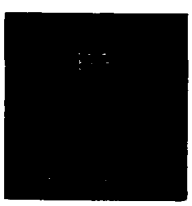
Measurement of CP Violation Parameters in $b \to c\bar{c}d$ Exclusive Decays at the *BABAR* Experiment

Justin E. Albert



A DISSERTATION PRESENTED TO THE FACULTY OF PRINCETON UNIVERSITY IN CANDIDACY FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE BY THE DEPARTMENT OF PHYSICS

November 2002

Abstract

The *BABAR* experiment at SLAC provides an opportunity for measurement of CP violation in B decays. A measurement of time-dependent CP violating asymmetries using exclusive B meson decays where the b quark decays to $c\bar{c}d$ (including $B^0 \to D^{*+}D^{*-}$ and $B^0 \to D^{*\pm}D^{\mp}$ decays) is presented here. This is the first measurement of CP violation in a mode sensitive to the Unitarity Triangle parameter $\sin 2\beta$ outside of decays containing charmonium. It provides a comparison to measurements of $\sin 2\beta$ using $b \to c\bar{c}s$, and permits an observation into potential new physics sources of CP violation, such as supersymmetry, via differences between these measurements and those of $B^0 \to J/\psi K_S^0$ as statistics of reconstructed neutral B decays to $D^{(*)+}D^{(*)-}$ increase. The measured value of the time-dependent CP violating asymmetries are:

$$S = 0.38 \pm 0.88(stat) \pm 0.12(syst)$$

$$C = -0.30 \pm 0.50(stat) \pm 0.13(syst)$$

for $B^0 \to D^{*-}D^{+}$,

$$S = -0.43 \pm 1.41(stat) \pm 0.23(syst)$$

$$C = 0.53 \pm 0.74(stat) \pm 0.15(syst)$$

for $B^0 \to D^{*+}D^{-}$, and

$$S = -0.05 \pm 0.45(stat) \pm 0.05(syst)$$

$$C = 0.12 \pm 0.30(stat) \pm 0.05(syst)$$

for $B^0 \to D^{*+}D^{*-}$, where S corresponds to CP violation in the interference of mixing and decay and C corresponds to CP violation in decay.

"thy fearful Symmetry"

— Blake

"All you can write is what you see."

— W. Guthrie

Contents

Acknowledgements

The analysis was done in concert with two groups: Dr. Gloria Vuagnin and Prof. Livio Lanceri from Universita de Trieste and Bram Lillard, Dr. Jochen Schieck, and Prof. Douglas Roberts from the University of Maryland. I am extremely gratefully indebted to them. Much of the work was shared between us and in particular the branching fraction and the *CP* fit for the $B^0 \to D^{*+}D^{*-}$ mode were largely the result of work from Bram, Jochen, and Doug and from Gloria and her Trieste collegues respectively. They have been wonderful people to work with and excellent physicists. Above all, I am especially happy about the prospect of continuing to work with them.

I am very much indebted to my parents for their emotional support without which this would have been difficult or impossible. I am indebted to my Princeton and *BABAR* colleagues who, needless to say, have been excellent people to work and talk with about physics.

Lastly, my advisor, Stew Smith, is the person without whom this work could never have taken place. He has truly been the foundation around which the physics has grown.

Introduction

The history of CP violation is one of experimental discovery overturning untested assumption. Observation of the $\theta - \tau$ puzzle in the early 1950s marks the inception of the discovery of the symmetry-violating properties of the weak interaction. Two spin-zero particles of the same mass and lifetime (now known to be charged kaons) were found to decay into different final states of opposite parity, one to two pions and the other to three, seemingly violating parity conservation. In 1956, Lee and Yang showed that parity conservation, while well-tested in strong and electromagnetic interactions, was not experimentally constrained for weak interactions, and proposed a list of experimental tests [1]. Shortly thereafter, C.S. Wu and collaborators performed one of these experiments, and showed that parity was not conserved in nuclear β decay, conclusively demonstrating the uniqueness of the weak interaction among the forces [2].[1] However, the combined CP transformation was still widely assumed to be a symmetry of nature due to the difficulty of explaining the weak interaction without it. The discovery eight years later of CP violation in neutral K mesons by Christenson, Cronin, Fitch and Turlay

[1]Note however that parity conservation is still poorly experimentally constrained in gravity.

provided the basis for both far-reaching insight (the Kobayashi-Maskawa prediction of a third family of quarks and leptons, a year before even the charm quark was discovered, and 4 years before the b) as well as uncertainty about the nature of this asymmetry that last to this day [3, 4].

Before 1964, no one realistically expected CP symmetry to be violated. Accomodating CP violating involves a significant increase in the complexity of weak interaction theory that was just not motivated at the time, as only the first 3 quarks had been discovered at that point. Fitch later remarked that "not many of our colleagues would have given credit for studying CP violation, but we did so anyway" [5]. The beam that was used contained pure K_2^0 ($= K_L^0$) mesons. A two-arm spectrometer was used in order to detect the decay products. After two months of data taking, a significant peak was indeed observed for a $\pi^+\pi^-$ decay of the K_2^0. If mass eigenstates and CP eigenstates were equivalent, K_2^0 would be the purely CP-odd (and slightly heavier) counterpart to the CP-even K_1^0 ($= K_S^0$). But the $\pi^+\pi^-$ final state is CP-even, thus the (presumably) CP-odd K_2^0 should not be able to decay into it *unless CP were violated.*

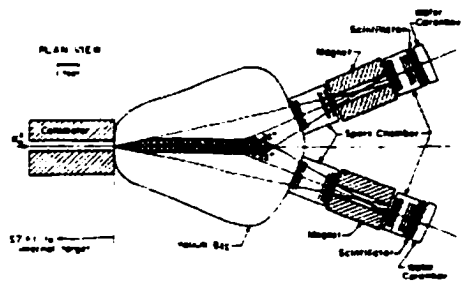


Figure 1: Plan view of the Fitch-Cronin experimental setup.

Fitch and Cronin observed a significant peak for a $\pi^+\pi^-$-hypothesis decay at the K_2^0 mass, consisting of 45 ± 9 two-pion events out of a total of 22,700 K_2^0 decays. Although the experiment did result in a slowly increasing acceptance in the physics community that CP was violated, immediately following

the measurement strong disbelief did exist. Alternative explanations were proposed, including regeneration of K_1^0, a non-bosonic version of the pion as the actual decay product, and violation of exponential decay laws. These alternatives were at least as unpleasant for theory as the violation of *CP* itself, and successive experiments refuted their possibility, eventually eliminating all but the Kobayashi-Maskawa quark mixing picture as the description of Cronin and Fitch's results.

Thirty-seven years of experimental study of the kaon sector after *CP* violation was discovered has yielded only recently the observation of direct *CP* violation, and has merely helped to confirm the counterintuitive picture of a small complex coefficient in a 3 x 3 unitary matrix as the source of the *CP* asymmetry. The smallness of *CP* violating effects in the kaon system is an impediment to progress in that sector, although the potential remains for measurements of the decays $K^+ \rightarrow \pi^+\nu\bar{\nu}$ and $K_L^0 \rightarrow \pi^0\nu\bar{\nu}$, which directly probe the imaginary part of the coefficient. The present and future lie in *B* decays, which, as shown last year by *BABAR* and Belle, exhibit significant *CP* asymmetries, as is predicted by the Standard Model. However, the predictions of the Standard Model regarding *CP* have yet to be fully examined experimentally, and, as seen above, one cannot take untested ideas for granted. Many well-motivated theoretical extensions of the Standard Model produce strikingly different predictions for *CP* violation; and the manifest baryon asymmetry of the universe poses great difficulty for reconciliation with the small amount of *CP* violation predicted by the Standard Model. Such tests are the primary purpose of the *BABAR* experiment — the goal is to either confirm or refute the Standard Model picture of *CP* violation.

Chapter 1 attempts to elucidate the Standard Model predictions for *CP* violation. and the ways. and reasons to expect. that such predictions may be disproven. Chapter 2 explains the details of the *BABAR* experiment. and. along with an overview of each of the detector subsystems. details contributions to the drift chamber (DCH) and to tracking software which were done by me (although I do not enjoy the necessity of defining personal contributions. the reality of the large collaboration forces one to detail what is one's own personal work as well as explaining the combined work of a very large group of physicists). which include determining the precise time sychronization of events. to drift chamber final assembly and contributions to detector commissioning. Chapter 3 contains the analysis of branching fractions for $B^0 \to D^{*+}D^{*-}$. $B^0 \to D^{*\pm}D^{\mp}$. and $B^{\pm} \to D^{*\pm}D^{*0}$. the latter two modes of which. as well as the initial discovery at *BABAR* of the first. are my analysis work. Chapter 4 contains the measurement of the time-dependent asymmetries in $B^0 \to D^{*+}D^{*-}$ and $B^0 \to D^{*\pm}D^{\mp}$. of which $B^0 \to D^{*\pm}D^{\mp}$ is my measurement. And Appendix C contains documentation on a kink-finder for *BABAR* which was a software contribution of my own.

The experimental study of *CP* violation within the framework of the Kobayashi-Maskawa quark mixing picture has proceeded for the past four decades. This thesis investigates the possibility of *CP* violation beyond those predictions. It inaugurates a new measurement sensitive to *CP* violation at mass scales higher than we can currently probe directly: and thus begins a new chapter in the quest for why the universe is made of matter.

Chapter 1

CP Violation in the B Meson System

1.1 Overview of CP Violation

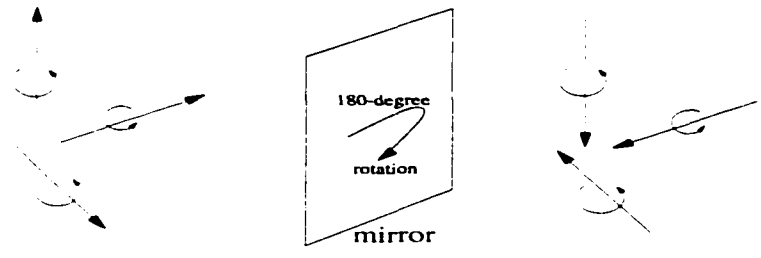
1.1.1 Discrete Symmetries

The set of operators on the Hilbert space of state functions on the quantum field contains both discrete and continuous transformations that preserve the Minkowski interval $t^2 - \boldsymbol{x}^2$. The set of continuous transformations that preserve this interval are the familiar Lorentz transformations, comprised of the product space of rotations, translations, and Lorentz boosts. The three independent discrete transformations that also preserve $t^2 - \boldsymbol{x}^2$ are the charge conjugation operator (C), the parity operator (P), and the time-reversal operator (T). These form a complete set of discrete Minkowski interval-preserving transformations of the Hilbert space. Although other

discrete interval-preserving transformations exist in the Standard Model, all can be formed from C, P, T, and the group of continuous Lorentz and gauge rotations.[1] The action of the three discrete transformations on, as an example, the special case of a spin 1/2 (Dirac) field, is discussed below.

Parity

The parity operator $\mathcal{P}$ reverses the signs of the 3 spatial elements of a four-vector: $(t, \boldsymbol{x}) \rightarrow (t, -\boldsymbol{x})$ and $(E, \boldsymbol{p}) \rightarrow (E, -\boldsymbol{p})$. One can easily visualize parity as a mirror-image plus an 180-degree rotation normal to the plane of the mirror (which works for any mirror angle) — this reverses the momentum of a particle but leaves its spin unchanged:



Consider the action of parity on the particle and antiparticle annihilation operators of the Dirac field $a^s_{\boldsymbol{p}}$ and $b^s_{\boldsymbol{p}}$. Parity should transform the states $a^s_{\boldsymbol{p}}|0\rangle$ and $b^s_{\boldsymbol{p}}|0\rangle$ to $a^s_{-\boldsymbol{p}}|0\rangle$ and $b^s_{-\boldsymbol{p}}|0\rangle$ as shown in the picture above. This implies

[1] For example, consider the set of discrete transformations $M_{\hat{n}}$, which take the mirror image of space with respect to a plane defined by a unit vector $\hat{n}$. This is simply parity combined with a rotation of π about $\hat{n}$. Minimal supersymmetry adds a single independent Lorentz-invariant transformation (R-parity), of which the symmetry is broken at observable energy levels, producing mass differences.

$$\mathcal{P}a^s_{\boldsymbol{p}}\mathcal{P}^{-1} = \eta_a a^s_{-\boldsymbol{p}} \qquad \text{and} \qquad \mathcal{P}a^s_{\boldsymbol{p}}\mathcal{P}^{-1} = \eta_b b^s_{-\boldsymbol{p}} \tag{1.1}$$

where η_a and η_b are phases. Since $\mathcal{P}^2 = \mathbf{1} \Rightarrow \eta_a . \eta_b$ must equal ± 1 (the parity group, as with the other two discrete operators, is idempotent, *ie.* $\mathcal{P}^{-1} = \mathcal{P}$, so the equation above could just as easily have been written $\mathcal{P}a^s_{\boldsymbol{p}}\mathcal{P}$, etc.) $\eta_a . \eta_b$ must equal ± 1. To find the matrix representation of $\mathcal{P}$ and the phases η_a and η_b, consider the action of $\mathcal{P}$ on $o(x)$. Decomposing o into eigenstates of spin and momentum gives:

$$\mathcal{P}o(x)\mathcal{P}^{-1} = \frac{1}{\sqrt{2E_{\boldsymbol{p}}}} \int \frac{d^3p}{(2\pi)^3} \sum_s \left(\eta_a a^s_{-\boldsymbol{p}} u^s(p) e^{-ipx} + \eta_b^* b^{s\dagger}_{-\boldsymbol{p}} v^s(p) e^{ipx} \right) \tag{1.2}$$

The key is to change variables to (not surprisingly) $p' = (p^0, -\boldsymbol{p}) \Rightarrow p \cdot x = p' \cdot (t, -\boldsymbol{x})$ and $p' \cdot \sigma$ (where σ is the four-vector of $2x2$ Pauli matrices) $= p \cdot \sigma^\dagger \gamma^0$ (where γ^0 is the 0-th Dirac matrix[2]) $= p \cdot \bar{\sigma}$, where

$$\bar{\sigma} \equiv \sigma^\dagger \gamma^0 \tag{1.3}$$

Thus the four-spinors $u(p)$ and $v(p)$ can be written as:

$$\begin{aligned} u(p) &= \begin{pmatrix} \sqrt{p \cdot \sigma}\varsigma \\ \sqrt{p \cdot \bar{\sigma}}\varsigma \end{pmatrix} = \begin{pmatrix} \sqrt{p' \cdot \bar{\sigma}}\varsigma \\ \sqrt{p' \cdot \sigma}\varsigma \end{pmatrix} = \gamma^0 u(p') \\ v(p) &= \begin{pmatrix} \sqrt{p \cdot \sigma}\varsigma \\ -\sqrt{p \cdot \bar{\sigma}}\varsigma \end{pmatrix} = \begin{pmatrix} \sqrt{p' \cdot \bar{\sigma}}\varsigma \\ -\sqrt{p' \cdot \sigma}\varsigma \end{pmatrix} = -\gamma^0 v(p') \end{aligned} \tag{1.4}$$

where ς is a generic two-component spinor. Thus (1.2) can be written as:

$$\mathcal{P}o(x)\mathcal{P}^{-1} = \frac{1}{\sqrt{2E_{\boldsymbol{p}'}}}\int \frac{d^3p'}{(2\pi)^3}\sum_s \Big(\eta_a a^s_{\boldsymbol{p}'}\gamma^0 u^s(p')e^{-ip'\cdot(t,-\boldsymbol{x})} - \eta_b^* b^{s\dagger}_{\boldsymbol{p}'}\gamma^0 v^s(p')e^{ip'\cdot(t,-\boldsymbol{x})}\Big) \tag{1.5}$$

But,

$$o(t,-\boldsymbol{x}) = \frac{1}{\sqrt{2E_{\boldsymbol{p}'}}}\int \frac{d^3p'}{(2\pi)^3}\sum_s \Big(a^s_{\boldsymbol{p}'}u^s(p')e^{-ip'\cdot(t,-\boldsymbol{x})} + b^{s\dagger}_{\boldsymbol{p}'}v^s(p')e^{ip'\cdot(t,-\boldsymbol{x})}\Big) \tag{1.6}$$

$\Rightarrow \eta_a = 1,\ \eta_b = -1,$ and

$$\mathcal{P}o(t,\boldsymbol{x})\mathcal{P}^{-1} = \gamma^0 o(t,-\boldsymbol{x}) \tag{1.7}$$

Now all that remains to be done is to see how the field bilinears (scalar, vector, tensor, pseudo-vector, pseudo-scalar) transform under parity.

Scalar The scalar bilinear, which can be represented as $\bar{o}o$, transforms as:

$$\begin{aligned}\mathcal{P}\bar{o}o\mathcal{P}^{-1} = \mathcal{P}\bar{o}\mathcal{P}^{-1}\mathcal{P}o\mathcal{P}^{-1} = (\mathcal{P}o^\dagger\mathcal{P}^{-1})\gamma^0\gamma^0(o(t,-\boldsymbol{x})) = \\ (\mathcal{P}o\mathcal{P}^{-1})^\dagger(o(t,-\boldsymbol{x})) = o(t,-\boldsymbol{x})^\dagger\gamma^0 o(t,-\boldsymbol{x}) = \\ +\bar{o}(t,-\boldsymbol{x})o(t,-\boldsymbol{x})\end{aligned} \tag{1.8}$$

thus, scalar quantities have "positive parity."

Pseudoscalar Pseudoscalars may be represened as $i\bar{\phi}\gamma^5\phi$, and transform as:

$$\mathcal{P}i\bar{\phi}\gamma^5\phi\mathcal{P}^{-1} = i\mathcal{P}\bar{\phi}\mathcal{P}^{-1}\gamma^5\mathcal{P}\phi\mathcal{P}^{-1} = i(\mathcal{P}\phi^\dagger\mathcal{P}^{-1})\gamma^0\gamma^5\gamma^0(\phi(t,-\boldsymbol{x})) =$$
$$-i(\mathcal{P}\phi\mathcal{P}^{-1})^\dagger\gamma^5(\phi(t,-\boldsymbol{x})) = -i\bar{\phi}(t,-\boldsymbol{x})\gamma^5\phi(t,-\boldsymbol{x}) \tag{1.9}$$

Pseudoscalars pick up an extra minus sign when undergoing parity transformation and thus have "negative parity."

Vector Vectors, which can be represented as $\bar{\phi}\gamma^\mu\phi$, transform as:

$$\mathcal{P}\bar{\phi}\gamma^\mu\phi\mathcal{P}^{-1} = \mathcal{P}\bar{\phi}\mathcal{P}^{-1}\gamma^\mu\mathcal{P}\phi\mathcal{P}^{-1} =$$
$$(\mathcal{P}\phi^\dagger\mathcal{P}^{-1})\gamma^0\gamma^\mu\gamma^0(\phi(t,-\boldsymbol{x})) = \bar{\phi}(t,-\boldsymbol{x})\gamma_\mu\phi(t,-\boldsymbol{x}) \tag{1.10}$$

Vectors thus have their spatial components sign-flipped, in the same manner as the spacetime four-vector itself.

Pseudovector Pseudovectors (axial vectors), which can be represented as $\bar{\phi}\gamma^\mu\gamma^5\phi$, transform as:

$$\mathcal{P}\bar{\phi}\gamma^\mu\gamma^5\phi\mathcal{P}^{-1} = \mathcal{P}\bar{\phi}\mathcal{P}^{-1}\gamma^\mu\gamma^5\mathcal{P}\phi\mathcal{P}^{-1} =$$
$$(\mathcal{P}\phi^\dagger\mathcal{P}^{-1})\gamma^0\gamma^\mu\gamma^5\gamma^0(\phi(t,-\boldsymbol{x})) = -\bar{\phi}(t,-\boldsymbol{x})\gamma_\mu\gamma^5\phi(t,-\boldsymbol{x}) \tag{1.11}$$

Pseudovectors have their time component sign-flipped, with the spatial components unchanged. Like pseudoscalars, they pick up an extra minus sign relative to their non-pseudo counterparts.

Tensor Tensors can be represented as $\bar{o}[\gamma^\mu, \gamma^\nu]o$ and transform as:

$$\mathcal{P}\bar{o}\gamma^\mu\gamma^\nu o\mathcal{P}^{-1} - \mathcal{P}\bar{o}\gamma^\nu\gamma^\mu o\mathcal{P}^{-1} =$$
$$(\mathcal{P}o^\dagger\mathcal{P}^{-1})\gamma^0\gamma^\mu\gamma^\nu\gamma^0(o(t, -\boldsymbol{x})) - (\mathcal{P}o^\dagger\mathcal{P}^{-1})\gamma^0\gamma^\nu\gamma^\mu\gamma^0(o(t, -\boldsymbol{x}))$$

Since γ^0 anticommutes with all γ matrices except itself, the matrix of sign-changes can be summarized as

$$\begin{pmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{pmatrix} \tag{1.12}$$

and thus

$$\mathcal{P}\bar{o}[\gamma^\mu, \gamma^\nu]o\mathcal{P}^{-1} = \bar{o}(t, -\boldsymbol{x})[\gamma_\mu, \gamma_\nu]o(t, -\boldsymbol{x}) \tag{1.13}$$

Derivative Operator It is also useful to note how the derivative operator ∂_μ transforms under parity — as spatial coordinates change sign. $(\partial/\partial t, \partial/\partial x, \partial/\partial y, \partial/\partial z) \rightarrow (\partial/\partial t, -\partial/\partial x, -\partial/\partial y, -\partial/\partial z) \Rightarrow$

$$\mathcal{P}\partial_\mu = \partial^\mu\mathcal{P} \tag{1.14}$$

Time Reversal

The time reversal operator reverses momentum and spin and also flips the sign of the time component of a state. Therefore we want the transformation of the Dirac particle and antiparticle annihilation operators to be:

$$\mathcal{T} a^s_{\boldsymbol{p}} \mathcal{T}^{-1} = \eta'_a a^{-s}_{-\boldsymbol{p}} \qquad \text{and} \qquad \mathcal{T} b^s_{\boldsymbol{p}} \mathcal{T}^{-1} = \eta'_b b^{-s}_{-\boldsymbol{p}} \tag{1.15}$$

We can start to compute the transformation of the fermion field ϕ:

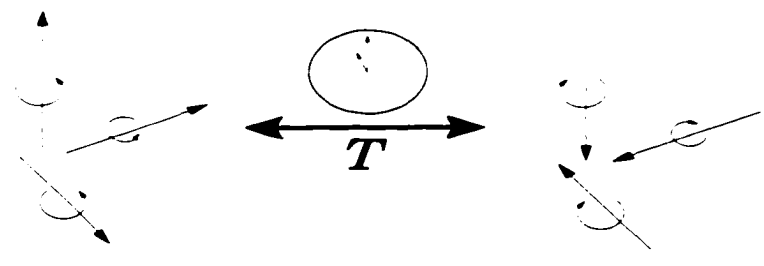
$$\mathcal{T}\phi(t,\boldsymbol{x})\mathcal{T}^{-1} = \frac{1}{\sqrt{2E_{\boldsymbol{p}}}} \int \frac{d^3p}{(2\pi)^3} \sum_s \mathcal{T}\left(a^s_{\boldsymbol{p}} u^s(p) e^{-ipx} + b^{s\dagger}_{\boldsymbol{p}} v^s(p) e^{ipx}\right) \mathcal{T}^{-1} \tag{1.16}$$

However, if $\mathcal{T}$ were to only act on the operators a and b, the situation would be the same as with parity and the spatial coordinates would flip sign instead of time (also the operators would reverse spin but not the spinors, which would be an unphysical nonlinearity). $\mathcal{T}$ therefore must act on more than just the operators.

The solution is to let $\mathcal{T}$ act on complex numbers in addition to operators. Let

$$\mathcal{T} z = z^* \mathcal{T} \quad \forall z \varepsilon \mathcal{C} \tag{1.17}$$

Thus (1.16) becomes



$$\frac{1}{\sqrt{2E_{\boldsymbol{p}}}}\int\frac{d^3p}{(2\pi)^3}\sum_s\left(\eta_a'^{*}a^{-s}_{-\boldsymbol{p}}(u^s(p))^{*}e^{ipx}+\eta_b'^{*}b^{-s\dagger}_{-\boldsymbol{p}}(v^s(p))^{*}e^{-ipx}\right) \tag{1.18}$$

We need to find a constant matrix $\mathcal{M}$ such that $\mathcal{M}u^{-s}(p') = (u^s(p))^*$ (and similarly for $v^s(p)$) — then we can change variables to p' and $(-t, \boldsymbol{x})$ so that we can obtain an answer for the action of the transformation in terms of $o(-t, \boldsymbol{x})$.

We can see that:

$$(u^s(p))^* = \begin{pmatrix}\sqrt{p\cdot\sigma^*}\varsigma^{s*}\\ \sqrt{p\cdot\bar{\sigma}^*}\varsigma^{s*}\end{pmatrix} = \begin{pmatrix}\sigma^2 & 0\\ 0 & \sigma^2\end{pmatrix}\begin{pmatrix}\sigma^2 & 0\\ 0 & \sigma^2\end{pmatrix}\begin{pmatrix}\sqrt{p\cdot\sigma^*}\varsigma^{s*}\\ \sqrt{p\cdot\bar{\sigma}^*}\varsigma^{s*}\end{pmatrix} =$$
$$-\gamma^1\gamma^3\begin{pmatrix}-i\sigma^2\sqrt{p\cdot\sigma^*}\varsigma^{s*}\\ -i\sigma^2\sqrt{p\cdot\bar{\sigma}^*}\varsigma^{s*}\end{pmatrix} \tag{1.19}$$

and can then use the identity

$$\sigma^2\sqrt{p\cdot\sigma^*} = \sqrt{p'\cdot\sigma}\,\sigma^2 \tag{1.20}$$

and the fact that

$$-i\sigma^2\varsigma^{s*} = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}\varsigma^{1*}\\ \varsigma^{2*}\end{pmatrix} = \begin{pmatrix}\varsigma^{2*}\\ -\varsigma^{1*}\end{pmatrix} = \varsigma^{-s} \tag{1.21}$$

to obtain for (1.19):

$$-\gamma^1\gamma^3\begin{pmatrix}\sqrt{p'\cdot\sigma}(-i\sigma^2\varsigma^{s*})\\ \sqrt{p'\cdot\bar{\sigma}}(-i\sigma^2\varsigma^{s*})\end{pmatrix} = -\gamma^1\gamma^3\begin{pmatrix}\sqrt{p'\cdot\sigma}\varsigma^{-s}\\ \sqrt{p'\cdot\bar{\sigma}}\varsigma^{-s}\end{pmatrix} = -\gamma^1\gamma^3u^{-s}(p') \tag{1.22}$$

and similarly for $(v^s(p))^*$. Thus (1.18) becomes

$$-\gamma^1\gamma^3\frac{1}{\sqrt{2E_{\boldsymbol{p}'}}}\int\frac{d^3p'}{(2\pi)^3}\sum_s \quad \Big(\eta_a'^* a^{-s}_{\boldsymbol{p}'}u^s(p')e^{-ip'\cdot(-t,\boldsymbol{x})}$$
$$+\eta_b'^* b^{-s\dagger}_{\boldsymbol{p}'}v^s(p')e^{ip'\cdot(-t,\boldsymbol{x})}\Big) \Rightarrow$$
$$\mathcal{T}o(t,\boldsymbol{x})\mathcal{T}^{-1} = -\gamma^1\gamma^3 o(-t,\boldsymbol{x}) \tag{1.23}$$

Scalar The scalar bilinear $\bar{o}o$ transforms as:

$$\mathcal{T}\bar{o}o\mathcal{T}^{-1} = \mathcal{T}o^\dagger\mathcal{T}^{-1}\gamma^0\mathcal{T}o\mathcal{T}^{-1} = -(\mathcal{T}o\mathcal{T}^{-1})^\dagger\gamma^0\gamma^1\gamma^3(o(-t,\boldsymbol{x})) =$$
$$-o^\dagger(-t,\boldsymbol{x})\gamma^1\gamma^3\gamma^0\gamma^1\gamma^3(o(-t,\boldsymbol{x})) = -\bar{o}(-t,\boldsymbol{x})\gamma^1\gamma^3\gamma^1\gamma^3 o(-t,\boldsymbol{x}) =$$
$$+\bar{o}(-t,\boldsymbol{x})o(-t,\boldsymbol{x}) \tag{1.24}$$

Pseudoscalar The pseudoscalar transforms as:

$$\mathcal{T}i\bar{o}\gamma^5 o\mathcal{T}^{-1} = -i\mathcal{T}o^\dagger\mathcal{T}^{-1}\gamma^0\gamma^5\mathcal{T}o\mathcal{T}^{-1} = i\bar{o}(-t,\boldsymbol{x})\gamma^1\gamma^3\gamma^5\gamma^1\gamma^3 o(-t,\boldsymbol{x}) =$$
$$i\bar{o}(-t,\boldsymbol{x})\gamma^5\gamma^1\gamma^3\gamma^1\gamma^3 o(-t,\boldsymbol{x}) = -i\bar{o}(-t,\boldsymbol{x})o(-t,\boldsymbol{x}) \tag{1.25}$$

Vector The vector transforms as:

$$\mathcal{T}\bar{o}\gamma^\mu o\mathcal{T}^{-1} = \mathcal{T}o^\dagger\mathcal{T}^{-1}\gamma^0\gamma^{\mu*}\mathcal{T}o\mathcal{T}^{-1} = -\bar{o}(-t,\boldsymbol{x})\gamma^1\gamma^3\gamma^{\mu*}\gamma^1\gamma^3 o(-t,\boldsymbol{x}) =$$
$$-\bar{o}(-t,\boldsymbol{x})\gamma_\mu\gamma^1\gamma^3\gamma^1\gamma^3 o(-t,\boldsymbol{x}) = \bar{o}(-t,\boldsymbol{x})\gamma_\mu o(-t,\boldsymbol{x}) \tag{1.26}$$

Pseudovector The pseudovector transforms as:

$$\mathcal{T}\bar{o}\gamma^{\mu}\gamma^{5}o\mathcal{T}^{-1} = \mathcal{T}o^{\dagger}\mathcal{T}^{-1}\gamma^{0}\gamma^{\mu *}\gamma^{5}\mathcal{T}o\mathcal{T}^{-1} = -\bar{o}(-t,\boldsymbol{x})\gamma^{1}\gamma^{3}\gamma^{\mu *}\gamma^{5}\gamma^{1}\gamma^{3}o(-t,\boldsymbol{x}) =$$
$$-\bar{o}(-t,\boldsymbol{x})\gamma_{\mu}\gamma^{5}\gamma^{1}\gamma^{3}\gamma^{1}\gamma^{3}o(-t,\boldsymbol{x}) = \bar{o}(-t,\boldsymbol{x})\gamma_{\mu}\gamma^{5}o(-t,\boldsymbol{x}) \qquad (1.27)$$

Tensor Tensors transform as:

$$\mathcal{T}i\bar{o}[\gamma^{\mu},\gamma^{\nu}]o\mathcal{T}^{-1} = -\bar{o}(-t,\boldsymbol{x})\gamma^{1}\gamma^{3}(\gamma^{\mu}\gamma^{\nu} - \gamma^{\nu}\gamma^{\mu})^{*}\gamma^{1}\gamma^{3}o(-t,\boldsymbol{x}) =$$
$$\bar{o}(-t,\boldsymbol{x})[\gamma_{\mu},\gamma_{\nu}]\gamma^{1}\gamma^{3}\gamma^{1}\gamma^{3}o(-t,\boldsymbol{x}) = -\bar{o}(-t,\boldsymbol{x})[\gamma_{\mu},\gamma_{\nu}]o(-t,\boldsymbol{x}) \quad (1.28)$$

Derivative Operator Time reversal negates the time component, so:

$$\mathcal{T}\partial_{\mu}\mathcal{T}^{-1} = -\partial^{\mu} \qquad (1.29)$$

Charge Conjugation

The charge conjugation operator is defined to be the transformation of a particle into its antiparticle without changing momentum or spin. Thus,

$$\mathcal{C}a^{s}_{\boldsymbol{p}}\mathcal{C}^{-1} = \eta''_{a}b^{s}_{\boldsymbol{p}} \qquad \text{and} \qquad \mathcal{C}b^{s}_{\boldsymbol{p}}\mathcal{C}^{-1} = \eta''_{b}a^{s}_{\boldsymbol{p}} \qquad (1.30)$$

so the transformation of the Dirac field is

$$\mathcal{C}o(x)\mathcal{C}^{-1} = \frac{1}{\sqrt{2E_{\boldsymbol{p}}}}\int\frac{d^{3}p}{(2\pi)^{3}}\sum_{s}\left(b^{s}_{\boldsymbol{p}}u^{s}(p)e^{-ipx} + a^{s\dagger}_{\boldsymbol{p}}v^{s}(p)e^{ipx}\right) \qquad (1.31)$$

We want to find what this is in terms of $\bar{\phi} = \phi^\dagger\gamma^0$, so we need a relation between $u^s(p)$ and $v^{s*}(p)$, and between $v^s(p)$ and $u^{s*}(p)$:

$$u^{s*}(p) = \begin{pmatrix} \sqrt{p\cdot\sigma^*}\varsigma^{s*} \\ \sqrt{p\cdot\bar{\sigma}^*}\varsigma^{s*} \end{pmatrix} = \begin{pmatrix} \sqrt{p\cdot\sigma^*}\varsigma^{s*} \\ \sqrt{p\cdot\bar{\sigma}^*}\varsigma^{s*} \end{pmatrix} = \begin{pmatrix} i\sqrt{p\cdot\sigma^*}\sigma^2\varsigma^{-s} \\ i\sqrt{p\cdot\bar{\sigma}^*}\sigma^2\varsigma^{-s} \end{pmatrix} \tag{1.32}$$

However, from the identity (1.20) we can see that:

$$\begin{aligned} \sqrt{p\cdot\sigma^*}\sigma^2 &= \sigma^2\sqrt{p\cdot\bar{\sigma}} \qquad \text{and} \\ \sqrt{p\cdot\bar{\sigma}^*}\sigma^2 &= \sigma^2\sqrt{p\cdot\sigma} \end{aligned} \tag{1.33}$$

Thus,

$$u^{s*}(p) = \begin{pmatrix} i\sigma^2\sqrt{p\cdot\bar{\sigma}}\varsigma^{-s} \\ i\sigma^2\sqrt{p\cdot\sigma}\varsigma^{-s} \end{pmatrix} = \begin{pmatrix} 0 & -i\sigma^2 \\ i\sigma^2 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{p\cdot\sigma}\varsigma^{-s} \\ -\sqrt{p\cdot\bar{\sigma}}\varsigma^{-s} \end{pmatrix} = -i\gamma^2 v^s(p). \tag{1.34}$$

Similarly, $v^{s*}(p) = -i\gamma^2 u^s(p)$, so (1.31) becomes:

$$\begin{aligned} \mathcal{C}\phi(x)\mathcal{C}^{-1} &= \frac{1}{\sqrt{2E_{\boldsymbol{p}}}}\int\frac{d^3p}{(2\pi)^3}\sum_s\left(i\gamma^2 b^s_{\boldsymbol{p}} v^{s*}(p)e^{-ipx} + i\gamma^2 a^{s\dagger}_{\boldsymbol{p}} u^{s*}(p)e^{ipx}\right) \\ &= i\gamma^2\phi^*(x) = i(\bar{\phi}\gamma^0\gamma^2)^T \end{aligned} \tag{1.35}$$

Scalar

$$\mathcal{C}\bar{\phi}(x)\mathcal{C}^{-1} = (\mathcal{C}\phi(x)\mathcal{C}^{-1})^\dagger\gamma^0 = -i(\gamma^0\gamma^2\phi(x))^T \tag{1.36}$$

Therefore,

$$
\begin{aligned}
\mathcal{C}\bar{o}(x)o(x)\mathcal{C}^{-1} &= (\bar{o}\gamma^0\gamma^2\gamma^0\gamma^2 o)^T \\
&= -(\bar{o}\gamma^2\gamma^0\gamma^0\gamma^2 o)^T = +(\bar{o}(x)o(x))^T \\
&= +\bar{o}(x)o(x) \qquad (1.37)
\end{aligned}
$$

Pseudoscalar

$$
\begin{aligned}
i\mathcal{C}\bar{o}(x)\gamma^5 o(x)\mathcal{C}^{-1} &= i(\bar{o}\gamma^0\gamma^2\gamma^5\gamma^0\gamma^2 o)^T \\
&= i(\bar{o}\gamma^5\gamma^0\gamma^2\gamma^0\gamma^2 o)^T = i(\bar{o}\gamma^5 o)^T \\
&= +i\bar{o}(x)\gamma^5 o(x) \qquad (1.38)
\end{aligned}
$$

Vector

$$
\begin{aligned}
\mathcal{C}\bar{o}(x)\gamma^\mu o(x)\mathcal{C}^{-1} &= (\bar{o}\gamma^0\gamma^2\gamma^\mu\gamma^0\gamma^2 o)^T \\
&= \begin{cases} (\bar{o}\gamma^\mu o)^T & \text{for } \mu = 1 \text{ or } 3 \\ -(\bar{o}\gamma^\mu o)^T & \text{for } \mu = 0 \text{ or } 2 \end{cases} \qquad (1.39) \\
&= -\bar{o}\gamma^\mu o \quad (\text{since } \gamma^1 \text{ and } \gamma^3 \text{ are antisymmetric})
\end{aligned}
$$

Pseudovector

$$
\begin{aligned}
\mathcal{C}\bar{o}(x)\gamma^\mu\gamma^5 o(x)\mathcal{C}^{-1} &= (\bar{o}\gamma^0\gamma^2\gamma^\mu\gamma^5\gamma^0\gamma^2 o)^T \\
&= \begin{cases} (\bar{o}\gamma^\mu\gamma^5 o)^T & \text{for } \mu = 1 \text{ or } 3 \\ -(\bar{o}\gamma^\mu\gamma^5 o)^T & \text{for } \mu = 0 \text{ or } 2 \end{cases} \qquad (1.40) \\
&= +\bar{o}\gamma^\mu\gamma^5 o \quad (\text{since } \gamma^0\gamma^5 \text{ and } \gamma^2\gamma^5 \text{ are antisymmetric})
\end{aligned}
$$

Tensor

$$
i\mathcal{C}\bar{o}[\gamma^\mu, \gamma^\nu]o\mathcal{C}^{-1} = i(\bar{o}\gamma^0\gamma^2[\gamma^\mu, \gamma^\nu]\gamma^0\gamma^2 o)^T
$$

$$
\begin{aligned}
&= \begin{cases} i(\bar{\mathcal{O}}[\gamma^\mu, \gamma^\nu]\mathcal{O})^T & \text{for } \mu + \nu = 2n \\ -i(\bar{\mathcal{O}}[\gamma^\mu, \gamma^\nu]\mathcal{O})^T & \text{for } \mu + \nu = 2n + 1 \end{cases} \\
&= -i\bar{\mathcal{O}}[\gamma^\mu, \gamma^\nu]\mathcal{O} \quad ([\gamma^\mu, \gamma^\nu], \text{ where } \mu + \nu = 2n, \text{ is antisymmetric})
\end{aligned} \tag{1.41}
$$

Derivative Operator The derivative operator is clearly unaffected by charge conjugation

$$
\mathcal{C}\partial_\mu \mathcal{C}^{-1} = \partial_\mu \tag{1.42}
$$

CPT

The combination $\mathcal{CPT}$ operator has a rather special property: it is guaranteed to be a fundamental symmetry of nature, with only the basic assumptions of Lorentz invariance, locality, and the spin-statistics relation.[2] A proof for the restricted case of the Dirac field follows.

As shown in the previous pages, and summarized in the table below, scalars, pseudoscalars, vectors, pseudovectors, and tensors are affected by the discrete symmetries as follows:

The Lagrangian $\mathcal{L}$ is a Lorentz scalar, and as we can see above, any contraction of indices to form a Lorentz scalar must result in an eigenstate with a $+1$ $\mathcal{CPT}$ eigenvalue.

[2]Note that the spin-statistics relation itself is implied from Lorentz invariance, positive energies, positive norms, and causality.

1.1.2 CP

The CP transformation properties of the fermion field bilinears are listed in the column next to CPT. As we can see, if we restrict our attention to scalars, pseudoscalars, vectors, and the derivative operator, a Lagrangian formed from only such quantities must remain CP-invariant. Thus a massless spin 1/2 field with real coupling constants cannot violate CP. This is in fact true for quantum fields of any spin. Charge conjugation ensures that the fields themselves transform to their Hermitian conjugates (we have seen this above for the special case of spin 1/2). However, particle masses and coupling constants do not transform under CP (as complex numbers such as these are only transformed by, of the discrete operators, $\mathcal{T}$, as seen previously). If any of these quantities is not purely real, it will suffer a phase shift relative to the quantities that are transformed by CP, thus potentially violating CP symmetry.

Such phase differences must be robust against gauge modifications in order to manifest themselves as CP violation. If simple redefinitions of the phases of any of the fields can remove overall phases in each field coupling, the theory remains CP-conserving. As will be shown in Section 1.4, if only two fermion generations are present, such a redefinition always exists, hence the Kobayashi-Maskawa prediction of a third generation. The effect of irreducible CP-violating phases will be elucidated in the following sections.

1.2 Mixing and Time Evolution of Neutral Mesons

The four pairs of conjugate neutral mesons that decay weakly, K^0, D^0, B^0, and B_s, can each mix with their respective antiparticle via a pair of box diagrams:

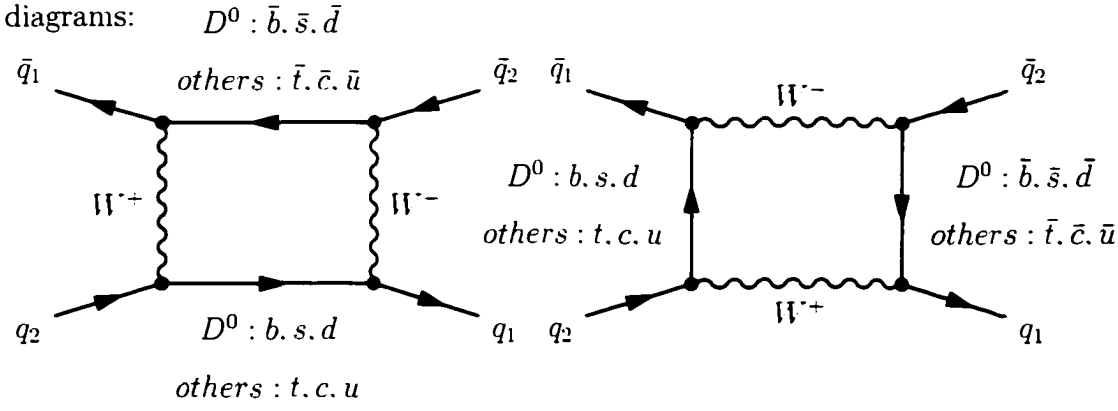


Figure 1.1: The two main mixing diagrams

The ability to mix implies that the flavor eigenstates may not be equivalent to the mass eigenstates: the observed presence of mixing (into conjugate flavor-specific decays) implies that the mass and flavor eigenstates are in fact different.

Lack of CP symmetry implies a third set of eigenstates, CP eigenstates, which can differ from the mass and flavor eigenstates, as will be seen below.

1.2.1 Mixing of a "Generic" Neutral Meson

Consider a weakly-decaying neutral meson X^0 (which could be any of K^0, D^0, B^0, or B_s). An arbitrary linear combination of the flavor eigenstates

$$\alpha|X^0\rangle + \beta|\bar{X}^0\rangle \tag{1.43}$$

mixes according to the time-dependent Schrödinger equation

$$i\frac{\partial}{\partial t}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \mathcal{H}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \equiv \begin{pmatrix} m_{11} - \frac{1}{2}i\gamma_{11} & m_{12} - \frac{1}{2}i\gamma_{12} \\ m_{21} - \frac{1}{2}i\gamma_{21} & m_{22} - \frac{1}{2}i\gamma_{22}\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \tag{1.44}$$

The m and γ parts represent the mixing and decay parts, respectively, of the time dependence. Each of the off-diagonal elements can be complex: the angle in the complex plane of m_{12} represents the phase of the mixing, and γ_{12} represents the (complex) coupling to common decay modes of X^0 and $\bar{X}^0$ (for example, $B^0/\bar{B}^0 \to J/\psi K_S^0$ or $\pi^+\pi^-$). We can see that CPT invariance guarantees that $m_{11} = m_{22}$ and $\gamma_{11} = \gamma_{22}$, and that $m_{21} = m_{12}^*$ and $\gamma_{21} = \gamma_{12}^*$ — the CPT-conjugate pairs of equations are:

$$\begin{aligned} i\frac{\partial\alpha}{\partial t} &= (m_{11} - \frac{1}{2}i\gamma_{11})\alpha + (m_{12} - \frac{1}{2}i\gamma_{12})\beta \\ i\frac{\partial\beta}{\partial t} &= (m_{21} - \frac{1}{2}i\gamma_{21})\alpha + (m_{22} - \frac{1}{2}i\gamma_{22})\beta \qquad \text{and} \qquad (1.45) \\ i\frac{\partial\beta}{\partial t} &= (m_{11} - \frac{1}{2}i\gamma_{11})\beta + (m_{12}^* - \frac{1}{2}i\gamma_{12}^*)\alpha \\ i\frac{\partial\alpha}{\partial t} &= (m_{21}^* - \frac{1}{2}i\gamma_{21}^*)\beta + (m_{22} - \frac{1}{2}i\gamma_{22})\alpha \qquad (1.46) \end{aligned}$$

which must be equivalent. Thus, setting m_{11} and m_{22} to m and γ_{11} and γ_{22} to γ, we have:

$$i\frac{\partial}{\partial t}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix} m - \frac{1}{2}i\gamma & m_{12} - \frac{1}{2}i\gamma_{12} \\ m_{12}^* - \frac{1}{2}i\gamma_{12}^* & m - \frac{1}{2}i\gamma\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \tag{1.47}$$

The mass eigenstates are the eigenvectors of the Hamiltonian:

$$|X_L\rangle = p|X^0\rangle + q|\bar{X}^0\rangle$$
$$|X_H\rangle = p|X^0\rangle - q|\bar{X}^0\rangle \tag{1.48}$$

where $|X_L\rangle$ and $|X_H\rangle$ are the lighter and heavier mass eigenstates.

$$q = \sqrt{\frac{m_{12}^* - \frac{1}{2}i\gamma_{12}^*}{m_{12} - \frac{1}{2}i\gamma_{12}}}p \tag{1.49}$$

$$\text{and} \qquad |p|^2 + |q|^2 = 1. \tag{1.50}$$

The difference in the magnitude of q/p from 1 is reponsible for CP-violation that is purely due to mixing — this will be discussed in section 1.3.2. The mass difference $\Delta m = m_H - m_L$ and decay width difference $\Delta\Gamma = \Gamma_H - \Gamma_L$ can also be obtained by diagonalizing the "mixing matrix" shown in Eqn. 1.44. Let

$$\alpha = |m_{12}|^2 - \frac{1}{4}|\gamma_{12}|^2, \qquad \beta = \mathcal{R}e(m_{12}\gamma_{12}^*) \tag{1.51}$$

Then,

$$\Delta m = \sqrt{2\alpha - 2\sqrt{\alpha^2 - \beta^2}} \quad \text{and} \tag{1.52}$$

$$\Delta\Gamma = 4\beta/\Delta m \tag{1.53}$$

An initially pure $|X^0\rangle$ state will, therefore, time evolve as a superposition of the mass eigenstates $|X_L\rangle$ and $|X_H\rangle$. Equation 1.49 may thus also be expressed as

$$q = \left(\frac{\Delta m - \frac{i}{2}\Delta\Gamma}{2(m_{12} - \frac{1}{2}i\gamma_{12})} \right) p \tag{1.54}$$

1.2.2 The Neutral K System

Mixing between the two neutral K weak eigenstates K^0 and $\bar{K}^0$ was first predicted in 1955 by Gell-Mann and Pais [7]. The two physical states, $|K_1\rangle = \frac{1}{\sqrt{2}}(K^0 + \bar{K}^0)$ and $|K_2\rangle = \frac{1}{\sqrt{2}}(K^0 + \bar{K}^0)$, would thus be CP eigenstates with eigenvalues $+1$ and -1. The dominant decay of neutral K mesons is $\pi^+\pi^-$, due to helicity constraints and the fact the 3-body phase space is strongly suppressed at these mass scales (due to the well-known $(\Delta m)^5$ scaling rule). However, $\pi^+\pi^-$ is itself a CP eigenstate with eigenvalue $+1$. Thus, if CP were exactly conserved, *only the $|K_1\rangle$ physical state could decay into it.*

The limited phase space to decays other than $\pi^+\pi^-$ forces the lifetime of the eigenstate with opposite CP, K_2, to be far larger (3 orders of magnitude) than the lifetime of the K_1, thus the nomenclature K_S^0 and K_L^0 (for short and long lifetimes) is used. The lifetime difference is very convenient since it allows for simple experimental separation of the two physical states.

In 1964, Fitch and Cronin made their discovery that K_L^0 can in fact decay into $\pi^+\pi^-$ with a branching fraction of 2×10^{-3} (see the Introduction). Since CP is thus not strictly conserved, the general formalism detailed in the previous subsection must be used. Thus we have

$$\begin{aligned} |K_S\rangle &= p|K^0\rangle + q|\bar{K}^0\rangle \\ |K_L\rangle &= p|K^0\rangle - q|\bar{K}^0\rangle \end{aligned} \tag{1.55}$$

where p and q are commonly parametrized as:

$$p = \frac{1+\epsilon}{\sqrt{2(1+|\epsilon|^2)}}; \qquad q = \frac{1-\epsilon}{\sqrt{2(1+|\epsilon|^2)}} \tag{1.56}$$

The real part of ϵ is a measure of CP violation purely in mixing whereas the imaginary part is a measure of CP violation in the interference between mixing and decay (see the following section). The former is simplest to measure experimentally and was the effect seen in the orginal 1964 discovery. Since, in the K system, $\Delta\Gamma$ is of the same order as Δm, these effects are of similar magnitude, quite unlike the neutral B system, where the latter is far more prevalent.

