and Methods A **343**, 292 (1994).

[29] *BABAR* Collaboration, B. Aubert *et al.*, Nucl. Instr. and Methods A **479**, 1 (2002).

[30] M. Benkebil *et al.*, Nucl. Instr. and Methods A **442**, 364 (2000).

[31] D. Boutigny *et al.*, The *BABAR* Techical Design Report, SLAC-R-457.

[32] R. Santonico, R. Cardarelli, Nucl. Instr. and Methods A **187**, 377 (1981).

[33] A. Zallo *et al.*, *The BABAR RPC System* in Proceedings of the 5th International Workshop on Resistive Palte Chambers and related Detectors, Bari (Italy) 1999.

[34] http://www.slac.stanford.edu/BFROOT/www/Detector/IFR/IFR.html

[35] http://www.slac.stanford.edu/BFROOT/www/Organization/CollabMtgs/2002/detJul2002/Mon1/karyotakis.pdf

[36] http://www.slac.stanford.edu/BFROOT/www/Detector/IFR/remediation

[37] http://www.vxworks.com

[38] http://www.sun.com

[39] F. Safai Tehrani, The *BABAR* Prompt Reconstruction Manager: A Real Life Example of a Constructive Approach to Software Development.
submitted to Comput. Phys. Commun. .

[40] http://www.kernel.org

[41] S.A. Lewis, Overview of the Experimental Physics and Industrial Control System: EPICS, (2000).

[42] Objectivity/DB, Objectivity Inc., 301B East Evelyn Avenue, Mountain View, 94041 CA, USA.

[43] C. Bozzi *et al.*, Nucl. Instr. and Methods A **447**, 20 (2000).

[44] L. Bosisio, preliminary *BABAR* internal results.

[45] G. Della Ricca *et al.*, Nucl. Instr. and Methods A **409**, 258 (1998).

[46] V. Re *et al.*, Nucl. Instr. and Methods A **409**, 354 (1998).

[47] F. Lanni and F. Palombo, Nucl. Instr. and Methods A **379**, 399 (1996).

[48] A. Drescher, *et al.*, Nucl. Instrum. Methods **A237**, 464 (1985).

[49] Ralph Sinkus and Thomas Voss, Nucl. Instrum. Methods **A391**, 360 (1997).

[50] CLEO Collaboration, S. Anderson, *et al.*, CLEO CONF 99-6 (1999).

[51] DELPHI Collaboration, P. Abreu *et al.*, Phys. Lett. B **426**, 231 (1998), and references therein.

[52] CLEO Collaboration, A. Anastassov *et al.*, Phys. Rev. Lett. **80**, 4127 (1998).

[53] CLEO Collaboration, P. Avery *et al.*, Phys. Lett. B **331**, 236 (1994).

[54] D. Kirkby *et al.*, "Semileptonic B sample for Lifetime and Mixing Analyses", *BABAR* Analysis Document 147.

[55] T. Brandt, private communication.

[56] E687 Collaboration, P. L. Frabetti *et al.*, Phys. Lett. B **331**, 217 (1994).

[57] D. Kirkby *et al.*, "Selection of $D^{*}\ell$ Events for B Lifetime and Mixing Analyses", *BABAR* Analysis Document 34.

[58] S. Brandt, "Statistical and Computational Methods in Data Analysis", North Holland Publishing.

[59] Gerhard Raven, presentation at September 2000 Collaboration meeting, http://www.slac.stanford.edu/BFROOT/www/Organization/CollabMtgs/2000/detSep2000/Thu3a/raven.ps

[60] Paul Avery, "Directly Determining Lifetime Using a 3-D Fit", http://www.phys.ufl.edu/ avery/fitting/lifetime.ps, December 15, 1999.

[61] Fernando Martinez-Vidal *et al.*, "The *BABAR* Vertexing", *BABAR* Analysis Document 102.

[62] F. Martinez-Vidal *et al.*, "Performance and Control Samples of the *BABAR* Vertexing", *BABAR* Analysis Document 130.

[63] U. Langenegger *et al.*, "Cut-based Electron Identification", *BABAR* Analysis Document 90.

[64] Electron ID AWG and PID Tables, http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/Pid/Electrons/index.html, http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/Pid/Electrons/pidtables.html

[65] L. Lista *et al.*, "Muon Identification in the *BABAR* Experiment", *BABAR* Analysis Document 60.

[66] M. Bona *et al.*, "Report of the Tracking Efficiency task Force for 2001", *BABAR* Analysis Document 324.

[67] O. Long, "Measurement of the slow π relative efficiency using halicity distributions", *BABAR* Analysis Document 54.

[68] S. Menary (CLEO Collaboration), "Measuring the Relative Slow Pion Efficiency in the Data and Monte Carlo", CBX 92-103.

[69] S. Grancagnolo, "http://babar-hn.slac.stanford.edu:5090/HyperNews/get/recoTracking/835/3.html".

[70] J.M. Blatt and V. F. Weisskopf, "Theoretical Nuclear Physics", published by John Wiley & Sons, 1952.

[71] L. Aspinwall *et al.*, "Studies on π^0 reconstruction: status report", *BABAR* Analysis Document 20.

[72] E691 Collaboration, J. C. Anjos *et al.*, Phys. Rev. D **48**, 56 (1993).

[73] T. Brandt, private communication.

[74] C. Borean *et al.*, "Measurement of $B^0 \to D^{*-}\ell^+\nu_\ell$ Branching Ratio using the 2001 Data Sample", *BABAR* Analysis Document 457.

[75] B.E. Berger, "A measurement of the semileptonic branching fraction $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell)$ at CLEO", Doctoral Dissertation, 2001.

[76] O. Buchmueller *et al.*, "Extraction of the First Hadronic Mass Moment from Semileptonic B Decays", *BABAR* Analysis Document 409.

[77] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **87**, 091801 (2001).

[78] F. James and M. Roos, "MINUIT, Function Minimization and Error Analysis", CERN D506.

[79] C. Hearty, "Measurement of the Number of $\Upsilon(4S)$ Mesons Produced in Run 1 (B counting)", *BABAR* Analysis Document 134.

[80] C. Borean *et al.*, "Measurement of $B^0 \to D^{*-}\ell^+\nu_\ell$ Branching Ratio using the 2000 Data Sample", *BABAR* Analysis Document 546.

[81] C. Borean *et al.*, "Measurement of $B^0 \to D^{*-}\ell^+\nu_\ell$ Branching Ratio using the 2002 Data Sample", *BABAR* Analysis Document 547.