1.2.3 The Neutral B System

In the case of neutral B mesons, in contrast with the neutral K system, the lifetime difference $\Delta\Gamma$ between the two mass eigenstates is *small* compared with the mixing frequency due to the difference in masses Δm. This difference in behavior of the K and B is due to the larger mass of the B meson and thus far greater phase space for flavor-specific decays in the B system, which dominate the partial width (in contrast to the K system) and give equivalent contributions (by CPT symmetry) to the width of both neutral B eigenstates. The resulting lack of decay suppression of either eigenstate implies nearly equivalent lifetimes.

Due to this simplification in formalism, the time evolution of neutral B mesons which are initially created (at time $t = 0$) as pure flavor eigenstates can be written as:

$$|B^0_{phys}(t)\rangle = f_+(t)|B^0\rangle + (q/p)f_-(t)|\overline{B}^0\rangle \tag{1.57}$$

$$|\overline{B}^0_{phys}(t)\rangle = f_+(t)|B^0\rangle + (q/p)f_-(t)|\overline{B}^0\rangle \tag{1.58}$$

where

$$f_+(t) = e^{-imt}e^{-\Gamma t/2}\cos(\Delta m t/2) \tag{1.59}$$

$$f_-(t) = e^{-imt}e^{-\Gamma t/2}i\sin(\Delta m t/2) \tag{1.60}$$

This approximation holds up to the condition that

$$\Delta\Gamma \ll \Delta m \tag{1.61}$$

Since $\Delta\Gamma = \mathcal{O}(10^{-3})\Delta m$ in the B system ($\Delta\Gamma_{B_d}$ is as yet unmeasured, but this is generally considered to be a safe and model-independent assumption[6]), corrections to it are not considered in CP asymmetry measurements with the current statistics (furthermore, *BABAR* will have the capability of measuring $\Delta\Gamma$ as statistics of reconstructed B decays increase).

1.3 Three Types of CP Violation

Three types of CP violation can potentially be observed at B physics experiments:[3] 1) CP violation in decay (often referred to as direct CP violation): this occurs when multiple amplitudes with different weak phases as well as different strong phases contribute to a given final state, the result is visible as differing magnitude of the amplitude to a decay versus its CP conjugate;

[3]There can be other manifestations of CP violation, *e.g.* CP violation in interaction, however observable CP violation at B-factories can all be classified into the 3 categories.

2) CP violation purely in mixing: this occurs when the mass eigenstates of a neutral meson are different from the CP eigenstates: 3) CP violation in the interference between decays of mixed and unmixed mesons: this occurs for decays which are common to a neutral meson and its antiparticle.

1.3.1 CP Violation in Decay (Direct CP Violation)

Direct CP violation manifests itself as a difference in the magnitude of the amplitude to a given decay as compared with its CP conjugate, thus resulting in differing rates to the two elements of the CP conjugate pair. It can occur for both neutral and charged decays.[4] Amplitudes from B^0 and $\bar{B}^0$ to a final state and its CP conjugate may be written as

$$A_f = \sum_i A_i e^{i(\phi_i+\delta_i)} \qquad \text{and} \qquad \bar{A}_{\bar{f}} = \eta_f \sum_i A_i e^{i(-\phi_i+\delta_i)} \tag{1.62}$$

where η_f is the CP eigenvalue (multiplied by a convention-dependent phase) if f is a CP eigenstate, ϕ are the weak phases, and δ are the strong phases. CP violation can only occur when the different weak phase contributions also have different strong phases (or else a simple rotation can remove the strong phase and thus the ratio would clearly have unit magnitude). It can also only occur when weak phases are nontrivial, *i.e.* when there exists a relative phase between them (that is therefore irreducibe by a rotation of the Lagrangian). Only when both different weak phases *and* different strong phases are present can one have the condition:

$$|\bar{A}_{\bar{f}}/A_f| \neq 1 \tag{1.63}$$

[4]For charged decays, it is the *only* potential manifestation of CP violation.

This is CP violation in decay. CP violation in decay has been observed in the kaon system (via ϵ'_K/ϵ_K measurements) but not yet in the B system. Since the strong phases that enter into measurements of CP violation in decay involve hadronic uncertainties, the relation of such measurements to CKM factors (see next section) cannot be calculated from first principles, but the strong phases may themselves be measured if the CKM factors are known from other measurements. These strong phase measurements can then be used as inputs to other measurements which have equivalent strong phases (thus allowing the extraction of other parameters), and thus measurements of CP violation in decay can (indirectly) provide a useful handle on fundamental quantities.

1.3.2 CP Violation Purely in Mixing

From section 1.2.1, recall that the mass eigenstates of the neutral meson system are the eigenvectors of the Hamiltonian

$$\begin{aligned} |X_L\rangle &= p|X^0\rangle + q|\bar{X}^0\rangle \\ |X_H\rangle &= p|X^0\rangle - q|\bar{X}^0\rangle \end{aligned} \tag{1.64}$$

where

$$q = \sqrt{\frac{m_{12}^* - \frac{1}{2}i\gamma_{12}^*}{m_{12} - \frac{1}{2}i\gamma_{12}}}p \tag{1.65}$$

If q and p have different magnitudes, the CP conjugates of the mass eigenstates clearly will differ from the mass eigenstates themselves by more than a trivial phase. Thus the mass eigenstates will not be CP eigenstates and CP violation will be manifest. CP violation from

$$|q/p| \neq 1 \tag{1.66}$$

is purely an effect of mixing and is independent of decay mode. Thus it may be referred to as CP violation purely in mixing.

In neutral B decays, as discussed in section 1.2.2, this effect is expected to be very small. Since

$$\Delta m = \mathcal{O}(10^3)\Delta\Gamma \tag{1.67}$$

this implies that

$$|m_{12}| \gg |\gamma_{12}| \tag{1.68}$$

and thus the factor in equation 1.54 simplifies to a near-phase. CP violation purely in mixing should thus only enter the neutral B system at the 10^{-3} level. An asymmetry in the measurements of the overall rate to flavor tagged B^0 vs. $\overline{B}^0$ would be a signature of CP violation purely in mixing. With greater statistics, evidence for this may be seen; at present, experimental limits exist. It has been clearly observed, however, in the neutral kaon system (where it is the prevalent effect); the discovery of CP violation in 1964 was a detection of CP violation purely in mixing (see Section 1.2.2).

1.3.3 CP Violation in Interference Between Decays of Mixed and Unmixed Mesons

Final states which may be reached from either B^0 or $\overline{B}^0$ decays can exhibit a third type of CP violation, which results from the interference between the decays of mixed and of unmixed neutral B mesons which both decay to the

final state. Consider the CP-violating asymmetry in rates between B^0 and $\overline{B}^0$ as a function of time:

$$a_{CP}(t) = \frac{\Gamma(B^0_{phys}(t) \to f) - \Gamma(\bar{B}^0_{phys}(t) \to f)}{\Gamma(B^0_{phys}(t) \to f) + \Gamma(\bar{B}^0_{phys}(t) \to f)} \tag{1.69}$$

To calculate each of the time-dependent rates $\Gamma(t)$, one can form the inner product of equations 1.55 and 1.56 with the final state f and then take the magnitude squared of the resulting amplitudes:

$$\begin{aligned}
&\Gamma(B^0(t) \to f) \propto \\
&|\langle f|\mathcal{H}|B^0(t)\rangle|^2 = \; e^{-\Gamma t} \; \Big\{\cos^2\left(\frac{\Delta m t}{2}\right)|\langle f|\mathcal{H}|B^0\rangle|^2 \\
&\qquad + \sin^2\left(\frac{\Delta m t}{2}\right)\left|\frac{q}{p}\right|^2 |\langle f|\mathcal{H}|\overline{B}^0\rangle|^2 \\
&\qquad - \frac{i}{2}\left|\frac{q}{p}\right| e^{-2i\phi_M}\sin(\Delta m t)\langle f|\mathcal{H}|B^0\rangle\langle f|\mathcal{H}|\overline{B}^0\rangle^* \\
&\qquad + \frac{i}{2}\left|\frac{q}{p}\right| e^{2i\phi_M}\sin(\Delta m t)\langle f|\mathcal{H}|B^0\rangle^*\langle f|\mathcal{H}|\overline{B}^0\rangle\Big\}
\end{aligned} \tag{1.70a}$$

$$\begin{aligned}
&\Gamma(\overline{B}^0(t) \to f) \propto \\
&|\langle f|\mathcal{H}|\overline{B}^0(t)\rangle|^2 = \; e^{-\Gamma t} \; \Big\{\cos^2\left(\frac{\Delta m t}{2}\right)|\langle f|\mathcal{H}|\overline{B}^0\rangle|^2 \\
&\qquad + \sin^2\left(\frac{\Delta m t}{2}\right)\left|\frac{p}{q}\right|^2 |\langle f|\mathcal{H}|B^0\rangle|^2 \\
&\qquad + \frac{i}{2}\left|\frac{p}{q}\right| e^{-2i\phi_M}\sin(\Delta m t)\langle f|\mathcal{H}|B^0\rangle\langle f|\mathcal{H}|\overline{B}^0\rangle^* \\
&\qquad - \frac{i}{2}\left|\frac{p}{q}\right| e^{2i\phi_M}\sin(\Delta m t)\langle f|\mathcal{H}|B^0\rangle^*\langle f|\mathcal{H}|\overline{B}^0\rangle\Big\}
\end{aligned} \tag{1.70b}$$

where $2\phi_M$ is the phase of q/p. Since, as shown above, for the B system $|q/p| \approx 1$, we can thus write

$$\langle f|\mathcal{H}|\bar{B}^0(t)\rangle = \eta e^{-2i\phi_D}|\lambda|\langle f|\mathcal{H}|B^0(t)\rangle \tag{71}$$

where ϕ_D is the phase of the decay, η is the CP eigenvalue of f, and

$$\lambda = \frac{q}{p}\frac{\langle f|\mathcal{H}|\bar{B}^0\rangle}{\langle f|\mathcal{H}|B^0\rangle} = |\lambda|e^{-2i(\phi_M+\phi_D)}. \tag{72}$$

the expressions simplify greatly:

$$|\langle f|\mathcal{H}|B^0(t)\rangle|^2 = A^2e^{-\Gamma t}\{1 - \mathrm{C}\cos(\Delta mt) - \mathrm{S}\sin(\Delta mt)\} \quad \text{and} \tag{73}$$

$$|\langle f|\mathcal{H}|\bar{B}^0(t)\rangle|^2 = A^2e^{-\Gamma t}\{1 + \mathrm{C}\cos(\Delta mt) + \mathrm{S}\sin(\Delta mt)\} \tag{74}$$

where $A^2 = |\langle f|\mathcal{H}|B^0\rangle|^2$ and

$$\mathrm{C} = \frac{1-|\lambda|^2}{1+|\lambda|^2} \qquad \text{and} \qquad \mathrm{S} = \eta\frac{-2\sin(2(\phi_M+\phi_D))}{1+|\lambda|^2} \tag{75}$$

Thus the time-dependent asymmetry

$$a_{CP}(t) = \frac{\Gamma(B^0_{phys}(t)\to f) - \Gamma(\bar{B}^0_{phys}(t)\to f)}{\Gamma(B^0_{phys}(t)\to f) + \Gamma(\bar{B}^0_{phys}(t)\to f)} = \mathrm{C}\cos(\Delta mt) + \mathrm{S}\sin(\Delta mt) \tag{76}$$

In the absence of CP violation, S and C must both go to zero, since they occur only when weak phases do not cancel. C is only nonzero when the ratio of the amplitude norms differs from unity, which is the signature of direct CP violation (detailed in section 1.3.1). S, however, represents a distinct type of CP violation that can occur even in the absence of CP violation purely in decay or in mixing. It results from the interference of the decays of mixed mesons with those of unmixed mesons: if the mixing contains a phase that is

not cancelled by the decay itself, this observable time-dependent asymmetry above will result. Unlike CP violation in decay, no nontrivial strong phases are required.

As will be seen in the next section, CP violation in interference between decays of mixed and unmixed mesons is a large effect in the Standard Model picture of the neutral B system. Since this is a measurement of an asymmetry rather than an absolute rate, many experimental and model-dependent uncertainties (such as reconstruction efficiency) that would otherwise contribute to experimental error, instead cancel out in the ratio. Thus it provides an excellent mechanism for precision measurements of CP violation and the study of the Standard Model picture of CPV.

1.4 CP Violation in the Standard Model

CP violation within the context of the Standard Model $SU(2) \times U(1)$ electroweak symmetry was introduced by Kobayashi and Maskawa in 1973 via the postulation of a third family of quarks. This occurred a year prior to the discovery of charm; only 3 quarks existed at the time, so the prediction was quite prescient. The b-quark was then first observed in 1977. The prediction of additional quarks did not occur entirely without precedent, however. Theoretical interpretation of quark mixing via the weak interaction has closely followed experimental result, and the development of the 3 x 3 CKM matrix and its CP violating phase was a steady and piecewise process.

1.4.1 Weak Interactions and the CKM Matrix

The observed suppression of flavor-changing neutral current decays indicates that the quark sector is separated into families, similar to the lepton sector. However, lepton flavor is conserved,[5] whereas quark generation is manifestly violated (*e.g.* in weak decays of kaons). However, strangeness-changing decays have an additional suppression compared with strangeness-conserving weak decays. This "Cabbibo factor" may be accounted for by considering that, similar to neutral mesons, the quark mass eigenstates differ from the weak eigenstates. Thus a mixing matrix describing transitions between quark generations is necessary.

Such a matrix must be unitary since quark number is manifestly conserved.[6] With 2 generations, a unitary matrix can be described by a single parameter Θ_c:

$$\begin{pmatrix} d_{mass} \\ s_{mass} \end{pmatrix} = \begin{pmatrix} \cos\Theta_c & \sin\Theta_c \\ -\sin\Theta_c & \cos\Theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} \tag{77}$$

where d_{mass} and s_{mass} are the mass eigenstates nearest to the flavor eigenstates d and s respectively.

The same matrix (experimentally) holds for the (u, c) quark pair (although the c quark was of course discovered afterwards in 1974, four years after its prediction via the GIM mechanism that required charm to explain the absence of weak flavor-changing neutral currents[10]). The Cabbibo angle Θ_c is thus a full description of 2-generation mixing.

[5]Discounting, for the purposes of this document, recently discovered neutrino oscillations and thus lepton mixing.

[6]in contrast with the number of neutral mesons

More generally, we can write the charged-current coupling j_{cc} with 2 generations as

$$j^{\mu}_{cc} = \begin{pmatrix} \bar{u} & \bar{c} \end{pmatrix} \gamma^{\mu}(1-\gamma^5) \begin{pmatrix} d_{mass} \\ s_{mass} \end{pmatrix} = \begin{pmatrix} \bar{u} & \bar{c} \end{pmatrix} \gamma^{\mu}(1-\gamma^5) V_{ij} \begin{pmatrix} d \\ s \end{pmatrix} \quad (78)$$

where V_{ij} is the 2 x 2 Cabbibo matrix parametrized by Θ_c above. With an arbitrary number of generations, the charged current ($W^{\pm}$) Lagrangian becomes:

$$\mathcal{L}_W = \frac{g}{\sqrt{2}} \left\{ \bar{u}^L_i \gamma^{\mu} W^{+}_{\mu} V_{ij} d^L_j + \bar{d}^L_i \gamma^{\mu} W^{-}_{\mu} V^{*}_{ij} u^L_j \right\} \quad (79)$$

with u^L_i representing the vector of up-type quarks and d^L_i representing the down-type quarks. Applying the CP operation to the Lagrangian, one obtains:

$$\mathcal{L}_W = \frac{g}{\sqrt{2}} \left\{ \bar{d}^L_i \gamma^{\mu} W^{-}_{\mu} V_{ij} u^L_j + \bar{u}^L_i \gamma^{\mu} W^{+}_{\mu} V^{*}_{ij} d^L_j \right\} \quad (80)$$

which is exactly the same except for the complex conjugation of V. Thus, if we can find a basis for which V (as well as the quark masses) are real, then CP is a symmetry.

Unitary matrices of dimension N form a group, the Lie group $SO(N)$. Elements of $SO(N)$ may be specified by $N^2 - 2N + 1$ real parameters. With 2 quark generations, V is defined by a single real parameter, the Cabbibo angle Θ_c. However, with 3 quark generations, 4 parameters are required. The real rotations may be taken to be the 3 Euler angles, but this leaves an extra parameter. The extra parameter is an *irreducible* complex phase. If

this phase is nonzero, one can no longer find a basis for which V is real. Thus CP would cease to be a symmetry, and indeed that is the case in nature.

1.4.2 Unitarity Conditions and the Unitarity Triangle

Unitarity of the CKM matrix V requires that

$$V^\dagger V = VV^\dagger = 1 \quad \Rightarrow \quad \sum_j V_{ji}^* V_{jk} = \sum_j V_{ij} V_{kj}^* = \delta_{ik} \tag{81}$$

With a 3-generation CKM matrix

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{82}$$

this results in 9 independent equations, 3 of which (for the diagonal of the product unit matrix) equal one and 6 of which equal zero. The equations for the off-diagonal elements, each containing a sum of 3 complex numbers which equals 0, will each describe a triangle in the complex plane:

$$V_{cd}V_{ud}^* + V_{cs}V_{us}^* + V_{cb}V_{ub}^* = 0 \tag{1.83a}$$

$$V_{cd}V_{td}^* + V_{cs}V_{ts}^* + V_{cb}V_{tb}^* = 0 \tag{1.83b}$$

$$V_{ud}V_{td}^* + V_{us}V_{ts}^* + V_{ub}V_{tb}^* = 0 \tag{1.83c}$$

$$V_{us}^*V_{ud} + V_{cs}^*V_{cd} + V_{ts}^*V_{td} = 0 \tag{1.83d}$$

$$V_{ub}^*V_{us} + V_{cb}^*V_{cs} + V_{tb}^*V_{ts} = 0 \tag{1.83e}$$

$$V_{ub}^*V_{ud} + V_{cb}^*V_{cd} + V_{tb}^*V_{td} = 0 \tag{1.83f}$$

The differences between these 6 triangles are purely empirical. There is no theoretical motivation at present for the fact that 4 of them are nearly degenerate and only 2 describe triangles that have each of their sides being the same order of magnitude in length — the 4 parameters that describe the CKM matrix are not predicted by the Standard Model and can only be determined experimentally. It is emprically the case that only equations 1.83c and 1.83f above describe triangles which are not nearly degenerate. Of these, the last equation, 1.83f, is the one that is typically used to pictorially represent the irreducible CP violating phase and is referred to as the Unitarity Triangle.

The empirical fact that 4 of the triangles are nearly degenerate allows for a convenient parametrization of the CKM matrix via an expansion around the order parameter $\lambda \equiv \Theta_c (= 0.2205 \pm 0.0018)$.[7] This approximate parameterization, first proposed by Wolfenstein, is

$$V = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \qquad (84)$$

with (λ, A, ρ, η) as the 4 real parameters describing the CKM matrix, the latter 3 being of order 1. ρ and η then describe the x and y position of the upper vertex of the Unitarity Triangle, as shown in Figure 1.2.

[7]Note that this Cabbibo parameter $\lambda \equiv \Theta_c$ differs from the time-dependent asymmetry parameter λ detailed in Section 1.3.3.

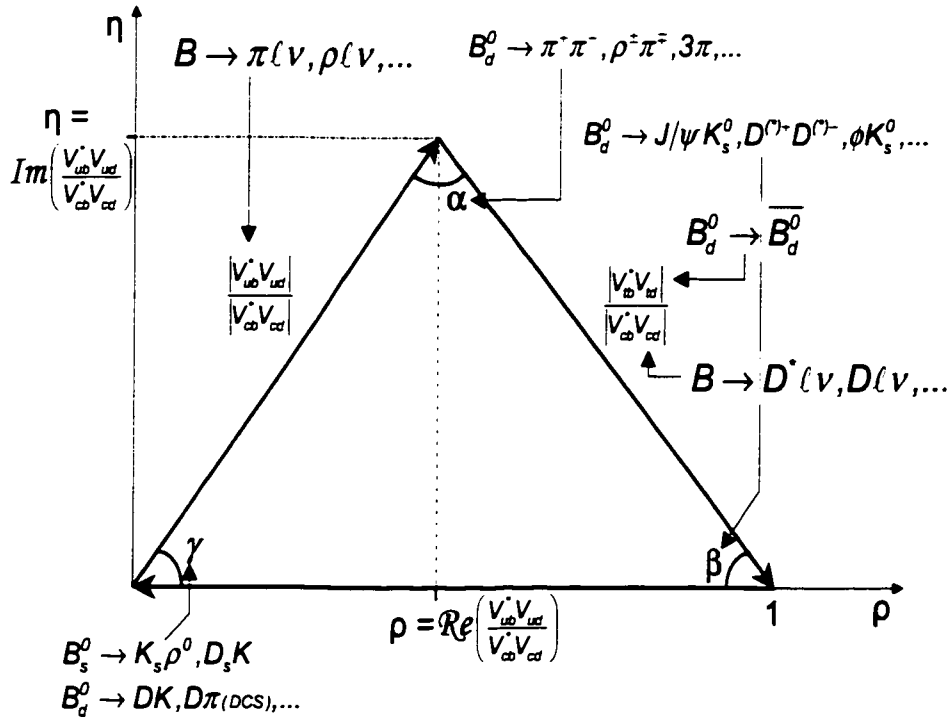


Figure 1.2: The Unitarity Triangle: In the B system, angles are measureable via time-dependent asymmetries in the modes listed, sides are measureable via semileptonic branching fractions and mixing frequency (with some theoretical error in the CKM extraction).

1.4.3 Measurement of CKM Parameters Using the B System

The time-dependent asymmetry parameter

$$\lambda = \frac{q}{p}\frac{\bar{A}_f}{A_f} = |\lambda| e^{-2i(\phi_M + \phi_D)} \tag{85}$$

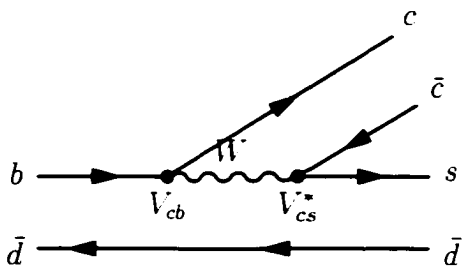


Figure 1.3: The $B^0 \to J/\psi K_S^0$ decay diagram.

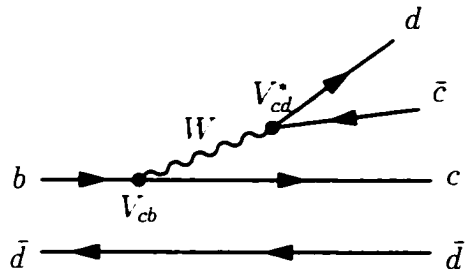


Figure 1.4: The tree diagram for $B^0 \to D^{(*)+} D^{(*)-}$.

(see also Section 1.3.3 [eqn. 72]) is of course dependent on the CKM matrix elements involved in the decay. In the case of $B^0 \to J/\psi K_S^0$, the quark subprocess is $b \to c\bar{c}s$, which is dominated by a color-suppressed tree diagram (see Figure 1.3). $\bar{B}^0$ decays to $J/\psi\ \bar{K}^0$ whereas B^0 decays to $J/\psi\ K^0$. Before the effects of $K - \bar{K}$ mixing, one has

$$\frac{A_{\bar{B}^0 \to J/\psi \bar{K}^0}}{A_{B^0 \to J/\psi K^0}} = \frac{V_{cb} V_{cs}^*}{V_{cb}^* V_{cs}}. \tag{86}$$

$K - \bar{K}$ mixing is necessary for interference to occur between B^0 and $\bar{B}^0$, since otherwise $B^0 \to K^0$ and $\bar{B}^0 \to \bar{K}^0$ and there would be no shared final state. Thus this neutral kaon mixing adds another phase factor:

$$\left(\frac{p}{q}\right)_K = \frac{V_{cs} V_{cd}^*}{V_{cs}^* V_{cd}} \tag{87}$$

and the $\frac{q}{p}$ factor from $B - \bar{B}$ mixing is also needed:

$$\left(\frac{q}{p}\right)_B = \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} \tag{88}$$

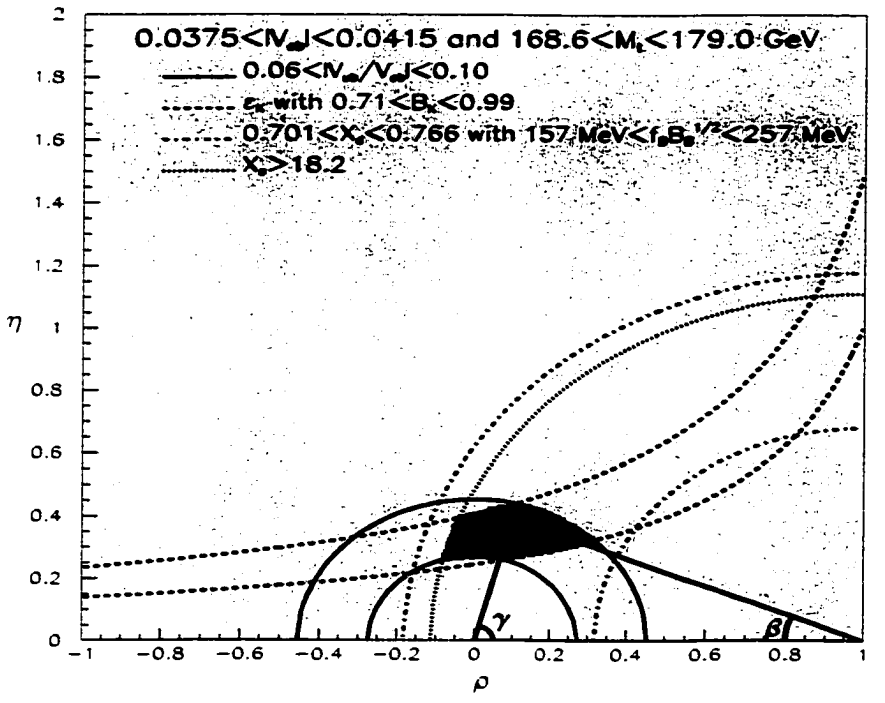


Figure 1.5: Experimental contraints on the Unitarity Triangle from measurements other than CP violating asymmetries in the B system. Ranges correspond to bounds from theoretical models: the errors for these constraints are dominated by theory.

so that

$$\begin{aligned} \lambda(B^0 \to J/\psi K_S^0) &= -\left(\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}\right)\left(\frac{V_{cb} V_{cs}^*}{V_{cb}^* V_{cs}}\right)\left(\frac{V_{cs} V_{cd}^*}{V_{cs}^* V_{cd}}\right) \\ &\Rightarrow \mathcal{I}m(\lambda_{J/\psi K_S^0}) = \sin 2\beta \end{aligned} \tag{89}$$

where β is the angle of the Unitarity Triangle. The minus sign is due to -1 CP eigenvalue of the J/ψ K_S^0 final state.

Decays with quark subprocess $b \to c\bar{c}d$ proceed similarly (see Figure 1.4).

although here there is a penguin contribution to the final state with a different weak phase which cannot be completely neglected. The tree contribution gives

$$\begin{aligned} \lambda(B^0 \to D^{(*)+}D^{(*)-})_{tree} &= \left(\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}\right)\left(\frac{V_{cd}^* V_{cb}}{V_{cd} V_{cb}^*}\right) \\ &\Rightarrow \mathcal{I}m(\lambda_{D^{(*)+}D^{(*)-}})_{tree} = -\sin 2\beta \end{aligned} \tag{90}$$

since the CP eigenvalue of $D^+ D^-$ is $+1$. Penguin contributions to the time-dependent asymmetry in $B^0 \to D^{(*)+}D^{(*)-}$, as well as the angular analysis which is necessary for the vector-vector final state $D^{*+} D^{*-}$, will be discussed in Chapter 4. Other decays probe other angles of the Unitarity Triangle: for example the time-dependent asymmetry in the tree contribution to $B^0 \to \pi^+\pi^-$ measures $\sin 2\alpha$.

A priori predicted values of time dependent CP violating asymmetries in modes such as $B^0 \to J/\psi K_S^0$, $B^0 \to D^{(*)+}D^{(*)-}$, and $B^0 \to \pi^+\pi^-$ can only be known via measurements sensitive to ρ and η (see Figure 1.5), which are very poorly constrained without such time dependent measurements from B decays. Hadronic uncertainties dominate the errors for other measurements of CP-violating CKM parameters such as V_{ub} and V_{td}. However, within the context of theoretical models of the strong interaction, constraints can be made on the (ρ, η) plane to bound an expected region where the asymmetries should lie. Figure 1.5 shows the constraints from semileptonic branching fraction ratios ($|V_{ub}/V_{cb}|$), CP asymmetries in the K system (ϵ_K), B_d^0 mixing (χ_d), and limits on B_s^0 mixing (χ_s). The bounds of this region, however, are large and are dominated by theoretical model uncertainties rather than on the experimental precision of the measurements.

1.5 Potential Beyond-Standard-Model Effects on *CP* Violation

The observed baryon asymmetry of the universe poses great difficulty for reconcilation with a purely Standard Model picture of *CP* violation. Although cosmological model-dependence does play a role in calculations of asymmetric baryogenesis, recent calculations show that the amount of *CP* violation predicted by Unitarity Triangle constraints is approximately 12 orders of magnitude below what is required to achieve the observed baryon asymmetry. Although there is a distinct possibility that the *CP* violation necessary to achieve consistency with observation lies at energies above what can be probed at *BABAR*, several well-motivated models, such as supersymmetry, predict extra *CP* violating phases that can potentially be examined indirectly through virtual loop-mediated processes. Alternative models which have other favorable features, such as explanations for the mass hierarchy of the quark and lepton sectors, can also predict observable differences from the Standard Model in the nature of *CP* violation. Thus it is important to test and constrain the Standard Model predictions for *CP* violation via tests that are sensitive to predictions from alternative models.

1.5.1 *CP* Violation and Cosmology

Sakharov showed in 1967 that 3 conditions are necessary for baryon asymmetry to devlop from an initially symmetric state [11]:

1. Baryon number must not be conserved.

2. There must be a departure from thermal equilibrium, *i.e.* the characteristic rate of cooling must be greater than the rate of the baryon number-violating process.

3. There must be CP violation (and C violation).

As shown in 1985 by Kuzmin, Rubakov, and Shaposhnikov, the first two conditions can be satisfied through the electroweak phase transition [12]. To satisfy condition 2, departure from thermal equilibrium may be effected by a first-order phase transition shockwave at the surface walls of expanding bubbles of low-temperature broken electroweak $SU(2)$ gauge symmetry as the universe cools below the phase transition. To satisfy condition 1, baryon number may be violated at temperatures *above* the phase transition by spontaneous production of baryon-lepton pairs via the (very counterintuitive) Adler-Bell-Jackiw anomaly in the axial vector vertex coupling, as shown by t'Hooft in 1976 [13, 14, 15].[8] The interior of the expanding bubbles then freezes in this asymmetry via exponential suppression of the baryon number-violating A-B-J current anomaly.

However, that whole very interesting process still requires a handedness to break the symmetry between particle and antiparticle behavior. CP violation provides this handedness and thus a favorable direction for the baryon number asymmetry to develop. CP violation is also accessible purely within the Standard Model and thus (qualitatively, at least) no new physics beyond the Standard Model is required to satisfy all 3 of the Sakharov conditions. Things are, however, not quite so simple as that: there is not enough CP violation

[8]Note that this high-temperature baryon number violation is *entirely within the context of the Standard Model*—no beyond standard model physics is required.

within the Standard Model, according to the above mechanism, to generate the observed baryon number / entropy ratio in the universe. Observable *CP* violation due to a complex phase in the CKM matrix is moderated by the λ^3 Cabbibo suppression, which reduces the maximal amount of *CP* violation in weak decays by $\lambda^6 \approx 10^{-4}$. On top of that there is the suppression of the weak interaction itself, which contributes a factor of $G_F^2 \approx 10^{-10}$. After all is considered, the *CKM* picture of *CP* violation is too inefficient by a factor of $\sim 10^{12}$ to produce observed baryon asymmetries. Furthermore, Gavela *et al.* have shown that the mechanism of *any* possible *CP* violating extension to the Standard Model must operate *above* the electroweak phase transition in order to avoid a thermally-stabilizing loss of quantum coherence within the expanding phase transition bubble wall [16].

Thus we need an extension to the Standard Model to provide us with more *CP* violation in order to agree with observed baryon asymmetry. Two such candidates[9] are supersymmetry and multiple Higgs-doublet models.

1.5.2 Supersymmetry and Multiple Higgs-Doublet Models

Supersymmetry can contribute additional *CP* violation via the addition of additional phases ϕ_{A_i} (where i runs over families) and ϕ_B to complement the Standard Model phases δ_{KM} ($= \gamma$ in the Unitarity Triangle) and Θ_{QCD} (≈ 0).[10] Tight constraints on ϕ_{A_i} and ϕ_B occur from measurements of the electric dipole moment of the neutron as well as from measurement of ϵ_K.

[9] which are not mutually exclusive

[10] the fact that Θ_{QCD} is zero or close to zero is the well-known "strong *CP* problem"

The small values of d_N and ϵ_K constrain the values of these supersymmetric phases.

Although these two constraints on supersymmetric *CP* violation are referred to as the "supersymmetric *CP* problem" and the "supersymmetric ϵ_K problem," it is well-known that many classes of supersymmetric models account for the smallness of these phases at experimental energies in relatively natural ways. For example, if Λ_S, the scale at which the soft supersymmetry-breaking terms are generated ($\approx$ the mass of the lightest supersymmetric particle) is much less than Λ_F, the scale at which flavor dynamics takes place ($\approx$ the masses of the gluinos), then both supersymmetric *CP* problems are solved. If $\Lambda_S \approx \Lambda_F$, one can still accommodate the smallness of the ϕ phases in models which contain horizontal symmetries (a horizontal symmetry is a symmetry which commutes with the Standard Model symmetries) if, for example, *CP* is a symmetry of the Lagrangian that is spontanously broken by the gauge field coupling that breaks the horizontal symmetry. Such models will, fortunately, potentially have observable consequences for physics at *BABAR*.

The supersymmetric contribution to flavor-changing neutral current amplitudes and phases can be large in such models. This will tend to introduce significant differences in the *CP* asymmetries which are prodicted to be nearly equal, such as, for example, in $B^0 \to J/\psi K_S^0$ and in $B^0 \to D^{(*)+} D^{(*)-}$. Comparison of the *CP* asymmetries in these decays presents a significant constraint on the parameters (and sensibility) of such models. Although deviations from Standard Model predictions for rates of FCNC processes have so far not been found, phases remain a completely open question, and the

comparison of CP violating asymmetries in these decays provides significant information.

In Multi-Higgs-Doublet Models (MHDM) with n Higgs doublets, there are $2(n-1)$ charged and $2n-1$ neutral scalars that remain after spontaneous symmetry breaking. These form an extended Higgs sector and can affect the couplings to weak vector bosons. The dominant effects would be in penguin decays (due to the fact that the Yukawa couplings of the scalars to tree-level decays are negligible)and could potentially affect both rates and CP asymmetries in $b \to s\gamma$, $B^0 \to \phi K_S^0$, and $B^0 \to D^{(*)+}D^{(*)-}$. The rates are constrained, but the phases in these modes are completely unmeasured.

In summary, there is ample evidence from cosmology that CP-violating processes that the Standard Model cannot accommodate are needed, and natural extensions to the SM that can potentially accomodate it have been proposed. The CP asymmetries in $B^0 \to D^{(*)+}D^{(*)-}$ provide a sensitive test of these models, especially when compared with CP asymmetries in $B^0 \to J/\psi K_S^0$, which should be nearly equivalent in the SM, but can differ in SUSY and MHDM models.

	$\mathcal{C}$	$\mathcal{P}$	$\mathcal{T}$	$\mathcal{CP}$	$\mathcal{CPT}$
Scalar	$+1$	$+1$	$+1$	$+1$	$+1$
Pseudoscalar	$+1$	-1	-1	-1	$+1$
Vector	$\mathbf{-1}$	$\begin{pmatrix} +1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$	$\begin{pmatrix} +1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$	$\begin{pmatrix} -1 \\ +1 \\ +1 \\ +1 \end{pmatrix}$	$\mathbf{-1}$
Pseudovector	$\mathbf{+1}$	$\begin{pmatrix} -1 \\ +1 \\ +1 \\ +1 \end{pmatrix}$	$\begin{pmatrix} +1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$	$\begin{pmatrix} -1 \\ +1 \\ +1 \\ +1 \end{pmatrix}$	$\mathbf{-1}$
Tensor	-1	$\begin{pmatrix} +1 & -1 & -1 & -1 \\ -1 & +1 & +1 & +1 \\ -1 & +1 & +1 & +1 \\ -1 & +1 & +1 & +1 \end{pmatrix}$	$\begin{pmatrix} -1 & +1 & +1 & +1 \\ +1 & -1 & -1 & -1 \\ +1 & -1 & -1 & -1 \\ +1 & -1 & -1 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & +1 & +1 & +1 \\ +1 & -1 & -1 & -1 \\ +1 & -1 & -1 & -1 \\ +1 & -1 & -1 & -1 \end{pmatrix}$	$+1$
Derivative Operator	$\mathbf{+1}$	$\begin{pmatrix} +1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$	$\begin{pmatrix} -1 \\ +1 \\ +1 \\ +1 \end{pmatrix}$	$\begin{pmatrix} +1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$	$\mathbf{-1}$

Chapter 2

The *BABAR* Detector

2.1 Overview — B-Factories

Exploring CP violation in the B system and its potential impact on the Standard Model, baryogenesis, and cosmology, requires copious production of B mesons, accurate measurement of the B flight time and flavor, and reasonably low background for reconstruction. There are several potential options for experiments which can fulfill these criteria:

1. Hadron colliders ($\overset{(-)}{p}p$): The cross section for $B\overline{B}$ production at TeV hadron colliders is very high compared with e^+e^- B factories, approximately 100 μb vs. 1.2 nb. This large advantage does compete with several disadvantages, however. Hadronic collisions have far more background, making reconstruction of final states which do not contain a J/ψ very challenging. Purely hadronic final states with non-negligible background in e^+e^- colliders at the $\Upsilon(4S)$, such as $D^{(*)+}D^{(*)-}$ or $\pi^0\pi^0$, may be extremely difficult at a hadronic collider and it is not clear

that it will be possible to reconstruct such decays. Nevertheless, these experiments do have a statistical advantage and also have the potential for observing CP violation in the B_s system, which is beyond the reach of $\Upsilon(4S)$ experiments. Two are currently under construction, BTeV at Fermilab and LHC-b at CERN.

2. Fixed target proton beam experiments: Fixed-target experiments also offer the potential of a higher rate of B production, but have even greater levels of backgrounds, superimposed interactions, and boost which compresses all tracks in a small solid angle. A significant effort was undertaken at DESY to build such an experiment, HERA-B, which, after years of construction, ended in what may be described as unmitigated failure. The high backgrounds prevented reconstruction of even a statistically significant J/ψ K_S^0 signal.

3. e^+e^- colliders at the Z-pole: The Z-pole presents a relatively clean environment for B-physics with a relatively large cross section ($\sim$ 6 nb). However, the luminosities achieved at this energy are low, the only two colliders in the world which can reach it, LEP and SLD, are both dismantled, and the cost of building new experiments at this energy prevents this from being a viable option.

4. Symmetric and asymmetric e^+e^- B-factories: The $\Upsilon(4S)$ resonance provides a very clean environment for B reconstruction. Asymmetric e^+ and e^- beams provide a boost to the B meson pair that is produced, allowing for reconstruction of B flavor as a function of flight time through the separation of the B vertices in the lab frame, Δz. Sta-

tistical limitations, of which luminosity is the critical factor, are the dominant source of error for time-dependent *CP* asymmetries. Two asymmetric *B*-factories have been built and are currently producing physics: PEP-II/*BABAR* and KEK-B/Belle. Previously, the symmetric *B*-factory CLEO (at the CESR ring at Cornell) was able to produce precision *B* physics results, however the symmetric design precluded measurement of time-dependent *CP*-violating asymmetries.

Figures 2.1 and 2.3 show the *BABAR* and Belle detectors. The experiments are very similar, with the following important differences: the KEK *B* factory has a nonzero beam crossing angle (4.2 mr) at the interaction point (IP), whereas the PEP-II/*BABAR* *B* factory has a more traditional collinear IP. The KEK design potentially allows a greater number of beam bunches to be stored in the ring, due to absence of parasitic crossings at $\pm$ 1m, as are present in the PEP-II design. However KEK-B is a highly non-traditional design: concerns over higher-order mode resonances at the IP led the PEP-II *B* factory to use a collinear crossing. So far, both KEK-B and PEP-II have performed well. At the time of writing, PEP-II has integrated 81.4 fb^{-1} and KEK-B has integrated 63.6 fb^{-1} (with KEK-B currently getting higher average luminosity). The particle identification method also differs between *BABAR* and Belle: as will be described in Section 2.6, *BABAR* uses quartz

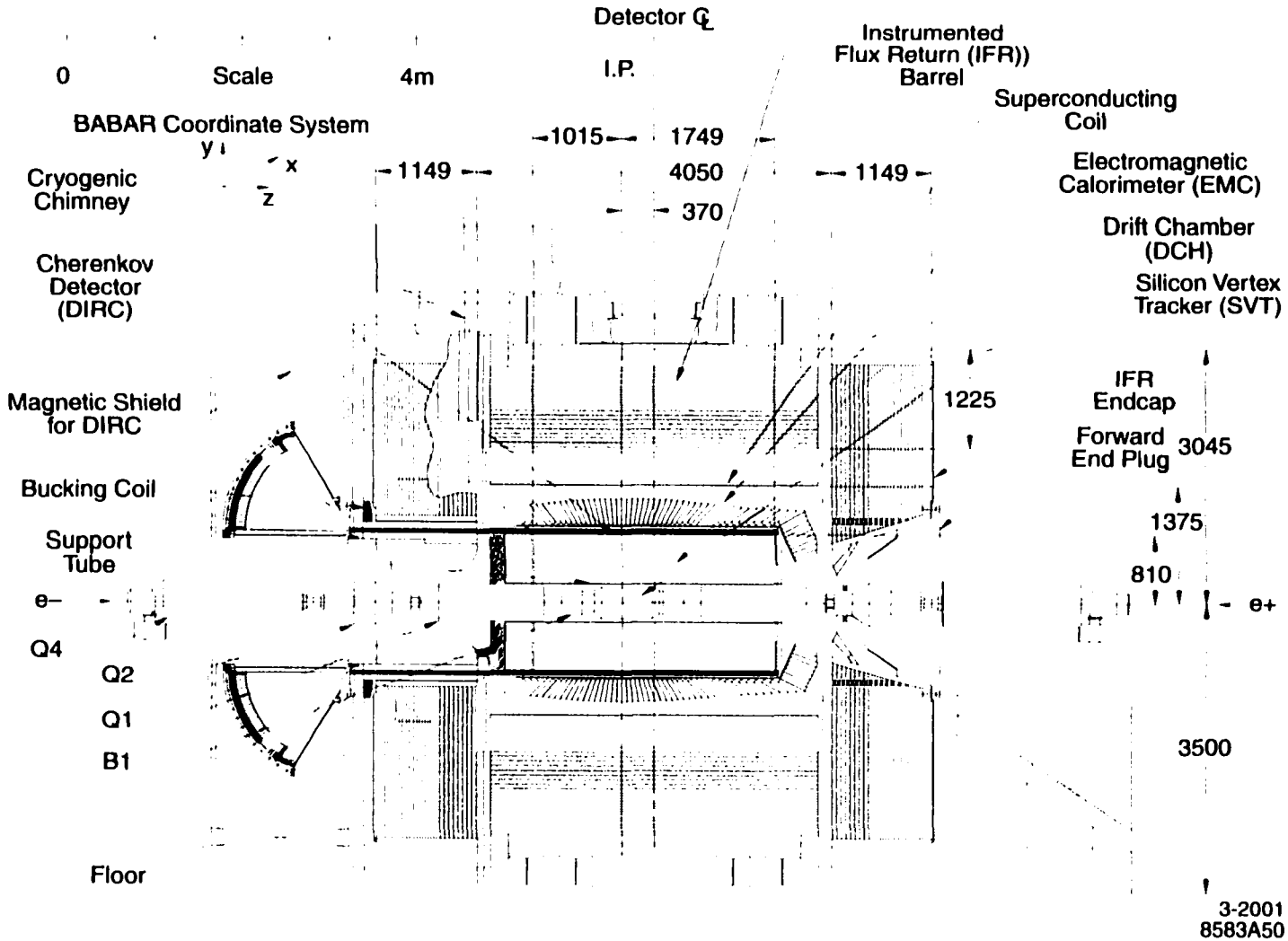


Figure 2.1: *BABAR* detector longitudinal section.

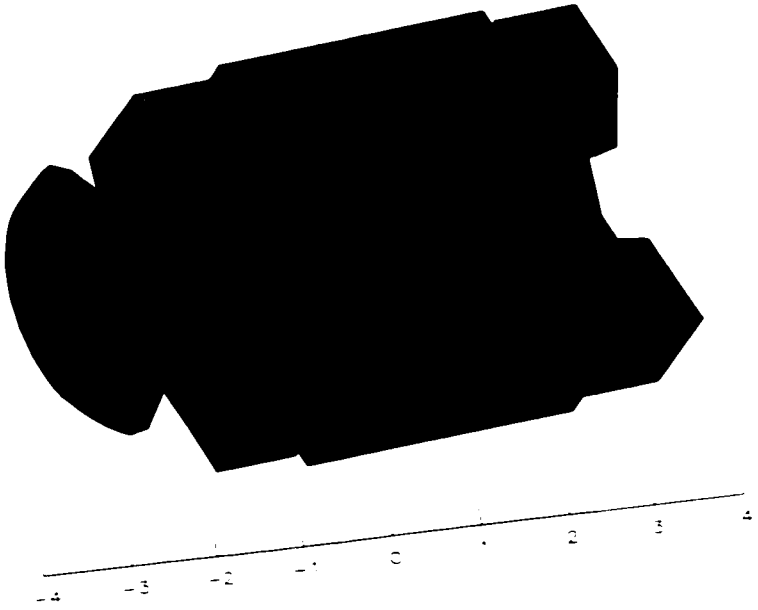


Figure 2.2: *BABAR* detector cutout diagram.

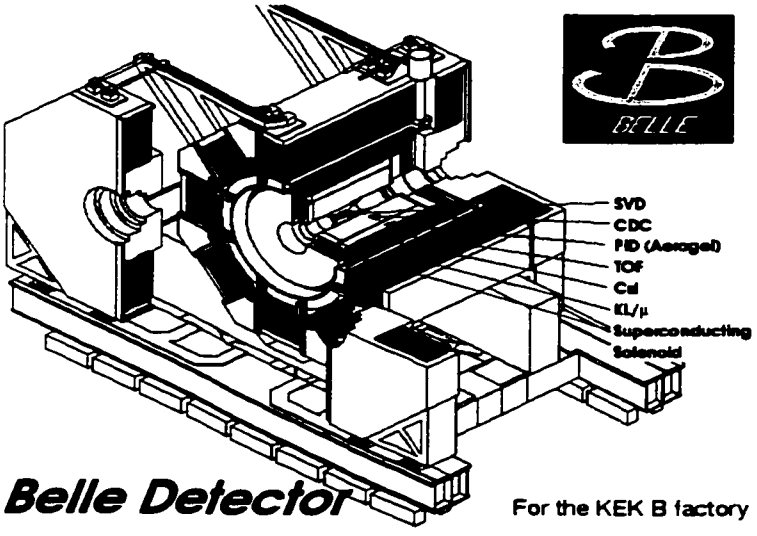


Figure 2.3: Belle detector cutout diagram.

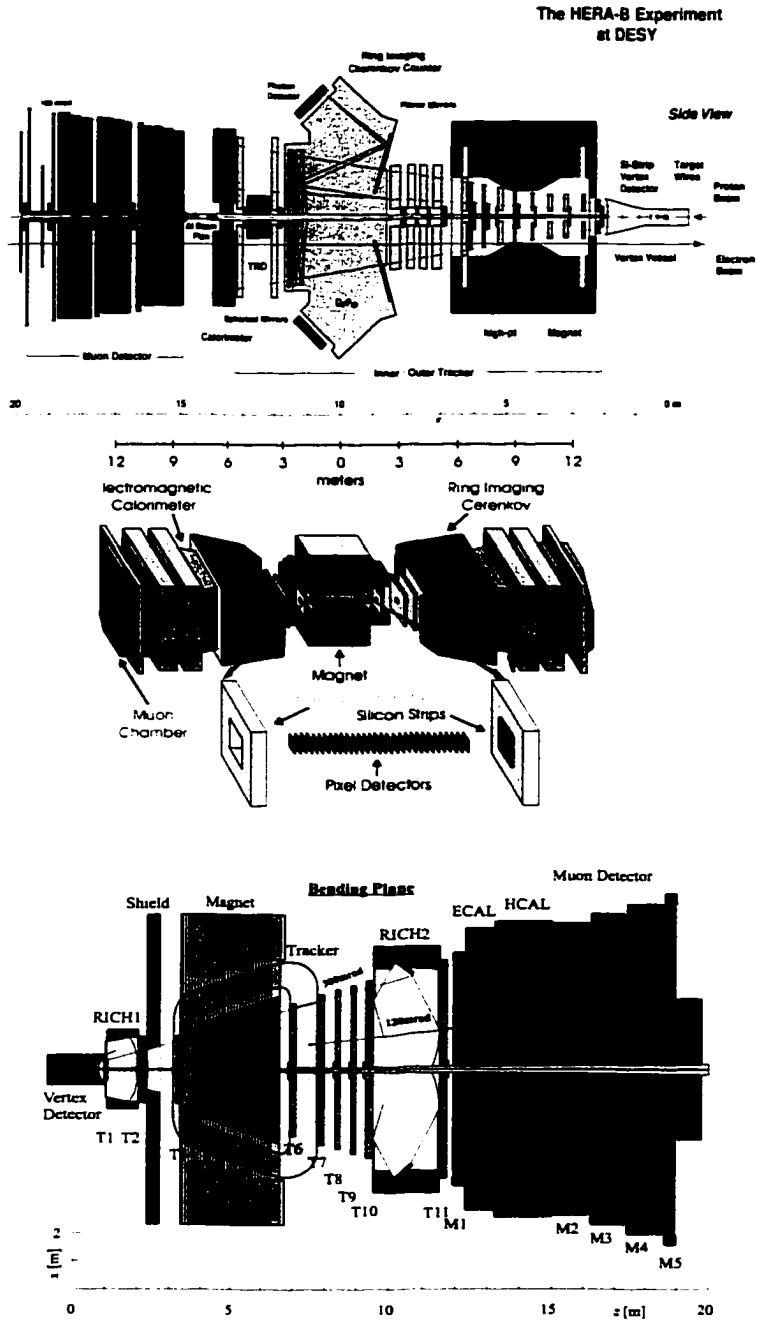


Figure 2.4: Diagrams of: (*top*) the HERA-B detector (at DESY, first beam in 2000), (*middle*) the BTeV detector (at Fermilab, to be completed in 2004), and (*bottom*) the LHC-b detector (CERN, to be completed in 2007).

bars to internally reflect Cerenkov light to a backward-mounted detector (the DIRC), whereas Belle uses an aerogel Cerenkov detector. In addition, *BABAR* has a 5-layer silicon vertex detector (SVT, see section 2.3) that can do standalone tracking (important for $D^{(*)+}D^{(*)-}$), whereas Belle uses a 3-layer silicon vertex detector.

Figure 2.4 shows the design of the HERA-B, BTeV, and LHC-b experiments. Each of these experiments uses hadron beams, with, in the case of HERA-B, a fixed (tungsten wire) target in the beam halo, and, for BTeV and LHC-b, colliding proton beams. HERA-B was not a successful experiment. The future looks more promising for BTeV and LHC-b, since colliding beams give a higher signal to background ratio; however, hadrons do present a challenging (but potentially very rewarding) environment for B physics.

2.2 The PEP-II Asymmetric Collider

The design of PEP-II is shown in figure 2.5. 9 GeV electrons and 3.1 GeV positrons are injected from the SLAC linac via bypass lines in the linac gallery. The beam parameters are listed in Table 1. PEP-II has surpassed design goals both in instantaneous and in average integrated luminosity.

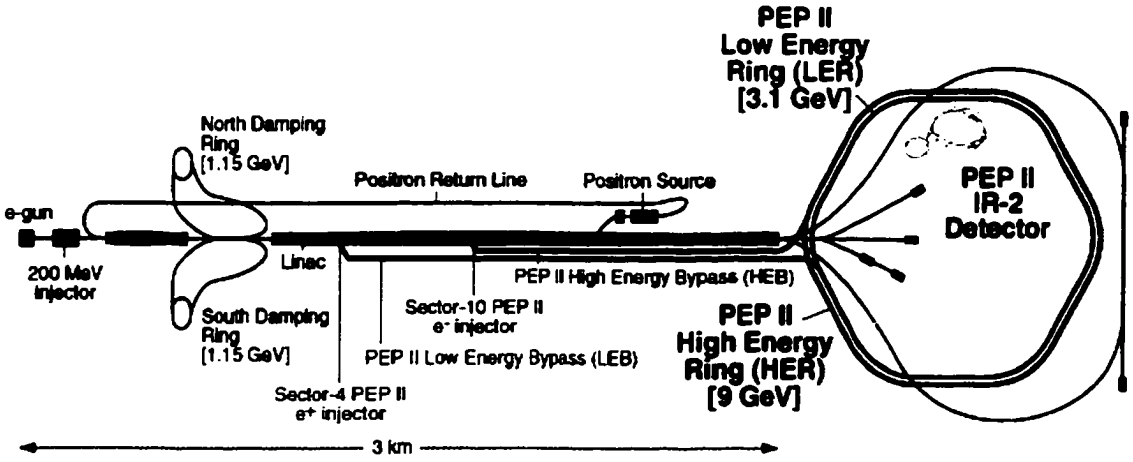


Figure 2.5: The PEP-II asymmetric storage ring and the SLAC linear accelerator. The SLAC linac is the injector for PEP-II. The single interaction point of PEP-II is at Interaction Region 2, where *BABAR* is situated.

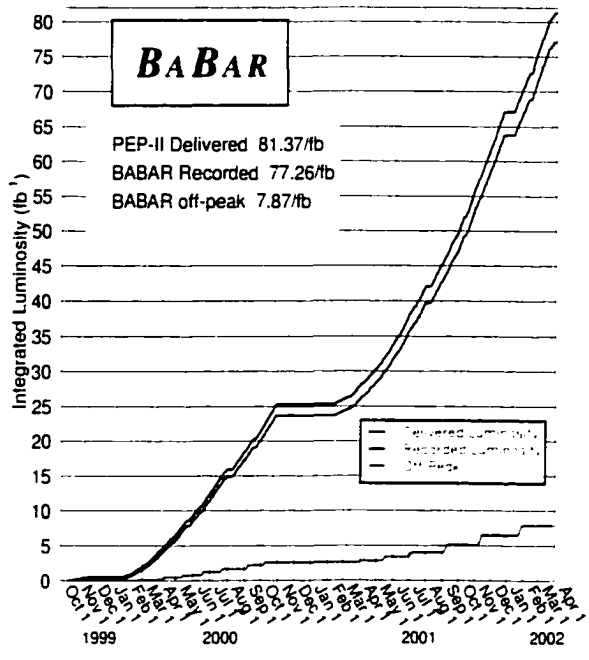


Figure 2.6: PEP-II -*BABAR* integrated luminosity since startup.

Parameters	Design	Typical
Energy HER/LER (GeV)	9.0/3.1	9.0/3.1
Current HER/LER (A)	0.75/2.15	1.0/1.8
# of bunches	1658	800-850
Bunch spacing (ns)	4.2	4.2-8.4
σ_{Lx} (μm)	110	110
σ_{Ly} (μm)	3.3	4.1
σ_{Lz} (mm)	9	9
Luminosity (10^{33} cm^{-2}s^{-1})	3	4.0
Luminosity (pb^{-1}/d)	135	270

Table 2.1: PEP-II beam parameters. Values are given for the design and for colliding beam operation at time of writing. HER and LER refer to the high energy e^- and low energy e^+ ring, respectively. σ_{Lx}, σ_{Ly}, and σ_{Lz} refer to the R.M.S. horizontal, vertical, and longitudinal bunch size at the IP.

Most of the data is taken at the $\Upsilon(4S)$ resonance (10.58 GeV), however approximately 12% are taken at 40 MeV below the resonance to allow studies of non-resonant background in data. A plot of PEP-II integrated luminosity as a function of time is in figure 2.6.

2.3 Overview of Experimental Technique at the $\Upsilon(4S)$

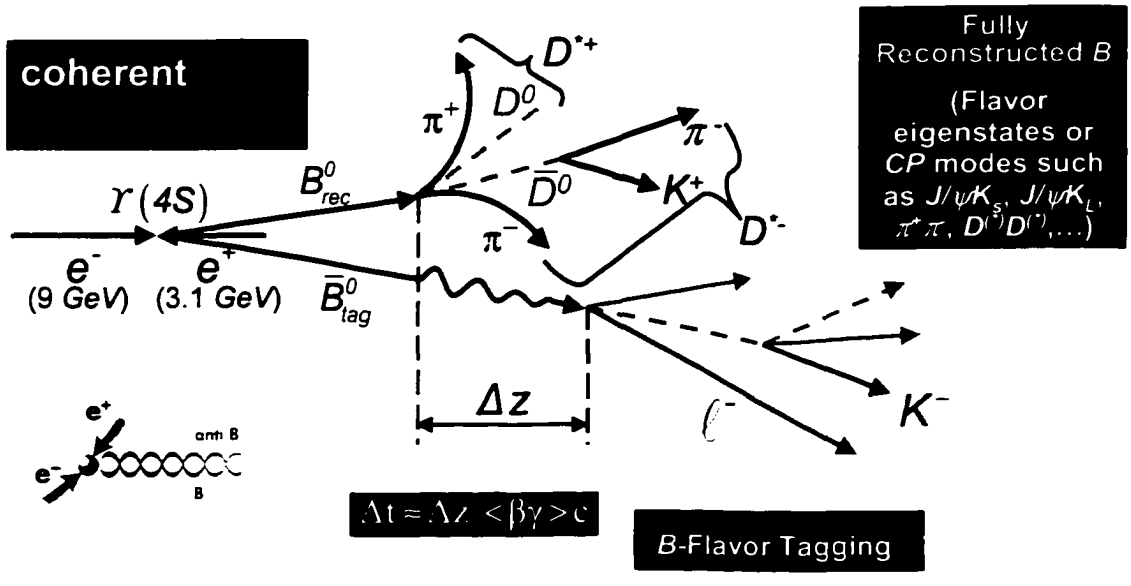


Figure 2.7: Experimental reconstruction technique used for measuring time-dependent CP-violating asymmetries at an $\Upsilon(4S)$ asymmetric collider. A coherent $B^0\overline{B}^0$ pair is produced from the $\Upsilon(4S)$ decay, which allows determination of reconstructed neutral B flavor as a function of decay time.

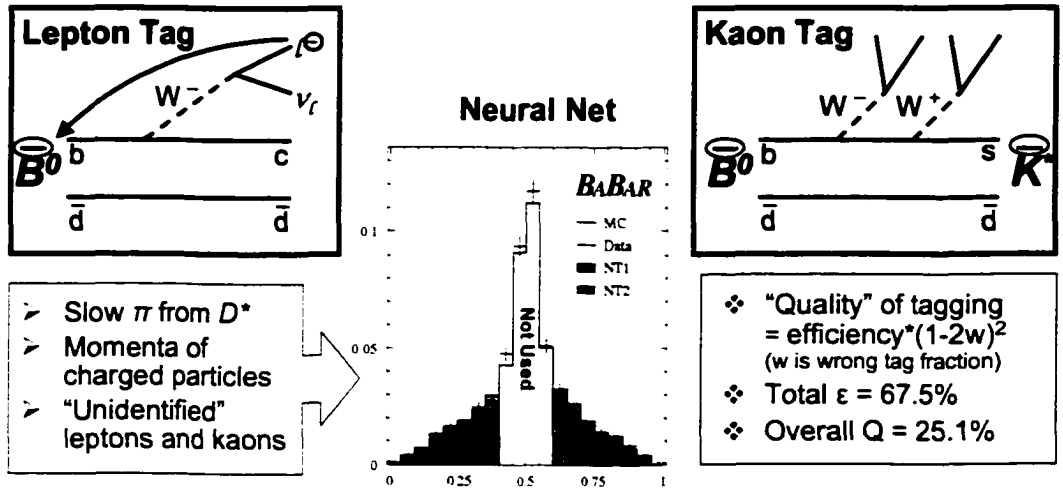


Figure 2.8: Technique used for tagging the flavor of the opposite-side B. Lepton and kaon charge is correlated with the flavor of the B. For events with no obvious lepton or kaon, a neural net is used to attempt to extract the flavor.

In order to measure time-dependent CP-violating asymmetries at the $\Upsilon(4S)$, one must (of course) first reconstruct a neutral B decay mode that can exhibit CP violation, such as $B^0 \to D^{(*)+}D^{(*)-}$ or $B^0 \to J/\psi K_S^0$. However, that is merely the first step. After signal event reconstruction, the additional tracks in the event (which correspond to the decay products of the other B [the "tag side B"]) must be used to determine whether the other B in the event was a B^0 or $\overline{B}^0$, due to the fact that the CP asymmetry is opposite for B^0 and $\overline{B}^0$ (see equations 1.73 and 1.74). After both the event reconstruction and the flavor tagging are completed, the difference in vertex z-position[1] between

[1]The z-axis in BABAR is along the direction of the beam line, with electrons (and the

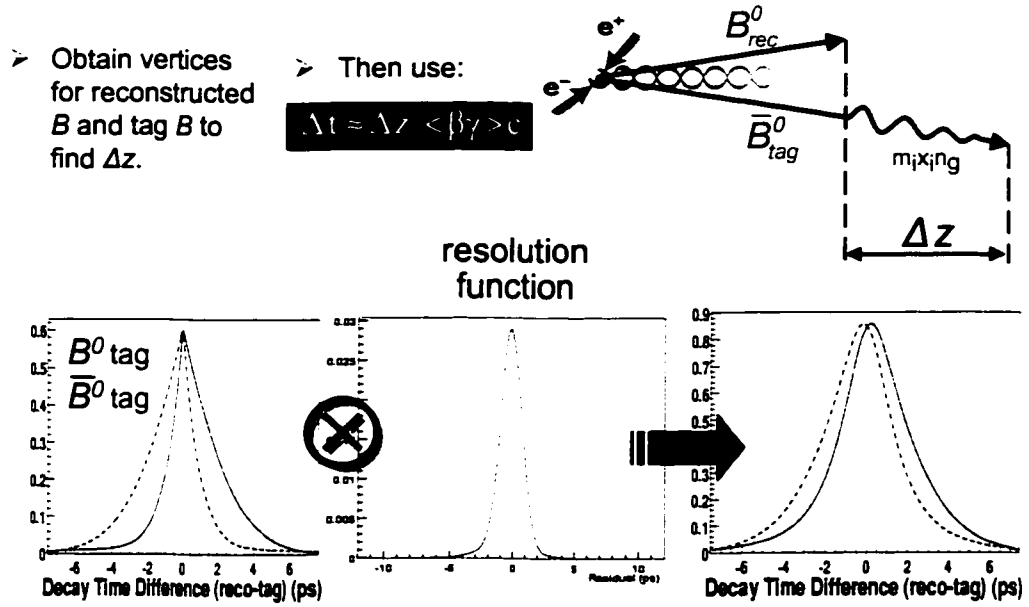


Figure 2.9: Δt measurement and resolution function. The difference in reconstructed z-position of the tag and reconstructed B decay vertices is used to determine the time difference Δt.

the reconstructed B vertex and the tag side B vertex must be determined. This difference, Δz, is (very nearly) proportional to the decay time difference Δt between the two B decays. Δt is the time measurement over which the CP-violating asymmetry can occur, and is input (as t) in equations 1.73 and 1.74. Figure 2.7 gives an overview of this reconstruction method.

Figure 2.8 briefly describes the technique used for flavor tagging. The sign of charged leptons and kaons in the event (which are not part of the reconstructed B) is correlated with the flavor of the tag side B. A cut-based selector using *BABAR*'s electron, muon, and kaon identification capabilities is

center-of-mass boost) pointing toward $+z$ in the lab frame.

used to select signal events with a lepton or kaon on the tag side, and from this determine the flavor of the tag side B. For events which are not cleanly tagged by the cut-based selector, a neural net algorithm is used to extract the flavor of the tag side B. The neural net uses information including slow pion charge, jettiness of the tag side tracks, and recovery of leptons and kaons which are not cleanly identified in order to reconstruct the tag side flavor. The overall efficiency of tagging is 67.5% and the fraction of tagged events which are given an incorrect tag is 19.5%. The error on time-dependent asymmetries is proportional to $Q = \epsilon(1 - 2w)^2$ where ϵ is the efficiency and w is the wrong-tag fraction. This quality factor Q is 25.1% for *BABAR*'s tagging algorithm.

Figure 2.9 briefly describes the Δt measurement and resolution function. A clustering algorithm is used to determine the vertex position for the tag side B decay: the error on this position dominates the resolution. *CP* violation evinces itself as a difference in Δt distribution depending on whether the flavor tag is B^0 or $\overline{B}^0$, but this decay time distribution is convoluted by the error of Δt. Fully reconstructed B events which have definite flavor (such as $B^0 \to D^{*+}\pi^-$ or $D^{*+}\rho^-$) are used to determine both the mistag fractions and the parameters of the resolution function (which is modelled as a triple gaussian) in data.

Figure 2.10: Fully assembled SVT. The silicon sensors of the outer layer are visible, as is the carbon-fiber space frame (black structure) that surrounds the silicon.

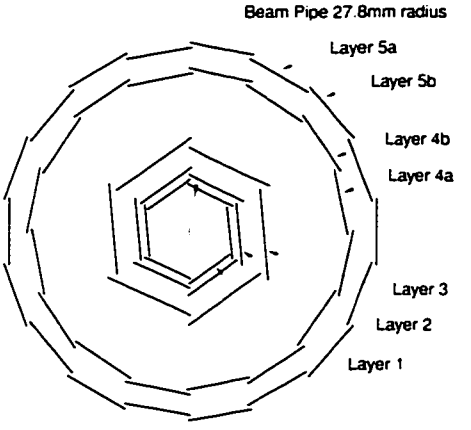


Figure 2.11: Transverse section of the SVT.

2.4 The Silicon Vertex Tracker (SVT)

The SVT contains 5 layers of silicon, double sided with conductive strip sensors. Strips on the opposite sides of each layer are orthogonal: ϕ strips run parallel to the beam axis and z strips run transverse to the beam direction.

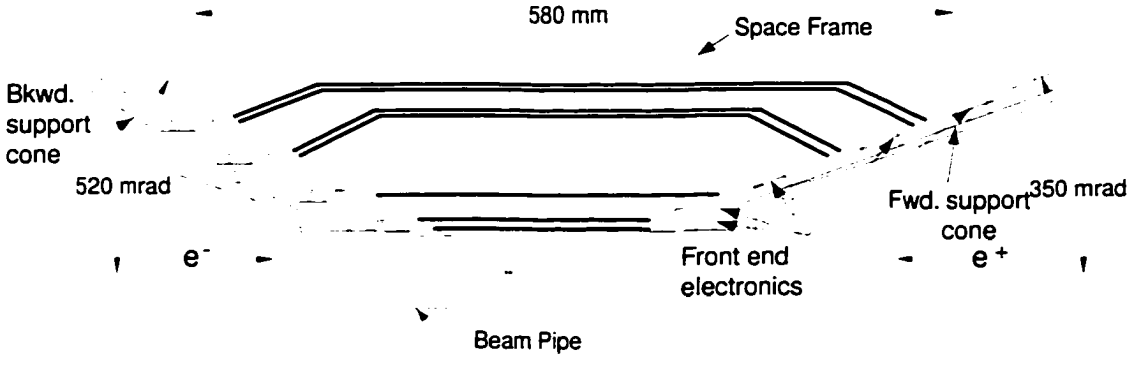


Figure 2.12: Longitudinal section of the SVT.

Together, the SVT and the central tracking drift chamber (DCH) form the charged particle tracking system (see also the following 2 sections). Precise and efficient measurement of track 4-momentum is necessary for full reconstruction of B meson decays, which tend to have multiple charged decay products. In addition, good vertex (and Δz) resolution and accurate extrapolation to the outer subdetectors is essential for reconstruction and background subtraction. Thus, accurate charged particle tracking and vertexing is required.

The 5 layers and relatively long radial separation between SVT detector layers provide both standalone track pattern recognition and refinement of drift chamber tracks via addition of SVT hits. The necessity of precise measurements close to the interaction point for Δz measurement and for background rejection using vertex quality, and for efficient reconstruction of low momentum tracks (such as slow pions from D^* decays), drive the requirements for the SVT.

The SVT silicon is composed of n-type substrate with p^+ and n^+ strips on opposite sides. The bias voltage ranges from 25-35 V. The layers of the SVT are divided radially into modules, shown as line segments in Figure 2.11. The modules in the inner 3 layers are straight along the z-axis, while those in layers 4 and 5 are arch-shaped, as shown in Figure 2.12. The arch design was selected to minimize the amount of silicon as well as increase the angle of incidence of tracks originating at the IP which cross the arch "lampshades" near the edges of acceptance. The total active silicon area is 0.96 m^2.

The strip pitch (width) varies from 50 to 210 μm depending on the layer (inner layers are more closely bonded). The strips are AC-coupled to the

electronic readout. Only approximately half the strips are read out; most have an unconnected "floating strip" between each pair of active strips (to reduce cost of readout electronics without adversely impacting performance). Digitization is performed by an ATOM ("A Time-Over-Threshold Machine") chip at the end of each set of 128 strips, which amplifies, digitizes, and buffers the signal from each channel. The ATOM chip compares the charge accumulated on each strip with an (adjustible) threshold of 0.95 fC, and records the time in clock intervals (30 MHz for the SVT) for which each strip is over threshold. This information is then delivered to a computer farm for further processing upon an accept signal from the Level 1 Trigger (see section 2.9).

A variety of monitoring checks and calibrations must be performed on the SVT to maintain data quality. Perhaps the most important of these from an avoidance-of-equipment-damage perspective is radiation protection. Twelve silicon PIN diodes surround the support cones and monitor both instantaneous radiation and accumulated dose. The beam is automatically aborted if radiation levels are above 1 Rad/ms threshold. So far, the SVT is well below the operational limit of 4 MRad integrated dose. The modules in the beam plane suffer the highest doses, and will be replaced in 2004-5. The others are expected to last for the duration of the experiment. For data quality calibration, channel gains and noise must be individually calibrated, and these are done online via an integrated pulse generator and calibration electronics. The offline reconstruction has the responsibility for calibration of the alignment of each SVT module. Alignment is critical for accuracy of vertexing and of tracking reconstruction, and is done in two steps. The local

SVT alignment uses dimuon and cosmic ray events to calibrate the relative position of each of the 340 wafers. The global alignment then determines the overall position and rotation of the SVT with respect to the DCH.

The SVT has performed according to design essentially since its inception. A combined hardware and software hit-finding efficiency of 97% is observed, excluding the 9 (out of 208) readout sections which are defective. Single hit resolution for tracks originating from the IP averages 20 μm in both z and ϕ for hits on the inner 3 layers and 40 μm in z and 20 in ϕ for hits in the outer 2 layers. The 9 out of 208 modules which are defective (due to a variety of causes including installation problems and defective connectors) do not significantly impact performance due to the redundancy which 5 layers provides (only four are needed for track reconstruction).

2.5 The Drift Chamber (DCH)

The DCH contains 40 layers of gold-coated tungsten-rhenium sense wires and gold-coated aluminum field wires in a mixture of 80% helium and 20% isobutane gas. There are a total of 7,104 sense wires and 21,664 field wires, with one wire per electronics channel. Wires are each tensioned (30 grams for sense wires, 155 grams for field wires) and pass through the aluminum endplates via feedthroughs made from Celenex insulating plastic around a copper wire jacket. The layers are grouped by four into 10 superlayers, with the wires in each superlayer oriented as either axial (directly parallel to the z-axis) or "stereo" (at a small angle in ϕ with respect to the z-axis, in order to obtain longitudinal position information). 6 of the 10 superlayers are stereo,

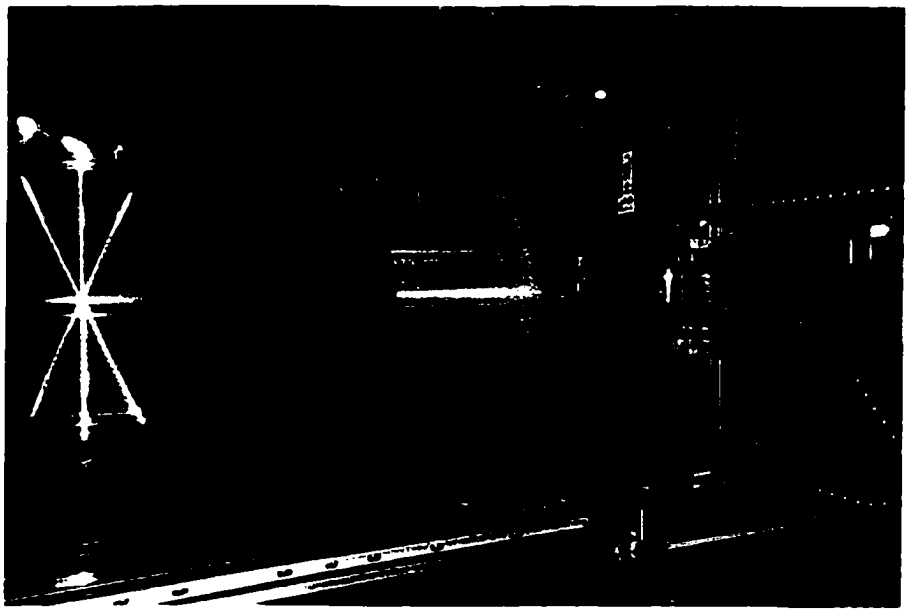


Figure 2.13: DCH wire stringing at TRIUMF (September 1997).

Figure 2.14: DCH installation (August 1998).

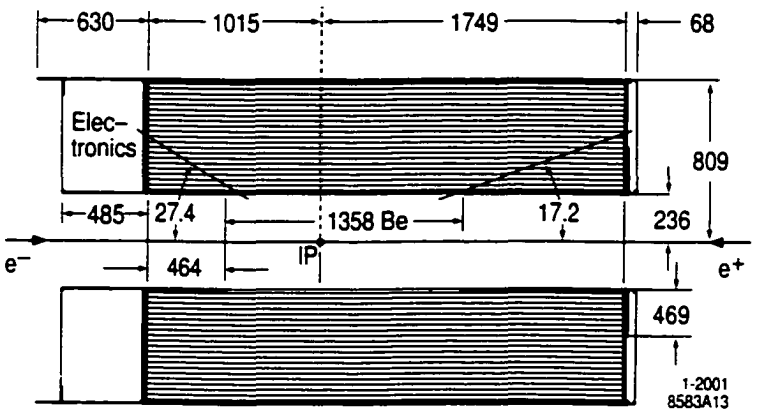


Figure 2.15: Longitudinal section of the drift chamber.

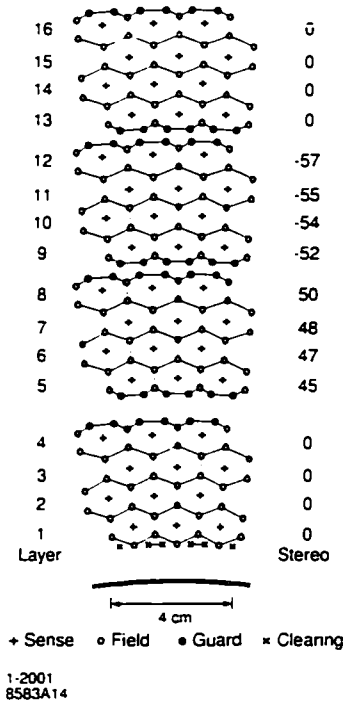


Figure 2.16: DCH cell pattern (for a section of the inner 16 [of 40] layers).

and the other 4 are axial.[a] The DCH is asymmetric in z about the interaction point, as shown in Figure 2.15, to

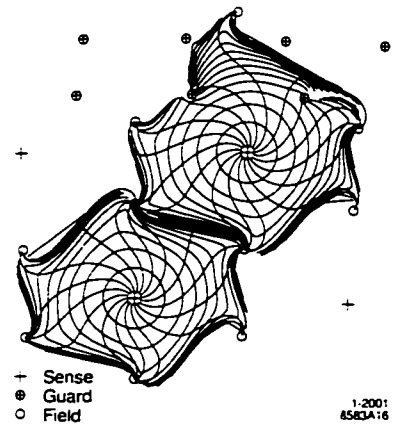


Figure 2.17: DCH cell drift isochrones for cells in layers 3 and 4 (axial). Isochrones are at 100 ns intervals.

[a]The arrangement is, from inner to outer, AUVAUVAUVA (A = axial, U = u stereo ($+o$), V = v stereo ($-o$)).

Figure 2.18: Partial view of drift chamber front-end electronics.

accommodate the forward boost of the center of mass of physics events. The endplates are 24mm thick aluminum. except for the outer 33.1 cm of the forward endplate. which is reduced to 12 cm to minimize the amount of material in front of the forward calorimeter endcap. The inner cylinder is fabricated from beryllium (to minimize the multiple scattering for the section of inner cylinder within the tracking fiducial volume) and aluminum (for the rest). The outer cylinder is 1.6 mm carbon fiber on 6 mm thick honeycomb Nomex core. The total thickness of the DCH is 1.08% X_0.

The cells are arranged in a hexagonal pattern, each with a sense wire at the center and field wires at the corners, as shown in figure 2.16. Cells on a superlayer boundary have a slightly different arrangement, with two guard wires taking the place of a single field wire. The nominal operating voltage is 1930 V. Isochrones and drift paths, calculated using the GARFIELD simulation, are shown in figure 2.17.

The DCH electronics are designed to provide accurate measurements of signal time and integrated charge (as well as providing information to the Level 1 Trigger, see Section 2.9). Service boards plug directly onto the wire feedthroughs on the rear endplate. These boards distribute the high voltages as well as pass signal and ground to the front-end electronics assemblies. The front-end assemblies (FEAs) plug into the service boards and amplify, digitize, and buffer (for 12.9 μs) the signals. A view of the front-end electronics including (enclosed) front-end assemblies and service boards below can be seen in figure 2.18. The digital data is sent, upon receipt of a level 1 trigger accept signal, via 59.5 MHz serial link to a data I/O module which transmits the signal to the external electronics via fiber-optic cable. Extraction of hit time and integrated charge from the digital waveform takes place in the readout modules (ROMs) in external electronics.

Online calibrations of channel gain and threshold are performed daily via internal pulse generation. The data are monitored online to check for FEA or other electronics failure or for miscalibrated output. Monitoring and control of high voltage, radiation protection (using silicon PIN diodes similar to the SVT, as well as RadFETs for integrated dose measurement), the gas system, and temperature are performed, similar to other subsystems, via a

slow control system based on EPICS.

Offline calibrations of the time-to-distance relation within cells, as well as of the deposited charge used for particle identification via dE/dx measurement, are performed. The time-to-distance relation is determined from two-prong events (Bhabha scattering events and dimuons) and is fit to a sixth-order Chebyshev polynomial for each cell layer, with separate fits to right and left sides of wires (to account for $E \times B$ asymmetries). A correction for time-to-distance variations as a function of track entrance angle to the cell is detemined via simulation (not calibration) and added to the calibrated entrance-angle-averaged relation. The energy loss per unit length of tracks, dE/dx, contains particle type information due to the dependence of dE/dx on particle velocity (Bethe-Bloch relation, see Figure 2.20), and is derived from measurements of integrated charge deposited in each cell along the track path. An overall multiplicative correction to the charge measurements

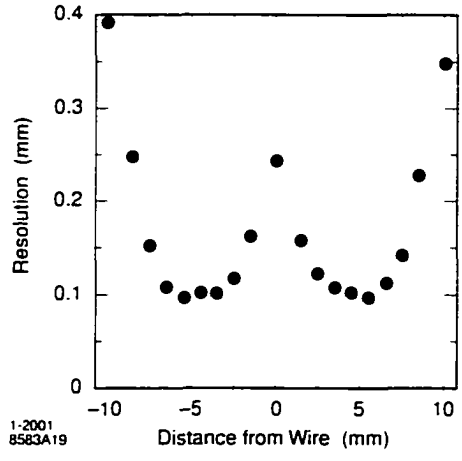


Figure 2.19: DCH drift distance resolution as a function of the drift distance in the cell.

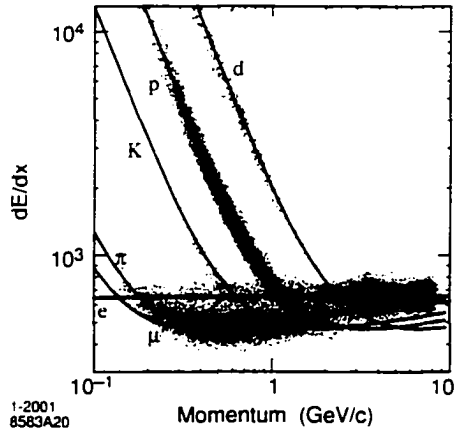


Figure 2.20: DCH particle identification as a function of momentum using dE/dx.

due to gas pressure and temperature variations is performed once per run: additional calibrations due to variations with track entrance angle in ϕ and in θ are performed only when high-voltage settings are changed.

The design goal for the average drift distance resolution was 140 μm. An average resolution of 110 μm is achieved. The drift distance resolution as a function of drift distance can be seen from the offline monitoring plot shown in figure 2.19. Particle identification using the drift chamber provides significant information up to high momenta. as can be seen in figure 2.20.

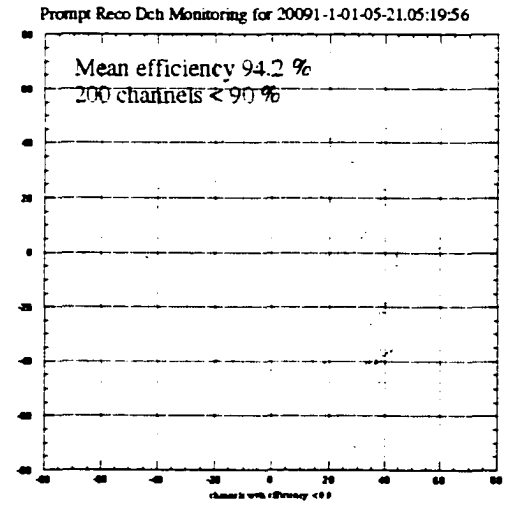


Figure 2.21: DCH channel efficiency.

Cell-by-cell channel efficiency is also monitored: typical efficiency is 90-95%, as may be seen in figure 2.21 (including a small region damaged from a high-voltage accident early in the commissioning phase, this can be seen towards the lower right of the figure as a higher concentration of points).

2.6 Track Reconstruction

Complete reconstruction of B decays (in addition to other major *BABAR* analysis techniques, such as tagging) requires precise and efficient charged particle tracking. As an example, $B^0 \to D^{*+}D^{*-}$ decays to reconstructible submodes have 7.85 charged tracks on average, thus single-track reconstruction inefficiency is effectively magnified by that power for reconstruction efficiency of $B^0 \to D^{*+}D^{*-}$. As will be seen in Chapters 3 and 4, separation of decays in these modes from combinatoric background requires precise determination of mass and energy, which in turn requires precision measurement of track momentum. The majority of other modes are just as dependent on charged particle tracking. Data from the DCH and SVT is combined to satisfy the stringent charged particle tracking requirements of *BABAR*.

Charged tracks are parametrized by the 5 variables $d_0, \phi_0, \omega, z_0, \tan\lambda$ and their error matrix. The parameters are defined as: d_0 is the distance of the track to the z-axis at the track's point of closest approach to the z-axis, z_0 is the distance along the z-axis of that point to the origin, ϕ_0 is the azimuthal angle of the track at that point, λ is the dip angle with respect to the transverse plane, and ω is the curvature of the track and is proportional to $1/p_t$. After tracks are recognized by the pattern recognition alorithms,

these 5 variables are fitted, and error matrices are extracted.

Offline track reconstruction begins with tracking and event time information from the Level 3 Trigger (see also section 2.9). Level 3 provides both tracks and an estimate of the time at which the interaction occured (with a resolution of approximately 5 ns), referred to as the $T0$.[2] Reconstruction continues by building track segments of 4 hits on four layers within a superlayer and using a linear fit to estimate and improve the level 3 $T0$ to a precision of about 3 ns. Next, the level 3 DCH tracks are refitted using the more precise offline time-to-distance calibration and placed on the list of reconstructed tracks. The $T0$ value is refined further (to 3 ns resolution) by finding the best $T0$ fit to the tracks themselves. Following that, two additional DCH track pattern recognition algorithms are run which select tracks with hits not used in L3 tracks. The $T0$ is improved again (to < 2 ns resolution) using these extra tracks. The DCH tracks are then fit using a Kalman filter algorithm. DCH tracks are then extrapolated into the SVT via a hit-adding algorithm, and then two standalone SVT track pattern recognition algorithms add any remaining SVT tracks. Finally, SVT tracks are extrapolated into the DCH to pick up any remaining DCH hits. This sequence will be discussed in detail in the following subsections.

2.6.1 $T0$ Reconstruction

Reconstruction of the event interaction time, or $T0$, is necessary for both fitting DCH tracks (since the DCH time-to-distance relation is necessary for position information of a track within a given cell, a time must be provided

[2]The e^+e^- interaction time is referred to as the "bunch $T0$," often shortened to "$T0$."

as input) and for rejecting out-of-time hits within the SVT to reduce background. $T0$ reconstruction takes place in several steps, iterated with track pattern recogition since the two are interdependent, during offline reconstruction.

The initial measurement of the event time is provided by the Level 1 Trigger in hardware electronics (see also section 2.9). The level 1 trigger looks for track segments in the DCH and clusters in the EMC, and sends an accept that includes all data in each subsystem's latency buffer to be read out. The trigger timing is tuned such that this buffer typically starts about 450 ns before the event, with a resolution of about 50 ns. That estimate, with an error of approximately 50 ns, forms the first event $T0$ estimate. The Level 3 Trigger uses the level 1 DCH segments and a fast, rudimentary time-to-distance function to fit tracks. The fit produces an event time measurement that is accurate to approximately 5 ns. This estimate is then given to offline reconstruction as input.

Offline $T0$ reconstruction begins with the DCH segment-based $T0$ finder. Four consecutive hits on adjacent layers within a superlayer form a DCH segment. A $T0$ value is fit for each segment such that the corresponding isochrone on each cell is tangent to a line segment passing through the superlayer. This requires a 3x3 matrix inversion corresponding to the two parameters of the line segment in addition to the $T0$. The event $T0$ estimate must then be obtained by combining these fitted segment $T0$s. There are several different segment cell patterns corresponding to different angles of the track through the superlayer, and segments are weighted according to type (highly-angled segments give lower-quality information than radial

ones, since they tend to be from lower-momentum tracks). The segment $T0$'s are sorted according to time and the weighted average is taken of the middle third of this list. Using only the middle third provides robustness against out-of-time background segments. The weighted average segment $T0$ is used as input to the Level 3 track converter, which then outputs a list of tracks to the event.

The tracks are then used to provide a more precise estimate of the $T0$ using the DCH track-based $T0$ finder. The DCH track-based $T0$ finder takes a list of tracks as input and finds the best fit $T0$ from the list of tracks. For each track, an average of the time residuals from each hit, weighted by the error on the residual taken from the time-to-distance resolution function, is calculated. This average is then added to the prior $T0$ estimate in order to obtain the best-fit $T0$ from that track, along with its associated error. The weighted average of these track $T0$s is then taken, with the track $T0$ with the largest χ^2 from the initial average dropped from the final average (for robustness against the occasional background track). This forms the track-based event $T0$.

The track-based $T0$ is calculated once after both the offline L3 track converter and the first DCH pattern recognition algorithm, the DchRadTrk-Finder, have run, and again after the two additional DCH pattern recognition algorithms, the DcxTrackFinder and DcxSparseFinder, have run. The final DCH $T0$ is used for all final track fits and is provided to the DRC for background rejection, after which the DRC is able to refine the $T0$ further.

2.6.2 Track Pattern Recognition

There are a total of 9 track pattern recognition algorithms run: 4 DCH track finders, 2 standalone SVT track finders, and 3 track hit-adders (one DCH-specific, one DCH → SVT, and one SVT → DCH). DCH track reconstruction occurs first, due to the use of an offline background filter that requires DCH tracks. The following paragraphs give some detail on each of the track pattern recognition algorithms.

The first track finder in the reconstruction sequence is the DchL3TrkConverter, which converts Level 3 Trigger tracks to offline reconstruction tracks. This is not a pattern recognition algorithm *per se*, since level 3 has already done the pattern recognition (with the use of Level 1 Trigger DCH segments), however it takes the level 3 tracks and refits them to a helix using the full offline time-to-distance calibration. Using the level 3 information is a dramatic speed advantage for track reconstruction, reducing the time spent by nearly 1/3. The next track finder in the sequence is the DchRadTrkFinder, which finds tracks which have small impact parameter from the origin and have not already been found by level 3. Like the level 3 algorithm, the DchRadTrkFinder is segment-based, taking the list of DCH segments and combining them to form tracks. This restricts the algorithm to tracks which are not too curved away from radial — only the outgoing parts of looper tracks (tracks under 200 MeV which curl inside the tracking volume) will be found by the DchRadTrkFinder.

To add hits onto the radial tracks found by level 3 and the DchRadTrkFinder, for example hits in the apsidal and incoming parts of looper tracks, a DCH-specific track hit-adding algorithm, the DcxHitAdder, is run. This

used a road-based algorithm to find and add hits, that are not necessarily part of segments, onto the ends of pre-existing tracks. Following the DcxHitAdder is a third DCH track pattern recognition algorithm, the DcxTrackFinder. Similar to the first two DCH track finders, the DcxTrackFinder is also segment-based, but it additionally searches for tracks with a larger impact parameter from the origin than the previous finders, which is necessary for the critical task of efficient K_S^0 recontruction (as well as for Λ's, etc.). However, a significant fraction of such tracks cannot be found by a segment-based algorithm at all (due to less than 100% efficiency, there is often a hit missing on a given layer within a superlayer, preventing segment

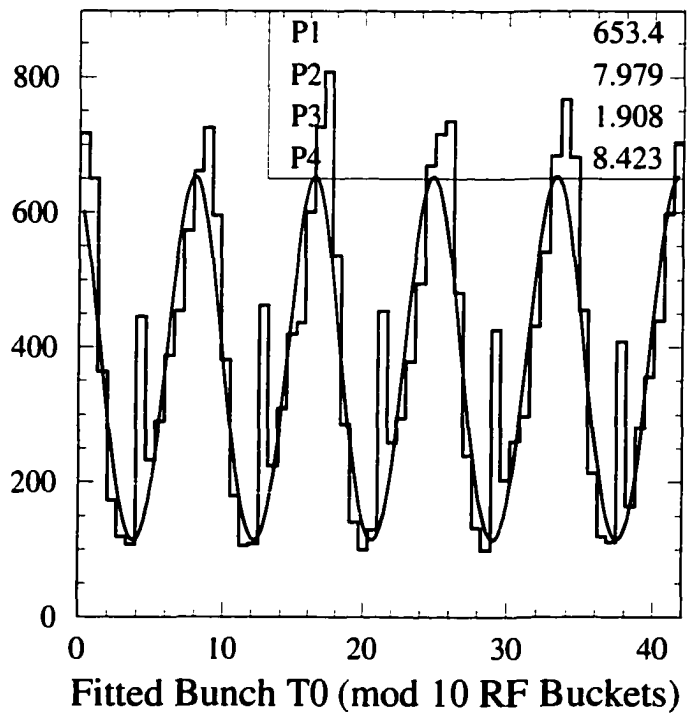


Figure 2.22: Fitted bunch T0 - PEP-II fiducial signal, modulo 42 ns (= 10 RF bucket intervals). Filled buckets contribute to peaks; the data is fit to a sum of gaussians and the resolution of the bunch T0 corresponds to the fitted width of the gaussians (= 1.9 ns in the above plot from a randomly chosen run). The spikes preceding the peaks correspond to events for which a $T0$ cannot be fit, such as events with no charged tracks.

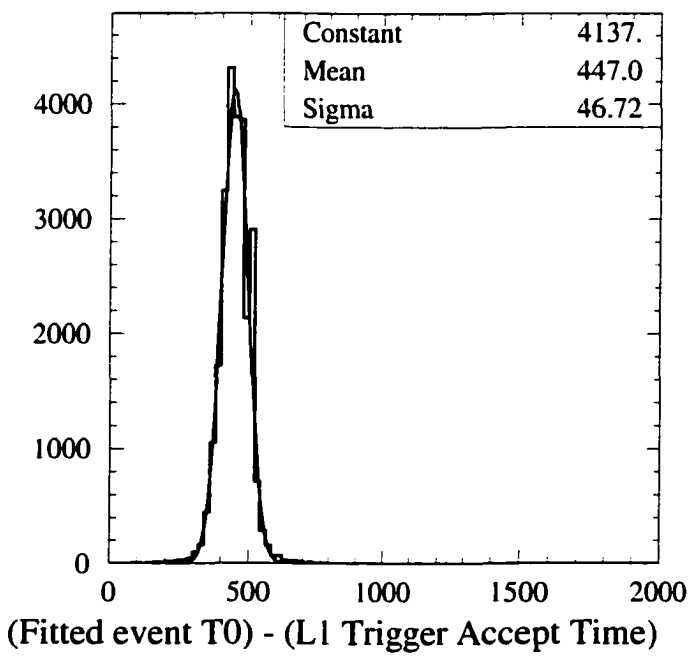


Figure 2.23: Fitted bunch T0 - Level 1 Trigger Accept time. The L1Accept averages approximately 450 ns prior to the $T0$, with a resolution of approximately 50 ns. The spike at 500 ns is for events for which a $T0$ cannot be fit, such as events with no charged tracks.

formation). Thus, a fourth DCH pattern recognition algorithm is run, the DcxSparseFinder. This is an arc-based algorithm that searches for hits along arcs of circles in the x-y plane.

Following the DCH-specific track finders, SVT tracking is performed. First the TrkSvtHitAdder is run, which adds SVT hits onto DCH tracks. This uses a Kalman-based algorithm to propagate tracks through SVT hits creating a "tree" of possible extensions to the track. Branches of the tree are pruned, leaving only the most favorable SVT extension as part of the track. The two standalone SVT track finders run after SVT hit-adding. The first of these, SvtTrackFinder, forms "space-points" from combinations of ϕ and z-hits from opposite sides of a wafer. At least three space-points are needed to form a helix. Since 9 of the 208 half-modules are defective, this can introduce an inefficiency since information from the working side of the module is not used if the other side is defective. To ameliorate this, and to find additional tracks, an orthogonal SVT standalone algorithm for track finding, SvtRPhiPR-SvtCircle2Helix, is used. This algorithm (SvtRPhiPR) uses ϕ-hits to search for circles in the x-y plane. Once a circle is found, z-hits are then added (SvtCircle2Helix)

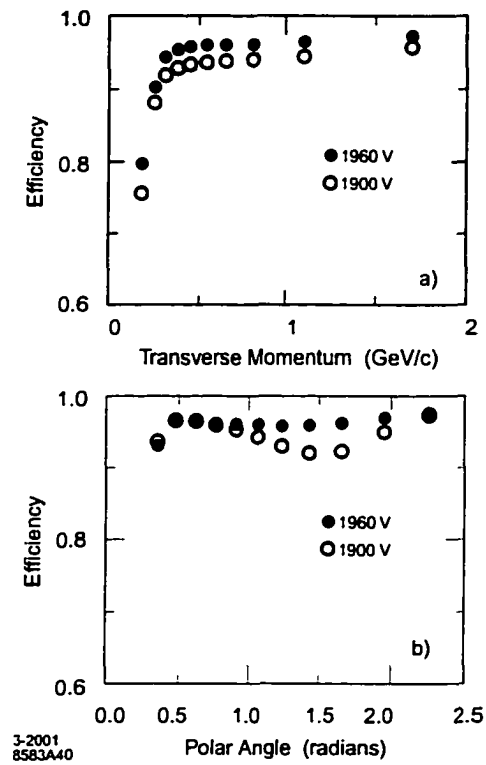


Figure 2.24: DCH tracking efficiency at operating voltages of 1900 V and 1960 V as a function of (a) transverse momentum and (b) polar angle. Efficiency is the fraction of all SVT tracks projecting into the DCH which have the DCH portion reconstructed. The current default operating voltage is 1930 V, which provides an efficiency that is between 1900 V and 1960 V, but closer to 1960 V.

to form a helix. Finally, a hit-adding algorithm very similar to TrkSvtHitAdder, TrkDchHitAdder, is used to add DCH hits onto standalone SVT tracks.

2.6.3 Track Fitting and the Kalman Algorithm

BABAR tracks would be exact helices if not for 3 effects: multiple scattering, energy loss in material, and inhomogeneities in the magnetic field. Although these effects are fairly small in *BABAR* due to the small amount of material in the tracking volume and the $< 5\%$ inhomogeneities in the field, they nevertheless are important for the level of precision needed for accurate Δz and vertex fit quality. Thus a Kalman filter is used to propagate tracks accounting for each of those three effects and create a piecewise helix trajectory. For initial fits (and for input to the Kalman algorithm), a simple helix fit will suffice. Track fitting is done using both helix fits (for initial fitting) and a piecewise-helix Kalman fit algorithm (for the final fit). The DCH and SVT standalone track fitters (and DCH-specific hit-adder) use a helix fit for the initial fit which is replaced by a Kalman fit later in reconstruction, whereas the DCH $\rightarrow$ SVT and SVT $\rightarrow$ DCH hit-adders are integrated with the Kalman fitter to perform a piecewise helix fit as the hits are added.

The helix fit algorithm performs a least-squares fit to a list of hits. It assumes the weight matrix is diagonal, *i.e.* that the hits are uncorrelated. The fit iterates and removes the hit with largest "pull" (residual divided by error) on each iteration if it lies more than 3 sigma from the fit. Removed hits which return to within 3 sigma after an iteration are added back. Convergence occurs either when the change in total χ^2 is less than 0.01 or if the fit

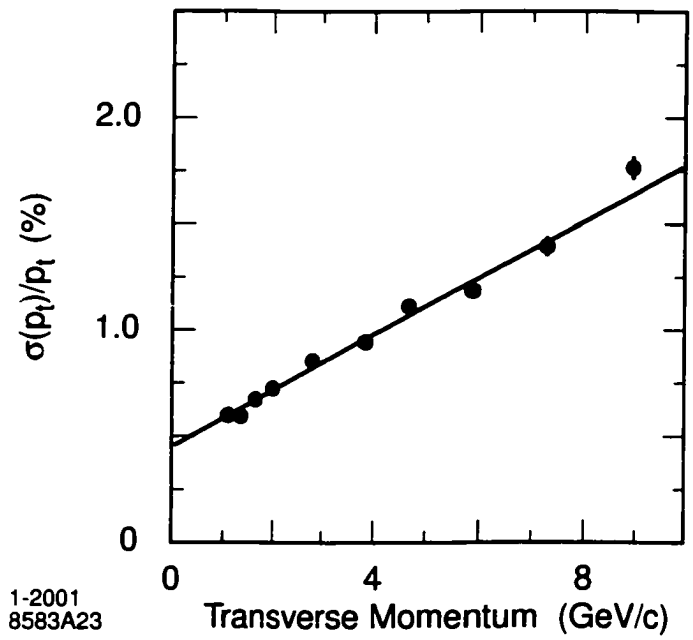


Figure 2.25: Transverse momentum (p_t) resolution as a function of transverse momentum, using cosmic ray muons.

reaches 12 iterations. If the latter occurs before the former, it is considered a failed fit.

The Kalman fit takes into account the effects of material and inhomogeneous magnetic field to create a piecewise helix fit. The parameters of a track can be transformed into "weight space" where:

$$\gamma \equiv C^{-1}, \quad \beta \equiv \gamma P \tag{1}$$

where P is the vector of 5 track parameters (taken as input from a preliminary helix fit) and C is the corresponding covariance matrix. The effect of adding a hit on these parameters is:

$$\gamma_H \equiv L^T w^2 L, \quad \beta_H \equiv L^T w(LR - r) \tag{2}$$

where r is the residual of the hit, R is the hit's position, and $L \equiv \frac{\delta r}{\delta P}|_R$. These act as additive corrections to the weights:

$$\gamma_{new} = \gamma_{old} + \gamma_H, \quad \beta_{new} = \beta_{old} + \beta_H \tag{3}$$

These are the Kalman filter equations.

Performing the fit in weight space also allows for simple equations desribing magnetic field inhomogenaities, energy loss, and multiple scattering. The materials and magnetic field map are kept in the *BABAR* conditions database (the database used for storage of detector calibration constants) allowing for their use directly in the fit. Both an inwards and an outwards fit are done, with the final weights β and γ being the sum of the inner anout outer fits (this is referred to as "smoothing"). A plot showing the results of *BABAR* track fitting on momentum measurement may be seen in figure 2.25. A resolution of $\sigma(p_t)/p_t = 0.13\% \times p_t + 0.45\%$ is obtained.

2.7 The DIRC

BABAR has stringent requirments for $\pi - K$ separation over a large momentum range. At the lower end of the range, primarily at momenta < 1 GeV, flavor tagging using kaons from cascade decays is an efficient way of determining B flavor. At the high end of the range, reconstructing $B^0 \rightarrow \pi^+\pi^-$ and $B^0 \rightarrow K^\pm\pi^\mp$ requires separation at momenta up to 4.2 GeV in the lab frame. At intermediate energies, reducing background in charm decays such

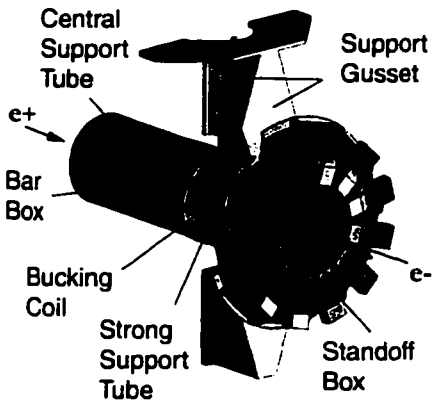


Figure 2.26: View of DIRC mechanical stucture.

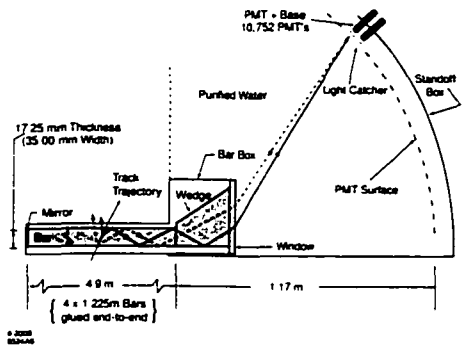


Figure 2.27: DIRC schematic showing the principle behind PID measurements. The Cerenkov angle is preserved through specular internal reflection.

as $D^0 \to K\pi$ is necessary for $B^0 \to D^{(*)+}D^{(*)-}$ reconstruction. The particle identification device must exhibit sufficient $\pi - K$ separation throughout this wide range of momentum with a minimum of material in order to avoid adversely impacting calorimeter resolution.

The DIRC (Detector of Internally Reflected Cerenkov light) principle uses internal reflection within quartz bars to propagate Cerenkov light to readout phototubes while preserving the Cerenkov angle. This requires extremely flat surfaces in order to avoid dispersing the reflected angles. Fused, synthetic silica quartz is used due to the excellent optical surface it allows through polishing, as well as other favorable properties such as long attenuation length, low chromatic dispersion, small radiation length, and radiation hardness. As

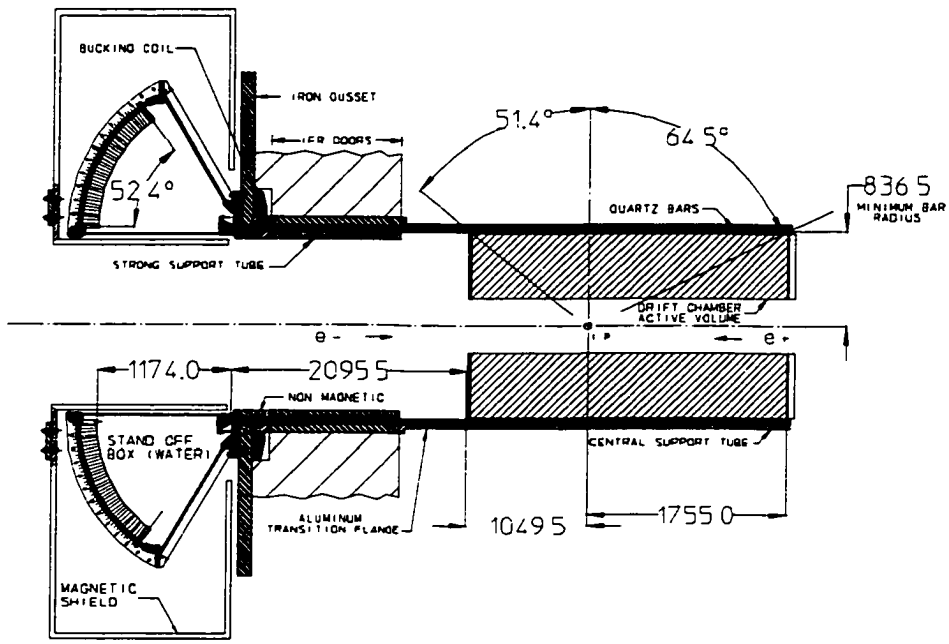


Figure 2.28: Longitudinal section of the DIRC. Length units are mm.

shown in figure 2.27, the light is internally reflected down to a wedge to reflect photons into a water-filled "standoff box." The standoff box is enclosed by an array of 10752 photomultiplier tubes, which are each 29 mm. in diameter. The Cerenkov light from a particle passing through the DIRC forms a ring (essentially a conic section) imaged on the phototubes. The opening angle of this conic section contains information on particle type via the typical relation $\cos\theta_c = 1/n\beta$, with β being the particle velocity normalized to the speed of light, and n being the mean index of refraction (= 1.473 for fused silica).

Both efficiency and the timing of the electronics are critical for DIRC performance. Timing is critical for two reasons: one, for background hit

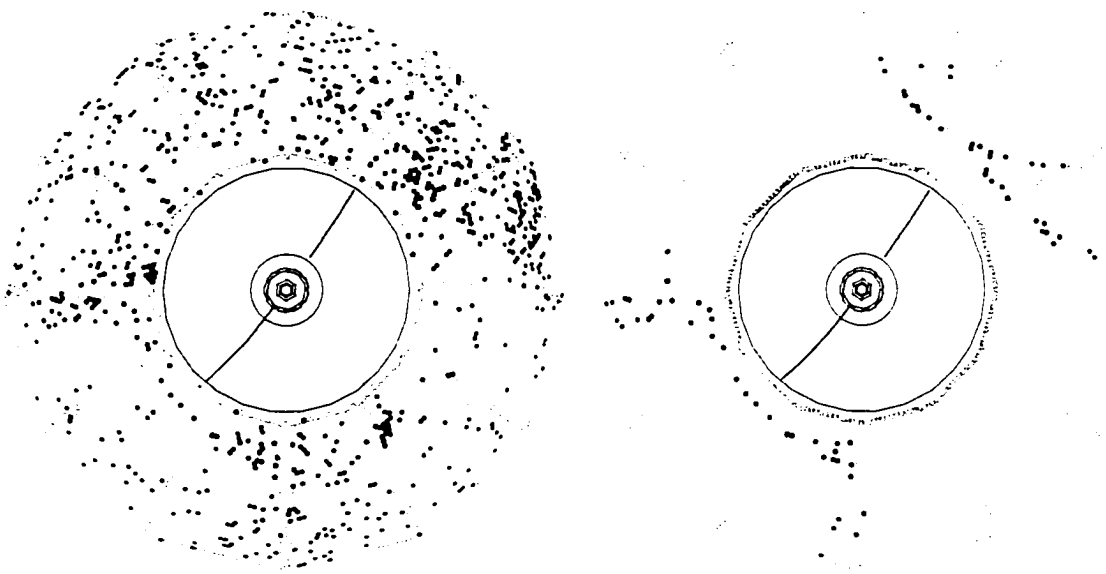


Figure 2.29: Display of an $e^+e^- \rightarrow \mu^+\mu^-$ event reconstructed in *BABAR* with two different time cuts. On the left, all DIRC PMTs with signals within the ± 300 ns trigger window are shown. On the right, only those PMTs with signals within 8 ns of the expected Cherenkov photon arrival time are displayed.

rejection, resolving ambiguities, and separation of hits from differing tracks within an event; and two, timing gives information on the photon propagation angles, allowing an independent measurement of the Cerenkov angle. The intrinsic timing resolution of the PMTs is limited to 1.5 ns by transit time spread. Data from the phototubes is read out to front-end electronics, which performs the amplification, digitization, and buffering. Reduction of data from out-of-time or noisy PMTs is performed in in the external electronics and reduces the data volume by 50% using rough timing cuts. Online calibration of PMT efficiency, timing response, and electronics delays uses

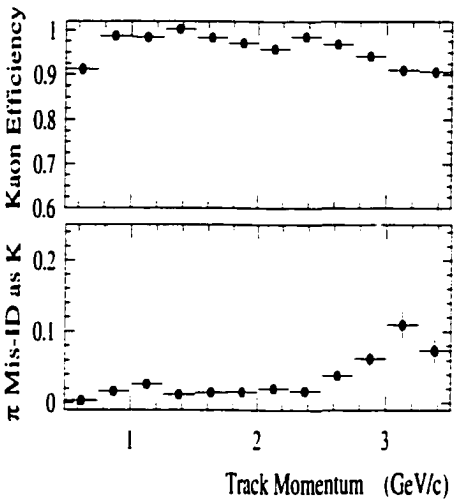


Figure 2.30: Kaon selection efficiency and misidentification probabilities as a function of momentum.

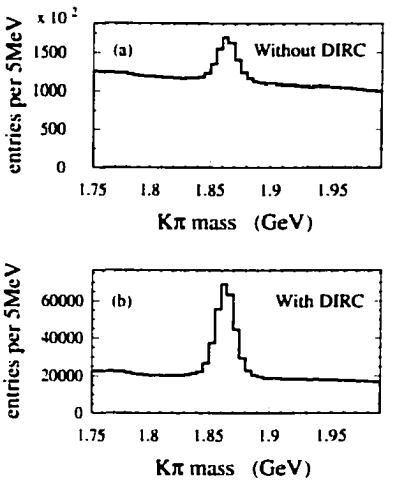


Figure 2.31: Inclusive $K\pi$ invarant mass spectrum (a) without and (b) with the use of the DIRC for kaon identification.

a light pulser system which generates precise 1 ns flashes from blue LEDs inside the SOB.

The DIRC has performed well throughout *BABAR*'s operational lifetime: 99.7% of PMTs are operating with design performance. The measured time resolution is 1.7 ns, very close to the intrinsic resolution of the PMTs. The Cerenkov angle resolution for dimuon events is 2.5 mrad, close to the design goal of 2.2 mrad. This results in $\pi - K$ separation at 3 GeV of 4.2σ. The mean kaon selection efficiency and pion misidentification for a "loose" selection are 96.2% and 2.1% respectively, as can be seen in figure 2.30. This results in dramatic background rejection with little signal loss for charm reconstruction, as may be seen in figure 2.31.

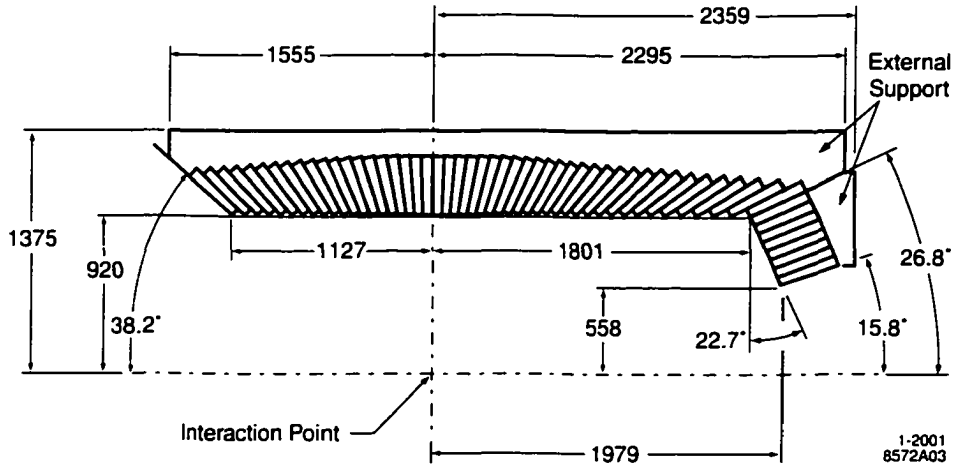


Figure 2.32: Longitudinal section of the EMC. Length units are mm.

2.8 The Electromagnetic Calorimeter (EMC)

The design parameters for the *BABAR* EMC are driven by the requirements of precisely measuring energies over a spectrum from 20 MeV to 9 GeV in a 1.5 T magnetic field and a high radiation environment. At the high end of the spectrum, measurements of QED processes such as Bhabha and two-photon scattering, as well as (at slightly lower energies) photons from the critical physics processes $B^0 \to \pi^0\pi^0$ and $B^0 \to K^*\gamma$ decays, present the motivating incentive. The need for efficient detection of photons from high multiplicity B decays containing π^0's (such as $B \to D^{(*)}\overline{D}^{(*)}$), as well as slow photons from D_s^{*+} decays, determines the requirement for the low end of the spectrum. *BABAR* uses a thallium-doped cesium iodide (CsI(Tl)) crystal calorimeter in order to achieve the necessary energy and angular resolution to meet these physics requirements.

The EMC contains a cylindrical barrel and a conical endcap containing a total of 6,580 CsI(Tl) crystals. The crystals have nearly square front and rear faces with a trapezoidal longitudinal cross-section. They range in length from 29.6 to 32.4 cm with a typical front face dimension of 4.7 x 4.7 cm. A diagram can be seen in figure 2.32. The crystals are mounted in thin (300 μm) carbon-fiber composite housings which are mounted on an aluminum strong-back (see figure 2.33). Although light incident on the crystal boundary is internally reflected, the small part that is emitted is reflected back with a coating of white reflective TYVEK on the outer surface. Surrounding that are thin layers of aluminum and mylar to act as RF shielding and electrical insulation respectively. On the rear face of the crystal, two 1 cm^2 silicon PIN diodes with quantum efficiency of 85% for CsI(Tl) scintillation light are mounted via transparent polystyrene substrate.

Each diode is connected to a low-noise preamplifier which shapes and amplifies the signal by a factor between 1 and 32. The signal is then transmitted to mini-crates at the end of the barrel (see figure 2.33) where a digitization CARE ("custom auto-range encoding") chip provides an additional variable amplification factor. Unlike other subsystems (except for the IFR), the EMC does not buffer the data on front-end electronics; rather it outputs the full digital datastream to the read-out modules in external electronics, which perform, on receipt of a level 1 accept signal, a parabolic fit to the digitally filtered datastream to derive energy and time measurements.

A variety of online calibrations and checks is performed, including a neutron source which produces a monoenergetic 6.13 MeV calibration signal and a xenon flash light pulser system. Offline energy calibrations are necessary

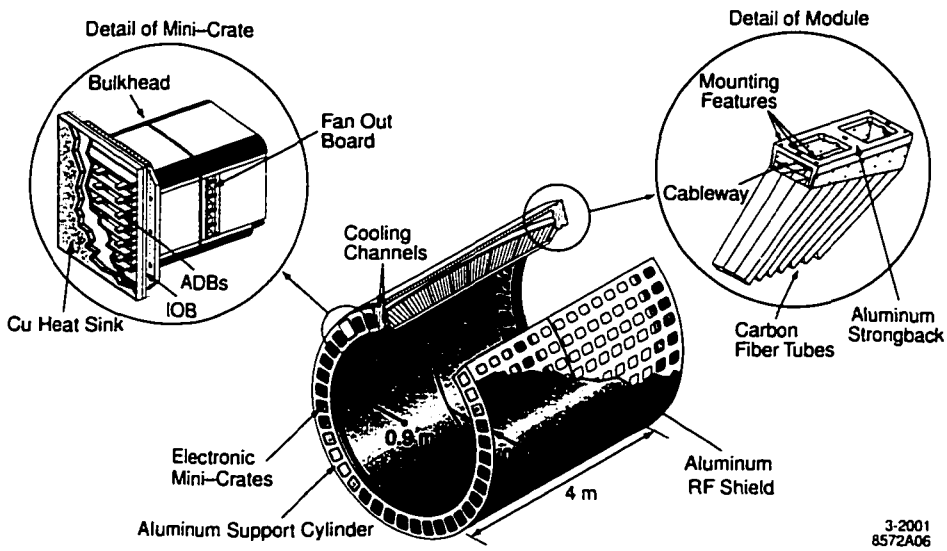


Figure 2.33: The EMC barrel support structure, modules, and mini-crates.

for higher energy corrections. The relation between polar angle and energy of Bhabha and radiative Bhabha scattering events is used to calibrate the 0.8-9 GeV range. The middle range is covered by π^0 calibration, which constrains the mass of a sample of π^0's to the known value, extracting correction coefficients.

The clustering pattern recognition uses a seed crystal algorithm to establish energy clusters. Local energy maxima within a cluster are used (if there are more than 1) to separate the cluster into bumps. Charged particle tracks are associated with bumps using a χ^2 consistency cut. In an average hadronic event, 15.8 clusters are detected, of which 10.2 are not associated to a track.

Energy resolution is determined using $\chi_{c1} \rightarrow J/\psi\ \gamma$ and Bhabha scattering events to be

$$\frac{\sigma_E}{E} = \frac{(2.32 \pm 0.30)\%}{\sqrt[4]{E(\,\mathrm{GeV})}} \oplus (1.85 \pm 0.12)\% \tag{4}$$

and angular resolution is determined using π^0 and η decays to be

$$\left(\frac{3.87 \pm 0.07}{\sqrt{E(\text{GeV})}} + 0.00 \pm 0.04\right) \text{mrad.} \tag{5}$$

As can be seen in figure 2.34, the reconstructed π^0 average width is 6.9 MeV. The separation of electrons from charged hadrons using the ratio of shower energy to track momentum (E/p) and other variables may be seen in figure 2.35.

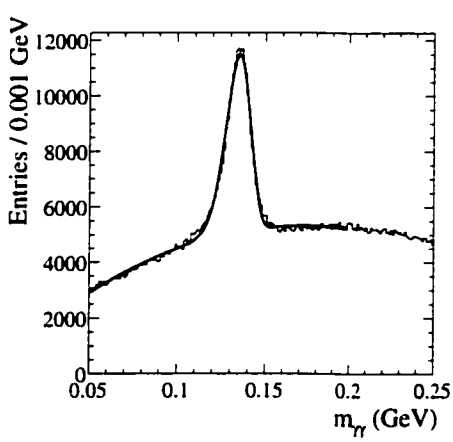


Figure 2.34: Two photon invariant mass, using photons between 30 and 300 MeV.

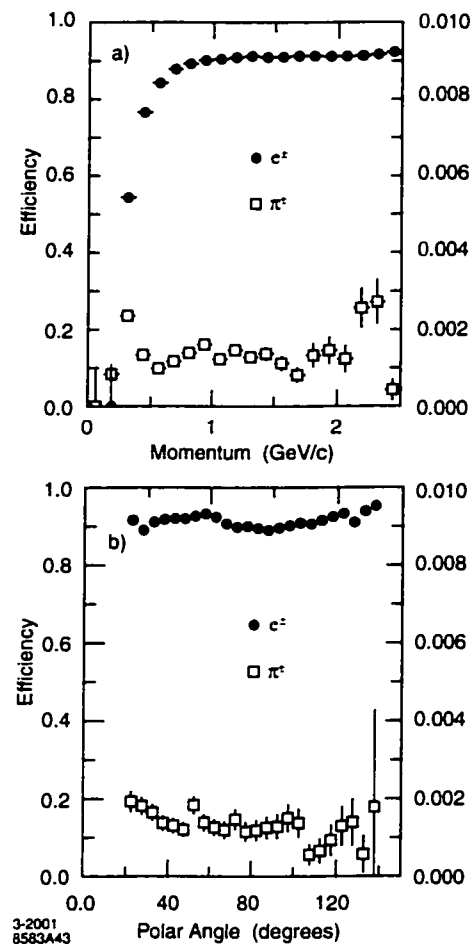


Figure 2.35: Electron efficiency and misidentification as a function of (a) momentum, and (b) polar angle.

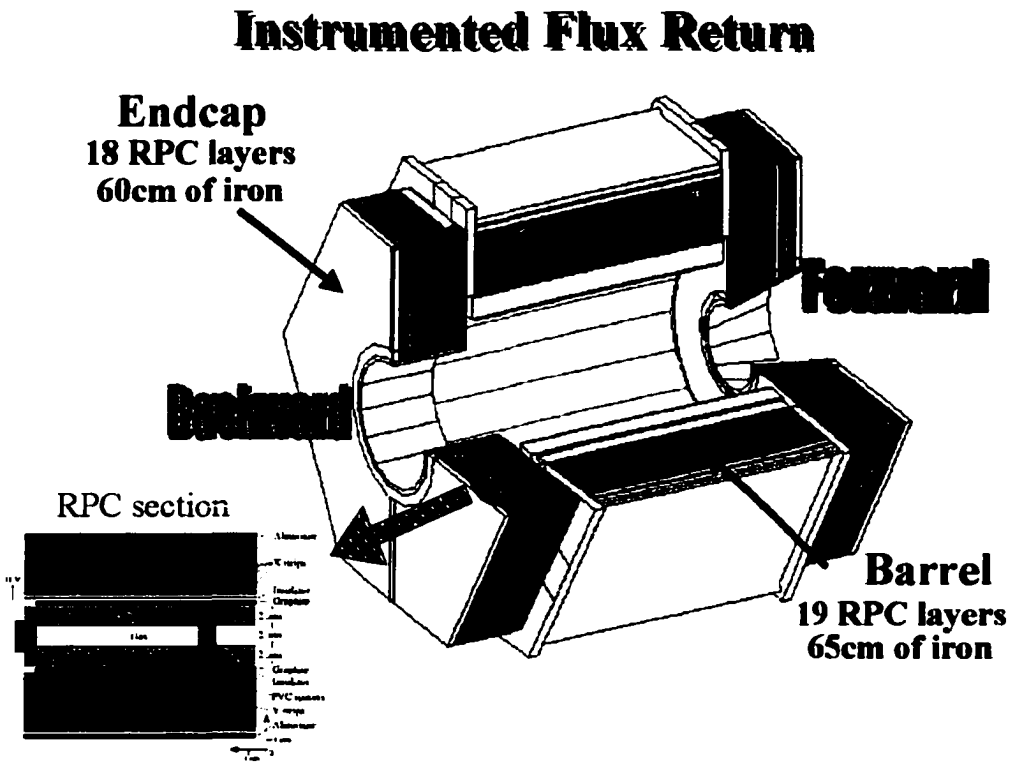


Figure 2.36: The instrumented flux return modules, structure, and RPCs.

2.9 The Instrumented Flux Return (IFR)

Detection of neutral hadrons (primarily K_L^0's) and muons is necessary for several *BABAR* analyses and analysis techniques. Muons provide a very clean B flavor tag, extremely useful for increased efficiency in tagging the opposite-side B for time-dependent CP violation measurements. Muons are also necessary for reconstructing $J/\psi \rightarrow \mu^+\mu^-$, as well as for measuring semileptonic branching fractions, required for extracting magnitudes of CKM elements. K_L^0 reconstruction is critical for the $B^0 \rightarrow J/\psi\, K_L^0$ channel. *BABAR* uses an Instrumented Flux Return (IFR) composed of layers of resistive plate chambers

(RPCs) and steel plates in order to provide enough material to separate pions and kaons from muons and to efficiently detect the presence and position of both μ and K_L^0 over a large solid angle.

As shown in figure 2.36, the IFR consists of layers of planar RPCs in a barrel and endcap (red lines) as well as 2 layers of cylindrical RPCs (green lines) between the EMC and the magnet. Between the RPC layers are steel plates of thickness varying between 2 cm. (inner plates) to 10 cm (outer plates). The total mass of the IFR is 870 metric tons. Planar RPCs contain a 2 mm Bakelite gap with $\sim$ 8 kV across it. Ionizing particles which cross the gap create streamers of ions and electrons in the gas mixture (which is typically 56.7% Argon, 38.8% Freon, and 4.5% isobutane) which in turn creates signals via capacitive coupling on the "x-strips" and "y-strips" on opposite sides of the RPC. Strip width varies between 16 mm and 38.5 mm. The 2 mm gap is kept constant using polycarbonate spacers spread at 10 cm intervals and glued to the Bakelite. The Bakelite surface is smoothed with an application of linseed oil. Cylindrical RPCs are composed of a special thin and flexible plastic, rather than Bakelite, and have no linseed oil or other surface treatment. They are laminated to cylindrical fiberglass boards.

The strips are connected to front-end cards (FECs) containing the amplification and digitization electronics, which are primarily sandwiched between the iron plates. FECs shape the signal and perform a comparison with an adjustible threshold. Similar to the EMC, the IFR does not buffer its data on the detector, so the full digital datastream is output to front-end crates (located beside the detector) which contain TDC timing as well as buffering and calibration electronics. Data is read out to the ROMs on receipt of a

level 1 accept signal. Online calibration is performed using test pulse generators integrated in the front-end crates. Offline efficiency calibration is also necessary for muon ID (in order to determine the expected hits for the muon hypothesis), and this is done using cosmic rays.

Reconstruction of clusters proceeds via two methods: a standalone method where groups of hits along 1 dimension within a module are joined to form one-dimensional clusters, which are then combined with opposite-side hits to form two-dimensional clusters, and then with other modules to form 3D clusters; and a "swimmer" method, where charged tracks are propagated to the IFR — 1D clusters within 12 cm. of the expected path are combined to form 2D or 3D clusters. Clusters which are not matched to a charged track are considered as neutral clusters. Muon identification uses variables such as number of expected vs. actual interaction lengths tranversed and the χ^2 match to the charged track. Muon identification performance may be seen in figure 2.38. K_L^0 efficiency roughly increases linearly with momentum and varies between 20% and 40% between 1 and 4 GeV.

The IFR has been having serious losses in efficiency since the start of

operation. Although this is still a topic of study, it seems that, for a significant number of RPCs, the linseed oil failed to cure properly and has caused a nonuniformity in the spacing between the plates, amongst other problems. A plot of efficiency as a function of time may be seen in figure 2.37. Replacement RPCs are currently undergoing fabrication and testing, as well as possible design modifications to the structure (adding more material to suppress hadron punch-through).

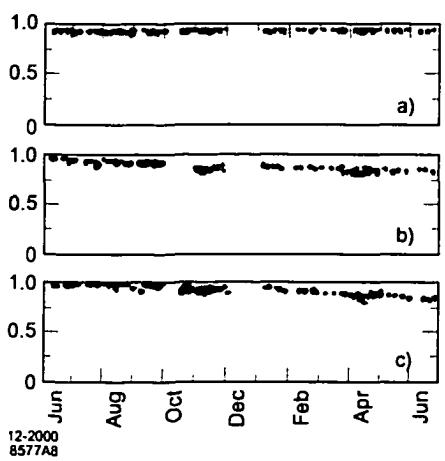


Figure 2.37: Average efficiencies for 12 months (beginning June 1999) for 3 types of RPCs: (a) efficient and stable, (b) continuous slow decrease in efficiency, (c) faster decrease in efficiency beginning near December 1999.

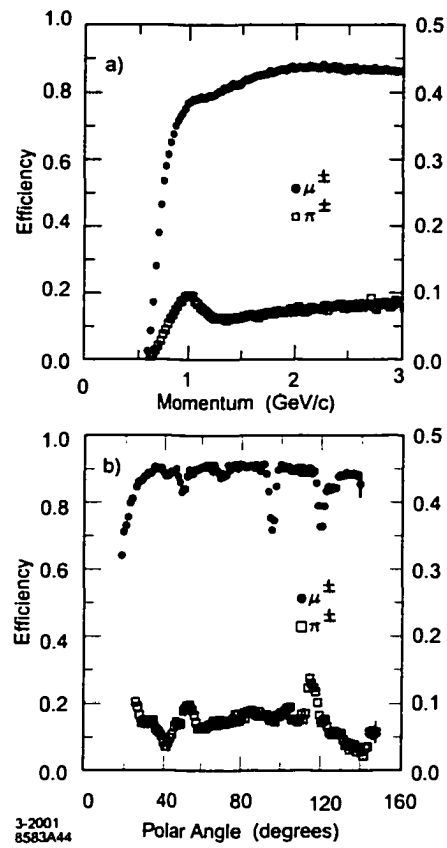


Figure 2.38: Muon efficiency and misidentification (for a "loose" selection) as a function of (a) momentum, and (b) polar angle.

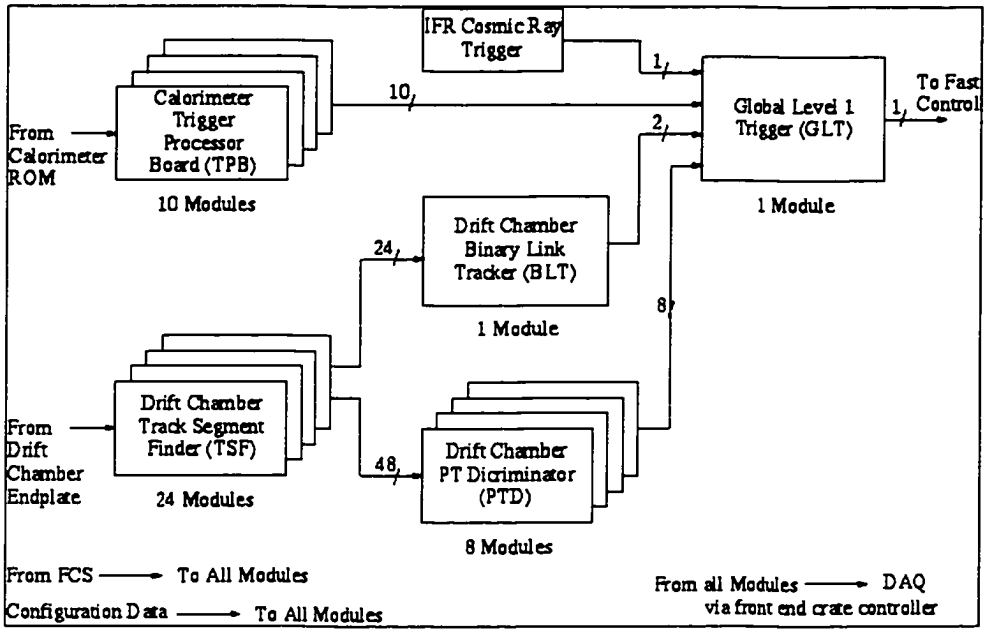


Figure 2.39: Simplified schematic diagram for the Level 1 Trigger.

2.10 The Trigger

The *BABAR* trigger needs to provide high efficiency that is well-understood and stable for physics events. Since the events which pass the trigger must be fully reconstructed in the offline event reconstruction, the output rate must be no higher than 120 Hz to satisfy computing limitations of the offline processing farm. Since events with either a DCH track or a > 100 MeV EMC cluster occur at 20 kHz, the trigger is responsible for scaling this rate down by a factor of > 150 while accepting over 99% of B events, over 95% of hadronic continuum, and over 90% of $\tau^+\tau^-$ events. It also must be flexible enough to deal with changing background conditions, as this can happen at any given time at *BABAR*, without impact on physics or increase in online

dead time (which must be < 1%). The *BABAR* trigger is implemented in two levels, a level 1 hardware trigger (L1), and a level 3 software trigger (L3). Although development of an additional level 2 hardware trigger would have been possible, it is not necessary under current background and luminosity conditions; the level 1 and level 3 triggers together suffice.

The level 1 trigger, as can be seen above in figure 2.39, is divided into subsystem compontents: the L1 drift chamber trigger (L1DCT), L1 calorimeter trigger (L1EMT), an IFR trigger used for calibration (L1IFT), and global electronics for producing the final L1 accept signal (L1GLT). The L1DCT is further subdivided azimuthally into track segment finders (TSF), a binary link tracker for producing tracks from the segments (BLT), and a p_t discriminator (PTD). The 24 TSF modules sample each DCH cell in axial layers for signals every 269 ns (64 clock ticks of the PEP-II-*BABAR* 4.2 ns clock interval). The axial DCH cells are arranged into 1776 "pivot groups" (see figure 2.40) and segments are constructed from hits within a pivot group. Only axial layers are used to avoid the complication of accounting for stereo layer within hardware electronics.[3] Each cell contributes a 2 bit quantity (containing very rough information on the time, as well as the presence, of a hit segment.

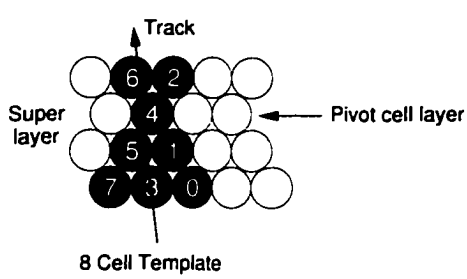


Figure 2.40: Level 1 drift chamber TSF pivot group (grey and black cells). The four grey cells, one per layer, are hit by the track, forming a

[3]However, stereo information will be incorporated in the level 1 "ZPD" upgrade (see the last paragraph in this subsection).

on that cell) per sample, thus each
pivot group outputs 16 bits. The TSF
look-up table then determines if a given 16-bit quantity corresponds to a valid segment or not. The binary link tracker (BLT) receives this information and detemines whether segments lie in a road defined by "supercells," which are sectors of a superlayer covering to 1/32 of the DCH in ϕ. Patterns of segment-containing supercells that appear to correspond to tracks (according to the BLT look-up table) are output to the L1 global trigger. In parallel with the BLT, the p_t discriminator (PTD) checks TSF segments in axial superlayers to see if they are consistent with a track having a greater p_t than a configurable miniumum cutoff value. This information is also output to the GLT.

The level 1 calorimeter trigger (EMT) divides the EMC into 280 "towers" of 24 crystals each (22 for the endcap). All crystal energies within a tower which are above a 20 MeV threshold are summed and supplied to the EMT trigger processor boards (TPBs). The TPBs digitally filter the energy deposition (to smooth the output waveform of noise) and compare neigboring towers to look for clusters which span more than one tower. Trigger line "primitives" (bytes corresponding to trigger type and information) are output to the GLT corresponding to the energy and placement of found clusters.

The global level 1 trigger (GLT) receives the trigger line primitives from the EMT and DCT, along with information from an IFR trigger (IFT) which is used for cosmic ray and dimuon calibration purposes, and performs timing alignment to reduce jitter. The GLT does some rudimentary matching between DCT tracks and EMT clusters, and performs a logical AND of the

input trigger primitives, which defines the output trigger line. The combined L1 trigger efficiency is > 99.9% for generic $B\overline{B}$ events, 99% for continuum, and 94.5% for $\tau^+\tau^-$ events.

The level 3 trigger combines DCT tracks and EMT clusters with the full DCH and EMC information. Level 3 runs on a farm of 64 Sun CPUs (which will shortly be upgraded to faster Pentium CPUs running Linux). The level 3 DCH algorithm fits L1 tracks to helices and is able to determine the z_0 of tracks, which is important information for rejecting background. The dominant source of events accepted by level 1 is beam-gas or beam-wall interaction background, as well as cosmic rays, which can be separated from physics events using the point of closest approach of tracks to the interaction point(IP). As can be seen in figure 2.42, avoiding tracks outside of the IP peak can result in significant rejection of background. In addition, based on information from the level 3 DCH algorithm and the level 3 EMC algorithm, the Bhabha event rate may be prescaled from ~130 Hz to 30 Hz, and $\gamma\gamma$ events down to 10 Hz. Configurable filters for special calibration modes which might otherwise be rejected run as part of L3. The total L3 output rate is 120 Hz.

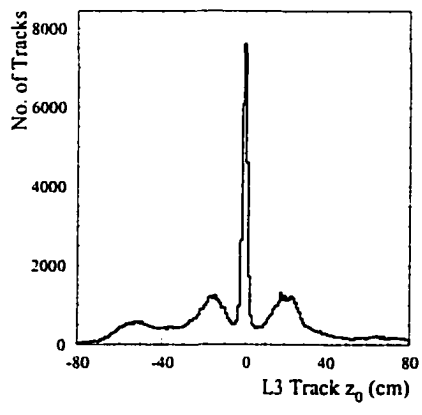


Figure 2.42: Track z_0 for L1 tracks, reconstructed by L3.

The *BABAR* trigger system has performed well since commissioning. It allows for significant flexibility while efficiently rejecting background. Future luminosity upgrades will significantly impact the trigger system, as level 1 is near output threshold rate.

An upgrade to the L1DCT, the ZPD (Z-Pt-Discriminator) will move some of the functions currently in L3 (z_0 rejection) to L1 to reduce the higher L1 accept rate expected at higher luminosity. This will, however, not become necessary until 2003, and will be added at that time.

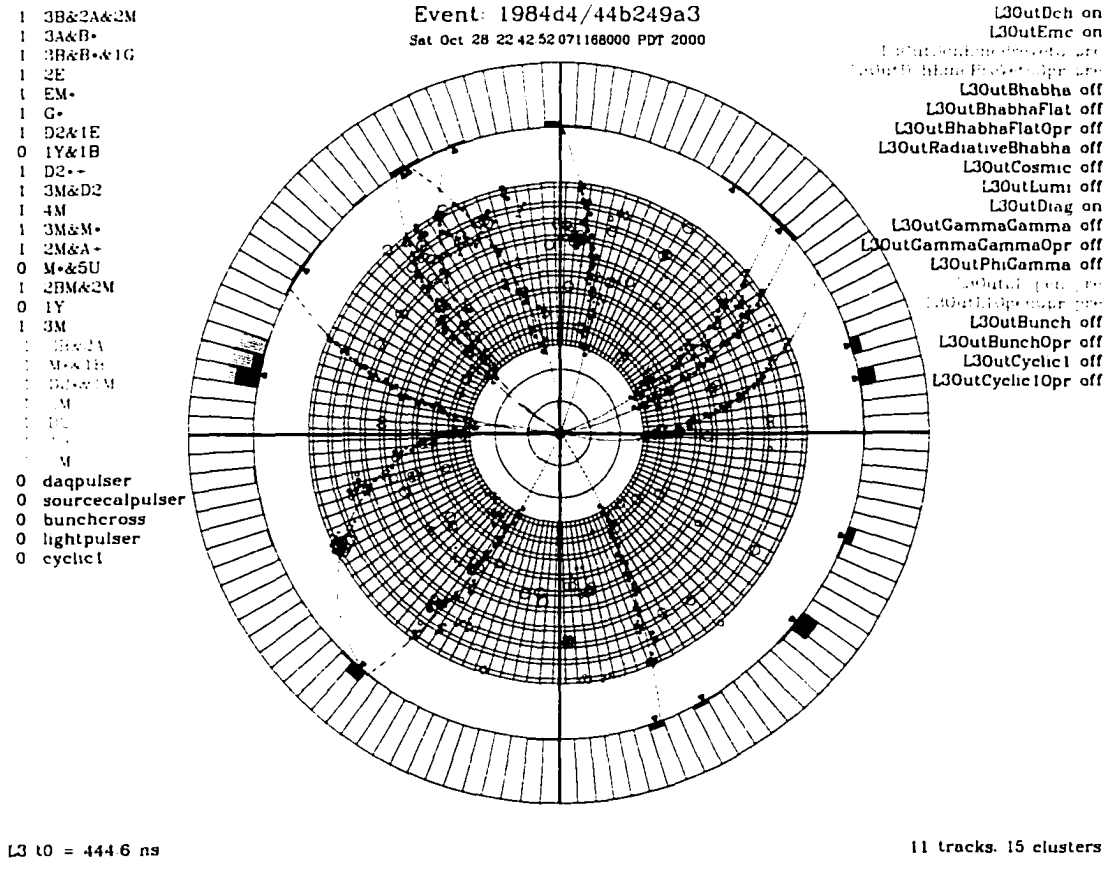


Figure 2.41: A level 3 trigger event display for a multihadron event.

Chapter 3

Measurement of $B \to D^{(*)}\overline{D}^{(*)}$ Branching Fractions

3.1 Introduction

This chapter contains measurements of branching fractions for the decays $B^0 \to D^{*+}D^{*-}$, $B^0 \to D^{*\pm}D^{\mp}$, and $B^{\pm} \to D^{*\pm}D^{*0}$. The primary motivation for analysis of $B^0 \to D^{(*)+}D^{(*)-}$ modes, such as the first two decays, is that they are sensitive to the Unitarity Triangle parameter $\sin 2\beta$, providing a complementary measurement to that derived from charmonium decays such as $B^0 \to J/\psi K^0_S$. Analysis of time-dependent CP-violating asymmetries in $B^0 \to D^{(*)+}D^{(*)-}$ will be detailed and performed in the following chapter. Although only the $B^0 \to D^{(*)+}D^{(*)-}$ modes can be used for $\sin 2\beta$ measurement, $B^{\pm} \to D^{(*)\pm}\overline{D}^{(*)0}$ modes, such as $B^{\pm} \to D^{*\pm}D^{*0}$, can be used for measuring direct CP violating charge asymmetries, as well as for calibration and checks on resolutions and fits in the $\sin 2\beta$ analysis. $B^{\pm} \to D^{(*)\pm}\overline{D}^{(*)0}$ modes are also interesting due to the fact that they

have not been previously observed, nor are there any previous limits on their branching fractions. This chapter contains a summary of the measurement. For details of the techniques used, Appendix A of this dissertation may be consulted.

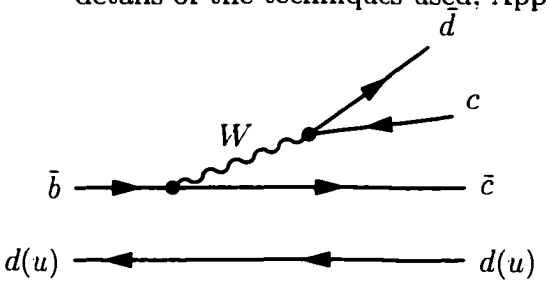


Figure 3.1: The main $B^0 \to D^{(*)+}D^{(*)-}$ ($B^+ \to D^{(*)+}\overline{D}^{(*)0}$) decay diagram.

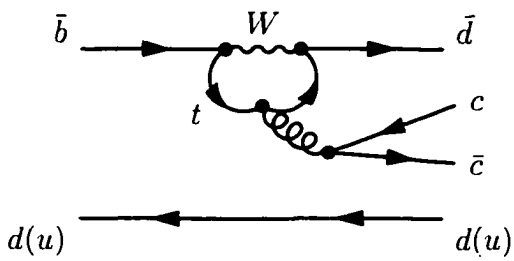


Figure 3.2: The color-suppressed penguin $B^0 \to D^{(*)+}D^{(*)-}$ ($B^+ \to D^{(*)+}\overline{D}^{(*)0}$) diagram.

The rate for the Cabibbo-suppressed decays $B \to D^{(*)}\overline{D}^{(*)(0)}$ (see figures 3.1 and 3.2) can be estimated from the measured rate of the Cabibbo-favored decays $B \to D_S^{(*)}\overline{D}^{(*)(0)}$:

$$\mathcal{B}(B \to D^{(*)}\overline{D}^{(*)(0)}) \approx \left(\frac{f_{D^{(*)}}}{f_{D_S^{(*)}}}\right) \tan^2\theta_C \mathcal{B}(B \to D_S^{(*)}\overline{D}^{(*)(0)}) \qquad (1)$$

From this it follows that the $B \to D^{(*)}\overline{D}^{(*)(0)}$ branching fractions are of the order of 0.1%.

Previous measurements of branching fractions and upper limits on these modes are summarized in Table 3.1.

Table 3.1: Summary of previous branching fraction and upper limit measurements [20]. Upper limits are quoted at the 90% confidence level.

Decay*	Branching Fraction ($\times 10^{-4}$)
$B^0 \to D^{*+}D^{*-}$	$9.9^{+4.2}_{-3.3}(stat) \pm 1.2(syst)$
$B^0 \to D^{*\pm}D^{\mp}$	< 6.3
$B^0 \to D^{+}D^{-}$	< 9.4

Of these modes, $B^0 \to D^{*+}D^{*-}$ contains the least background, due to the background rejection gained from the two slow charged pions. Although it cannot be described as a "golden mode" like $B^0 \to J/\psi K^0_S$, background is typically suppressed by nearly an order of magnitude compared with signal. $B^0 \to D^{*\pm}D^{\mp}$ and $B^{\pm} \to D^{*\pm}D^{*0}$ have more background, due to having only a single slow $\pi^{\pm}$; signal:background ratios for these modes are $\approx$ 1:1.

3.1.1 Data Sample

The data sample used represents an integrated luminosity of $22.9\,\mathrm{fb}^{-1}$, with $20.7\,\mathrm{fb}^{-1}$ collected on the $\Upsilon(4S)$ resonance. The total number of $B\overline{B}$ pairs produced in this sample was $N_{B\overline{B}} = (22.515 \pm 0.360) \times 10^6$.

*Note that $B^{\pm} \to D^{(*)\pm}\overline{D}^{(*)0}$ modes have neither previous branching fraction measurements nor limits.

3.2 Measurement of the $B^0 \to D^{*+}D^{*-}$ Branching Fraction

B^0 mesons are exclusively reconstructed by combining two charged D^* candidates reconstructed in a number of D^* and D decay modes. Events are pre-selected by requiring that there be three or more charged tracks and that the normalized second Fox-Wolfram moment [21] of the event be less than 0.6. We also require that the cosine of the angle between the reconstructed B direction and the thrust axis of the rest of the event be less than 0.9.

Charged kaon candidates are required to be inconsistent with the pion hypothesis, as inferred from the Cherenkov angle measured by the DIRC and the specific ionization measured by the SVT and DCH. No particle identification requirements are made for the kaon from the decay $D^0 \to K^-\pi^+$.

$K_S^0 \to \pi^+\pi^-$ candidates are required to have an invariant mass within 25 MeV/c^2 of the nominal K_S^0 mass [24]. The opening angle between the flight direction and the momentum vector of the K_S^0 candidate is required to be less than 200 mrad, and the transverse flight distance from the primary event vertex must be greater than 2 mm.

Neutral pion candidates are formed from pairs of photons in the EMC with energy above 30 MeV, an invariant mass within 20 MeV/c^2 of the nominal π^0 mass, and a summed energy greater than 200 MeV. A mass-constraint fit is then applied to these π^0 candidates. The π^0 from $D^{*+} \to D^+\pi^0$ decays ("soft" π^0), however, is required to have an invariant mass within 35 MeV/c^2 of the nominal π^0 mass and momentum in the $\Upsilon(4S)$ frame of $70 < p^* < 450$ MeV/c, with no requirement on the summed photon energies.

The decay modes of the D^0 and D^+ used in this analysis were selected by an optimization of $S^2/(S+B)$ (where S is the expected number of signal events and B is the expected number of background events) based on Monte Carlo simulations. The D^0 and D^+ modes used and their branching fractions are summarized in Table 3.2. D^0 (D^+) meson candidates are required to have an invariant mass within 20 MeV/c^2 of the nominal D^0 (D^+) mass.

Table 3.2: D^0 and D^+ decay modes and branching fractions [24]. The branching fraction for $K_S^0 \to \pi^+\pi^-$ is included for modes containing a K_S^0.

Decay Mode	Branching Fraction (%)
$D^0 \to K^-\pi^+$	3.83 ± 0.09
$D^0 \to K^-\pi^+\pi^0$	13.9 ± 0.9
$D^0 \to K^-\pi^+\pi^+\pi^-$	7.49 ± 0.31
$D^0 \to K_S^0\pi^+\pi^-$	1.85 ± 0.14
Total D^0 Branching Fraction	27.1
Decay Mode	Branching Fraction (%)
$D^+ \to K^-\pi^+\pi^+$	9.0 ± 0.6
$D^+ \to K_S^0\pi^+$	0.99 ± 0.09
$D^+ \to K^-K^+\pi^+$	0.87 ± 0.07
Total D^+ Branching Fraction	10.9

The D^{*+} mesons are reconstructed in their decays $D^{*+} \to D^0\pi^+$ and $D^{*+} \to D^+\pi^0$. We include for this analysis the decay combinations $D^{*+}D^{*-}$decaying

Table 3.3: D^* and D^{*0} decay modes and branching fractions [24]. D^{*0} is used for the $B^\pm \to D^{*\pm}D^{*0}$ analysis described in Section 3.3.

Particle	Decay Mode	Branching Fraction (%)
D^{*+}	$D^{*+} \to D^0\pi^+$	67.7 ± 0.5
	$D^{*+} \to D^+\pi^0$	30.7 ± 0.5
Total Visible D^{*+} Branching Fraction		98.4
D^{*0}	$D^{*0} \to D^0\pi^0$	61.9 ± 2.9
	$D^{*0} \to D^0\gamma$	38.1 ± 2.9
Total D^{*0} Branching Fraction		100.0

to ($D^0\pi^+$, $D^0\pi^-$) or ($D^0\pi^+$, $D^-\pi^0$), but not ($D^+\pi^0$,$D^-\pi^0$) due to the smaller branching fraction and larger expected backgrounds. The branching fractions for these modes are summarized in Table 3.3. D^0 and D^+ candidates are subjected to a mass-constraint fit and then combined with soft pion candidates. A vertex fit is performed that includes the position of the beam spot to improve the angular resolution of the soft pion.

To select B^0 candidates with well reconstructed D^* and D mesons, we construct a χ^2 that includes all measured D^* and D masses:

$$\chi^2_{Mass} = \left(\frac{m_D - m_{D_{PDG}}}{\sigma_{m_D}}\right)^2 + \left(\frac{m_{\bar{D}} - m_{\bar{D}_{PDG}}}{\sigma_{m_{\bar{D}}}}\right)^2 + \left(\frac{\Delta m_{D^*} - \Delta m_{D^*_{PDG}}}{\sigma_{\Delta m_{D^*}}}\right)^2 + \left(\frac{\Delta m_{\bar{D}^*} - \Delta m_{D^*_{PDG}}}{\sigma_{\Delta m_{D^*}}}\right)^2$$

where the subscript PDG refers to the nominal value, and Δm_{D^*} is the $D^* - D$ mass difference. For σ_{m_D} we use values computed for each D candidate, while for $\sigma_{\Delta m_{D^*}}$ we use fixed values of $0.83\,\mathrm{MeV}/c^2$ for $D^{*+} \to D^0\pi^+$ and $1.18\,\mathrm{MeV}/c^2$ for $D^{*+} \to D^+\pi^0$. A requirement that $\chi^2_{Mass} < 20$ is applied to all B^0 candidates. In events with more than one B^0 candidate, we choose the candidate with the lowest value of χ^2_{Mass}.

A B meson candidate is characterized by two kinematic variables. We use the energy-substituted mass, m_{ES}, defined as

$$m_{\mathrm{ES}} \equiv \sqrt{E^{*2}_{Beam} - {p^*_B}^2}$$

and the difference of the B candidate's energy from the beam energy, ΔE ,

$$\Delta E \equiv E^*_B - E^*_{Beam}$$

where E^*_B (p^*_B) are the energy (momentum) of the B candidate in the center-of-mass frame and E^*_{Beam} is one-half of the center-of-mass energy. The signal region in the ΔE *vs.* m_{ES} plane is defined to be $|\Delta E| < 25\,\mathrm{MeV}$ and $5.273 < m_{\mathrm{ES}} < 5.285\,\mathrm{GeV}/c^2$. The width of this region corresponds to approximately $\pm 2.5\sigma$ in both ΔE and m_{ES}.

These values on χ^2_{Mass}, m_{ES}, and ΔE were chosen based on an optimization of $S^2/(S+B)$, where (again) S is the expected number of signal events and B is the expected number of background events. The optimization process was done entirely with samples of signal and generic $B\overline{B}$ and $c\bar{c}$ Monte Carlo where the background distribution is taken from a sideband region, defined as

$$|\Delta E| < 200 \ \mathrm{MeV}$$

$$5.200 \ \mathrm{GeV}/c^2 < m_{\mathrm{ES}} < 5.26 \ \mathrm{GeV}/c^2$$

and

$$50 \text{ MeV} < |\Delta E| < 200 \text{ MeV}$$

$$5.26 < m_{\text{ES}} < 5.290 \text{ GeV}/c^2$$

These values were chosen based on a maximization of $S^2/(S+B)$ with a tendency towards looser cut values to reduce any possible systematic error incurred due to the differences in the reconstructed mass resolutions between data and Monte Carlo.

To determine the number of signal events in the signal region, we must estimate the expected contribution from background. This is done by scaling the number of events seen in the data sideband (defined above) with a scaling factor which gives a measure of the relative areas of the signal region to the sideband region. We parameterize the shape of the background in the ΔE *vs.* m_{ES} plane as the product of an ARGUS function [22] in m_{ES} and a first-order polynomial in ΔE . Based on this parameterization we estimate that the ratio of the number of background events in the signal region to the number in the sideband region is $(1.72 \pm 0.10) \times 10^{-2}$. The uncertainty is derived from the observed variation of this ratio under alternative assumptions for the background shape in m_{ES} and ΔE .

Figure 3.3 shows the events in the ΔE *vs.* m_{ES} plane after all selection criteria have been applied. The small box in the figure indicates the signal region defined above, and the sideband is the entire plane excluding the region bounded by the larger box outside the signal region. There are a total of 38 events located in the signal region, with 363 events in the sideband region. The latter, together with the effective ratio of areas of the signal region to the sideband region, implies an expected number of background events in the signal region of $6.24 \pm 0.33(stat) \pm$

0.36($syst$). The quoted systematic uncertainty comes from the background shape variation discussed previously. Figure 3.4 shows a projection of the data on the $m_{\rm ES}$ axis after requiring $|\Delta E| < 25$ MeV.

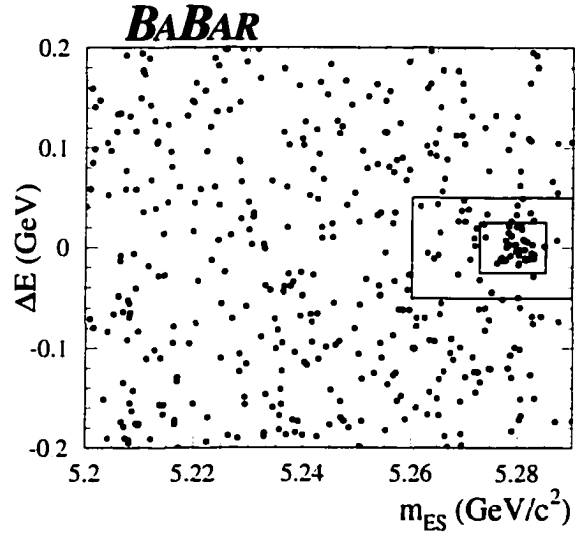


Figure 3.3: Distribution of $B^0 \to D^{*+}D^{*-}$ events in the ΔE *vs.* $m_{\rm ES}$ plane. The small box indicates the signal region, while the sideband region is everything outside the larger box.

A detailed Monte Carlo simulation of the *BABAR* detector is used to determine the efficiency for reconstructing the signal. This, together with the total number of $B\overline{B}$ pairs produced during data collection, allows for the determination of a branching fraction for $B^0 \to D^{*+}D^{*-}$:

$$\mathcal{B}(B^0 \to D^{*+}D^{*-}) = (8.0 \pm 1.6(stat) \pm 1.2(syst)) \times 10^{-4}$$

The dominant systematic uncertainty in this measurement comes from our level of understanding of the charged particle tracking efficiency (9.4%). The

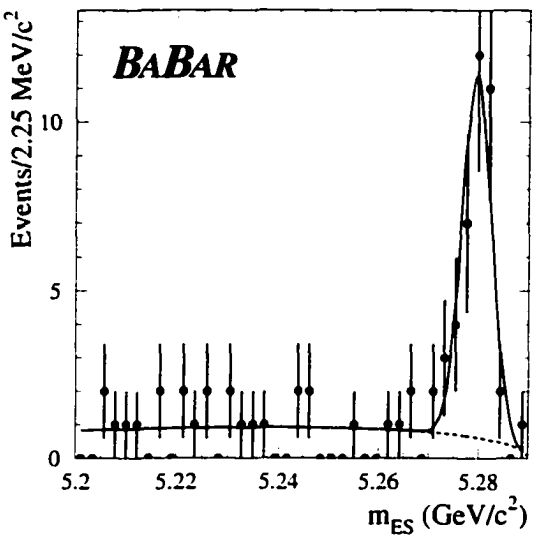


Figure 3.4: Distribution of $B^0 \to D^{*+}D^{*-}$ events in $m_{\rm ES}$ plane with a cut of $|\Delta E| < 25$ MeV applied. The curve represents a fit to the distribution of the sum of a Gaussian to model the signal and an ARGUS function [22] to model the background shape.

high charged particle multiplicity in this decay mode makes this measurement particularly sensitive to tracking efficiency. Systematic errors were assigned on a per track basis for π, K and slow π, and were added linearly due to large correlations. The imprecisely known partial-wave content of the $B^0 \to D^{*+}D^{*-}$ final state is another potential source of systematic biases. Monte Carlo events in each of the two extremes of transversity amplitudes $(A_{//}, \sqrt{2}A_0, A_\perp) = (1., 0., 0.)$ and $(0., 1., 0.)$ were generated and reconstructed [23]. Although both mixtures correspond to $R_t = 0$, the resulting p_t distributions of the slow pion represent the two extreme cases of possible p_t distributions. The change in the reconstruction efficiency of these final angular states is quoted as systematic error (6.6%). Other

significant systematic biases arise due to the uncertainties on the D^{*+}, D^0 and D^+ branching fractions (5.6%) and the differences in mass resolutions between Monte Carlo and data (4.1%). Possible contributions from peaking backgrounds was found to be negligable.The total systematic uncertainty from all sources is 14.5%.

3.3 Measurement of the Branching Fractions of $B^0 \to D^{*\pm}D^{\mp}$ and $B^{\pm} \to D^{*\pm}D^{*0}$

The decays $B^0 \to D^{*\pm}D^{\mp}$ and $B^{\pm} \to D^{*\pm}D^{*0}$ are studied following a method largely similar to that described in Sec. 3.2. Here, only those aspects of the analyses that differ significantly from that of the $B^0 \to D^{*+}D^{*-}$ analysis are discussed in some detail.

For $B^0 \to D^{*+}D^-$, B^0 mesons are exclusively reconstructed by combining a $D^{*\pm}$ and a $D^{\mp}$ candidate that are reconstructed in a number of $D^{*\pm}$ and $D^{\mp}$ decay modes. For $B^{\pm} \to D^{*\pm}D^{*0}$ the exclusive reconstruction combines a $D^{*\pm}$ and a D^{*0}. The kaon flavor of the D^{*0} is checked to make sure that a D^{*+} is paired only with a $\overline{D}^{*0}$ and a D^{*-} is only paired with a D^{*0}. The selection of $D^{\pm}$ and $D^{*\pm}$ candidates, and the K_S^0 and π^0 candidates that are used to compose them, is identical to that described for the $B^0 \to D^{*+}D^{*-}$ analysis.

The decay modes of the D and D^* used in these analyses are selected by an optimization of $S^2/(S+B)$ based on Monte Carlo simulations. $D^{*\pm}$ mesons are reconstructed in their decays $D^{*+} \to D^0\pi^+$ and $D^{*+} \to D^+\pi^0$, and D^{*0} mesons are reconstructed in their decays $D^{*0} \to D^0\pi^0$ and $D^{*0} \to D^0\gamma$. Modes used and their branching fractions are summarized in Tables A.3 and A.2. As in the

$B^0 \to D^{*+}D^{*-}$ analysis, we construct χ^2_{Mass} variables that include all measured $D^{*\pm}$, D^{*0}, and D masses. For $B^\pm \to D^{*\pm}D^{*0}$, χ^2_{Mass} contains 4 terms:

$$\chi^2_{Mass} = \left(\frac{m_D - m_{D_{PDG}}}{\sigma_{m_D}}\right)^2 + \left(\frac{m_{\bar{D}} - m_{\bar{D}_{PDG}}}{\sigma_{m_{\bar{D}}}}\right)^2 + \left(\frac{\Delta m_{D^*} - \Delta m_{D^*_{PDG}}}{\sigma_{\Delta m_{D^*}}}\right)^2 + \left(\frac{\Delta m_{D^{*0}} - \Delta m_{D^{*0}_{PDG}}}{\sigma_{\Delta m_{D^{*0}}}}\right)^2$$

For $B^0 \to D^{*+}D^-$, χ^2_{Mass} contains 3 terms:

$$\chi^2_{Mass} = \left(\frac{m_D - m_{D_{PDG}}}{\sigma_{m_D}}\right)^2 + \left(\frac{m_{D_{D^*}} - m_{D_{PDG}}}{\sigma_{m_{D_{D^*}}}}\right)^2 + \left(\frac{\Delta m_{D^*} - \Delta m_{D^*_{PDG}}}{\sigma_{\Delta m_{D^*}}}\right)^2$$

The major difference between these analyses and the $B^0 \to D^{*+}D^{*-}$ analysis is that the χ^2_{Mass} cut values for these analyses are set individually for each submode instead of having a global χ^2_{Mass} value for all submodes, to better take into account the fact that the amount of background is quite different for each of the different submodes in these analyses. The procedure used to optimize the χ^2_{Mass} value chooses the individual values of χ^2_{Mass} that together maximize the global $S^2/(S+B)$. Submodes which have an optimal value of less than 2 times the number of degrees of freedom of the χ^2_{Mass} (8 for $B^\pm \to D^{*\pm}D^{*0}$, 6 for $B^0 \to D^{*+}D^-$) are rejected on the grounds that the resulting sensitivity is poor, and that the tightness of the χ^2_{Mass} value makes such modes more susceptible to systematic errors in measured yield.

These values were tuned with samples of signal and generic $B\overline{B}$ and $c\bar{c}$ Monte Carlo where the background distribution is taken from a sideband region. For $B^\pm \to D^{*\pm}D^{*0}$, the sideband region is the same as that for $B^0 \to D^{*+}D^{*-}$. For $B^0 \to D^{*+}D^-$, the region where decays (such as $B^0 \to D^{*+}D^{*-}$) can feed down

into the ΔE -$m_{\rm ES}$ plane must be eliminated from the sideband. As $B^0 \to D^{*+}D^{*-}$ contains a reconstructed $D^{*\pm}$ and D, it is only separated from $B^0 \to D^{*+}D^-$ events due to the missing energy of the slow pion from the second $D^{*\pm}$. This missing energy manifests itelf as a negative shift in ΔE . These events accumulate in the area below the $B^0 \to D^{*+}D^-$ signal region, and in order to remove them, the region defined by:

$$\Delta E < -50\,\text{MeV}$$

$$m_{\rm ES} > 5.26\,\text{GeV}/c^2$$

is removed. The $m_{\rm ES}$ and ΔE distributions for events reconstructed in the channels $B^0 \to D^{*+}D^-$ and $B^\pm \to D^{*\pm}D^{*0}$ are shown in Fig. 3.5 and in Fig. 3.6.

For the $B^0 \to D^{*+}D^-$ channel we reconstruct a total of 31 events, of which $10.5 \pm 1.7(stat)$ are background (see Figure 3.5) The probability that the visible signal is a statistical fluctuation of the background is 9.7×10^{-7} ($> 4.3\sigma$). For $B^\pm \to D^{*\pm}D^{*0}$ we reconstruct a total of 39 events, of which $20.3 \pm 0.5(stat)$ are background (see Figure 3.6). The probability that the visible signal is a statistical fluctuation of the background is 2.9×10^{-6} ($> 4.1\sigma$).

Detailed Monte Carlo simulation of the *BABAR* detector is used to determine the efficiency for reconstructing $B^0 \to D^{*\pm}D^\mp$ and $B^\pm \to D^{*\pm}D^{*0}$. This, together with the total number of $B\overline{B}$ pairs produced during data collection, allows for the determination of branching fractions for $B^0 \to D^{*\pm}D^\mp$ and $B^\pm \to D^{*\pm}D^{*0}$:

$$\mathcal{B}(B^0 \to D^{*\pm}D^\mp) = (6.7^{+2.0}_{-1.7}(stat) \pm 1.1(syst)) \times 10^{-4}$$

$$\mathcal{B}(B^\pm \to D^{*\pm}D^{*0}) = (10.5^{+3.3}_{-2.8}(stat) \pm 2.0(syst)) \times 10^{-4}$$

As with $B^0 \to D^{*+}D^{*-}$, a significant systematic uncertainty in these measurements comes from our level of understanding of the charged particle tracking

efficiency (9.4% for $B^0 \to D^{*\pm}D^{\mp}$ and 8.9% for $B^{\pm} \to D^{*\pm}D^{*0}$). The imprecisely known partial-wave content of the $B^{\pm} \to D^{*\pm}D^{*0}$ final state is another source of potential systematic biases. The systematic for this (6.6%) was taken to be the same as for $B^0 \to D^{*+}D^{*-}$ due to the fact that the two decays should have the same partial wave content if final state interactions are neglected. Other significant systematic biases arise due to the uncertainties on the D^{*+}, D^{*0}, D^0, and D^+ branching fractions (7.4% for $B^0 \to D^{*\pm}D^{\mp}$ and 9.1% for $B^{\pm} \to D^{*\pm}D^{*0}$) and the differences in mass resolutions between Monte Carlo and data (4.1% for both $B^0 \to D^{*\pm}D^{\mp}$ and $B^{\pm} \to D^{*\pm}D^{*0}$). The total systematic uncertainty from all sources is 17.1% for $B^0 \to D^{*\pm}D^{\mp}$ and 18.7% for $B^{\pm} \to D^{*\pm}D^{*0}$.

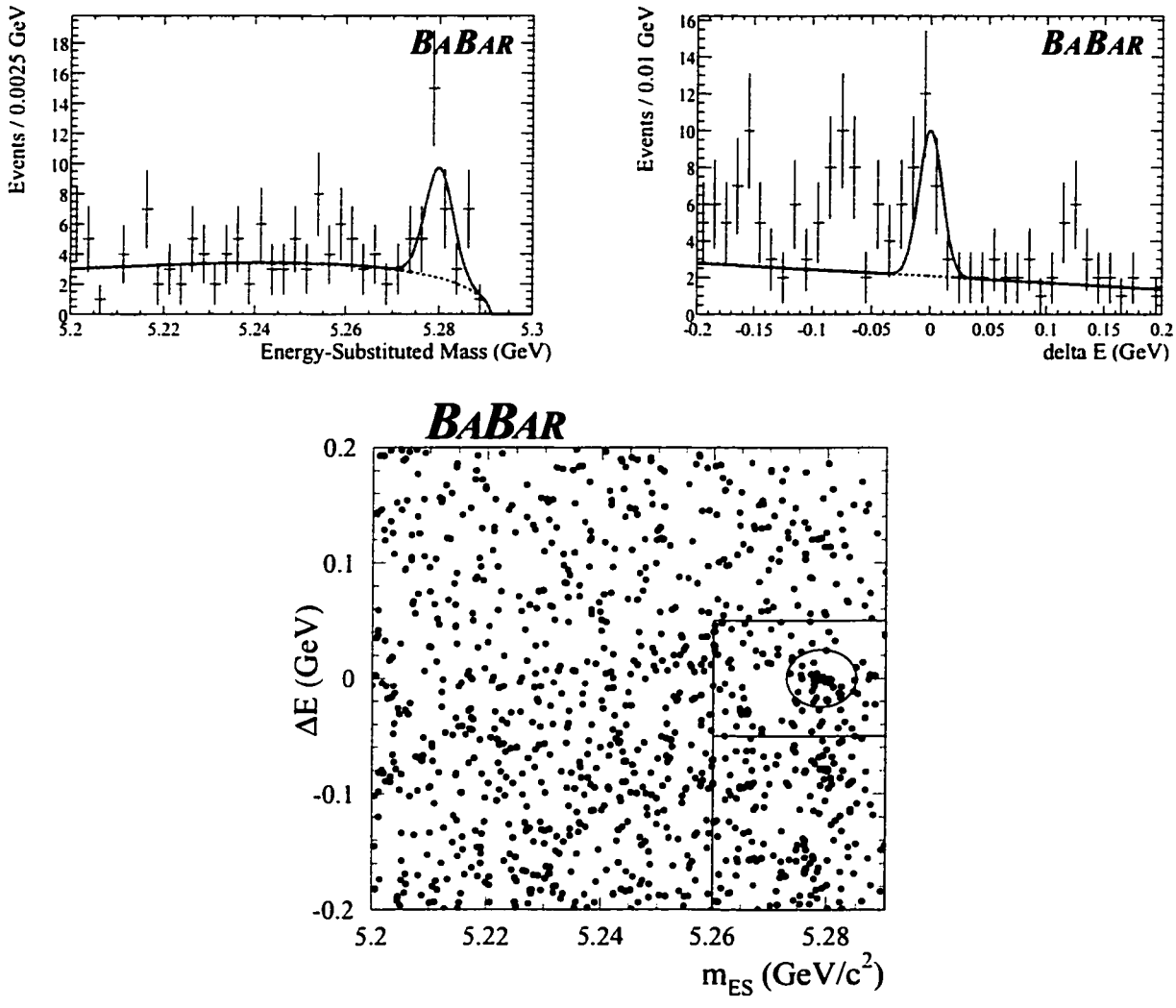


Figure 3.5: Top Left: $m_{\rm ES}$ projection of the $B^0 \to D^{*+}D^-$ event population, in the ΔE signal band ($-0.025 < \Delta E < 0.025\,$GeV). The crosses are the data. The dashed line represents the extrapolation, to the ΔE signal band, of a two-dimensional background fit in the ΔE -$m_{\rm ES}$ sidebands. The solid line is the sum of this background extrapolation and of a fitted, Gaussian-shaped signal centered on the B mass. Top Right: ΔE projection of the $B^0 \to D^{*+}D^-$ event population, in the $m_{\rm ES}$ signal band ($5.273 < m_{\rm ES} < 5.285\,$GeV/$c^2$). The crosses are the data. The dashed line represents the extrapolation, to the $m_{\rm ES}$ signal band, of the above-mentioned background fit. The solid line is the sum of this background extrapolation, and of a gaussian-shaped signal centered on $\Delta E = 0$. In the case of both the upper plots, the points in the feed-down region ($\Delta E < -0.05\,$GeV) are excluded from the fits. Bottom: Two-dimensional distribution of the $B^0 \to D^{*+}D^-$ events in the ΔE *vs.* $m_{\rm ES}$ plane. The small ellipse indicates the signal region, while the sideband region is everything that is outside the box that surrounds the signal region and also outside the (feed-down) box below the signal region.

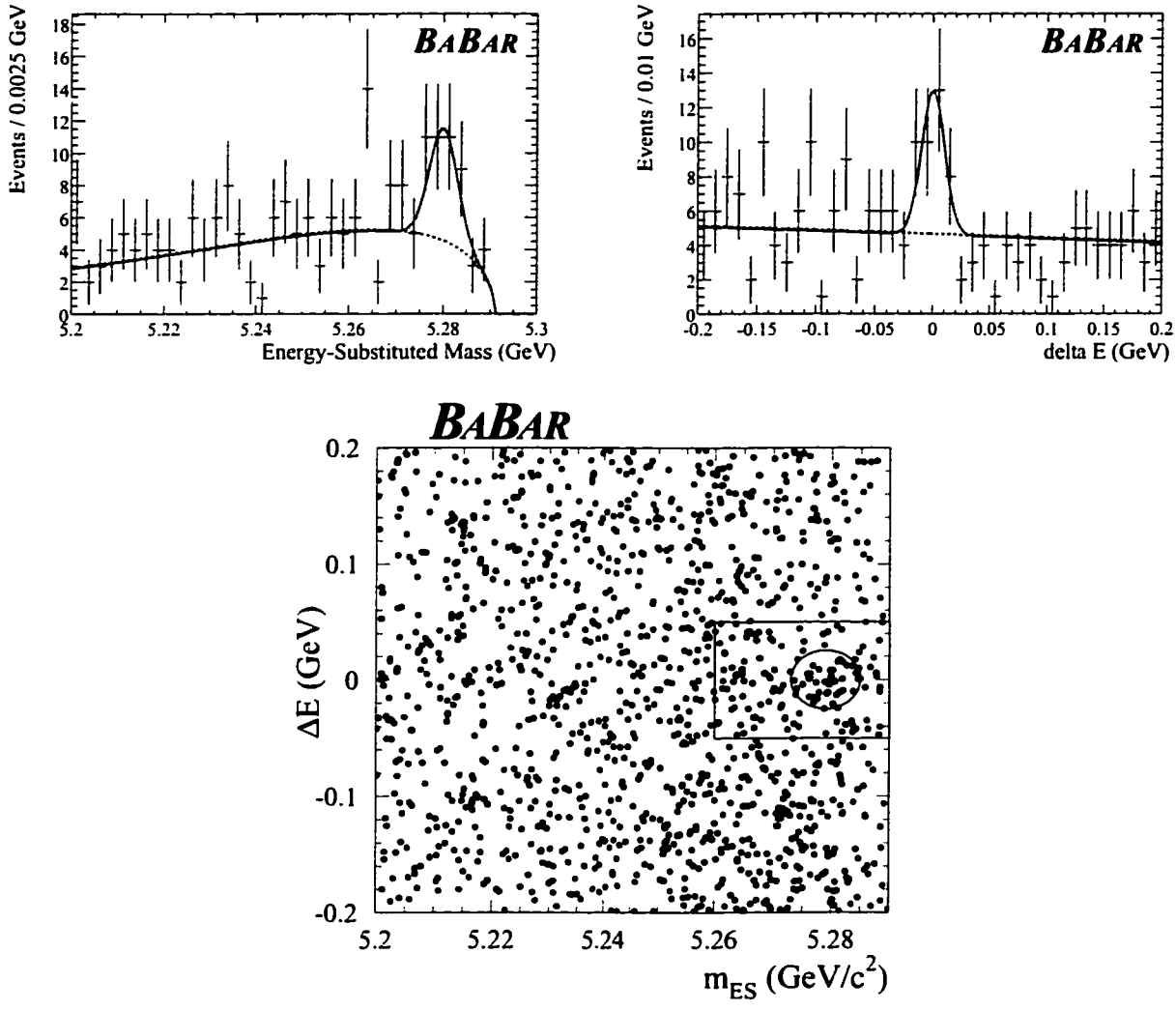


Figure 3.6: Top Left: $m_{\rm ES}$ projection of the $B^{\pm} \to D^{*\pm}D^{*0}$ event population, in the ΔE signal band ($-0.025 < \Delta E < 0.025\,$GeV). The crosses are the data. The dashed line represents the extrapolation, to the ΔE signal band, of a two-dimensional background fit in the ΔE -$m_{\rm ES}$ sidebands. The solid line is the sum of this background extrapolation and of a fitted, Gaussian-shaped signal centered on the B mass. Top Right: ΔE projection of the $B^{\pm} \to D^{*\pm}D^{*0}$ event population, in the $m_{\rm ES}$ signal band ($5.273 < m_{\rm ES} < 5.285\,$GeV/$c^2$). The crosses are the data. The dashed line represents the extrapolation, to the $m_{\rm ES}$ signal band, of the above-mentioned background fit. The solid line is the sum of this background extrapolation, and of a Gaussian-shaped signal centered on $\Delta E = 0$. Bottom: Two-dimensional distribution of the $B^{\pm} \to D^{*\pm}D^{*0}$ events in the ΔE *vs.* $m_{\rm ES}$ plane. The small ellipse indicates the signal region, while the sideband region is everything that is outside the box that surrounds the signal region.

Chapter 4

Measurement of $\sin 2\beta$ in Exclusive $b \to c\bar{c}d$ Channels

4.1 $B^0 \to D^{*+}D^{*-}$: Fit Method

4.1.1 $B^0 \to D^{*+}D^{*-}$: Measurement of CP asymmetry

The D^{*+} D^{*-} final state is a vector-vector mode, with different partial waves contributing different CP parities. Therefore, time-dependent CP asymmetry measurements are diluted, unless a combined time and angular analysis is performed in order to separate the CP-even and CP-odd components.

Moreover, although factorization models predict a rather small penguin contamination of the weak phase $\Im(\lambda_f) = -\sin 2\beta$, a sizeable penguin diagram contribution cannot a priori be excluded, leading to different $\lambda_f = \frac{q}{p}\frac{\bar{A}(f)}{A(f)}$ values for the three transversity amplitudes, because of different penguin to tree ratios.

In the following two subsections the time-dependent angular distributions in the transversity basis are first recalled, for the limiting case of negligible penguin contribution; a parametrization of the penguin effect is then given for the time-dependent transversity distribution, that is used to fit the CP asymmetry in our D^{*+} D^{*-} sample.

Time dependent angular distributions

From the angular distribution in the transversity basis [38], with implicitely time-dependent transversity amplitudes $A_0, A_\parallel, A_\perp$, we can write the time dependent angular distribution as:

$$\begin{aligned}
\frac{1}{\Gamma}\frac{d^3\Gamma}{d\cos\theta_l d\cos\theta_{tr} d\phi_{tr}} = \quad & \frac{9}{32\pi}\frac{1}{|A_0|^2+|A_\parallel|^2+|A_\perp|^2} \\
& \{4|A_0|^2\cos^2\theta_l\sin^2\theta_{tr}\cos^2\phi_{tr} \\
& +2|A_\parallel|^2\sin^2\theta_l\sin^2\theta_{tr}\sin^2\phi_{tr} \\
& +2|A_\perp|^2\sin^2\theta_l\cos^2\theta_{tr} \\
& +\sqrt{2}\Re(A_\parallel^* A_0)\sin 2\theta_l\sin^2\theta_{tr}\sin 2\phi_{tr} \\
& -\sqrt{2}\Im(A_0^* A_\perp)\sin 2\theta_l\sin 2\theta_{tr}\cos\phi_{tr} \\
& -2\Im(A_\parallel^* A_\perp)\sin^2\theta_l\sin 2\theta_{tr}\sin\phi_{tr}\} \ . \qquad (1)
\end{aligned}$$

Integration over $d\cos\theta_l$ and $d\phi_{tr}$ [38] yields:

$$\frac{1}{\Gamma}\frac{d\Gamma(\overset{(-)}{B^0}\to D^{*+}D^{*-})}{d\cos\theta_{tr}} = \frac{9}{32\pi}\{(|\overset{(-)}{A_\parallel}|^2+|\overset{(-)}{A_0}|^2)G_+(\theta_{tr})+|\overset{(-)}{A_\perp}|^2 G_-(\theta_{tr})\} \quad (2)$$

where the CP even and CP odd angular terms are:

$$G_+(\theta_{tr}) = \frac{8\pi}{3}\sin^2\theta_{tr}$$

$$G_-(\theta_{\rm tr}) \;=\; \frac{16\pi}{3}\cos^2\theta_{\rm tr}. \tag{3}$$

Inserting the explicit time dependence of amplitudes:

$$\begin{aligned} A_{0,\|}(t) &= A^0_{0,\|}e^{-imt}\, e^{-\Gamma t/2}\left(\cos\frac{\Delta m t}{2} + i\lambda\sin\frac{\Delta m t}{2}\right) \\ A_\perp(t) &= A^0_\perp e^{-imt}\, e^{-\Gamma t/2}\left(\cos\frac{\Delta m t}{2} - i\lambda\sin\frac{\Delta m t}{2}\right) \end{aligned} \tag{4}$$

and:

$$\begin{aligned} \bar{A}_{0,\|}(t) &= \bar{A}^0_{0,\|}e^{-imt}\, e^{-\Gamma t/2}\left(\cos\frac{\Delta m t}{2} + \frac{i}{\lambda}\sin\frac{\Delta m t}{2}\right) \\ \bar{A}_\perp(t) &= -\bar{A}_\perp{}^0 e^{-imt}\, e^{-\Gamma t/2}\left(\cos\frac{\Delta m t}{2} - \frac{i}{\lambda}\sin\frac{\Delta m t}{2}\right) \end{aligned} \tag{5}$$

under the assumption of equal $\lambda = \frac{q}{p}\frac{\bar{A}}{A}$, that is assuming negligible penguin contributions, and $|\frac{q}{p}| = 1$, so that $|\bar{A}^0|^2 = |\lambda|^2|A^0|^2$, the decay rate can be written in a form convenient for comparisons with the general case, discussed in the following subsection:

$$\begin{aligned} \frac{1}{\Gamma}\frac{d\Gamma(\overset{(-)}{B^0}\to D^{*+}D^{*-})}{d\cos\theta_{\rm tr}} \;=\; & \frac{9}{32\pi}e^{-\Gamma t}\{ \\ & [(|A^0_\||^2 + |A^0_0|^2)G_+(\theta_{\rm tr}) + |A^0_\perp|^2 G_-(\theta_{\rm tr})] \\ \overset{(-)}{+}\ \tfrac{(1-|\lambda|^2)}{(1+|\lambda|^2)} & [(|A^0_\||^2 + |A^0_0|^2)G_+(\theta_{\rm tr}) + |A^0_\perp|^2 G_-(\theta_{\rm tr})]\cos\Delta m t \\ \overset{(+)}{-}\ \tfrac{2\Im(\lambda)}{(1+|\lambda|^2)} & [(|A^0_\||^2 + |A^0_0|^2)G_+(\theta_{\rm tr}) - |A^0_\perp|^2 G_-(\theta_{\rm tr})]\sin\Delta m t\} \end{aligned} \tag{6}$$

where terms are grouped by CP parity content and time dependence, and $\eta_{CP} = +1$ for A_0, $A_\|$. The amplitudes are normalized so that:

$$|A^0_0|^2 + |A^0_\||^2 + |A^0_\perp|^2 = 1 \tag{7}$$

The usual CP dilution factor K is defined as:

$$\begin{aligned} K &= \frac{|A_0^0|^2 + |A_\parallel^0|^2 - |A_\perp^0|^2}{|A_0^0|^2 + |A_\parallel^0|^2 + |A_\perp^0|^2} \\ |A_0^0|^2 + |A_\parallel^0|^2 &= \frac{1}{2}(1+K) \\ |A_\perp^0|^2 &= \frac{1}{2}(1-K). \end{aligned} \tag{8}$$

Parametrization of penguin contributions

If the possibility of penguin contributions, leading to different λ_0, $\lambda_\parallel$, $\lambda_\perp$, is taken into account, then the time-dependent decay amplitudes (eq. 4) can be rewritten as:

$$\begin{aligned} A_{0,\parallel}(t) &= \frac{A_{0,\parallel}^0}{\sqrt{(1+|\lambda_{0,\parallel}|^2)}} e^{-imt}\, e^{-\Gamma t/2} \left(\cos\frac{\Delta m t}{2} + i\lambda_{0,\parallel} \sin\frac{\Delta m t}{2}\right) \\ A_\perp(t) &= \frac{A_\perp^0}{\sqrt{(1+|\lambda_\perp|^2)}} e^{-imt}\, e^{-\Gamma t/2} \left(\cos\frac{\Delta m t}{2} - i\lambda_\perp \sin\frac{\Delta m t}{2}\right). \end{aligned} \tag{9}$$

As a result, the time- and transversity-dependent decay rate (6) can be rewritten in a similar form, where however a common value of $|\lambda|$ or $\Im(\lambda)$ can no longer be factorized:

$$\begin{aligned} \frac{1}{\Gamma}\frac{\mathrm{d}\Gamma(\overset{(-)}{B^0} \to D^{*+}D^{*-})}{\mathrm{d}\cos\theta_{\mathrm{tr}}} &= \frac{9}{32\pi} e^{-\Gamma t} \{ \\ & [O_+ G_+(\theta_{\mathrm{tr}}) + O_- G_-(\theta_{\mathrm{tr}})] \\ & \overset{(-)}{+} [C_+ G_+(\theta_{\mathrm{tr}}) + C_- G_-(\theta_{\mathrm{tr}})] \cos\Delta m t \\ & \overset{(+)}{-} [S_+ G_+(\theta_{\mathrm{tr}}) - S_- G_-(\theta_{\mathrm{tr}})] \sin\Delta m t \} \end{aligned} \tag{10}$$

where the angular dependence of $G_+(\theta_{\mathrm{tr}})$ and $G_-(\theta_{\mathrm{tr}})$ was defined in eqn. 3, and the six coefficients O_+, O_-, C_+, C_-, S_+ and S_- are given by:

$$O_+ = |A_\parallel^0|^2 + |A_0^0|^2$$

$$
\begin{aligned}
O_- &= |A_\perp^0|^2 \\
C_+ &= \frac{(1-|\lambda_\parallel|^2)}{(1+|\lambda_\parallel|^2)}|A_\parallel^0|^2 + \frac{(1-|\lambda_0|^2)}{(1+|\lambda_0|^2)}|A_0^0|^2 \\
C_- &= \frac{(1-|\lambda_\perp|^2)}{(1+|\lambda_\perp|^2)}|A_\perp^0|^2 \\
S_+ &= \frac{2\Im(\lambda_\parallel)}{(1+|\lambda_\parallel|^2)}|A_\parallel^0|^2 + \frac{2\Im(\lambda_0|)}{(1+|\lambda_0|^2)}|A_0^0|^2 \\
S_- &= \frac{2\Im(\lambda_\perp)}{(1+|\lambda_\perp|^2)}|A_\perp^0|^2. \qquad (11)
\end{aligned}
$$

For the initial measurement of CP asymmetry in $B^0 \to D^{*+}D^{*-}$ different fitting strategies are possible and will be discussed in more detail in subsection 4.1.2. In particular, one can perform:

1) a two-parameter fit of the coefficients in front of $\cos\Delta mt$ and $\sin\Delta mt$ terms in eq. 10, after integration over $\cos(\theta_{\rm tr})$;

2) a five-parameter fit of the $O_i, C_i, S_i, (i = +, -)$ terms in eq 10. The normalization of amplitudes (eq. 7) reduce the number of independent parameters to 5 in this case.

Neither of these two strategies requires any assumption on the presence or magnitude of penguin contibutions.

In the first approach a smaller number of free parameters is fitted, but, due to the fact that the angular information is not used, the $\sin\Delta mt$ coefficient is diluted by the mixture of CP even and CP odd contributions. It should be noted that the direct CP asymmetry ($\cos\Delta mt$) term is not diluted. In this note this first approach is studied in detail.

In the second strategy angular information is used, to fit five free parameters. If we define for the CP even amplitudes "average values"

$$\langle\frac{(1-|\lambda_+|^2)}{(1+|\lambda_+|^2)}\rangle \qquad \text{and} \qquad \langle\frac{\Im(\lambda_+)}{(1+|\lambda_+|^2)}\rangle$$

as:

$$\begin{aligned}
\langle\frac{(1-|\lambda_+|^2)}{(1+|\lambda_+|^2)}\rangle &= \frac{\frac{(1-|\lambda_\parallel|^2)}{(1+|\lambda_\parallel|^2)}|A_\parallel^0|^2+\frac{(1-|\lambda_0|^2)}{(1+|\lambda_0|^2)}|A_0^0|^2}{|A_\parallel^0|^2+|A_0^0|^2} \\
\langle\frac{\Im(\lambda_+)}{(1+|\lambda_+|^2)}\rangle &= \frac{\frac{\Im(\lambda_\parallel)}{(1+|\lambda_\parallel|^2)}|A_\parallel^0|^2+\frac{\Im(\lambda_0)}{(1+|\lambda_0|^2)}|A_0^0|^2}{|A_\parallel^0|^2+|A_0^0|^2}, \qquad (12)
\end{aligned}$$

equation (10) can be rewritten in terms of the five relevant physical quantities $|\lambda_+|$, $\Im(\lambda_+)$, $|\lambda_\perp|$, $\Im(\lambda_\perp)$ and K, as:

$$\begin{aligned}
\frac{1}{\Gamma}\frac{\mathrm{d}\Gamma(\overset{(-)}{B^0}\to D^{*+}D^{*-})}{\mathrm{d}\cos\theta_{\mathrm{tr}}} &= \frac{9}{32\pi}e^{-\Gamma t}\{ \\
&\quad [\frac{1}{2}(1+K)G_+(\theta_{\mathrm{tr}})+\frac{1}{2}(1-K)G_-(\theta_{\mathrm{tr}})] \\
&\overset{(-)}{+} [\frac{1}{2}(1+K)\langle\frac{(1-|\lambda_+|^2)}{(1+|\lambda_+|^2)}\rangle G_+(\theta_{\mathrm{tr}}) \\
&\quad +\frac{1}{2}(1-K)\frac{(1-|\lambda_\perp|^2)}{(1+|\lambda_\perp|^2)}G_-(\theta_{\mathrm{tr}})]\cos\Delta mt \\
&\overset{(+)}{-} [\frac{1}{2}(1+K)\langle\frac{2\Im(\lambda_+)}{(1+|\lambda_+|^2)}\rangle G_+(\theta_{\mathrm{tr}}) \\
&\quad -\frac{1}{2}(1-K)\frac{2\Im(\lambda_\perp)}{(1+|\lambda_+|^2)}G_-(\theta_{\mathrm{tr}})]\sin\Delta mt.\}
\end{aligned} \qquad (13)$$

In the limit of negligible penguin contributions, and taking into account the definition (8) of the dilution factor K, one recovers eq.(6).

4.1.2 $B^0 \to D^{*+}D^{*-}$: likelihood fit method

The CP asymmetry measurement is based on a maximum likelihood fit. The probability density function (PDF) is given by the angle and time dependent distribution (eqn. 13) discussed in the previous subsections, combined with a description of incorrect B flavor tags, of the experimental time resolution, and of the relevant background properties that are discussed in the following subsections.

Mistag modelling, time resolution and background

The strategy used to model the mistag and time resolution and to describe the background is the same as in the $\sin 2\beta$ measurement with $J/\ \psi K_s$ [41].

In this section a summary of the relevant equations is given in the case of the fit of the coefficients of the $\sin \Delta m \Delta t$ and $\cos \Delta m \Delta t$ terms.

The same model is adopted for the formally more complicated case of a five-parameter fit that uses also the angular information. For the present analysis the angular resolution effects are considered as negligible, and the corresponding resolution function is not included in the pdf.

The decay distributions for a B_{tag} tagged as a B^0 (+ sign) or a $\bar{B}^0$ (− sign) can be written as :

$$\begin{aligned} f_{\pm}(\Delta t_{\text{true}}) &= \frac{\Gamma}{2} e^{-\Gamma|\Delta t_{\text{true}}|} \left\{ (1 - \tfrac{1}{2}\Delta\mathcal{D}) \right. \\ &\mp \left. \langle \mathcal{D} \rangle \ [S \sin(\Delta m_d \Delta t_{\text{true}}) + C \cos(\Delta m_d \Delta t_{\text{true}})] \right\} \end{aligned} \quad (14)$$

where the C and S coefficients have the physical meaning:

$$C = (|A_{\|}^0|^2 + |A_0^0|^2) \langle \frac{(1 - |\lambda_+|^2)}{(1 + |\lambda_+|^2)} \rangle + |A_{\perp}^0|^2 \frac{(1 - |\lambda_{\perp}|^2)}{(1 + |\lambda_{\perp}|^2)}$$

$$S = (|A_{\parallel}^0|^2 + |A_0^0|^2)\langle\frac{2\Im(\lambda_+)}{(1+|\lambda_+|^2)}\rangle - |A_{\perp}^0|^2\frac{2\Im(\lambda_{\perp})}{(1+|\lambda_+|^2)}. \quad (15)$$

For the dilution factor due to incorrect B flavor tags $\mathcal{D} = 1 - 2w$, where w is the mistag rate, the formalism that takes into account any possible mistag differences $\Delta\mathcal{D}$ between B^0 and $\overline{B}^0$ is used. Due to the fact that the probabilities of mistagging a B^0 or $\overline{B}^0$ may differ slightly, separate mistag dilutions $\mathcal{D} = 1 - 2w$ and $\bar{\mathcal{D}} = 1 - 2\bar{w}$ are fitted for B^0 and $\overline{B}^0$, in the same manner as the charmonium $\sin 2\beta$ fit [41].

To account for the finite resolution of the detector, the time-dependent distributions $f_{\pm}$ for B^0 and $\overline{B}^0$ tagged events are convoluted with a time resolution function

$$\mathcal{R}(\delta_t = \Delta t - \Delta t_{\text{true}}; \hat{a})$$

to yield the experimentally measured decay time PDF

$$\mathcal{F}_{\pm}(\Delta t; \Gamma, \Delta m_d, \mathcal{D}, C, S, \hat{a}) = f_{\pm}(\Delta t; \Gamma, \Delta m_d, \mathcal{D}, C, S) \otimes \mathcal{R}(\delta_t; \hat{a}), \quad (16)$$

where $\hat{a}$ represents the set of parameters $(f_{\text{core,tail}}, \sigma_{\text{core,tail}}, \delta_{\text{core,tail}}, f_{\text{outlier}}, \sigma_{\text{outlier}})$ that describe the resolution function:

$$\mathcal{R}(\delta_t; \hat{a}) = \sum_{k=1}^{2} \frac{f_{\text{core,tail}}}{\sigma_{\text{core,tail}}\sqrt{2\pi}} \exp\left(-\frac{(\delta_t - \delta_{\text{core,tail}})^2}{2\sigma_{\text{core,tail}}}\right) + \frac{f_{\text{outlier}}}{\sigma_{\text{outlier}}\sqrt{2\pi}} \exp\left(-\frac{\delta_t^2}{2\sigma_{\text{outlier}}{}^2}\right). \quad (17)$$

The default resolution function is a triple Gaussian model, where two Gaussians (core and tail) have a width proportional to the event by event error and the third (outlier) has a width fixed to 8 ps. The bias of the outlier

Gaussian is fixed to 0, while the core and tail biases scale with the event by event error.

Events are separated into four different tagging categories (see Section 2.3), each of which has a different mean mistag fraction, w_i, determined individually for each category [39]. The value of the free parameters C and S are extracted using the tagged $B^0 \to D^{*+}D^{*-}$ sample by maximizing the likelihood function

$$\begin{aligned} \ln \mathcal{L}_{CP} &= \sum_{i}^{\text{tagging}} \left[\sum_{B^0\text{tag}} \ln \mathcal{F}_+(\Delta t; \Gamma, \Delta m_d, \hat{a}, \mathcal{D}, C, S) \right. \\ &+ \left. \sum_{\bar{B}^0\text{tag}} \ln \mathcal{F}_-(\Delta t; \Gamma, \Delta m_d, \hat{a}, \mathcal{D}, C, S) \right], \end{aligned} \tag{18}$$

where the outer summation is over tagging categories i. The fit for C and S will be performed on the combined flavor-eigenstate and $B^0 \to D^{*+}D^{*-}$ samples, in order to determine the parameters, the mistag fractions for each tagging category, and the vertex resolution parameters $\hat{a}_i$. Additional terms are included in the likelihood to account for backgrounds and their time dependence.

Background modelling

The background parametrizations can be allowed to differ for each tagging category. Each event belongs to a particular tagging category, i. In addition, the event is classified by whether B_{tag} was a B^0 $(+)$ or a $\bar{B}^0$ $(-)$. The distribution, specified for each possibility $(+/-, i)$ can be written as:

$$\begin{aligned} \mathcal{F}_{\pm,i} &= f^{CP}_{i,\text{sig}} \mathcal{F}_{\pm}(\Delta t; \Gamma, \Delta m_d, w_i, C, S, \hat{a}_i) \\ &+ f^{CP}_{i,\text{peak}} \mathcal{B}^{CP}_{\pm,i,\text{peak}}(\Delta t; \hat{a}_i) \end{aligned}$$

$$+ \sum_{\beta=\text{bkgd}} f^{CP}_{i,\beta} \mathcal{B}^{CP}_{\pm,i,\beta}(\Delta t; \hat{b}_i). \tag{19}$$

where the $f_{i,\beta}$ are the fractions of background events for each source and tagging category as a function of m_{ES}. The peaking and combinatorial background PDFs, $\mathcal{B}_{\pm,i,\text{peak}}$ and $\mathcal{B}_{\pm,i,\beta}$, provide an empirical description the Δt distribution of the background events in the sample, including resolution functions (see the previous subsection) parametrized by $\hat{a}_i$ and $\hat{b}_i$ respectively. The probability that a B^0 candidate is a signal or a background event is determined from a separate fit to the observed m_{ES} distributions, describe as the sum of a single Gaussian distribution for the signal plus an ARGUS parametrization for the background.

Based on this fit, we find that the event-by-event signal and background probabilities appear as the relative weights for the various signal and background terms in eq. (19).

Backgrounds arise from many different sources. Rather than attempting to determine the various physics contributions, we make for an empirical description in the likelihood fit, allowing for background components with various time dependences.

The possible background contributions considered are

$$\begin{aligned}
\mathcal{B}^{CP}_{\pm,i,1} &= (1/2)\delta(\Delta t) \otimes \mathcal{R}(\delta_t; \hat{b}_i), \\
\mathcal{B}^{CP}_{\pm,i,2} &= (\Gamma_{i,2}/4)(1 \pm \mathcal{D}^{CP}_{i,2} \sin \Delta m_d \Delta t)\, \mathrm{e}^{-\Gamma^{CP}_{i,2}|\Delta t|} \otimes \mathcal{R}(\delta_t; \hat{b}_i), \\
\mathcal{B}^{CP}_{\pm,i,\text{peak}} &= (\Gamma_{i,\text{peak}}/4)(1 \pm \mathcal{D}^{CP}_{i,\text{peak}} \sin \Delta m_d \Delta t)\, \mathrm{e}^{-\Gamma^{CP}_{\text{peak}}|\Delta t|} \otimes \mathcal{R}(\delta_t; \hat{a}_i), \quad (20)
\end{aligned}$$

corresponding to prompt and CP background components as well as a peaking contribution. The background resolution function parameters $\hat{b}_i$ are common with the background resolution function of the B_{flav} sample. The like-

lihood fit includes as free parameters the fraction of each time component, as well as apparent lifetimes, resolutions and dilutions that best describe the events with high weights for being background [39].

The actual use of these options in the present analysis is described in the sections (4.4, 4.5, 4.6) devoted to results and systematics.

Angular acceptance

A study of the angular acceptance effects, mainly due to the efficiency threshold in the detection of slow pions, was performed at generator level [38] for the measurement of the odd CP fraction $R_t = \frac{1}{2}(1 - K)$ (where K was defined in eqn. 13). For the present CP asymmetry analysis, a toy Monte Carlo study showed that a simultaneous fit of the angular and time dependence is able to resolve the effective content of the even and odd CP components in the selected sample, reproducing the input $\Im(\lambda)$ value, independently of the angular acceptance, that only affects the "effective" fitted R_t value.

4.2 $B^0 \to D^{*\pm}D^{\mp}$: Fit Method

4.2.1 $B^0 \to D^{*\pm}D^{\mp}$: Measurement of CP asymmetry

D^*D is not (*a priori*) a CP eigenstate. The two $B^0 \to D^{*\pm}D^{\mp}$ final states, $D^{*+}D^-$ and $D^{*-}D^+$, are linear combinations of the CP eigenstates

$$|D^{*+}D^- + D^{*-}D^+\rangle \qquad \text{and} \qquad |D^{*+}D^- - D^{*-}D^+\rangle.$$

The CP-violating parts of the PDFs for $B^0 \to D^{*+}D^-$ and $B^0 \to D^{*-}D^+$ are thus linearly coupled and may be solved simultaneously.

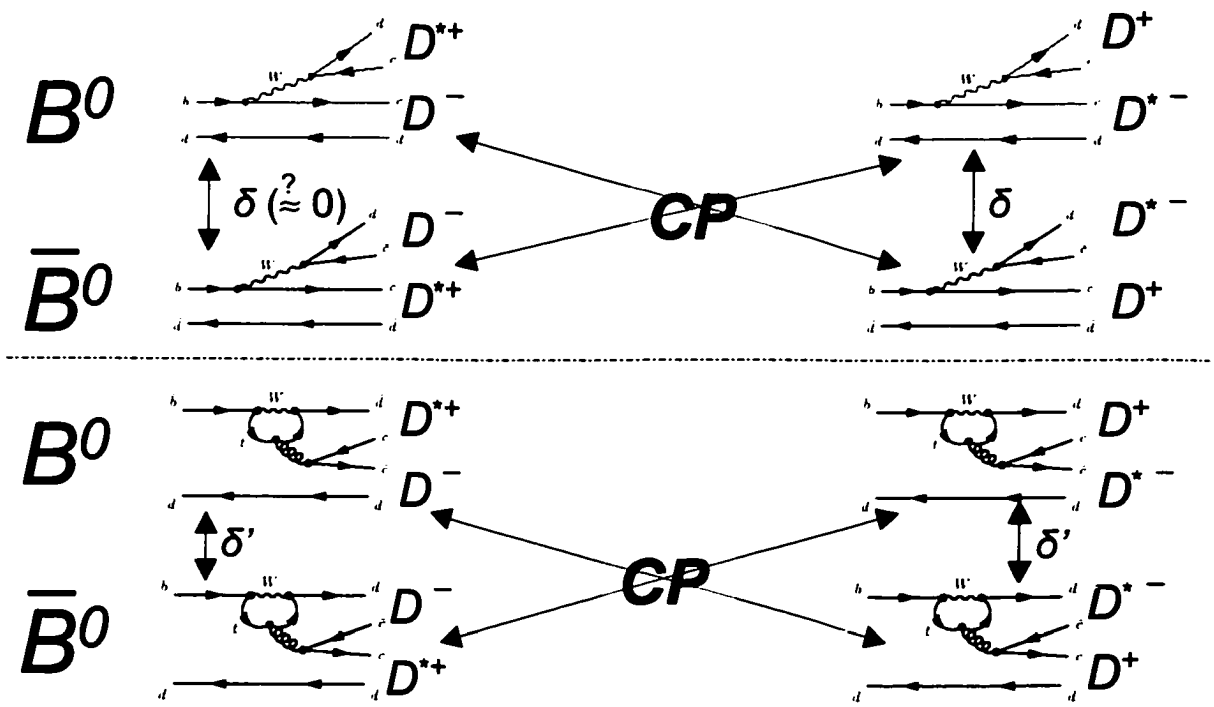


Figure 4.1: The four pairs of Feynman diagrams for $B^0 \to D^{*\pm}D^{\mp}$. The upper diagrams represent the tree amplitude to the D^*D final state. Unlike CP-eigenstates, the final states ($D^{*+}D^-$ and $D^{*-}D^+$) are not self-conjugate; the final states are instead conjugate with each other in the manner shown. A strong phase difference δ may separate B^0 and $\overline{B}^0$ amplitudes to each final state; this phase can differ (δ') for the penguin amplitudes. In HQET the strong phase difference is zero, thus the phases are expected to at least be fairly small (nevertheless we will fit for a floating phase, as explained below).

There are 4 pairs of Feynman diagrams for $\overset{(-)}{B^0} \to D^{*\pm}D^{\mp}$, as may be seen in Figure 4.1. The CP parameters λ for such non-CP-eigenstates are as follows [42]:

$$\lambda_{D^{*+}D^-} = \frac{q}{p}\frac{A(\bar{B}^0 \to D^{*+}D^-)}{A(B^0 \to D^{*+}D^-)} = e^{-2i\phi}\frac{|\bar{A}_{D^{*+}D^-}|}{|A_{D^{*+}D^-}|} = e^{-i(2\phi+\delta)}|\lambda_{D^{*+}D^-}| \quad (21)$$

and

$$\lambda_{D^{*-}D^+} = \frac{q}{p}\frac{A(\bar{B}^0 \to D^{*-}D^+)}{A(B^0 \to D^{*-}D^+)} = e^{-2i\phi}\frac{|\bar{A}_{D^{*-}D^+}|}{|A_{D^{*-}D^+}|} = e^{-i(2\phi-\delta)}|\lambda_{D^{*-}D^+}| \quad (22)$$

where ϕ is the weak phase and δ is a possible strong phase difference in B^0 and $\bar{B}^0$ amplitudes to the D^*D final states. The equalities hold under the typical assumption that CP-violation purely in mixing may be neglected.

Formalism if penguins may be neglected

If one only considers the tree amplitude to the D^*D final states (which is not so typical an assumption in a type II CP mode such as D^*D, where penguins could potentially be significant), then ϕ is identical to the CKM Unitarity Triangle angle β, allowing for direct measurement of $\sin 2\beta$. The magnitudes of the amplitude ratios of $\overset{(-)}{B^0} \to D^{*\pm}D^{\mp}$, are then inversely related, *ie.*

$$|\lambda_{D^{*-}D^+}| = \frac{1}{|\lambda_{D^{*+}D^-}|} \quad (23)$$

Thus the PDFs for $D^{*+}D^-$ and $D^{*-}D^+$ are

$$\Gamma\left(\overset{(-)}{B^0} \to D^{*+}D^-\right) = e^{-\Gamma t}A^2 \left\{1 \overset{(-)}{+} (1-2w)R\cos(\Delta m_d t) \right.$$
$$\left. \overset{(+)}{-} (1-2w)D\sin(2\beta+\delta)\sin(\Delta m_d t)\right\} \quad (24)$$

and

$$\Gamma\left(\overset{(-)}{B^0} \to D^{*-}D^+\right) = e^{-\Gamma t}\bar{A}^2 \left\{1 \overset{(+)}{-} (1-2w)R\cos(\Delta m_d t) \right.$$
$$\left. \overset{(+)}{-} (1-2w)D\sin(2\beta-\delta)\sin(\Delta m_d t)\right\} \quad (25)$$

where

$$R = \frac{1-|\lambda_{D^*+D^-}|^2}{1+|\lambda_{D^*+D^-}|^2} \qquad D = \frac{2|\lambda_{D^*+D^-}|}{1+|\lambda_{D^*+D^-}|^2} \quad (26)$$

One thus fits for the 3 parameters $\sin(2\beta+\delta)$, $\sin(2\beta-\delta)$, and $|\lambda_{D^*+D^-}|$. The quantities $\sin(2\beta+\delta)$ and $\sin(2\beta-\delta)$ can then be combined to solve for $sin(2\beta)$ (up to a discrete ambiguity) via:

$$\sin^2(2\beta) = \frac{1}{2}\left[1 + S\bar{S} \pm \sqrt{(1-S^2)(1-\bar{S}^2)}\right] \quad (27)$$

where $S = \sin(2\beta+\delta)$ and $\bar{S} = \sin(2\beta-\delta)$.

In an HQET approximation, where final state interactions are neglected (as well as penguins), $|\lambda_{D^*+D^-}| = 1$ and the strong phase difference $\delta = 0$, *ie.* both decays measure $\sin(2\beta)$ exactly as a CP-eigenstate would, without needing to have two coupled PDFs as above (although the above would of course still be valid) [43]. Of the two discrete ambiguities above, one (minus in front of the square root) occurs when $|\delta| < \pi/2$ and the other occurs when δ is obtuse, $|\delta| > \pi/2$. Clearly the former is very strongly favored — although HQET is not expected to hold exactly, such a large deviation as would cause the latter solution would be highly unexpected. Thus the doubling of uncertainty that one might expect from the addition of an extra discrete ambiguity is not in fact the case; one of the ambiguity sets is very strongly disfavored.

General formalism — multiple weak phases

If one no longer constrains oneself to a single weak phase from the tree amplitude, a more general formalism is required. Penguins with a t and with a u quark in the loop have a different weak phase than the tree — their phase cancels with the mixing phase, leaving 0 rather than β. Such penguins may potentially have some significance since they are not Cabbibo-suppressed relative to the tree amplitude, although color-suppression and small QCD coefficients indicate that they should in principle be smaller than the tree. In addition, hypothetical new physics (for example, SUSY particles in the penguin loop) could potentially have a still different CP-violating phase.

In general, when multiple CP-violating phase contributions are considered, the ratios of amplitudes

$$\frac{A(\bar{B}^0 \to D^{*+}D^-)}{A(B^0 \to D^{*+}D^-)} \quad \text{and} \quad \frac{A(\bar{B}^0 \to D^{*-}D^+)}{A(B^0 \to D^{*-}D^+)} \tag{28}$$

equal

$$\lambda_{D^{*+}D^-} = \frac{e^{-i\beta} - \sum_{i=1}^{n} |R_i| e^{i(\delta_i - \phi_i)}}{R_0' e^{i(\beta+\delta_0')} - \sum_{i=1}^{n} |R_i'| e^{i(\delta_i' + \phi_i)}} \tag{29}$$

and

$$\lambda_{D^{*-}D^+} = \frac{R_0' e^{-i(\beta+\delta_0')} - \sum_{i=1}^{n} |R_i'| e^{i(\delta_i' - \phi_i)}}{e^{i\beta} - \sum_{i=1}^{n} |R_i| e^{i(\delta_i + \phi_i)}} \tag{30}$$

(assuming no CP-violation purely in mixing) where the $|R_i|$ and $|R_i'|$ are the norms of the ratios of the amplitudes of the additional contributions over the primary contribution (including CKM factors), the ϕ_i are the weak phases of the additional contributions, δ_0' is the difference, if any, in strong phase between $B^0 \to D^{*+}D^-$ and $\bar{B}^0 \to D^{*+}D^-$, and the δ_i and δ_i' are the differences in strong phase between the secondary (penguin) and primary (tree) CP-violating contributions to the $\bar{B}^0$ and B^0 decays respectively.

Let

$$\overset{(-)}{a} = cos(\beta + \delta_0'/2) - \sum |R_i| cos(\langle\delta_i\rangle - \phi_i \overset{(-)}{+} \Delta\delta_i) \quad (4.31a)$$

$$\overset{(-)}{b} = -sin(\beta + \delta_0'/2) - \sum |R_i| sin(\langle\delta_i\rangle - \phi_i \overset{(-)}{+} \Delta\delta_i) \quad (4.31b)$$

$$\overset{(-)}{c} = R_0' cos(\beta + \delta_0'/2) - \sum |R_i'| cos(\langle\delta_i\rangle + \phi_i \overset{(+)}{-} \Delta\delta_i) \quad (4.31c)$$

$$\overset{(-)}{d} = R_0' sin(\beta + \delta_0'/2) - \sum |R_i'| sin(\langle\delta_i\rangle + \phi_i \overset{(+)}{-} \Delta\delta_i) \quad (4.31d)$$

where

$$\langle\delta_i\rangle = \frac{\delta_i + \delta_i'}{2} \quad \text{and} \quad \Delta\delta_i = \frac{\delta_i - \delta_i'}{2}. \quad (4.31e)$$

Then the 3 parameters that are fit as in the single weak phase case, $abs(\lambda)$, $sin(2\phi + \delta)$, and $sin(2\phi - \delta)$, instead become

$$|\lambda| = \sqrt{\frac{a^2 + b^2}{c^2 + d^2}} = \sqrt{\frac{\bar{a}^2 + \bar{b}^2}{\bar{c}^2 + \bar{d}^2}} \quad (32)$$

$$\text{“}sin(2\phi - \delta)\text{”} \rightarrow \frac{ad - bc}{c^2 + d^2} \quad \text{and}$$

$$\text{“}sin(2\phi + \delta)\text{”} \rightarrow \frac{\bar{a}\bar{d} - \bar{b}\bar{c}}{\bar{c}^2 + \bar{d}^2} \quad (33)$$

One can instead use an alternate fitting parametrization, fitting for the magnitudes of the cosine and sine terms of the PDF for each of $D^{*+}D^-$ and $D^{*-}D^+$. One thus fits for 4 parameters ($C_{D^{*+}D^-}$, $S_{D^{*+}D^-}$, $C_{D^{*-}D^+}$, and $S_{D^{*-}D^+}$) rather than the 3 in the case of no penguins, via the following PDFs:

$$\Gamma\left(\overset{(-)}{B^0} \rightarrow D^{*+}D^-\right) = e^{-\Gamma t} A^2 \left\{1 \overset{(-)}{+} (1 - 2w) C_{D^{*+}D^-} cos(\Delta m_d t) \overset{(+)}{-} (1 - 2w) S_{D^{*+}D^-} sin(\Delta m_d t)\right\} \quad (34)$$

and

$$\Gamma\left(\overset{(-)}{B^0} \to D^{*-}D^+\right) = e^{-\Gamma t}\bar{A}^2 \left\{1 \overset{(+)}{-} (1-2w)C_{D^{*-}D^+}cos(\Delta m_d t) \right.$$
$$\left. \overset{(+)}{-} (1-2w)S_{D^{*-}D^+}sin(\Delta m_d t)\right\} \quad (35)$$

Since an association between the 4 parameters and Standard Model Unitarity Triangle values is *not* made, such a parametrization is by nature general. In the case of the multiple weak phase formalism, these parameters are:

$$C_{D^{*+}D^-} = \frac{1-a^2-b^2}{1+a^2+b^2}$$

$$S_{D^{*+}D^-} = \frac{2(ad-bc)}{a^2+b^2+c^2+d^2} \times \sqrt{\frac{a^2+b^2}{c^2+d^2}}$$

$$C_{D^{*-}D^+} = \frac{1-\bar{a}^2-\bar{b}^2}{1+\bar{a}^2+\bar{b}^2}$$

$$S_{D^{*-}D^+} = \frac{2(\bar{a}\bar{d}-\bar{b}\bar{c})}{\bar{a}^2+\bar{b}^2+\bar{c}^2+\bar{d}^2} \times \sqrt{\frac{\bar{a}^2+\bar{b}^2}{\bar{c}^2+\bar{d}^2}} \quad (36)$$

Both parametrizations (3 and 4 parameter) may be fit using the D^*D maximum likelihood fit and results using both parametrizations will be presented.

4.2.2 Likelihood fit method

To account for finite Δz resolution of the detector, the PDFs above are convoluted with a time-resolution function $\mathcal{R}(\delta_t = \Delta t - \Delta t_{true}; \hat{a})$:

$$
\begin{aligned}
&{}^{3-par}\mathcal{F}^{D^{*+}D^{-}}_{\substack{(-)\\B^0}}(\Delta t;\Gamma,\Delta m_d,w,|\lambda|,sin(2\beta+\delta),\hat{a}) = \\
&\qquad {}^{3-par}\Gamma^{D^{*+}D^{-}}_{\substack{(-)\\B^0}}(\Delta t;\Gamma,\Delta m_d,w,|\lambda|,sin(2\beta+\delta)) \;\otimes\; \mathcal{R}(\delta_t;\hat{a}) \quad (37)
\end{aligned}
$$

$$
\begin{aligned}
&{}^{3-par}\mathcal{F}^{D^{*-}D^{+}}_{\substack{(-)\\B^0}}(\Delta t;\Gamma,\Delta m_d,w,|\lambda|,sin(2\beta-\delta),\hat{a}) = \\
&\qquad {}^{3-par}\Gamma^{D^{*-}D^{+}}_{\substack{(-)\\B^0}}(\Delta t;\Gamma,\Delta m_d,w,|\lambda|,sin(2\beta-\delta)) \;\otimes\; \mathcal{R}(\delta_t;\hat{a}) \quad (38)
\end{aligned}
$$

and for the 4-parameter PDF parametrizations:

$$
\begin{aligned}
&{}^{4-par}\mathcal{F}^{D^{*+}D^{-}}_{\substack{(-)\\B^0}}(\Delta t;\Gamma,\Delta m_d,w,|\lambda|,C_{D^{*+}D^{-}},S_{D^{*+}D^{-}},\hat{a}) = \\
&\qquad {}^{4-par}\Gamma^{D^{*+}D^{-}}_{\substack{(-)\\B^0}}(\Delta t;\Gamma,\Delta m_d,w,|\lambda|,C_{D^{*+}D^{-}},S_{D^{*+}D^{-}}) \;\otimes\; \mathcal{R}(\delta_t;\hat{a}) \quad (39)
\end{aligned}
$$

$$
\begin{aligned}
&{}^{4-par}\mathcal{F}^{D^{*-}D^{+}}_{\substack{(-)\\B^0}}(\Delta t;\Gamma,\Delta m_d,w,|\lambda|,C_{D^{*-}D^{+}},S_{D^{*-}D^{+}},\hat{a}) = \\
&\qquad {}^{4-par}\Gamma^{D^{*-}D^{+}}_{\substack{(-)\\B^0}}(\Delta t;\Gamma,\Delta m_d,w,|\lambda|,C_{D^{*-}D^{+}},S_{D^{*-}D^{+}}) \;\otimes\; \mathcal{R}(\delta_t;\hat{a}) \quad (40)
\end{aligned}
$$

The values of the floating parameters are extracted by maximizing the likelihood functions:

$$
\begin{aligned}
\ln \mathcal{L}^{3-par}_{CP} \quad = \quad & \sum_{i}^{tagging} \Bigg[\\
& \sum_{B^0 tag} \ln {}^{3-par}\mathcal{F}^{D^{*+}D^{-}}_{B^0}(\Delta t;\Gamma,\Delta m_d,w_i,|\lambda|,sin(2\beta+\delta),\hat{a}) + \\
& \sum_{\bar{B}^0 tag} \ln {}^{3-par}\mathcal{F}^{D^{*+}D^{-}}_{\bar{B}^0}(\Delta t;\Gamma,\Delta m_d,w_i,|\lambda|,sin(2\beta+\delta),\hat{a}) + \\
& \sum_{B^0 tag} \ln {}^{3-par}\mathcal{F}^{D^{*-}D^{+}}_{B^0}(\Delta t;\Gamma,\Delta m_d,w_i,|\lambda|,sin(2\beta-\delta),\hat{a}) + \\
& \sum_{\bar{B}^0 tag} \ln {}^{3-par}\mathcal{F}^{D^{*-}D^{+}}_{\bar{B}^0}(\Delta t;\Gamma,\Delta m_d,w_i,|\lambda|,sin(2\beta-\delta),\hat{a}) \Bigg] \quad (41)
\end{aligned}
$$

for the 3-parameter parametrization and

$$\begin{aligned}
\ln \mathcal{L}_{CP}^{4-par} \quad = \quad & \sum_{i}^{tagging} \Bigg[\\
& \sum_{B^0 tag} \ln {}^{4-par}\mathcal{F}_{B^0}^{D^{*+}D^-}(\Delta t; \Gamma, \Delta m_d, w_i, C_{D^{*+}D^-}, S_{D^{*+}D^-}, \hat{a}) + \\
& \sum_{\bar{B}^0 tag} \ln {}^{4-par}\mathcal{F}_{\bar{B}^0}^{D^{*+}D^-}(\Delta t; \Gamma, \Delta m_d, w_i, C_{D^{*+}D^-}, S_{D^{*+}D^-}, \hat{a}) + \\
& \sum_{B^0 tag} \ln {}^{4-par}\mathcal{F}_{B^0}^{D^{*-}D^+}(\Delta t; \Gamma, \Delta m_d, w_i, C_{D^{*-}D^+}, S_{D^{*-}D^+}, \hat{a}) + \\
& \sum_{\bar{B}^0 tag} \ln {}^{4-par}\mathcal{F}_{\bar{B}^0}^{D^{*-}D^+}(\Delta t; \Gamma, \Delta m_d, w_i, C_{D^{*-}D^+}, S_{D^{*-}D^+}, \hat{a}) \Bigg] \quad (42)
\end{aligned}$$

for the 4-parameter, where (in both cases) the outer summation is over tagging categories i. The fits for the parameters are done on the combined flavor-eigenstate and D^*D samples in order to determine the parameters which are left floating in the fit. Additional terms included in the likelihood are also required in order to account for backgrounds and their time dependence, as explained in the following subsection.

In order to properly incoporate the correlations between the CP-violation parameters of physical interest and the dilution and resolution function parameters which must also be determined, the fit is performed (for both the 3-parameter and 4-parameter fits) using both the mixing and D^*D samples by simultaneously maximizing the sum:

$$\ln \mathcal{L}_{CP} + \ln \mathcal{L}_{mix} \quad (43)$$

The values of B^0 lifetime and the mixing frequency Δm_d are kept fixed in the fit. The value of the mixing frequency Δm_d is obtained with an unbinned maximum likelihood fit using the tagged flavor-eigenstate B^0 sample alone.

$B^0 \to D^{*\pm}D^{\mp}$: Background modeling

In the presence of backgrounds, the PDFs $\mathcal{F}_{\substack{(-)\\B^0}}$ above must be extended to include a term for each of the four significant background sources: 1) continuum, 2) B^+B^-, 3) combinatoric $B^0\overline{B}^0$, and 4) peaking $B^0\overline{B}^0$. The dominant background in $B^0 \to D^{*\pm}D^{\mp}$ is combinatoric $B\bar{B}$ background, however there are known to be both a small continuum contribution and a small peaking (non-CP) $B^0\overline{B}^0$ contribution. Furthermore, these backgrounds may differ depending on tagging category. Thus a distribution must be specified for each possibility

$$\mathcal{F}_{\substack{(-)\\i,B^0}} = f_{i,sig}\mathcal{F}_{\substack{(-)\\B^0}} + f_{i,peak}\mathcal{B}_{i,peak}(\Delta t, \hat{a}_i) + \sum_{\beta=bkgd} f_{i,\beta}\mathcal{B}_{i,\beta}(\Delta t, \hat{b}_i) \qquad (44)$$

where β runs over i) continuum, ii) B^+B^-, and iii) combinatoric $B^0\overline{B}^0$ contributions. This method of background modeling is very similar to what is used for the charmonium $\sin 2\beta$ measurement (for $\eta_{CP} = -1$ modes). The background PDFs $\mathcal{B}$ provide a description of the Δt distribution for each of the four types of background contributions (3 combinatoric + 1 peaking). The probability breakdown for a given event between signal and each of the four types of background is determined by an $m_{\rm ES}$ fit to data (to separate probability of combinatoric background vs. signal and peaking background), from Monte Carlo (to determine the relative probabilities of signal vs. peaking background), and from the data Δt distribution (to determine the relative probabilities of combinatoric $B^0\overline{B}^0$ vs. B^+B^- vs. continuum). In the $m_{\rm ES}$ fit, the signal (& peaking background) is described by a gaussian, and combinatoric backgrounds by an ARGUS function. This provides a parametrization

of the combinatoric background probability as a function of $m_{\rm ES}$; the further subdivision of probabilities using Monte Carlo utilizes scale factors (which are not functions of $m_{\rm ES}$), *i.e.*

$$f_{i,\rm sig}(m_{\rm ES}) = \frac{(1-\delta_{\rm peak})\mathcal{S}(m_{\rm ES})}{\mathcal{S}(m_{\rm ES})+\mathcal{A}(m_{\rm ES})} \tag{45}$$

$$f_{i,\rm peak}(m_{\rm ES}) = \frac{\delta_{\rm peak}\mathcal{S}(m_{\rm ES})}{\mathcal{S}(m_{\rm ES})+\mathcal{A}(m_{\rm ES})} \tag{46}$$

$$\sum_{\beta=\rm bkgd} f_{i,\beta}(m_{\rm ES}) = \frac{\mathcal{A}(m_{\rm ES})}{\mathcal{S}(m_{\rm ES})+\mathcal{A}(m_{\rm ES})} \tag{47}$$

where $\mathcal{S}$ and $\mathcal{A}$ are the fitted signal gaussian and combinatoric ARGUS functions respectively. In addition, the background resolution function parameters $\hat{b}_i$ and background effective dilutions are included as free parameters within the fit, as detailed in the following section.

Combinatoric $B^0\overline{B}^0$ backgrounds to $B^0 \to D^{*\pm}D^{\mp}$ can potentially add an effective cosine term in the asymmetry due to differing rates for $B^0 \to D^{*-}X$ and $\overline{B}^0 \to D^{*-}X$ (the former is a Cabbibo-favored decay whereas the latter is not). This is corrected for by fitting the large sidebands ($|\Delta E| < 200\,\rm MeV$) for an effective cosine term in each of the 4 tag categories and in each of $D^{*-}D^{+}$ and $D^{*+}D^{-}$ (thus 8 parameters). These parameters are then input to the fit as the effective cosine terms for the background. The systematic error due to this method is evaluated by splitting up the sideband into high and low ΔE halves as well as high and low $m_{\rm ES}$ halves and reevaluating the fit on each of these 4 subsets. The largest of the 4 deviations for each of the 8 parameters is taken to be the systematic error on that parameter. The results of the fits are:

stat syst

```
--> cos coeff D*-D+ bkgd (lepton)      0.4098 +/- 0.1663 +/- 0.1191
--> cos coeff D*-D+ bkgd (kaon)        0.2974 +/- 0.0630 +/- 0.0580
--> cos coeff D*-D+ bkgd (NT1)         0.2687 +/- 0.1573 +/- 0.2093
--> cos coeff D*-D+ bkgd (NT2)         0.1090 +/- 0.0910 +/- 0.0685
--> cos coeff D*-D+ bkgd (lepton)     -0.6384 +/- 0.2045 +/- 0.1420
--> cos coeff D*-D+ bkgd (kaon)       -0.1892 +/- 0.0636 +/- 0.0492
--> cos coeff D*-D+ bkgd (NT1)        -0.2855 +/- 0.1638 +/- 0.0675
--> cos coeff D*-D+ bkgd (NT2)        -0.1543 +/- 0.0907 +/- 0.0749
```

4.3 Blind analysis

For $D^{*+}D^{*-}$ we choose the strategy of fitting for the sine and cosine coefficients. These two parameters were blinded in order to eliminate possible experimenter's bias.

Similar to the charmonium sin2β measurement, the amplitude of the asymmetry $\mathcal{A}_{CP}(\Delta t)$ from the fits are hidden by a one-time random choice of sign flip and arbitrary offset based on a user-specified key word. The sign flip hides whether a change in the analysis increases or decreases the resulting asymmetry. However, the magnitude of the change is not hidden.

With this technique, systematic studies may be performed while keeping the numerical values of the physical parameters hidden. $D^{*+}D^{*-}$ and $D^{*}D$ are blinded separately using separate blinding strings, and both of these modes are separate from other sin2β modes.

For $D^{*}D$, the values of $sin(2\beta+\delta)$, $sin(2\beta-\delta)$, and $|\lambda|$ (for the 3 parameter fit) and $C_{D^{*+}D^{-}}$, $S_{D^{*+}D^{-}}$, $C_{D^{*-}D^{+}}$, and $S_{D^{*-}D^{+}}$ (for the 4 parameter fit) are hidden.

	Number of events	lepton	Kaon	NT1	NT2
all events	223	16	71	13	33
signal	90.0	8.0	29.8	8.6	10.6
background	134.0	8.0	41.2	4.3	22.4

Table 4.1: $D^{*+}D^{*-}$ break down in each tagging categories.

4.4 Results

4.4.1 $B^0 \to D^{*+}D^{*-}$ results

Subdivision in tagging categories

In table 4.1 the data are subdivided in four tagging categories. The number of events for signal and background are evaluated from the unbinned fit to the $m_{\rm ES}$ distribution. For each category, the mean value μ and standard deviation σ of the signal and the Argus coefficient K of the background were fixed at the values obtained from a fit to all tagged and untagged events ($\mu = 5.2801$ GeV, $\sigma = 0.0023$ GeV, $K = -39.4$ respectively). The numbers of signal and background events where floated.

Fit to data

A combined fit to the $D^{*+}D^{*-}$ and $B_{\rm flav}$ samples was performed. For $D^{*+}D^{*-}$ we floated the sin and cos coefficients S and C and the fraction of prompt background and we fixed the lifetime of the background to the B^0 lifetime.

The *usual* parameters [40] for the $B_{\rm flav}$ sample were floated: $\mathcal{D}$, $\Delta\mathcal{D}$, the resolution function parameters for signal and background, and the back-

S	C	$f_{\tau=0}$
-0.045 ± 0.453	0.121 ± 0.303	0.359 ± 0.162
$(0.4463; -0.4623)$	$(0.2969; -0.3024)$	

Table 4.2: Results of the fit to $D^{*+}D^{*-}$ data. S and C are the sine and cosine coefficients respectively and $f_{\tau=0}$ is the fraction of prompt background. The numbers in parenthesis in the second row represent their almost symmetric errors from the fit.

ground fractions. The results obtained are summarized in table 4.2. The global correlation coefficient for the sine coefficient S is 12.3%, for the cosine coefficient C it is 13.3%.

4.4.2 $B^0 \to D^{*\pm}D^{\mp}$ results

Blind fit on data

A blind fit on the $B^0 \to D^{*\pm}D^{\mp}$ data sample (combined with data Breco events) has been performed. The four physical parameters corresponding to the cosine and sine terms for each of D*+D- and D*-D+ were each separately blinded with a blinding width of 0.7.

Using the central values of the background cosine terms (see Section 4.2.2) in a blind fit on data with 13% peaking background gives:

```
===== Final value for Scale (core) - sig              1.1316 +/-    0.0597
===== Final value for Scale (core) - bgd              1.4335 +/-    0.0192
===== Final value for d(dt) (core) [ps] - lepton     -0.0111 +/-    0.0440
===== Final value for d(dt) (core) [ps] - kaon       -0.1706 +/-    0.0308
===== Final value for d(dt) (core) [ps] - NT1        -0.1118 +/-    0.0557
```

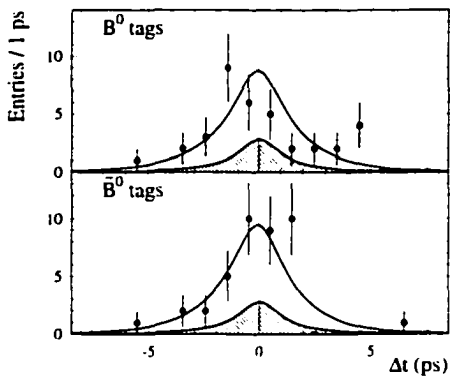


Figure 4.2: Flavor tagged $B^0 \rightarrow D^{*+}D^{*-}$ Δt distributions with the fit likelihood superimposed. The shaded portion represents the background contribution.

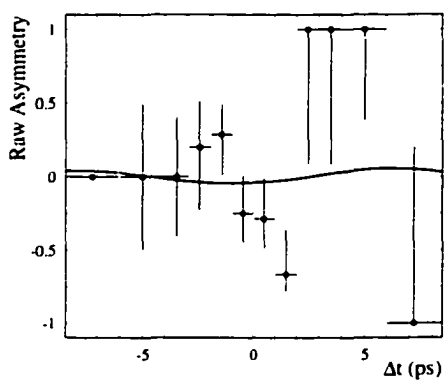


Figure 4.3: Raw $B^0 \rightarrow D^{*+}D^{*-}$ time dependent asymmetry with the likelihood superimposed.

```
===== Final value for d(dt) (core) [ps] - NT2          -0.1667 +/-    0.0484
===== Final value for d(dt) (core) [ps] - bkgd         -0.0262 +/-    0.0093
===== Final value for d(dt) (tail) [ps] - sig          -0.8584 +/-    0.2790
===== Final value for fraction in tail - sig            0.0905 +/-    0.0246
===== Final value for outlier fraction - sig            0.0031 +/-    0.0016
===== Final value for outlier fraction - bgd            0.0120 +/-    0.0018
===== Final value for dilu ave - sig, lepton            0.8264 +/-    0.0189
===== Final value for dilu ave - sig, kaon              0.6392 +/-    0.0145
===== Final value for dilu ave - sig, NT1               0.5585 +/-    0.0312
===== Final value for dilu ave - sig, NT2               0.2553 +/-    0.0270
===== Final value for dilu ave - t=0 bkgd, lepton       0.6573 +/-    0.1213
===== Final value for dilu ave - t=0 bkgd, kaon         0.4845 +/-    0.0191
===== Final value for dilu ave - t=0 bkgd, NT1          0.3150 +/-    0.0497
===== Final value for dilu ave - t=0 bkgd, NT2          0.0994 +/-    0.0264
===== Final value for dilu ave - t>0 no mix, lepton     0.1390 +/-    0.1212
===== Final value for dilu ave - t>0 no mix, kaon       0.1610 +/-    0.0424
===== Final value for dilu ave - t>0 no mix, NT1        0.0588 +/-    0.0980
===== Final value for dilu ave - t>0 no mix, NT2        0.0431 +/-    0.0672
===== Final value for dilu diff - sig, lepton          -0.0076 +/-    0.0315
===== Final value for dilu diff - sig, kaon             0.0194 +/-    0.0225
===== Final value for dilu diff - sig, NT1             -0.0401 +/-    0.0488
===== Final value for dilu diff - sig, NT2              0.1046 +/-    0.0403
===== BLINDED Final value for cos coeff D*-D+ signal    0.5440 +/-    0.5181 (+0.5142  -0.5101)
===== BLINDED Final value for sin coeff D*-D+ signal    1.2697 +/-    0.9191 (+0.8892  -0.9160)
===== BLINDED Final value for cos coeff D*+D- signal    0.3777 +/-    0.7731 (+0.7670  -0.7369)
===== BLINDED Final value for sin coeff D*+D- signal   -0.1120 +/-    1.4577 (+1.4238  -1.3850)
===== Final value for f(tau=0) D*D                      0.3442 +/-    0.0948
===== Final value for f(tau=0) breco lepton             0.4886 +/-    0.0581
===== Final value for f(tau=0) breco kaon               0.7108 +/-    0.0146
===== Final value for f(tau=0) breco NT1                0.6924 +/-    0.0297
===== Final value for f(tau=0) breco NT2                0.7563 +/-    0.0175
```

The <u>unblind</u> results of the fit are:

```
===== Final value for Scale (core) - sig                1.1315 +/-    0.0597
===== Final value for Scale (core) - bgd                1.4335 +/-    0.0192
===== Final value for d(dt) (core) [ps] - lepton       -0.0111 +/-    0.0440
```

```
===== Final value for d(dt) (core) [ps] - kaon          -0.1706 +/-    0.0309
===== Final value for d(dt) (core) [ps] - NT1           -0.1117 +/-    0.0557
===== Final value for d(dt) (core) [ps] - NT2           -0.1666 +/-    0.0484
===== Final value for d(dt) (core) [ps] - bkgd          -0.0262 +/-    0.0093
===== Final value for d(dt) (tail) [ps] - sig           -0.8581 +/-    0.2789
===== Final value for fraction in tail - sig             0.0906 +/-    0.0246
===== Final value for outlier fraction - sig             0.0031 +/-    0.0016
===== Final value for outlier fraction - bgd             0.0120 +/-    0.0018
===== Final value for dilu ave - sig, lepton             0.8263 +/-    0.0189
===== Final value for dilu ave - sig, kaon               0.6392 +/-    0.0145
===== Final value for dilu ave - sig, NT1                0.5585 +/-    0.0312
===== Final value for dilu ave - sig, NT2                0.2553 +/-    0.0270
===== Final value for dilu ave - t=0 bkgd, lepton        0.6573 +/-    0.1213
===== Final value for dilu ave - t=0 bkgd, kaon          0.4845 +/-    0.0191
===== Final value for dilu ave - t=0 bkgd, NT1           0.3150 +/-    0.0497
===== Final value for dilu ave - t=0 bkgd, NT2           0.0994 +/-    0.0264
===== Final value for dilu ave - t>0 no mix, lepton      0.1390 +/-    0.1212
===== Final value for dilu ave - t>0 no mix, kaon        0.1610 +/-    0.0424
===== Final value for dilu ave - t>0 no mix, NT1         0.0588 +/-    0.0980
===== Final value for dilu ave - t>0 no mix, NT2         0.0431 +/-    0.0672
===== Final value for dilu diff - sig, lepton           -0.0075 +/-    0.0315
===== Final value for dilu diff - sig, kaon              0.0193 +/-    0.0225
===== Final value for dilu diff - sig, NT1              -0.0403 +/-    0.0488
===== Final value for dilu diff - sig, NT2               0.1049 +/-    0.0403
===== Final value for cos coeff D*-D+ signal            -0.2988 +/-    0.4974 (+    0.4884    -0.4934)
===== Final value for sin coeff D*-D+ signal             0.3751 +/-    0.8754 (+    0.8719    -0.8436)
===== Final value for cos coeff D*+D- signal             0.5329 +/-    0.7418 (+    0.7362    -0.7076)
===== Final value for sin coeff D*+D- signal            -0.4344 +/-    1.4093 (+    1.3774    -1.3412)
===== Final value for f(tau=0) D*D                       0.3440 +/-    0.0948
===== Final value for f(tau=0) breco lepton              0.4886 +/-    0.0581
===== Final value for f(tau=0) breco kaon                0.7108 +/-    0.0146
===== Final value for f(tau=0) breco NT1                 0.6923 +/-    0.0297
===== Final value for f(tau=0) breco NT2                 0.7563 +/-    0.0175
```

The addition to the error due to the statistical errors of the cosine terms of the background (see Section 4.2.2) must be added in quadrature with these

statistical errors. (The systematic errors on the cosine terms are treated in the Systematics section, but the statistical errors on these quantities result in quadrature additions to the statistical error on the physical parameters.) Variation of the background cosine terms according to their statistical errors (Section 4.2.2) results in additional statistical errors of 0.066 for $S_{D^{*+}D^{-}}$, 0.036 for $C_{D^{*+}D^{-}}$, 0.022 for $S_{D^{*-}D^{+}}$, and 0.037 for $C_{D^{*-}D^{+}}$. Combining these in quadrature with the statistical errors from the global fit gives:

```
===== Final value for cos coeff D*-D+ signal      -0.2988 +/-   0.4988 (+   0.4898   -0.4948)
===== Final value for sin coeff D*-D+ signal       0.3751 +/-   0.8757 (+   0.8723   -0.8439)
===== Final value for cos coeff D*+D- signal       0.5329 +/-   0.7427 (+   0.7371   -0.7095)
===== Final value for sin coeff D*+D- signal      -0.4344 +/-   1.4105 (+   1.3786   -1.3428)
```

The above are the final results for the $B^0 \to D^{*\pm}D^{\mp}$ CP fit.

4.5 $D^{*+}D^{*-}$ systematics

Table 4.3 summarizes our estimate of the systematic errors δS and δC on the coefficients S and C. In this section we describe the individual contributions to the systematic error which are taken into account in the present analysis.

4.5.1 Signal Parameters

As for the charmonium CP events, the underlying assumption in the analysis of $B^0 \to D^{*+}D^{*-}$ events is that the resolution function and the dilutions are the same in the B flavor sample and in the CP sample. Possible deviations from these assumptions are accounted for in the systematic error.

Systematics source	δS	δC
signal Δt resolution function	0.008	0.003
tagging dilution	0.005	0.005
peaking background	0.003	0.009
CP background content	0.022	0.038
lifetime of background	0.034	0.005
B^0 lifetime variation	0.001	0.001
Δm_d variation	0.030	0.022
SVT misalignment	0.011	0.008
Boost uncertainty	0.002	0.001
Fit bias	0.001	0.004
TOTAL	0.053	0.046

Table 4.3: Summary of the systematics errors estimated for the coefficients S and C.

Tagging Dilutions

We extract and compare the tagging dilutions from large samples of B_{flav} and CP simulated samples. The results are shown in Table 4.4. We assign as systematic errors the observed shift on the fitted S and C parameters, fixing alternatively the mistag fractions at the values obtained from the two samples. In the $B^0 \to D^{*+}D^{*-}$ analysis, for the coefficient S of the sine term we assign a systematic error of 0.0052, for the coefficient C of the cosine term 0.0051.

Parameter	$D^{*+}D^{*-}$ MC	B_{flav} MC
w (Lepton)	0.059 ± 0.004	0.066 ± 0.003
w (Kaon)	0.164 ± 0.004	0.171 ± 0.003
w (NT1)	0.196 ± 0.009	0.213 ± 0.006
w (NT2)	0.348 ± 0.008	0.353 ± 0.006
$\Delta\ w$ (Lepton)	-0.028 ± 0.008	-0.007 ± 0.006
$\Delta\ w$ (Kaon)	-0.005 ± 0.008	-0.014 ± 0.006
$\Delta\ w$ (NT1)	0.023 ± 0.018	0.045 ± 0.013
$\Delta\ w$ (NT2)	-0.032 ± 0.016	-0.029 ± 0.011

Table 4.4: Values of mistag fractions in $D^{*+}D^{*-}$ and B_{flav} MC samples (truth level), used as alternative fixed parameters in the fit of the coefficients S and C of the sine and cosine terms.

Parameterization of the signal resolution function

A similar method is also applied to study the systematic uncertainty in the resolution function. We extract the Δt resolution function parameters from large samples of B_{flav} and CP simulated events. The fitted parameters values are shown in table B.6. Table 4.5 shows the fitted values of the S and C coefficients, obtained fixing four different sets of resolution function parameters, corresponding to MC truth and fitted values for the $D^{*+}D^{*-}$ and B_{flav} MC samples. The shift in the fitted S and C coefficients values between the two sets corresponding to $D^{*+}D^{*-}$ B_{flav} truth level is assigned as the systematic error.

Resol.F. Parameter set	S	C	δS	δC
$D^{*+}D^{*-}$ MC Truth	-0.2312 ± 0.0221	-0.0017 ± 0.0145	0	0
$D^{*+}D^{*-}$ MC fitted	-0.2328 ± 0.0223	-0.0044 ± 0.0145	0.016	0.0027
B_{flav} MC Truth	-0.2229 ± 0.0219	-0.0042 ± 0.0145	0.0083	0.0025
B_{flav} MC fitted	-0.2289 ± 0.0219	-0.0042 ± 0.0145	0.0023	0.0025

Table 4.5: Fitted values of the S and C coefficients obtained using different resolution function parameters corresponding to the $D^{*+}D^{*-}$ and B_{flav} MC samples (truth level, fitted). The shift in the fitted S and C values from the case of $D^{*+}D^{*-}$ (truth level), is also shown.

CP background peaking component

In section B.1.2 we estimated an upper limit for the amount of peaking background to be of the order of 2.9 events for 55 fb^{-1}, that corresponded to a fraction of peaking background $f_{\text{peak}} = 0.032$.

In the *CP* fit we fix the fraction of peaking background to be 0.0. The systematic assigned to the measurement is estimated by a new fit done with a fraction of peaking background fixed to f_{peak}. The difference on fitted values of sine and cosine coefficients between the two fits is 0.003 and 0.009 respectively.

CP content of background

We assume the *CP* content of the background events to be 0.0 in the *CP* fit. We estimate the systematic error due to this assumption by changing its value to ± 1. The difference on fitted CP parameters with respect the nominal

fit is 0.022 for the sine coefficient and 0.038 for the cosine coefficient.

Lifetime of the background

The Δt spectrum of the background events is modeled as the sum of a prompt component and a component with an effective lifetime τ_{bkg} that is fixed to the B^0 lifetime in the nominal fit. We evaluate the systematic error to be 0.034 for the sine coefficient and 0.005 by varying the fixed value from 1.0 to 2.5 ps.

Variation of Δm_d and τ_{B^0}

In the CP fit we fix the value of the mixing fraction to $\Delta m_d = 0.472$ and the lifetime of the B meson to $\tau_{B^0} = 1.548$. By varying them of $\pm 1\sigma$ (i.e. ± 0.017 for Δm_d and ± 0.032 for τ_{B^0}) we evaluate the systematic error due to the use of those fixed values. The shift in lifetime produce a variation of both δS and δC of 0.001. The variation of the mixing fraction of $\pm 1\sigma$ results in a variation of the sine coefficient of 0.030 and 0.022 for the cosine coefficient.

4.5.2 Detector effects

SVT misalignment

A subsample of D^*D^* signal SP4 Monte Carlo was processed with and without SVT misalignment (SVT "Misaligned set" called: diffBlend). In table 4.6 we compare the S and C fit results on reconstructed D^*D^* from the misaligned sample and on the D^*D^* candidates reconstructed in the same events, after reprocessing without misalignment. After the selection of the signal the

Resol.F. Parameter set	S	C	δS	δC
$D^{*+}D^{*-}$ "aligned SVT"	-0.1778 ± 0.0550	0.0173 ± 0.0360	0	0
$D^{*+}D^{*-}$ "misaligned SVT"	-0.1885 ± 0.0554	0.0252 ± 0.0363	0.011	0.008

Table 4.6: Fitted values of the sine and cosine coefficients S and C, obtained from $D^{*+}D^{*-}$ Montecarlo samples processed with different SVT alignment sets.

two sets have 4063 events in common (315 are present in the "aligned set" only, 228 in the "misaligned set"). The systematic error due to SVT misalignment, assigned to the S and C coefficients, is 0.011 and 0.008 respectively.

Uncertainty on Boost

The uncertainty on the values of the center-of-mass boost used for Δz to Δt conversion can result in a scaling of Δt measurement. In order to estimate the systematic error due to 0.6% uncertainty on the boost we scale the measured Δt and its error up and down by the same amount and repeat the nominal fit. The systematic error assigned to the sine coefficint S is 0.002 and to the cosine coefficient C is 0.001.

4.5.3 Monte Carlo correction (fit bias)

A fit was also done using the whole MC sample (26853 events) fixing the tagging dilutions and resolution function parameters at the fitted value of the B_{flav} MC sample and floating the sine and cosine coefficients. The input

value of the sine coefficient in the Monte Carlo is -0.231 corresponding to $Im\lambda = -0.7$. Inputting a flat angular distribution gives $K = 0.33$. The input cosine coefficient is zero corresponding to $|\lambda| = 1$. The result obtained is: $S = -0.232 \pm 0.022$, $C = -0.004 \pm 0.015$. The systematic error assigned to S is 0.001 and the one assigned to C is 0.004.

4.6 $D^{*\pm}D^{\mp}$ systematics

Table 4.7 summarizes our estimate of the systematic errors $\delta S_{D^{*+}D^-}$, $\delta C_{D^{*+}D^-}$, $\delta S_{D^{*-}D^+}$, and $\delta C_{D^{*-}D^+}$ on the 4 time-dependent asymmetry parameters for $D^{*\pm}D^{\mp}$. In this section we describe the individual contributions to the systematic error which are taken into account in the present analysis.

4.6.1 Signal Parameters

Potential deviations from the assumption that the resolution function and the dilutions in $B^0 \to D^{*\pm}D^{\mp}$ are the same in the B flavor sample, and variations on other parameters which are fixed in the fit but have known errors, are accounted for in the systematic errors below.

Tagging Dilutions

We extract and compare the tagging dilutions from large samples of B_{flav} and CP simulated samples. The results are shown in Table 4.8. We assign as systematic errors the observed shift on the fitted S and C parameters, fixing alternatively the tagging dilutions at the values obtained from the two samples. In the $B^0 \to D^{*\pm}D^{\mp}$ analysis, we assign systematic errors on the

Systematics source	$\delta S_{D^{*+}D^-}$	$\delta C_{D^{*+}D^-}$	$\delta S_{D^{*-}D^+}$	$\delta C_{D^{*-}D^+}$
signal Δt resolution function	0.057	0.049	0.010	0.011
tagging dilution	0.079	0.007	0.020	0.009
peaking bkgd. fraction	0.042	0.011	0.024	0.044
CP of peaking bkgd.	0.101	0.098	0.102	0.099
cos coeffs. of comb. bkgd.	0.044	0.024	0.015	0.025
lifetime of background	0.144	0.090	0.032	0.052
B^0 lifetime variation	0.045	0.035	0.013	0.007
Δm_d variation	0.055	0.028	0.004	0.010
SVT misalignment	0.011	0.008	0.011	0.008
Boost uncertainty	0.002	0.001	0.002	0.001
Fit bias	0.039	0.029	0.038	0.027
TOTAL	0.225	0.154	0.120	0.127

Table 4.7: Summary of the systematics errors estimated for the coefficients $S_{D^{*+}D^-}$, $C_{D^{*+}D^-}$, $S_{D^{*-}D^+}$, and $C_{D^{*-}D^+}$.

CP parameters of 0.079 for $S_{D^{*+}D^-}$, 0.007 for $C_{D^{*+}D^-}$, 0.020 for $S_{D^{*-}D^+}$, and 0.009 for $C_{D^{*-}D^+}$.

Parameterization of the signal resolution function

A similar method is also applied to study the systematic uncertainty due to the resolution function model. We extract the Δt resolution function parameters from large samples of B_{flav} and CP simulated events. The fitted parameters values are shown in table B.7. Table 4.9 shows the fitted values

Parameter	$D^{*\pm}D^{\mp}$ MC	B_{flav} MC
w (Lepton)	0.059 ± 0.005	0.066 ± 0.003
w (Kaon)	0.161 ± 0.004	0.171 ± 0.003
w (NT1)	0.194 ± 0.012	0.213 ± 0.006
w (NT2)	0.344 ± 0.011	0.353 ± 0.006
$\Delta\ w$ (Lepton)	-0.001 ± 0.011	-0.007 ± 0.006
$\Delta\ w$ (Kaon)	-0.017 ± 0.010	-0.014 ± 0.006
$\Delta\ w$ (NT1)	0.039 ± 0.023	0.045 ± 0.013
$\Delta\ w$ (NT2)	-0.057 ± 0.021	-0.029 ± 0.011

Table 4.8: Values of dilutions in $D^{*\pm}D^{\mp}$ and B_{flav} MC samples (truth level), used as alternative fixed parameters in the fit of the coefficients S and C of the sine and cosine terms.

of the S and C coefficients, obtained fixing four different sets of resolution function parameters, corresponding to MC truth and fitted values for the $D^{*\pm}D^{\mp}$ and B_{flav} MC samples. The shift in the fitted S and C coefficients values between the two sets corresponding to $D^{*\pm}D^{\mp}$ and B_{flav} truth level is assigned as the systematic error.

Fraction of peaking background

In section B.1.2 we estimate the number of peaking background events in the signal region to be 6.0 ± 4.0 events. This correponds to the fraction of peaking background being equal to $9.32 \pm 6.15\%$. The systematic assigned to the measurement is estimated by varying in the fit the fraction of peaking

Resol.F. Parameter set	$\delta S_{D^{*+}D^-}$	$\delta C_{D^{*+}D^-}$	$\delta S_{D^{*-}D^+}$	$\delta C_{D^{*-}D^+}$
$D^{*\pm}D^{\mp}$ MC Truth	0	0	0	0
$D^{*\pm}D^{\mp}$ MC fitted	−0.034	0.042	−0.016	0.027
B_{flav} MC Truth	0.057	−0.049	0.010	−0.011
B_{flav} MC fitted	0.048	−0.052	0.011	−0.013

Table 4.9: Differences in fitted values of the S and C coefficients obtained using different resolution function parameters corresponding to the $D^{*\pm}D^{\mp}$ and B_{flav} MC samples (truth level, fitted) from that obtained using $D^{*\pm}D^{\mp}$ MC truth values.

background accourding to the measured value and error while fitting for sine and cosine coefficients. The difference on fitted CP parameters with respect the nominal fit is 0.042 for $S_{D^{*+}D^-}$, 0.011 for $C_{D^{*+}D^-}$, 0.024 for $S_{D^{*-}D^+}$, and 0.044 for $C_{D^{*-}D^+}$.

CP content of the peaking background

We assume the CP content of the peaking background events to be 0.0 in the nominal fit. We estimate the systematic error due to this assumption by changing its value to ± 1. The difference on fitted CP parameters with respect the nominal fit is 0.101 for $S_{D^{*+}D^-}$, 0.098 for $C_{D^{*+}D^-}$, 0.102 for $S_{D^{*-}D^+}$, and 0.099 for $C_{D^{*-}D^+}$.

Cosine coefficients of the combinatoric background

As explained in Section 4.2.2, combinatoric $B^0\overline{B}^0$ backgrounds to $B^0 \to D^{*\pm}D^{\mp}$ can potentially add an effective cosine term due to differing rates

for $B^0 \to D^{*-}X$ and $\overline{B}^0 \to D^{*-}X$. The 8 effective cosine coefficients are fit on the large sideband with associated statistical and systematic errors on the fit. The statistical errors from these parameters are incorporated into the statistical error as detailed in Section 4.2.2. The systematic errors on the effective cosine coefficients result in systematic errors on the CP parameters of 0.044 for $S_{D^{*+}D^-}$, 0.024 for $C_{D^{*+}D^-}$, 0.015 for $S_{D^{*-}D^+}$, and 0.025 for $C_{D^{*-}D^+}$.

Variation of Δm_d and τ_{B^0}

In the nominal fit we fix the value of the mixing frequency to $\Delta m_d = 0.472$ and the lifetime of the B meson to $\tau_{B^0} = 1.548$. By varying them $\pm 1\sigma$ (*i.e.* ± 0.017 for Δm_d and ± 0.032 for τ_{B^0}) we evaluate the systematic error due to fixing the values in the fit. The variation of the lifetime results in systematic errors on the CP parameters of 0.045 for $S_{D^{*+}D^-}$, 0.035 for $C_{D^{*+}D^-}$, 0.013 for $S_{D^{*-}D^+}$, and 0.007 for $C_{D^{*-}D^+}$. The variation of the mixing frequency results in systematic errors on the CP parameters of 0.055 for $S_{D^{*+}D^-}$, 0.028 for $C_{D^{*+}D^-}$, 0.004 for $S_{D^{*-}D^+}$, and 0.010 for $C_{D^{*-}D^+}$.

4.6.2 Monte Carlo correction (fit bias)

A fit was also done using a large MC sample (15000 events) fixing the tagging dilutions and resolution function parameters at the fitted value of the B_{flav} MC sample and floating the 4 CP parameters. The input values of the $\sin 2\beta$ and $|\lambda|$ in the Monte Carlo are 0.7 and 1.0 respectively, corresponding to values of 0.7 for the sine parameters and 0 for the cosine parameters.

The result obtained is: $S_{D^{*+}D^-} = 0.7374 \pm 0.0385$, $C_{D^{*+}D^-} = 0.0247 \pm 0.0294$, $S_{D^{*-}D^+} = 0.7005 \pm 0.0380$, and $C_{D^{*-}D^+} = 0.0267 \pm 0.0267$. The

systematic error thus assigned are 0.039 for $S_{D^{*+}D^{-}}$, 0.029 for $C_{D^{*+}D^{-}}$, 0.038 for $S_{D^{*-}D^{+}}$, and 0.027 for $C_{D^{*-}D^{+}}$.

Chapter 5

Conclusions

Neglecting the effects of Standard Model penguins and assuming that HQET holds for $B \to D^{(*)}\overline{D}^{(*)}$ decays (the charm quark from the W makes it likely that significant corrections to HQET may in fact be necessary in actuality), one can simply take the weighted average of the sine terms from $B^0 \to D^{*\pm}D^{\mp}$ and the R_t-corrected sine term from $B^0 \to D^{*+}D^{*-}$ to obtain $\sin 2\beta$ from $b \to c\bar{c}d$ decays. With warning considered, this weighted average is:

"$\sin 2\beta$" from $B^0 \to D^{*\pm}D^{\mp}$:	0.15 ± 0.75
"$\sin 2\beta$" from $B^0 \to D^{*+}D^{*-}$:	-0.09 ± 0.81[1]
"combined" (weighted average) "$\sin 2\beta$" from $b \to c\bar{c}d$:	0.04 ± 0.55

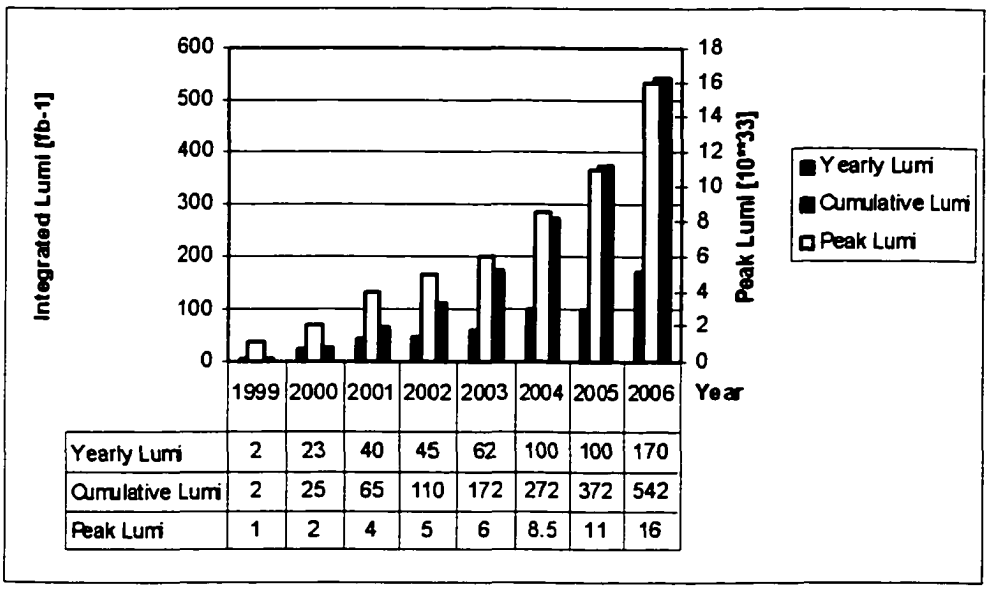
One should be forewarned, however, that theoretical corrections for penguins are needed. Grossman and Worah estimate the Standard Model penguin corrections to be < 0.1 on $\sin 2\beta$ for $b \to c\bar{c}d$, however this is model-dependent.

[1] using $R_t = 0.22 \pm 0.18$ [18]

It is likely that further phenomenological estimates of Standard Model penguin corrections to $b \to c\bar{c}d$ $\sin 2\beta$ will be forthcoming as experimental errors decrease.

The planned increases in luminosity at *BABAR* and Belle promise large improvements in statistical error for $D^{(*)+}D^{(*)-}$ as well as other CP violation measurements. The results presented here all have statistics-dominated errors, and the largest systematics also scale with statistics, so updated results in coming years should contain significant decreases in error bars. The prediction for luminosity increase over the next 4 years is summarized in figure 5.1. Statistics will quadruple from the present analysis by summer 2004, promising a factor of 2 decrease in errors. A factor of 3 should come by the end of 2006, when the data is predicted to have increased nine-fold. Lest one think that these predictions are likely to be optimistic, note figure 5.2, which shows a comparison of the model as predicted in March 2001 with data taken since that time. Note also that we are currently exceeding the PEP-II design integrated data rate by over a factor of two. This also does not include the likely future measurements of CP violation in $B \to D^{(*)}\bar{D}^{(*)}$ from Belle, which will create a still-smaller world-average error.

Supersymmetric models can have significant deviations in $\sin 2\beta$ from $b \to c\bar{c}d$ from that of $b \to c\bar{c}s$ decays such as $B^0 \to J/\psi K_S^0$. Grossmann and Worah predict that supersymmetric deviations of 0.3 can occur in measured $\sin 2\beta$ from $b \to c\bar{c}d$. With data collected over the next few years, significant light will be shed on supersymmetric contributions to CP violation, both using $b \to c\bar{c}d$ decays as well as $b \to s\bar{s}s$, which will soon have sufficient statistics for a first measurement of $\sin 2\beta$. At present, however, statistics do not allow



	1999	2000	2001	2002	2003	2004	2005	2006
Yearly Lumi	2	23	40	45	62	100	100	170
Cumulative Lumi	2	25	65	110	172	272	372	542
Peak Lumi	1	2	4	5	6	8.5	11	16

Figure 5.1: Projected luminosities over the next four years at PEP-II-*BABAR*.

us to say anything definitive contributions to CP violation in $b \to c\bar{c}d$ from beyond the Standard Model.

To summarize, a first measurement of time-dependent CP violating asymmetries in weak decays with quark process $b \to c\bar{c}d$ is presented. This represents the first measurement of $\sin 2\beta$ (neglecting penguins) outside of charmonium. It is the first measurement to use a new quark process, $b \to c\bar{c}d$, rather than $b \to c\bar{c}s$. Significant deviations to time-dependent CP violating asymmetries in this mode may be expected from extensions to the Standard Model, and as statistics increase, constraints on the CP violating sector of these models will be possible. CP violation in the Standard Model is inconsistent with being the sole source of the CP asymmetry in the universe, so with this measurement and its descendents, we gain a handle on why nature and our universe is composed of matter.

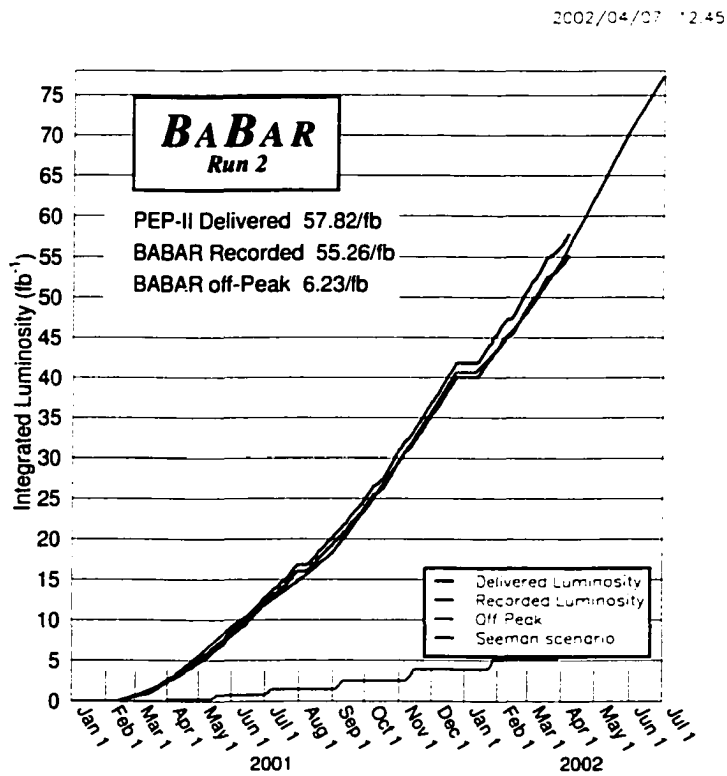


Figure 5.2: Data taken at *BABAR* over the past 15 months as compared with model (violet) predicted on March 29, 2001.

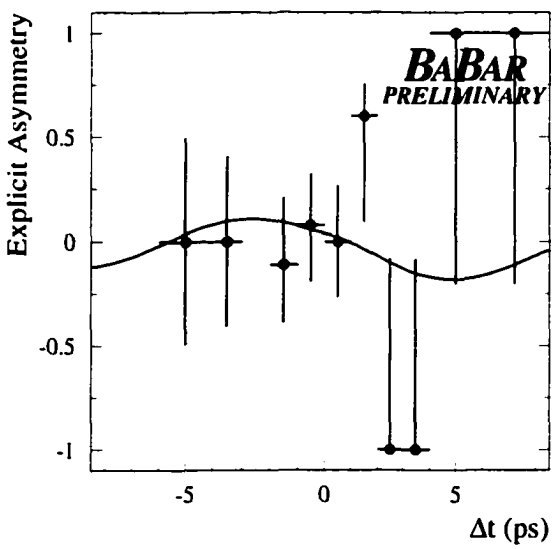


Figure 5.3: *CP* violating time-dependent asymmetry in $B^0 \to D^{*-}D^+$ using 56 fb^{-1} of *BABAR* data. Fit is normalized to relative numbers of B^0 and $\overline{B}^0$ seen in data in this mode.

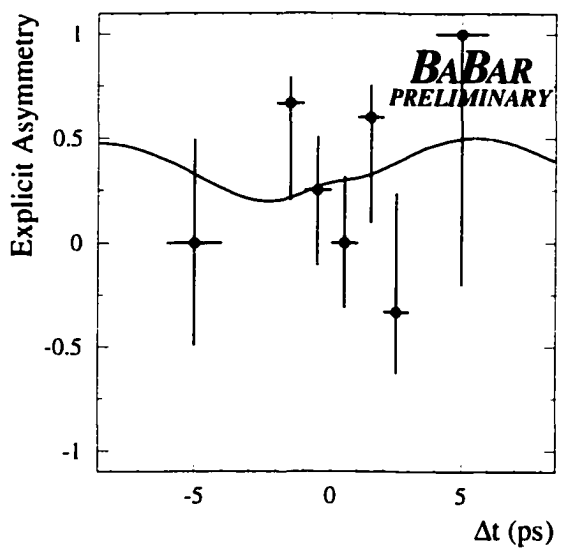


Figure 5.4: *CP* violating time-dependent asymmetry in $B^0 \longrightarrow D^{*+}D^-$ using 56 fb^{-1} of *BABAR* data. Fit is normalized to relative numbers of B^0 and $\overline{B}^0$ seen in data in this mode.

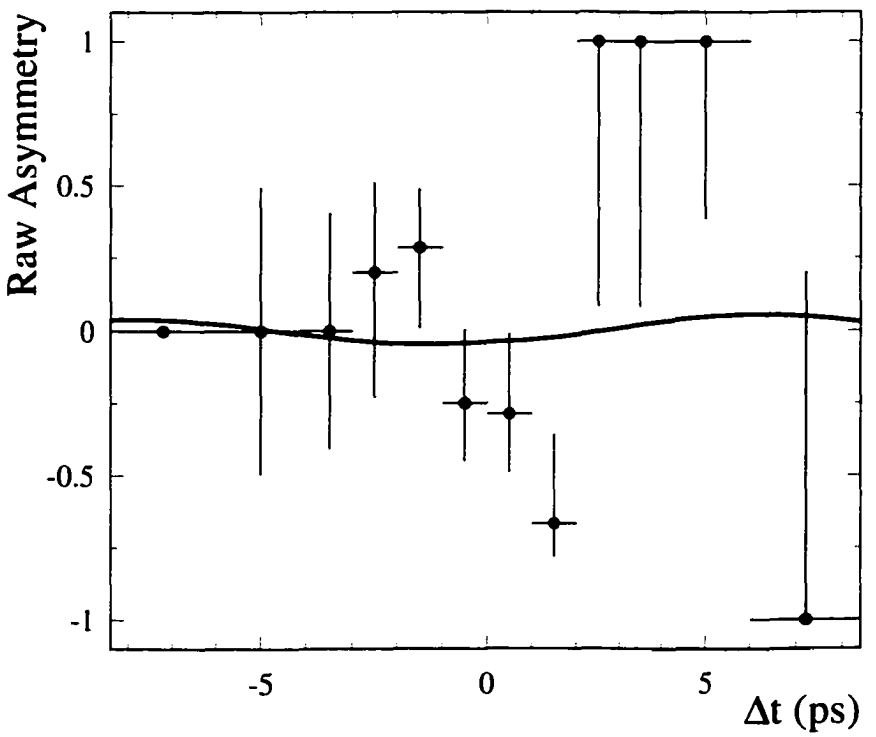


Figure 5.5: *CP* violating time-dependent asymmetry in $B^0 \to D^{*+}D^{*-}$ using 56 fb^{-1} of *BABAR* data.

Appendix A

Details of $B \to D^{(*)}\overline{D}^{(*)}$ Branching Fraction Measurements

A.1 Data and Monte Carlo Samples

A.1.1 Data Sample

The data sample used for this analysis corresponds to Run 1 of *BABAR*. A total of 3668 colliding beam runs are used, in the run range of 9931 to 17106. This data sample represents an integrated luminosity of $22.9\,\mathrm{fb}^{-1}$, with $20.7\,\mathrm{fb}^{-1}$ collected on the $\Upsilon(4S)$ resonance and $2.2\,\mathrm{fb}^{-1}$ collected at center-of-mass energy just below the $\Upsilon(4S)$. The total number of $B\overline{B}$ pairs produced in this sample is [25]

$$N_{B\bar{B}} = (22.515 \pm 0.360) \times 10^6$$

Note that we do not require the same event selection as used by the B-counting group and hence use the efficiency-corrected number of $B\overline{B}$.

One of the features that distinguishes different data samples is the operating high voltage used in the drift chamber. The primary effect is seen in the corrections applied to the Monte Carlo to account for differences seen in the tracking efficiency (see Section A.7). In our data sample, 54.4% of the luminosity was collected with a drift chamber voltage of 1900 V, the rest at a voltage of 1960 V[27].

A.1.2 Monte Carlo Samples

For optimizing and measuring the efficiency of our event selection criteria, we use Monte Carlo simulated events of our signal modes.

For $B^0 \to D^{*+}D^{*-}$, the Monte Carlo for signal was generated as a 'cocktail' of the various D decay modes. The first cocktail for $B^0 \to D^{*+}D^{*-}$ signal was comprised of events where both of the charged D^*s decay via $D^{*+} \to D^0\pi^+$. The D^0s were then allowed to decay into the following modes:

- $D^0 \to K^-\pi^+$
- $D^0 \to K^-\pi^+\pi^0$
- $D^0 \to K^-\pi^+\pi^-\pi^+$
- $D^0 \to K^0_S\pi^+\pi^-$
- $D^0 \to K^0_S\pi^0$
- $D^0 \to K^-K^+$

Each D^0 decay mode was generated with the same branching fraction.

For $B^0 \to D^{*\pm}D^{\mp}$, the Monte Carlo for signal was generated as a set of separate runs for each of the various D subdecay modes. These $B^0 \to D^{*+}D^-$ signal runs were comprised of events where the D^* decays via either $D^{*+} \to D^0\pi^+$ or $D^{*+} \to D^+\pi^0$. Charged D's were allowed to decay via the following 3 modes:

- $D^+ \to K^-\pi^+\pi^+$
- $D^+ \to K_S^0\pi^+$
- $D^+ \to K^-K^+\pi^+$

The D^0 from $D^{*+} \to D^0\pi^+$ was allowed to decay via the following 4 modes:

- $D^0 \to K^-\pi^+$
- $D^0 \to K^-\pi^+\pi^0$
- $D^0 \to K^-\pi^+\pi^-\pi^+$
- $D^0 \to K_S^0\pi^+\pi^-$

This gives a total of 21 subdecay modes generated: 3 for the bachelor charged D multiplied by $4 + 3 = 7$ for the $D^{*+} \to D^0\pi^+$ or $D^{*+} \to D^+\pi^0$.

The cocktail for $B^{\pm} \to D^{*\pm}D^{*0}$ signal was comprised of events where the charged D^* decays via either $D^{*+} \to D^0\pi^+$ or $D^{*+} \to D^+\pi^0$ (in the ratio 0.75:0.25) and the D^{*0} decays via either $D^{*0} \to D^0\pi^0$ or $D^{*0} \to D^0\gamma$ (in the ratio 0.667:0.333). D^0s were then allowed to decay into the following modes:

- $D^0 \to K^-\pi^+$

- $D^0 \to K^-\pi^+\pi^0$
- $D^0 \to K^-\pi^+\pi^-\pi^+$
- $D^0 \to K_S^0\pi^+\pi^-$
- $D^0 \to K_S^0\pi^0$
- $D^0 \to K^-K^+$

Each D^0 decay mode was generated with the same branching fraction. D^+s were allowed to decay into the following modes:

- $D^+ \to K^-\pi^+\pi^+$
- $D^+ \to K_S^0\pi^+$
- $D^+ \to K^-K^+\pi^+$
- $D^+ \to K_S^0\pi^+\pi^0$
- $D^+ \to K^-K^+\pi^+\pi^0$
- $D^+ \to K_S^0\pi^+\pi^-\pi^+$

Each D^+ decay mode was generated with the same branching fraction.

A summary of the total Monte Carlo sample used is given in Section A.7.

For studying background and optimizing event selection, Monte Carlo simulated generic $B^0\overline{B}^0$, B^+B^-, and $c\bar{c}$ events are also used. The total number of events simulated of each type and the corresponding effective luminosity are summarized in Table A.1.

Table A.1: Breakdown of Monte Carlo events used for background studies.

Event Type	Events Generated ($\times 10^6$)	Corresponding Luminosity (fb^{-1})
$B^0\overline{B}^0$	4.2	8.0
$B^+ B^-$	6.7	12.8
$c\bar{c}$	8.9	6.8

A.2 Event Pre-Selection

Pre-selection starts from events assigned to the `isPhysicsEvents` stream in online prompt reconstruction (OPR). This stream includes events assigned to any of the physics sub-streams by the background filter (BGF). We do not make any explicit requirement that events belong to the `BGFMultiHadron` sub-stream, but our requirements on charged track multiplicity and event shape are tighter than those for the `BGFMultiHadron` list (3 or more tracks, $R_2 < 0.6$). Finally, the angle between the reconstructed B and the thrust axis of the rest of the event is defined and we require $cos(\theta_{thrust}) < 0.9$.

We also do not require that events pass the B Counting tag bit. This tag bit is used by the B Counting group to obtain a statistical estimate of the total number of B's produced in the data set [25], rather than a direct count of all B's, so it is not necessary that signal events form a subset of the events which pass this tag.

Table A.2: D^* and D^{*0} decay modes and branching fractions [24].

Particle	Decay Mode	Branching Fraction (%)
D^{*+}	$D^{*+} \to D^0\pi^+$	67.7 ± 0.5
	$D^{*+} \to D^+\pi^0$	30.7 ± 0.5
Total D^{*+} Branching Fraction		98.4
Particle	**Decay Mode**	**Branching Fraction (%)**
D^{*0}	$D^{*0} \to D^0\pi^0$	61.9 ± 2.9
	$D^{*0} \to D^0\gamma$	38.1 ± 2.9
Total D^{*0} Branching Fraction		100.0

A.3 Charmed Meson Reconstruction

A.3.1 D^* Decay Modes

The D^{*+} meson is reconstructed in its decays into $D^{*+} \to D^0\pi^+$ and $D^{*+} \to D^+\pi^0$ and the D^{*0} meson is reconstructed in its decays into $D^{*0} \to D^0\pi^0$ and $D^{*0} \to D^0\gamma$. In$B^0 \to D^{*+}D^{*-}$, for decay modes with a D^{*+} in the final state, the D^{*+} meson is reconstructed in its decays into $D^{*+} \to D^0\pi^+$ and $D^{*+} \to D^+\pi^0$. We include for this analysis the decay combinations D^{*+} D^{*-} decaying to (D^0 π^+,$\overline{D}^0$ π^-) or (D^0 π^+,D^- π^0), but not (D^+ π^0,D^- π^0) due to the smaller branching fraction and expected high backgrounds. The branching fractions for these modes are summarized in Table A.2.

Table A.3: D^0 and D^+ decay modes and branching fractions [24]. The branching fraction for $K_S^0 \to \pi^+\pi^-$ is included for modes containing a K_S^0.

Decay Mode	Branching Fraction (%)
$D^0 \to K^-\pi^+$	3.83 ± 0.09
$D^0 \to K^-\pi^+\pi^0$	13.9 ± 0.9
$D^0 \to K^-\pi^+\pi^+\pi^-$	7.49 ± 0.31
$D^0 \to K_S^0\pi^+\pi^-$	1.85 ± 0.14
Total D^0 Branching Fraction	27.1
Decay Mode	**Branching Fraction (%)**
$D^+ \to K^-\pi^+\pi^+$	9.0 ± 0.6
$D^+ \to K_S^0\pi^+$	0.99 ± 0.09
$D^+ \to K^-K^+\pi^+$	0.87 ± 0.07
Total D^+ Branching Fraction	10.9

A.3.2 D^0 and D^+ Decay Modes

The decay modes of the D^0 and D^+ used in this analysis were selected by an optimization of $S^2/(S+B)$ based on Monte Carlo. A more detailed description of this procedure will be given in Section A.4. The D^0 and D^+ modes used and their branching fractions are summarized in Table A.3.

A.3.3 Charged Pion Selection

The selection of all charged pions, other than those from daughters of K_S^0, is based on the `GoodTracksVeryLoose` selection.

Charged pion daughters of K_S^0 are from the loosest selection, `ChargedTracks`, a list containing all tracks reconstructed in the SVT, DCH, or both subdetectors.

The tracks used for the `GoodTracksVeryLoose` list are a subset of the `ChargedTracks` list with the following:

1. a maximum momentum measured in the lab frame of 10 GeV/c, and
2. a distance of closest approach to the per-run nominal beam spot centroid of $\sqrt{\Delta x^2 + \Delta y^2} < 1.5$ cm, and $|\Delta z| < 10$ cm.

A.3.4 Charged Kaon Identification

The kaon selection, `KMicroNotPionGTL`, is used for the kaon daughter for the following D decays:

- $D^0 \to K^-\pi^+\pi^0$
- $D^0 \to K^-\pi^+\pi^-\pi^+$
- $D^+ \to K^-\pi^+\pi^+$

A particle passing the `KMicroNotPionGTL` selection must not have a Cerenkov ring consistent with a pion in the DIRC, or not have dE/dx consistent with a pion in the DCH and SVT. Details of this selection are given in BAD #116.

The tighter kaon selection, `KMicroTight`, is used for one of the kaon daughters for the decay $D^+ \to K^-K^+\pi^+$. The other kaon daughter is taken from the `KMicroNotPion` list.

The kaon from the decay $D^0 \to K\pi$ has no kaon ID applied — it is instead is taken from the `GoodTracksLoose` list since this mode is background-free enough not to require kaon ID. These tracks are a subset of the `GoodTracksVeryLoose` list and are required to satisfy the following additional cuts:

1. a minimum transverse momentum of 100 MeV/c, and
2. at least 12 hits recorded in the drift chamber, out of a possible maximum of 40 hits for tracks perpendicular to the beam pipe.

A.3.5 K_S^0 Selection

K_S^0 are taken from the standard `KsDefault` list, which contains K_S^0 reconstructed in the mode $K_S^0 \to \pi^+\pi^-$, where the pion candidates are taken from the `ChargedTrack` list. These candidates contain a vertexing constraint from the GeoKin vertex fit (requiring only that the fit converge). The invariant mass computed at this vertex location is required to lie within ± 25 MeV/c^2 of the nominal K_S^0 mass.

A.3.6 π^0 Selection

π^0 from D^0 and charged D decays are taken from the `piOAllDefault` list. π^0 from `piOAllDefault` may be either standard or "merged" π^0 (the photons from a merged π^0 share a cluster in the EMC). The π^0 candidates are formed by combining pairs of photon candidates from the `GoodPhotonLoose` list. The

photon energy threshold is set at 30 MeV and a cut $LAT < 0.8$ is applied to reject the background. The invariant mass of the photon pair is required to be within a window around the nominal π^0 mass of ± 35 MeV/c^2, and a mass constrained fit is applied for subsequent reconstruction of D candidates. Also, the π^0 energy is required to be greater than 200 MeV.

Slow π^0 s from $D^{*+} \rightarrow D^+\pi^0$ are taken from the `pi0SoftDefaultMass` list. These candidates are required to have $E(\gamma) > 30$ MeV and $LAT < 0.8$, as well as a composite $\gamma\gamma$ momentum in the center of mass frame of 70 MeV/c $< p^* <$ 450 MeV/c. Again, an invariant mass cut around the nominal π^0 mass of ± 35 MeV/c^2 and a mass-constrained fit are applied.

Further details about π^0 selection can be found in BAD #20.

A.3.7 D^0 and charged D Selection

D^0 candidates are taken from the `D0DefaultMass` list. The invariant mass for D^0 in this list is required to be within ± 20 MeV/c^2 of the nominal D^0 mass. This cut is used for all D^0 modes except $K^- \pi^+ \pi^0$, which has a looser cut of 35 MeV/c^2 due to the π^0 resolution. Charged D candidates are taken from `DcDefaultMass` list. Charged D candidates from this list are required to have an invariant mass within ± 20 MeV/c^2 of the nominal charged D mass. Also, a vertex fit is performed for both D^0 and charged D candidates using `GeoKin`, where a χ^2 probability greater than 0.1% is required.

A.3.8 $D^{*\pm}$ and D^{*0} Selection

$D^{*\pm}$ candidates are taken from the `DstarDefaultMass` list, where D^0 and charged D candidates are combined with a slow pion candidate to produce

D^* candidates. D^{*0} candidates are taken from the `DstarODefaultMass` list, where D^0 candidates are combined with a slow π^0 or γ to produce D^{*0}s. Before being combined with the slow pion a mass constrained fit is applied to the D^0 and charged D candidates. `GeoKin` is used to perform a vertex fit for the $D^{*\pm}$ using the constraint of the beam spot to improve the angular resolution for the soft pion or photon. The fit is not required to converge, and no cut is applied on the probability of χ^2. After fitting a $D^{*\pm}$, a cut of total width 11 MeV/c^2, centered on the nominal Δm (the $D^{*\pm}$-D mass difference), is applied to D^0 π candidates, and a cut of total width 30 MeV/c^2, again centered on the nominal Δm, is applied to charged D π^0 candidates. D^{*0} candidates (from either D^0 π^0 or D^0 γ) also have a cut applied of total width 30 MeV/c^2, centered on the nominal D^{*0} Δm. The bachelor D candidate (but not the D^*) is required to have $p^* > 1.3$ GeV/c (the `DstarHardDefaultMass` list). A mass constrained fit is also applied to these candidates for use in the subsequent reconstruction of B^0 and $B^\pm$.

A.4 B Meson Reconstruction

B meson candidates are constructed by combining a D^* candidate and a D^*, charged D, or D^{*0} which have passed the selection criteria described previously. The primary variables used to distinguish our signal from background are the energy-substituted mass,

$$m_{\rm ES} \equiv \sqrt{\left(\frac{\sqrt{s}}{2}\right)^2 - {p_B^*}^2}$$

and the difference of the B candidate energy from beam energy,

$$\Delta E \equiv E_B^* - \frac{\sqrt{s}}{2}$$

(used for $B^0 \to D^{*+}D^{*-}$) or

$$\Delta E' \equiv E_B^* - \frac{\sqrt{s}}{2} + m_{\rm ES} - m_B^{PDG}$$

(used for $B^0 \to D^{*\pm}D^{\mp}$ and $B^{\pm} \to D^{*\pm}D^{*0}$) where the star indicates variables evaluated in the center-of-mass frame and the last two of the four $\Delta E'$ terms are introduced for the $B^0 \to D^{*\pm}D^{\mp}$ and $B^{\pm} \to D^{*\pm}D^{*0}$ analyses as a correction to avoid ΔE -$m_{\rm ES}$ correlation (these analyses were done after $B^0 \to D^{*+}D^{*-}$ and included this additional correction, which is advantageous for analysis optimization but not critical). The benefits of this correction are shown in Figure A.1. Figures A.2, A.3, and A.4 show the distributions for $\Delta E'$ and $m_{\rm ES}$ for signal Monte Carlo (where each subdecay mode contribution is not weighted by individual branching ratios). Tables A.4, A.5, and A.6 give resolutions for $\Delta E'$ and $m_{\rm ES}$ for the individual submodes of each channel.

Instead of cutting on individual D and D^* masses, we form a χ^2-like quantity, χ^2_{Mass}, that includes all D and D^* masses. For $B^0 \to D^{*+}D^{*-}$ and $B^{\pm} \to D^{*\pm}D^{*0}$, χ^2_{Mass} contains 4 terms:

$$\chi^2_{Mass} = \left(\frac{m_D - m_{D_{PDG}}}{\sigma_{m_D}}\right)^2 + \left(\frac{m_{\bar{D}} - m_{\bar{D}_{PDG}}}{\sigma_{m_{\bar{D}}}}\right)^2 + \left(\frac{\Delta m_{D^*} - \Delta m_{D^*_{PDG}}}{\sigma_{\Delta m_{D^*}}}\right)^2 + \left(\frac{\Delta m_{D^{*(0)}} - \Delta m_{D^{*(0)}_{PDG}}}{\sigma_{\Delta m_{D^{*(0)}}}}\right)^2$$

For $B^0 \to D^{*\pm}D^{\mp}$, χ^2_{Mass} contains 3 terms:

$$\chi^2_{Mass} = \left(\frac{m_D - m_{D_{PDG}}}{\sigma_{m_D}}\right)^2 + \left(\frac{m_{D_{D^*}} - m_{D_{PDG}}}{\sigma_{m_{D_{D^*}}}}\right)^2 + \left(\frac{\Delta m_{D^*} - \Delta m_{D^*_{PDG}}}{\sigma_{\Delta m_{D^*}}}\right)^2$$

Table A.4: $B^0 \to D^{*+}D^{*-}$ resolutions for ΔE and $m_{\rm ES}$ in individual submodes (signal Monte Carlo).

mode	$m_{\rm ES}$ resolution (MeV)	ΔE resolution (MeV)
$(K\pi)$ $(K\pi)$	2.43 ± 0.16	7.34 ± 0.77
$(K\pi)$ $(K\pi\pi^0)$	2.36 ± 0.22	7.43 ± 0.52
$(K\pi)$ $(K3\pi)$	2.34 ± 0.11	6.83 ± 0.33
$(K\pi)$ $(K^0_S\pi\pi)$	2.67 ± 0.48	7.19 ± 1.73
$(K\pi)$ $(K\pi\pi)$	2.30 ± 0.05	7.08 ± 0.18
$(K\pi)$ $(K^0_S\pi)$	2.33 ± 0.07	7.22 ± 0.21
$(K\pi)$ $(KK\pi)$	2.24 ± 0.05	7.08 ± 0.25
$(K\pi\pi^0)$ $(K\pi\pi^0)$	2.47 ± 0.09	9.91 ± 0.42
$(K\pi\pi^0)$ $(K3\pi)$	2.35 ± 0.22	10.25 ± 0.89
$(K\pi\pi^0)$ $(K^0_S\pi\pi)$	2.10 ± 0.22	7.47 ± 0.82
$(K\pi\pi^0)$ $(K\pi\pi)$	2.27 ± 0.09	7.33 ± 0.26
$(K\pi\pi^0)$ $(K^0_S\pi)$	2.56 ± 0.16	10.24 ± 0.71
$(K\pi\pi^0)$ $(KK\pi)$	2.09 ± 0.08	8.62 ± 0.39
$(K3\pi)$ $(K3\pi)$	2.27 ± 0.07	6.52 ± 0.22
$(K3\pi)$ $(K^0_S\pi\pi)$	2.87 ± 0.33	8.11 ± 0.95
$(K3\pi)$ $(K\pi\pi)$	2.40 ± 0.11	7.47 ± 0.35
$(K3\pi)$ $(K^0_S\pi)$	2.40 ± 0.14	7.59 ± 0.53
$(K3\pi)$ $(KK\pi)$	2.44 ± 0.12	6.40 ± 0.35
$(K^0_S\pi\pi)$ $(K^0_S\pi\pi)$	2.17 ± 0.13	7.56 ± 0.62
$(K^0_S\pi\pi)$ $(K\pi\pi)$	2.27 ± 0.09	6.17 ± 0.24
$(K^0_S\pi\pi)$ $(K^0_S\pi)$	2.32 ± 0.18	7.46 ± 0.43
$(K^0_S\pi\pi)$ $(KK\pi)$	2.21 ± 0.11	7.28 ± 0.38

where the subscript PDG refers to the nominal value, and Δm is the $D^* - D$ or $D^{*0} - D$ mass difference. For σ_{m_D} we use errors calculated candidate by candidate. For $\sigma_{\Delta m_{D^*}}$ we use fixed values of 0.83 MeV for $D^{*+} \rightarrow D^0\pi^+$ and 1.18 MeV for $D^{*+} \rightarrow D^+\pi^0$, and for $\sigma_{\Delta m_{D^{*0}}}$ we use fixed values of 1.32 MeV for $D^{*0} \rightarrow D^0\pi^0$ and 6.73 MeV for $D^{*0} \rightarrow D^0\gamma$. These values are derived from fits to Monte Carlo Δm distributions shown in Figures A.5 and A.6. The fit function is the sum of two Gaussians; the resolution we use is a weighted average calculated from the square of sigmas obtained from the two individual Gaussians.

When reconstructing Monte Carlo events, we use the σ_{m_D} and $\sigma_{\Delta m}$ values as is, however for reconstructing data we scale σ_{m_D} by 1.2 to account for the different resolutions seen in data and Monte Carlo. No additional smearing is applied to the Δm resolutions. These smearing factors are taken from pull distributions found in Ref. [26].

Some events contain more than one B candidate. If several candidates within an event each satisfy the χ^2_{Mass} cut for their respective modes, we choose only the one with the smallest value of χ^2_{Mass} in order to prevent double counting.

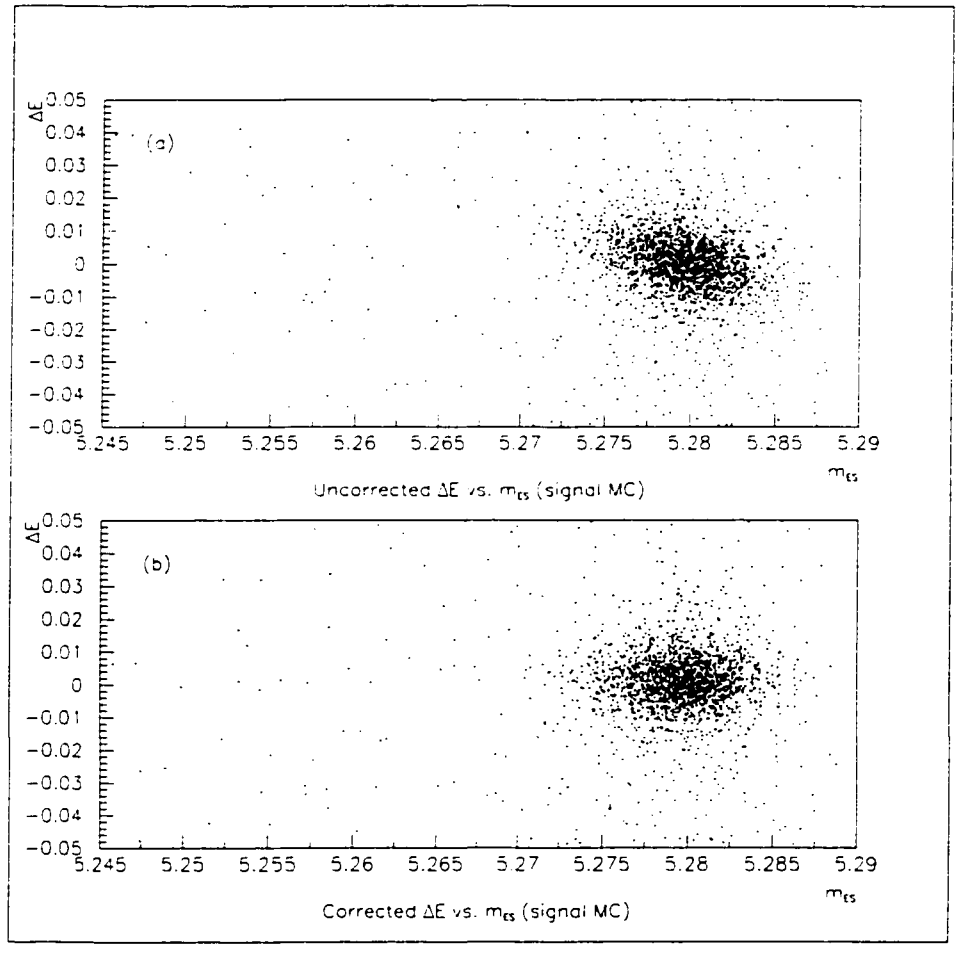


Figure A.1: Signal Monte Carlo plots of ΔE vs. $m_{\rm ES}$ for (a) ΔE $(\equiv E_B^* - \frac{\sqrt{s}}{2})$ and (b) $\Delta E'$ $(\equiv E_B^* - \frac{\sqrt{s}}{2} + m_{\rm ES} - m_B^{PDG})$ (where $m_B^{PDG} = 5.279 GeV$). Note the tilted distribution of signal for the uncorrected ΔE , indicating a non-optimal correlation with $m_{\rm ES}$. Adding the $\Delta E'$ correction removes this correlation, producing the untilted distribution shown in (b), and improves the ΔE resolution. This is especially necessary when using an elliptical cut with $\Delta E^{(\prime)}$ and $m_{\rm ES}$ as the ellipse axes, as described below.

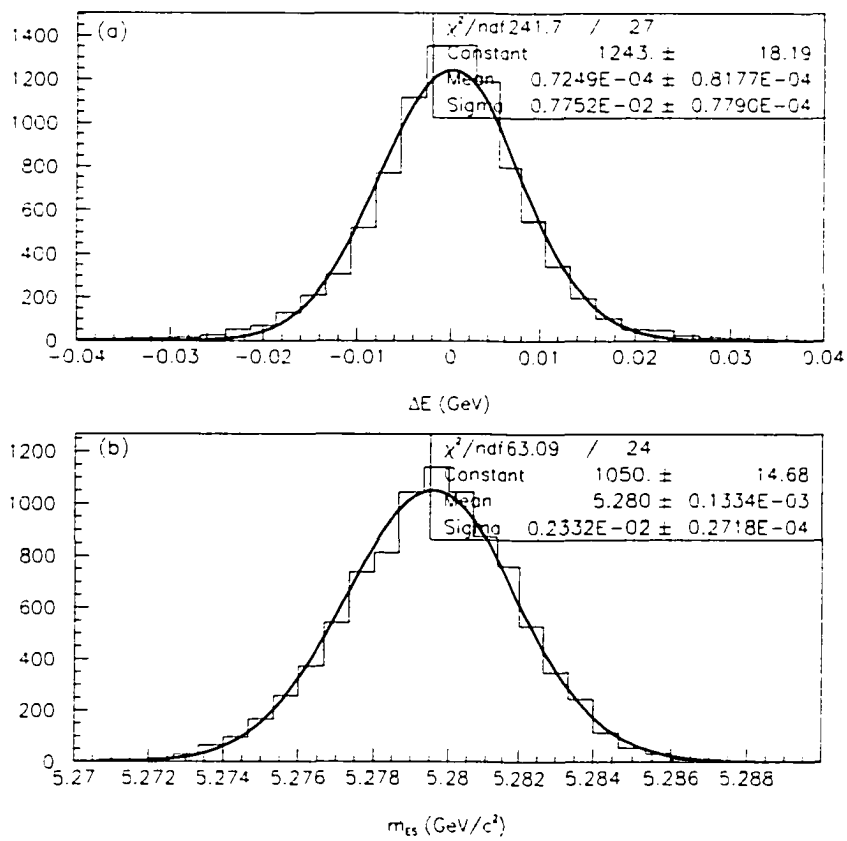


Figure A.2: Signal Monte Carlo distributions in $B^0 \to D^{*+}D^{*-}$ for (a) ΔE and (b) $m_{\rm ES}$. (Note: each subdecay mode contribution is not weighted by individual branching ratios).

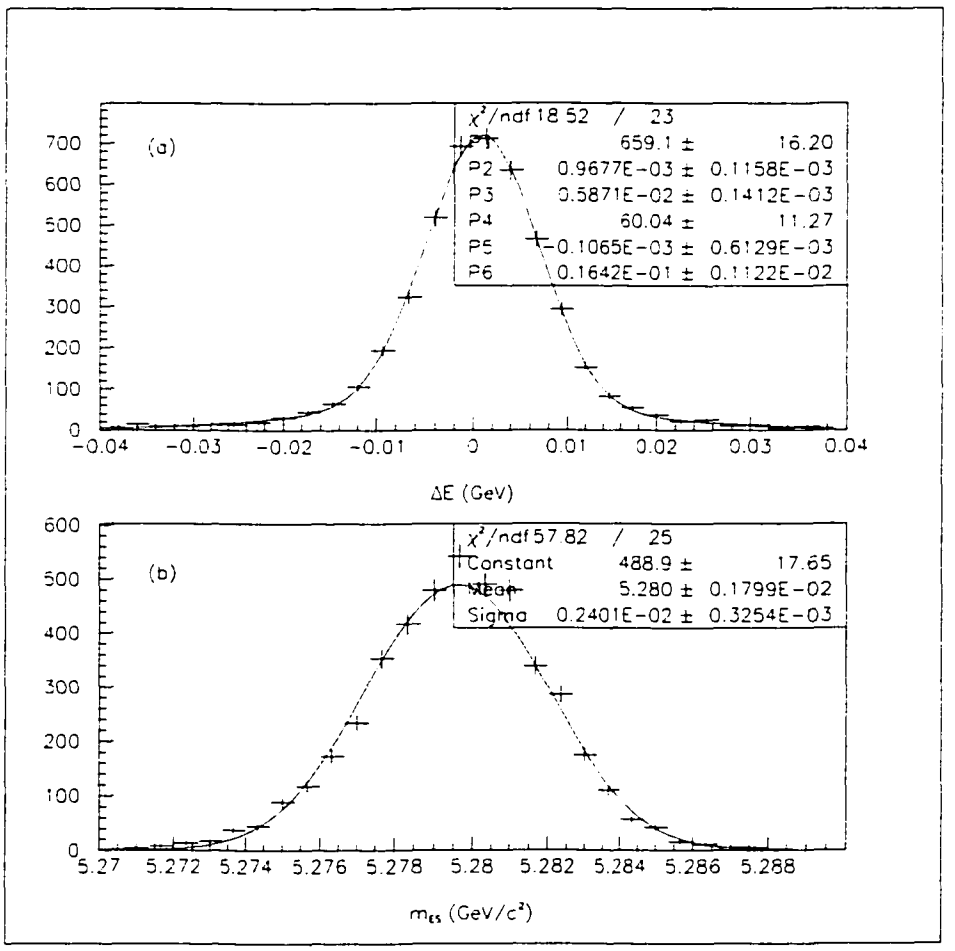


Figure A.3: Signal Monte Carlo distributions in $B^0 \to D^{*\pm}D^{\mp}$ for (a) $\Delta E'$ and (b) $m_{\rm ES}$ (note: each subdecay mode contribution is not weighted by individual branching ratios). $\Delta E'$ is best fit with a double gaussian (due both to differing resolutions from $D^{*0} \to D^0\pi^0$ and $D^{*0} \to D^0\gamma$ and to submode crossfeed).

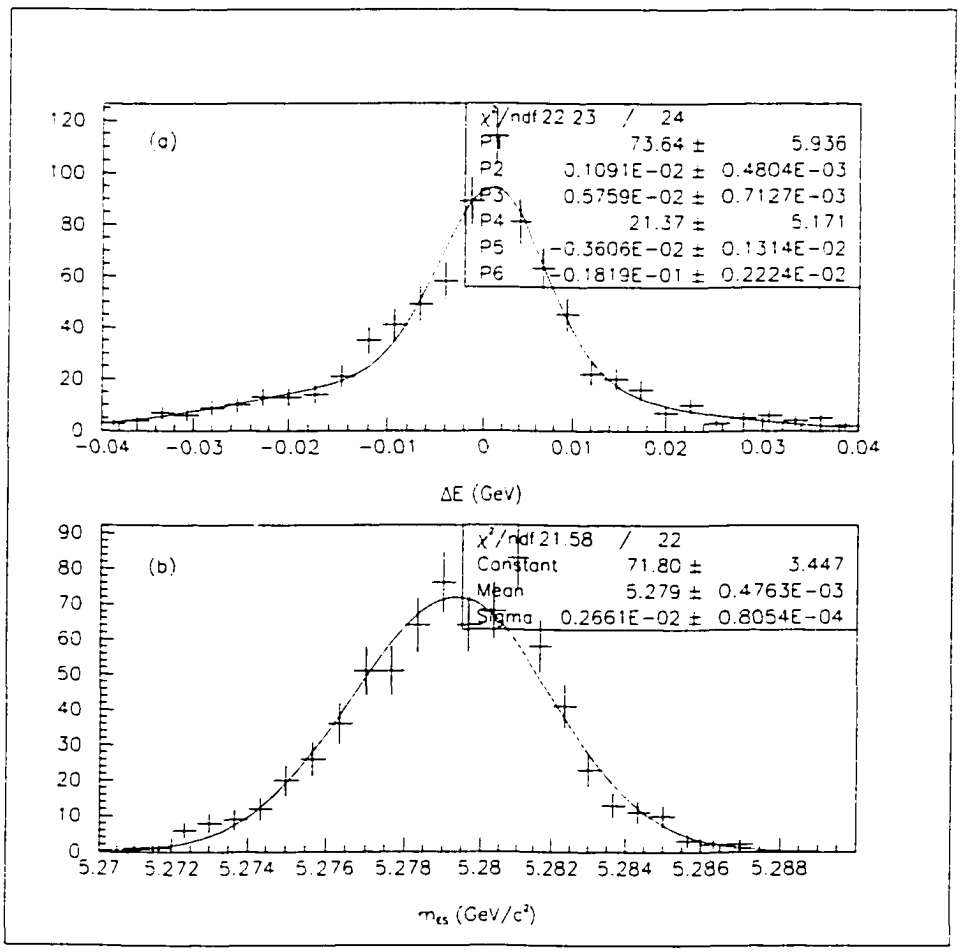


Figure A.4: Signal Monte Carlo distributions in $B^{\pm} \to D^{*\pm}D^{*0}$ for (a) $\Delta E'$ and (b) $m_{\rm ES}$ (note: each subdecay mode contribution is not weighted by individual branching ratios). $\Delta E'$ is best fit with a double gaussian (due both to differing resolutions from $D^{*0} \to D^0\pi^0$ and $D^{*0} \to D^0\gamma$ and to submode crossfeed).

Table A.5: $B^0 \to D^{*\pm}D^{\mp}$ resolutions for $\Delta E'$ and $m_{\rm ES}$ in individual submodes (signal Monte Carlo, for the submodes which contained enough statistics for an individual fit).

mode		$m_{\rm ES}$ resolution	$\Delta E'$ resolution
D^*	$D^{\pm}$	(MeV)	(MeV)
$(K\pi)\pi$	$K\pi\pi$	2.25 ± 0.08	6.56 ± 0.25
$(K\pi)\pi$	$K^0_S\pi$	2.34 ± 0.08	6.45 ± 0.23
$(K\pi)\pi$	$KK\pi$	2.27 ± 0.10	6.12 ± 0.21
$(K\pi\pi^0)\pi$	$K\pi\pi$	2.32 ± 0.09	8.46 ± 0.59
$(K\pi\pi^0)\pi$	$K^0_S\pi$	1.90 ± 0.25	9.10 ± 1.46
$(K3\pi)\pi$	$K\pi\pi$	2.29 ± 0.15	5.74 ± 0.29
$(K3\pi)\pi$	$K^0_S\pi$	2.18 ± 0.28	6.34 ± 0.61
$(K3\pi)\pi$	$KK\pi$	2.27 ± 0.20	6.54 ± 0.69
$(K^0_S\pi\pi)\pi$	$K\pi\pi$	2.40 ± 0.14	6.47 ± 0.39
$(K^0_S\pi\pi)\pi$	$K^0_S\pi$	2.72 ± 0.18	5.75 ± 0.43
$(K^0_S\pi\pi)\pi$	$KK\pi$	2.48 ± 0.37	7.10 ± 1.27
$(K\pi\pi)\pi^0$	$K\pi\pi$	2.27 ± 0.11	7.02 ± 0.53
$(K\pi\pi)\pi^0$	$K^0_S\pi$	2.47 ± 0.13	7.71 ± 0.58
$(K\pi\pi)\pi^0$	$KK\pi$	3.28 ± 0.89	4.40 ± 0.76
$(K^0_S\pi)\pi^0$	$K\pi\pi$	2.20 ± 0.13	7.07 ± 0.39
$(K^0_S\pi)\pi^0$	$K^0_S\pi$	2.34 ± 0.13	6.33 ± 0.40
$(K^0_S\pi)\pi^0$	$KK\pi$	2.14 ± 0.17	7.82 ± 0.86

Table A.6: $B^{\pm} \rightarrow D^{*\pm}D^{*0}$ resolutions for $\Delta E'$ and $m_{\rm ES}$ in individual submodes (signal Monte Carlo. for submodes which contained enough statistics for an individual fit).

mode		$m_{\rm ES}$ resolution	$\Delta E'$ resolution
D^*	D^{*0}	(MeV)	(MeV)
$(K\pi)\pi$	$(K\pi)\pi^0$	2.48 ± 0.21	8.40 ± 0.95
$(K\pi)\pi$	$(K\pi\pi^0)\pi^0$	3.60 ± 0.75	8.01 ± 1.76
$(K\pi)\pi$	$(K3\pi)\pi^0$	2.71 ± 1.04	6.97 ± 1.48
$(K\pi)\pi$	$(K\pi)\gamma$	3.35 ± 0.81	8.70 ± 2.15
$(K3\pi)\pi$	$(K\pi)\pi^0$	3.15 ± 0.84	16.20 ± 5.82
$(K^0_S\pi\pi)\pi$	$(K\pi)\pi^0$	3.20 ± 1.44	15.52 ± 5.56

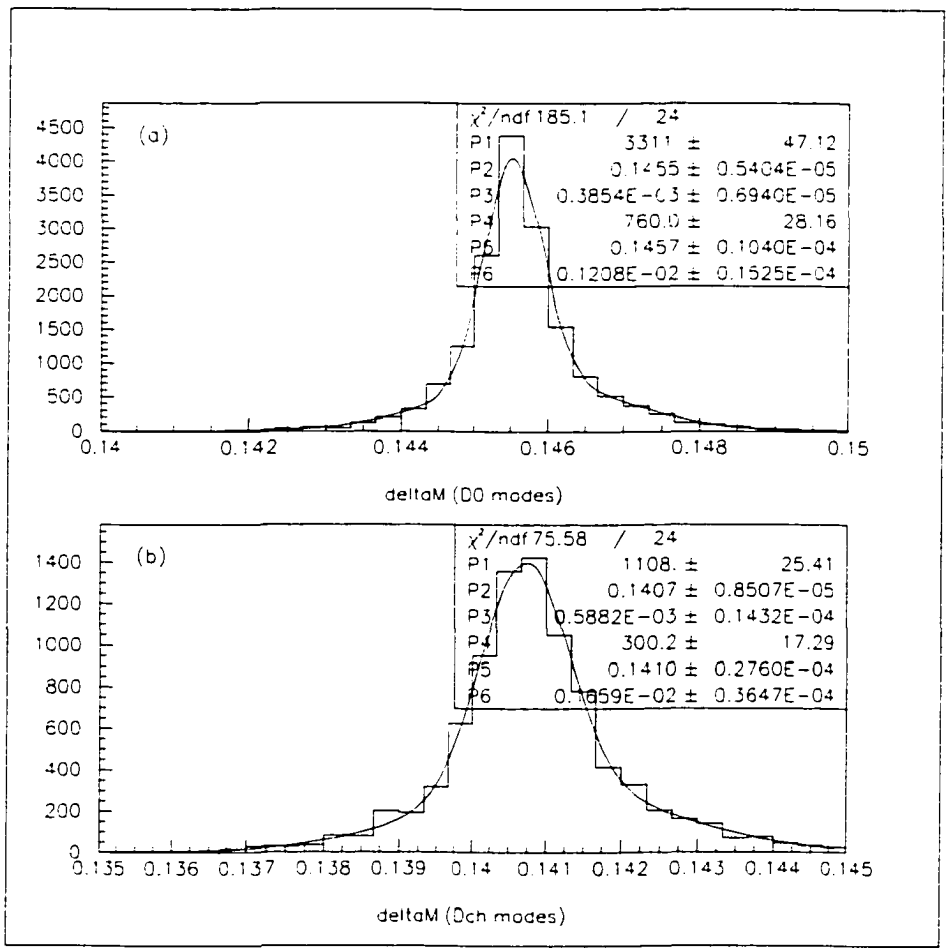


Figure A.5: Double-Gaussian fits to Δm for (a) $D^{*+} \to D^0\pi^+$ and (b) $D^{*+} \to D^+\pi^0$ from signal Monte Carlo (Note: contributions from each final state are not weighted by individual subdecay mode branching ratios).

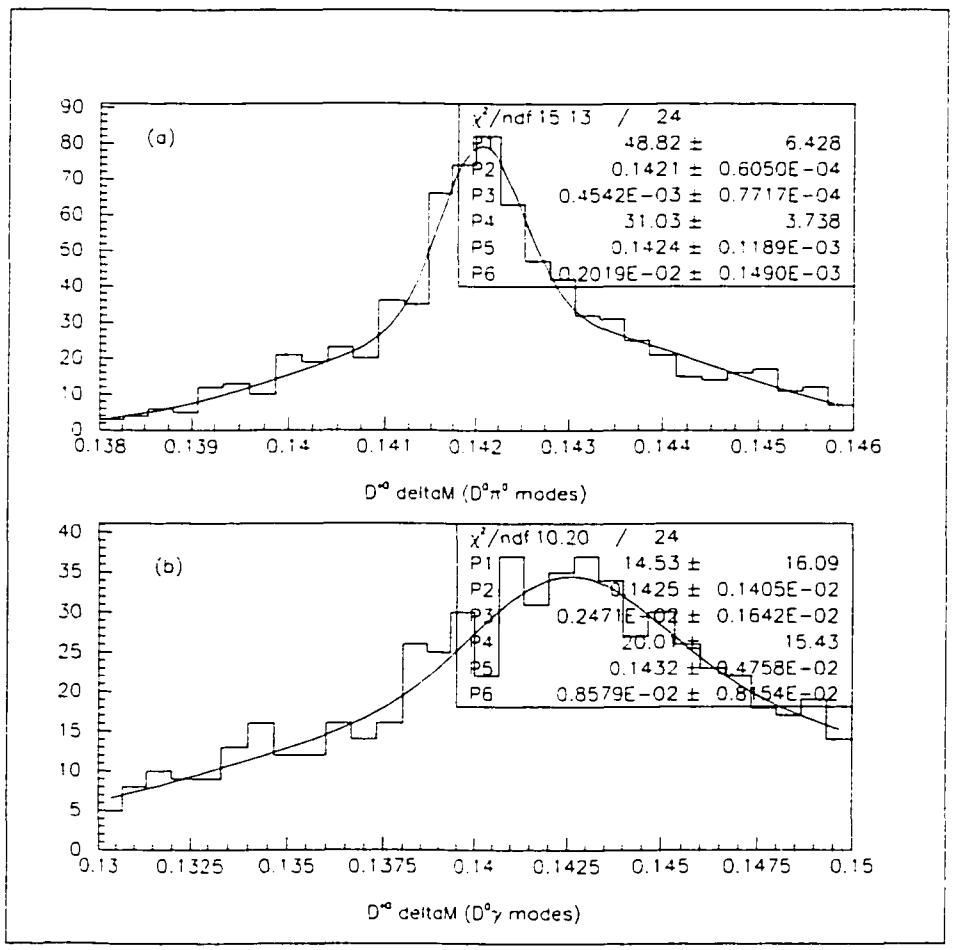


Figure A.6: Double-Gaussian fits to Δm for (a) $D^{*0} \rightarrow D^0\pi^0$ and (b) $D^{*0} \rightarrow D^0\gamma$ from signal Monte Carlo (Note: contributions from each final state are not weighted by individual subdecay mode branching ratios). The beamspot refit is not done for D^{*0} candidates, hence the wider Δm distributions than D^*.

A.5 Analysis Optimization Procedure for $B^0 \to D^{*+}D^{*-}$

A.5.1 Decay Mode Selection

We select which D decay mode combinations to use in this analysis using a $S^2/(S+B)$ optimization. We first determined, based on Monte Carlo, the expected $S-B$ for each decay mode combination individually. We then sort this list and successively add modes to compute an overall $S^2/(S+B)$ value. Once the overall $S^2/(S+B)$ value begins to decrease, we stop including modes.

Figure A.7 shows the results of this analysis. Table A.7 shows the ordering of decay mode combinations outlined previously, and corresponds to the x-axis values in Figure A.7. The last four mode combinations shown all contain, and were the only combinations to contain, the decay mode $D^+ \to K_S^0\pi^+\pi^0$. As a matter of simplicity in which modes are used, we therefore chose to exclude this D^+ mode from the final analysis but use all other combinations listed in Table A.7.

A.5.2 Cut Optimization

The optimization of the analysis cuts was performed by maximizing $S^2/(S+B)$ with respect to the cut in question after all other cuts have been applied. The process was iterated as each cut was re-optimized. The cut tuning was done entirely using the Monte Carlo samples described in Section A.1. For the signal estimation in $B^0 \to D^{*+}D^{*-}$, a branching ratio of 0.1% was assumed.

Table A.7: Ordering of mode combinations based on $S - B$ used for mode selection. (See Figure A.7.)

Rank	Mode Combination
1	$(K\pi\pi^0)(K\pi\pi^0)$
2	$(K\pi\pi^0)(K3\pi)$
3	$(K\pi)(K3\pi)$
4	$(K\pi)(K\pi\pi^0)$
5	$(K\pi\pi^0)(K\pi\pi)$
6	$(K\pi)(K\pi)$
7	$(K3\pi)(K3\pi)$
8	$(K\pi)(K\pi\pi)$
9	$(K3\pi)(K\pi\pi)$
10	$(K\pi)(K_S^0\pi\pi)$
11	$(K\pi\pi^0)(K_S^0\pi\pi)$
12	$(K_S^0\pi\pi)(K\pi\pi)$
13	$(K\pi)(K_S^0\pi)$
14	$(K\pi\pi^0)(KK\pi)$
15	$(K\pi)(KK\pi)$
16	$(K3\pi)(K_S^0\pi)$
17	$(K_S^0\pi\pi)(K_S^0\pi)$
18	$(K_S^0\pi\pi)(KK\pi)$
19	$(K\pi\pi^0)(K_S^0\pi)$
20	$(K3\pi)(K_S^0\pi\pi)$
21	$(K_S^0\pi\pi)(K_S^0\pi\pi)$
22	$(K3\pi)(KK\pi)$
23	$(K\pi)(K_S^0\pi\pi^0)$
24	$(K_S^0\pi\pi)(K_S^0\pi\pi^0)$
25	$(K3\pi)(K_S^0\pi\pi^0)$
26	$(K\pi\pi^0)(K_S^0\pi\pi^0)$

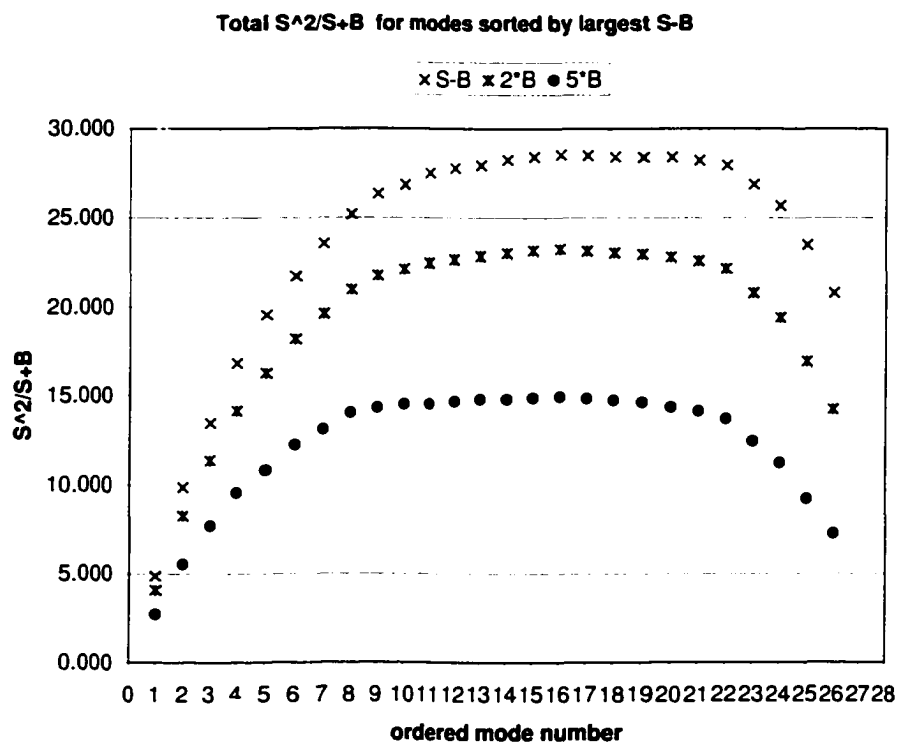


Figure A.7: Optimization of D modes to be reconstructed. The top curve shows the expected $S^2/(S+B)$ at nominal background ($B^0\overline{B}^0$ only), the middle curve at two times nominal background and the lower curve at five times nominal background. All numbers are weighted by appropriate subdecay mode branching ratios and scaled to $19.5 fb^{-1}$.

For the background estimation we use appropriately weighted samples from the $B\overline{B}$ generic and $c\overline{c}$ Monte Carlo. For the background sample we look at the distributions for events in a sideband of the ΔE vs. $m_{\rm ES}$ plane away from the signal region, defined as

$$|\Delta E| < 200 \text{ MeV}$$

$$5.200 \text{ GeV}/c^2 < m_{\text{ES}} < 5.26 \text{ GeV}/c^2$$

and

$$50 \text{ MeV} < |\Delta E| < 200 \text{ MeV}$$

$$5.26 \text{ GeV}/c^2 < m_{\text{ES}} < 5.290 \text{ GeV}/c^2$$

Figure A.8 shows the definition of the sideband region, as well as the signal region (to be described below).

In Figure A.9 we show the results of this optimization for the cut on χ^2_{Mass}. The top plot shows the χ^2_{Mass} distribution for signal Monte Carlo events, the middle plot shows the χ^2_{Mass} distribution for background Monte Carlo, and the bottom plot shows $S^2/(S+B)$ vs. the cut value on χ^2_{Mass}. Here all mode combinations are appropriately weighted by branching ratios and the sample is scaled to the total luminosity of the data sample. In choosing the cut value we take into consideration not only this plot, but also the fact that there are differences in the reconstructed mass resolutions between data and Monte Carlo. We therefore choose to lean towards looser cuts in order to reduce any possible systematic error that may be incurred due to these resolution differences. Our final choice for this cut is

$$\chi^2_{Mass} < 20$$

Figures A.10 and A.11 show the results of the optimization procedure for the ΔE and m_{ES} cuts used to define the signal region. Based on these studies, we choose our signal region to be

$$|\Delta E| < 25 \text{ MeV}$$

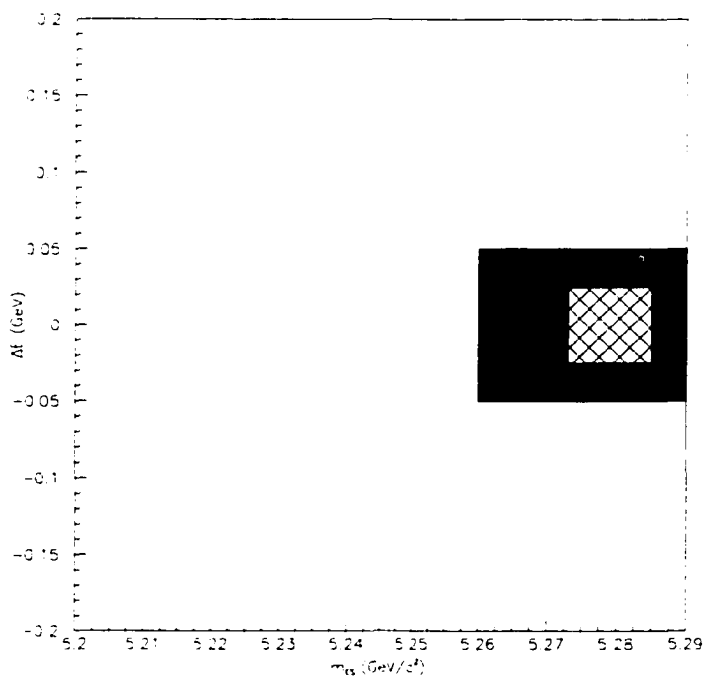


Figure A.8: Definition of sideband and signal regions used. The large unhatched area is the sideband. The smaller hatched box represents the signal region.

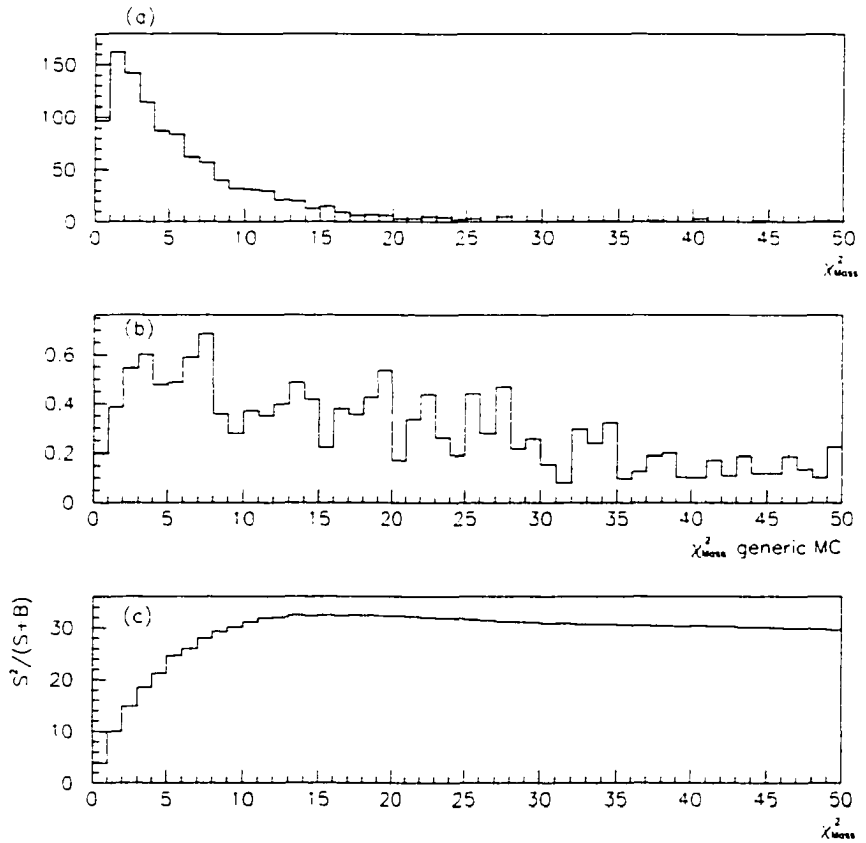


Figure A.9: Optimization of the cut on χ^2_{Mass}. (a) Distribution of χ^2_{Mass} for signal Monte Carlo (not weighted by sub-mode branching ratios). (b) Distribution of χ^2_{Mass} for background Monte Carlo in the ΔE , $m_{\rm ES}$ sideband. (c) $S^2/(S+B)$ vs. cut value for χ^2_{Mass} (includes appropriate weighting for subdecay mode branching ratios and scaled to the luminosity of the data sample).

$$5.273\ \text{GeV}/c^2 < m_{\text{ES}} < 5.285\ \text{GeV}/c^2$$

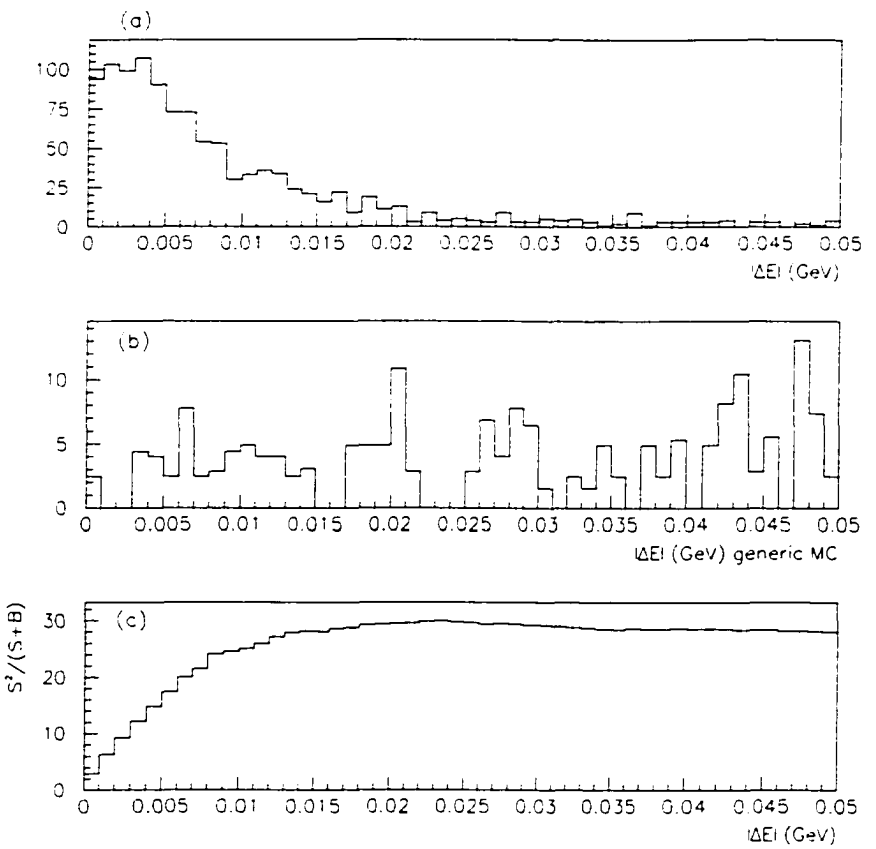


Figure A.10: Optimization of the cut on $\Delta E'$. (a) Distribution of $|\Delta E|$ for signal Monte Carlo (not weighted according to decay mode branching fractions). (b) Distribution of $|\Delta E|$ for background Monte Carlo. (c) $S^2/(S+B)$ vs. $|\Delta E|$ cut (includes appropriate weighting for subdecay mode branching ratios and is scaled to the luminosity of the data sample).

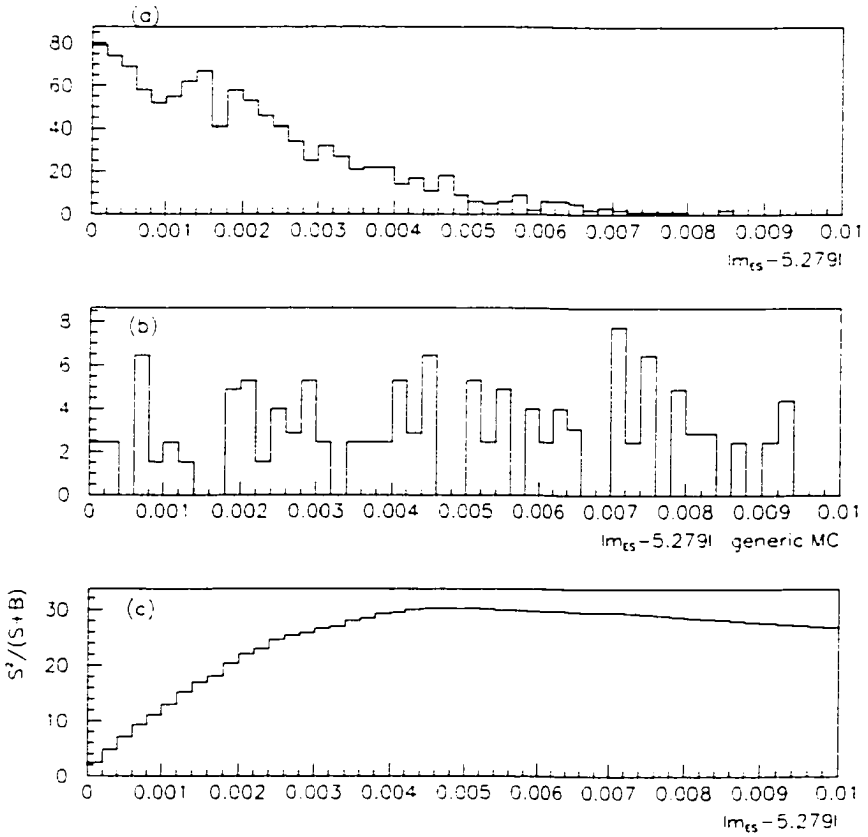


Figure A.11: Optimization of the cut on m_{ES}. (a) Distribution of $|m_{\mathrm{ES}} - m_B|$ for signal Monte Carlo (not weighted by submode branching ratios). (b) Same distribution for background Monte Carlo. (c) $S^2/(S+B)$ vs. cut on $|m_{\mathrm{ES}} - m_B|$ (includes appropriate weights for subdecay mode branching ratios and scaled to the luminosity of the data sample).

A.6 Analysis Optimization Procedure for $B^0 \to D^{*\pm}D^{\mp}$ and $B^{\pm} \to D^{*\pm}D^{*0}$

A.6.1 Submode-Specific Cut Optimization

Cuts are tuned on χ^2_{Mass} (the calculation of which is described in the previous section) for each subdecay mode used based on a $S^2/(S+B)$ optimization. Due to differing amounts of signal and background in each of the subdecay modes, the cuts on χ^2_{Mass} for all individual submodes are treated as parameters and tuned in Monte Carlo such that the overall $S^2/(S+B)$ is maximized (this is a more optimal refinement of the procedure used for $B^0 \to D^{*+}D^{*-}$ in the previous section). This tuning maximizes the predicted significance of the final result, but cannot affect or bias the central value; as such we are not dependent on the accuracy of the Monte Carlo simulation. In the case of $B^0 \to D^{*\pm}D^{\mp}$, we start with all 21 possible submode combinations:

and in the case of $B^{\pm} \to D^{*\pm}D^{*0}$, we start with all 40 possible submode combinations:

The cut tuning was done entirely using the Monte Carlo samples described in Section A.1. For the signal estimation in $B^0 \to D^{*\pm}D^{\mp}$, a branching fraction of 5×10^{-4} was assumed. For the signal estimation in $B^{\pm} \to D^{*\pm}D^{*0}$, a branching fraction of 0.1% was assumed. For the background estimation we use appropriately weighted samples from the $B\overline{B}$ generic and $c\bar{c}$ Monte Carlo. For the background sample we look at the distributions for events in a sideband of the $\Delta E'$ vs. $m_{\rm ES}$ plane away from the signal region, defined

Table A.8: Table of submodes used in χ^2_{Mass} cut optimization.

		$D^{\pm}$		
		$K\pi\pi$	$K^0_S\pi$	$KK\pi$
D^*	$(K\pi)\pi$	$\surd$	$\surd$	$\surd$
	$(K\pi\pi^0)\pi$	$\surd$	$\surd$	$\surd$
	$(K3\pi)\pi$	$\surd$	$\surd$	$\surd$
	$(K^0_S\pi\pi)\pi$	$\surd$	$\surd$	$\surd$
	$(K\pi\pi)\pi^0$	$\surd$	$\surd$	$\surd$
	$(K^0_S\pi)\pi^0$	$\surd$	$\surd$	$\surd$
	$(KK\pi)\pi^0$	$\surd$	$\surd$	$\surd$

as

$$|\Delta E'| < 200 \text{ MeV}$$

$$5.200 \text{ GeV}/c^2 < m_{\text{ES}} < 5.26 \text{ GeV}/c^2$$

and

$$50 \text{ MeV} < |\Delta E'| < 200 \text{ MeV}$$

$$5.26 \text{ GeV}/c^2 < m_{\text{ES}} < 5.290 \text{ GeV}/c^2$$

Figure A.12 shows the definition of the sideband region, as well as the signal region (to be described below).

Each submode's χ^2_{Mass} cut is tuned to maximize the overall $S^2/(S+B)$. We begin with a global χ^2_{Mass} cut of 20. The change in overall $S^2/(S+B)$

Table A.9: Table of submodes used in χ^2_{Mass} cut optimization.

		D^*				
		$(K\pi)\pi^\pm$	$(K\pi\pi^0)\pi^\pm$	$(K3\pi)\pi^\pm$	$(K^0_S\pi\pi)\pi^\pm$	$(K\pi\pi)\pi^0$
D^{*0}	$(K\pi)\pi^0$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
	$(K\pi\pi^0)\pi^0$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
	$(K3\pi)\pi^0$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
	$(K^0_S\pi\pi)\pi^0$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
	$(K\pi)\gamma$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
	$(K\pi\pi^0)\gamma$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
	$(K3\pi)\gamma$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
	$(K^0_S\pi\pi)\gamma$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$

is determined as a function of changes in the χ^2_{Mass} cuts for each of the 21 individual submodes (for $B^0 \to D^{*\pm}D^{\mp}$) and 40 individual submodes (for $B^\pm \to D^{*\pm}D^{*0}$). The value of the χ^2_{Mass} cut that maximizes the overall $S^2/(S+B)$ is selected for each individual submode. After each submode's cut is optimized, the process is then iterated, due to the fact that changes in cuts from other submodes can affect the optimal cut in a given submode. The process fully converges in 3 iterations (the coupling between submodes is not very strong). The final values of the cuts in each of the submodes are shown in Tables A.10 and A.11.

In Figures A.13 and A.9 we show the results of the $B^0 \to D^{*\pm}D^{\mp}$ opti-

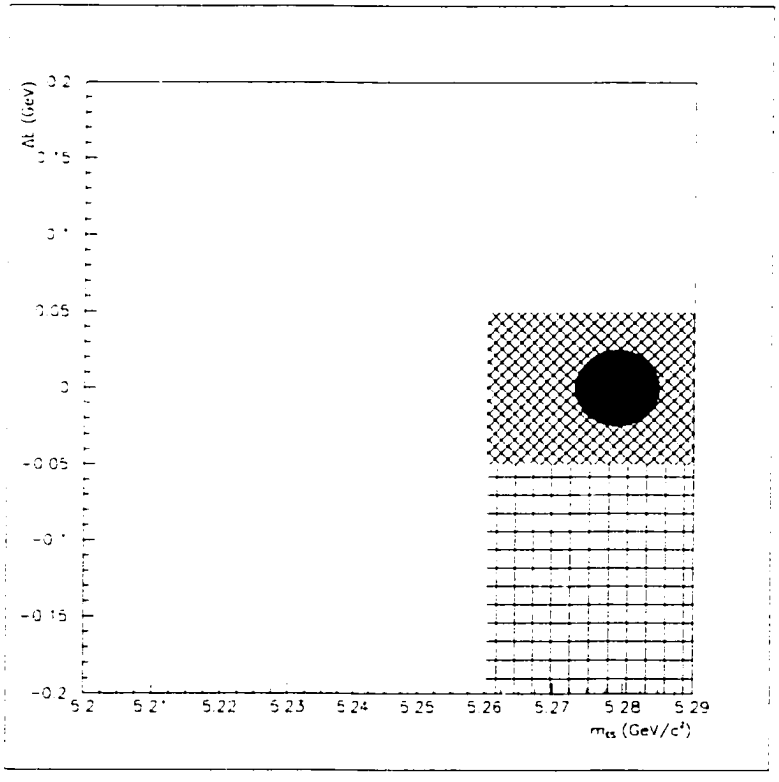


Figure A.12: Definition of signal, sideband, and feed-down regions used for $B^0 \rightarrow D^{*\pm}D^{\mp}$. The large unfilled area is the sideband. The small filled ellipse represents the signal region. The horizonally- and vertically-hatched region below the signal area is the feed-down region, which is not included in sideband background fits (see section 8) due to decays to higher resonances such as $B^0 \rightarrow D^{*+}D^{*-}$ peaking in this region. $B^{\pm} \rightarrow D^{*\pm}D^{*0}$ uses the exact same signal and sideband regions, but does not require the feed-down region, thus for $B^{\pm} \rightarrow D^{*\pm}D^{*0}$ the feed-down region above is included in the sideband.

Table A.10: Final values of the χ^2_{Mass} cuts used for each $B^0 \rightarrow D^{*\pm}D^{\mp}$ submode.

		$D^{\pm}$		
		$K\pi\pi$	$K^0_S\pi$	$KK\pi$
D^*	$(K\pi)\pi$	21.5	8.5	6.5
	$(K\pi\pi^0)\pi$	10.5	4.5*	0.5*
	$(K3\pi)\pi$	7.5	1.5*	3.5*
	$(K^0_S\pi\pi)\pi$	5.5*	5.5*	1.5*
	$(K\pi\pi)\pi^0$	8.5	8.5	1.5*
	$(K^0_S\pi)\pi^0$	7.5	7.5	2.5*
	$(KK\pi)\pi^0$	0.0*	0.0*	0.0*

mization for the cuts on χ^2_{Mass} for 4 (representative) modes of each of the respective decays.

The leftmost plots in each of the figures show the χ^2_{Mass} distribution for signal Monte Carlo events, the second-from-left plots shows the χ^2_{Mass} distribution for background Monte Carlo, and the third-from-left plots shows $S^2/(S+B)$ vs. the cut value on χ^2_{Mass}. Here all mode combinations are appropriately weighted by branching ratios and the sample is scaled to the total luminosity of the data sample. The rightmost plots show the change in global $S^2/(S+B)$ ($S^2/(S+B)$ for all modes combined, appropriately weighted by submode branching ratios) as a function of a change in the cut value for χ^2_{Mass}. As can be seen, for these plots all changes in the value of

Table A.11: Final values of the χ^2_{Mass} cuts used for each $B^\pm \to D^{*\pm}D^{*0}$ submode.

		D^*				
		$(K\pi)\pi^\pm$	$(K\pi\pi^0)\pi^\pm$	$(K3\pi)\pi^\pm$	$(K^0_S\pi\pi)\pi^\pm$	$(K\pi\pi)\pi^0$
D^{*0}	$(K\pi)\pi^0$	50.	22.	19.	50.	34.
	$(K\pi\pi^0)\pi^0$	17.	31.	17.	4.*	7.*
	$(K3\pi)\pi^0$	22.	11.	14.	11.	2.*
	$(K^0_S\pi\pi)\pi^0$	7.*	2.*	11.	0.*	2.*
	$(K\pi)\gamma$	50.	39.	37.	29.	4.*
	$(K\pi\pi^0)\gamma$	26.	8.	2.*	7.*	1.*
	$(K3\pi)\gamma$	23.	6.*	12.	7.*	0.*
	$(K^0_S\pi\pi)\gamma$	26.	0.*	0.*	2.*	0.*

the χ^2_{Mass} cut lead to negative changes in overall $S^2/(S+B)$, thus the modes have the optimal cut applied — these plots were made after the 3rd iteration when the cuts were fully optimized.

Submodes which have an optimized χ^2_{Mass} cut of less than 6 (for $B^0 \to D^{*\pm}D^{\mp}$) and 8 (for $B^\pm \to D^{*\pm}D^{*0}$) (*i.e.* less than 2 per degree of freedom) are rejected for this analysis as the statistical significance of their optimized cuts is low, and the systematic errors from these modes tend to be high.

*As described in the text, modes with a χ^2_{Mass} cut of less than 6 (*i.e.* less than 2 per degree of freedom) are rejected from the analysis.

Such submodes are removed from efficiency calculation and from data signal yield.

Figures A.15 and A.16 show the results of the optimization procedure for the $\Delta E'$ and $m_{\rm ES}$ cuts used to define the signal region from both $B^0 \to D^{*\pm}D^{\mp}$ and $B^{\pm} \to D^{*\pm}D^{*0}$. Based on these studies, the signal region is chosen to be the ellipse circumscribed by

$$|\Delta E'| < 25 \text{ MeV}$$

$$5.273 \text{ GeV}/c^2 < m_{\rm ES} < 5.285 \text{ GeV}/c^2$$

We looked at several other discriminating variables, such as $\cos(\theta_T)$ and the Dalitz weight in $D^0 \to K^-\pi^+\pi^0$ using this same procedure, and found no improvement in $S^2/(S+B)$ if a cut were to be made on these variables.

A.7 Reconstruction Efficiency Determination

The efficiency for reconstructing signal decay modes is determined using a detailed Monte Carlo simulation. Once all of the event selection criteria are established, the efficiency is simply the number of signal Monte Carlo events passing the selection cuts divided by the number of events generated. Due to differences in the Monte Carlo and data, there are a few corrections that need to be applied. The difference in PID efficiency is corrected using the standard PID correction procedure described in [28]. In addition to these differences for PID the following efficiency corrections are applied to the signal Monte Carlo:

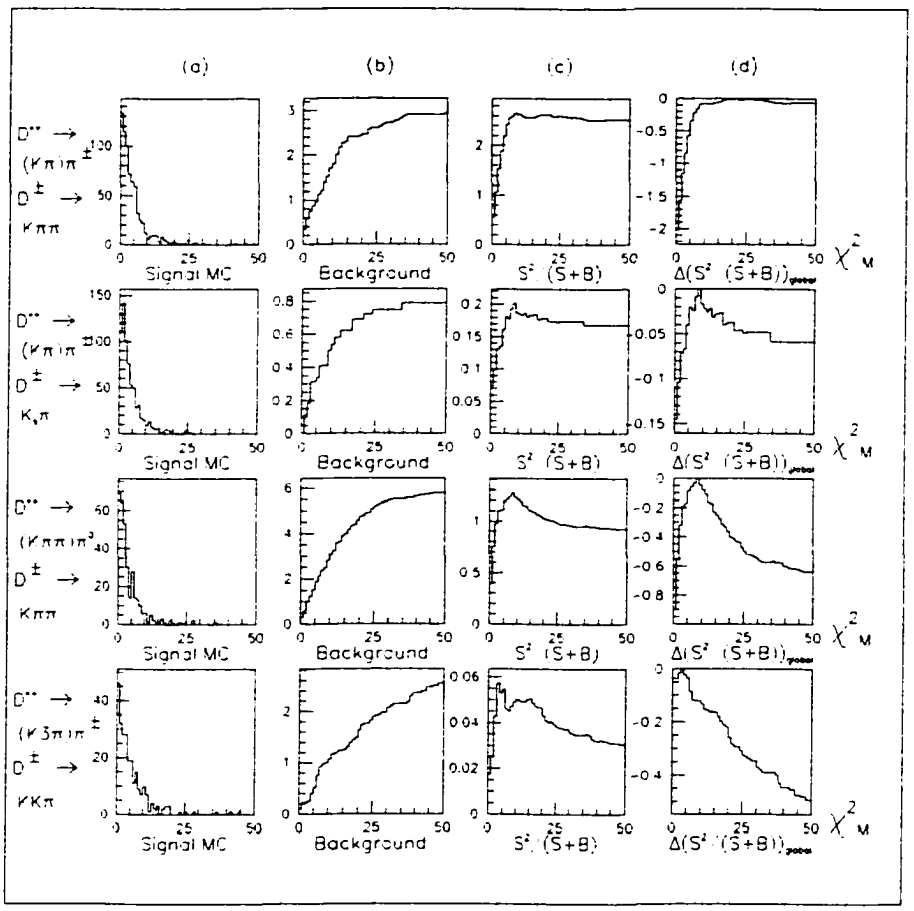


Figure A.13: $B^0 \rightarrow D^{*\pm}D^{\mp}$ optimization of the cuts on χ^2_{Mass}. (a) Distribution of χ^2_{Mass} for signal Monte Carlo for the 4 modes shown. (b) Distribution of χ^2_{Mass} for background Monte Carlo in the $\Delta E'$, $m_{\rm ES}$ sideband. (c) $S^2/(S+B)$ for the mode shown vs. cut value for χ^2_{Mass} (includes appropriate weighting for subdecay mode branching ratios and scaled to the luminosity of the data sample). (d) Change in global $S^2/(S+B)$ (for all modes combined, appropriately weighted by submode branching ratios) as a function of a change in the cut value for χ^2_{Mass}. As can be seen, for these plots all changes in the value of the χ^2_{Mass} cut lead to negative changes in overall $S^2/(S+B)$, thus the modes have the optimal cut applied these plots were made after the 3rd iteration when the cuts were fully optimized.

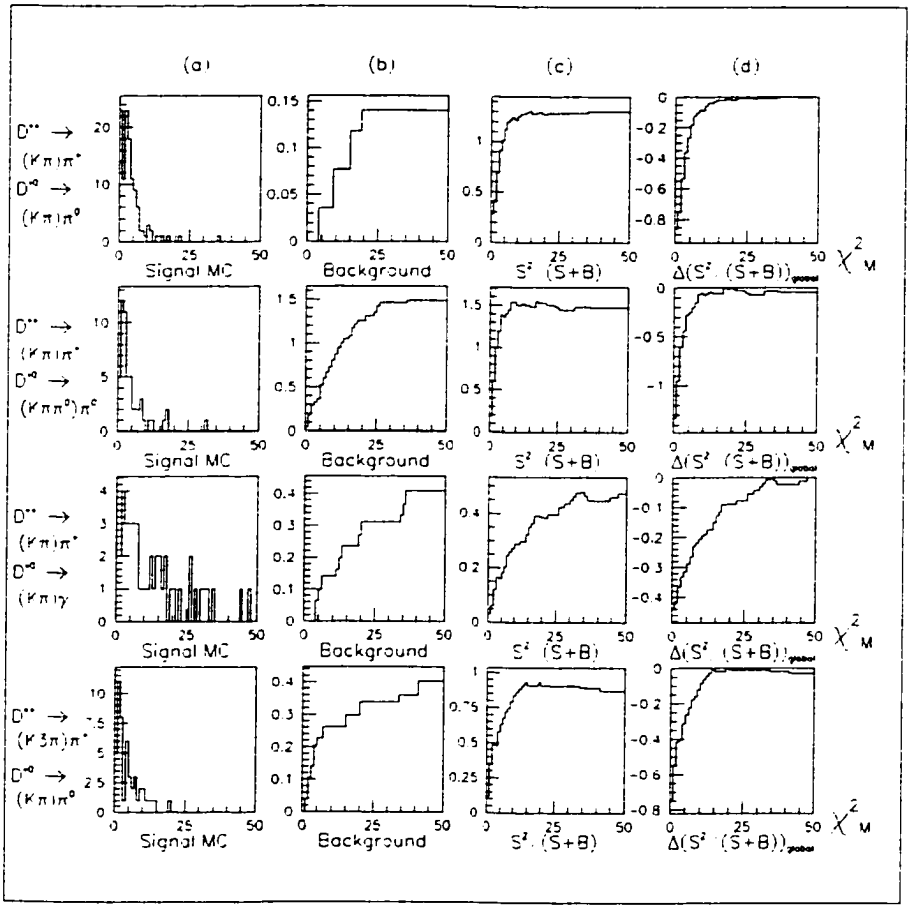


Figure A.14: $B^{\pm} \rightarrow D^{*\pm}D^{*0}$ optimization of the cuts on χ^2_{Mass}. (a) Distribution of χ^2_{Mass} for signal Monte Carlo for the 4 modes shown. (b) Distribution of χ^2_{Mass} for background Monte Carlo in the $\Delta E'$, $m_{\rm ES}$ sideband. (c) $S^2/(S+B)$ for the mode shown vs. cut value for χ^2_{Mass} (includes appropriate weighting for subdecay mode branching ratios and scaled to the luminosity of the data sample). (d) Change in global $S^2/(S+B)$ (for all modes combined, appropriately weighted by submode branching ratios) as a function of a change in the cut value for χ^2_{Mass}. As can be seen, for these plots all changes in the value of the χ^2_{Mass} cut lead to negative changes in overall $S^2/(S+B)$, thus the modes have the optimal cut applied these plots were made after the 3rd iteration when the cuts were fully optimized.

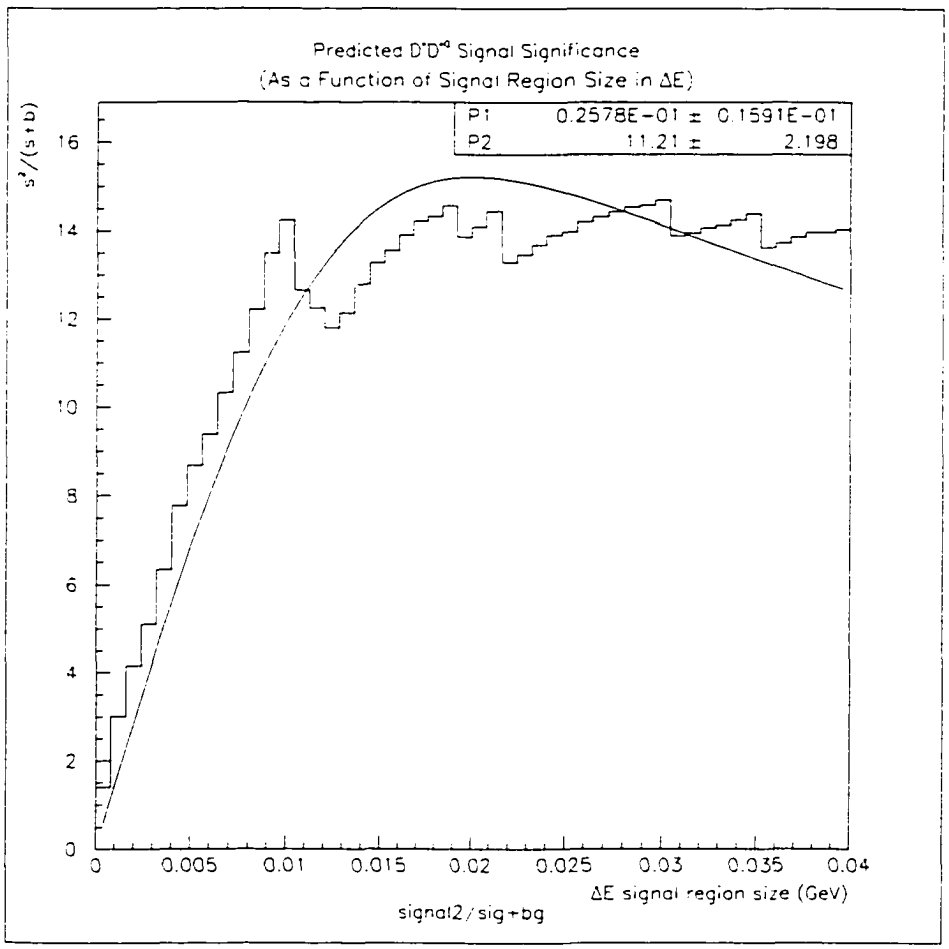


Figure A.15: Optimization of the cut on $\Delta E'$. Plot shows $S^2/(S+B)$ vs. $|\Delta E'|$ elliptical signal region size (includes appropriate weighting for subdecay mode branching ratios and is scaled to the luminosity of the data sample). The fit is to a $S^2/(S+B)$ function if signal were gaussian and background were flat in $\Delta E'$. Although the maximum is at slightly less than 25 MeV, the cut was placed there as there is not much difference in overall $S^2/(S+B)$ in this region, for consistency with the $B^0 \rightarrow D^{*+}D^{*-}$ analysis, and because we wished to remain equal to or above 2 RMS widths in $\Delta E'$.

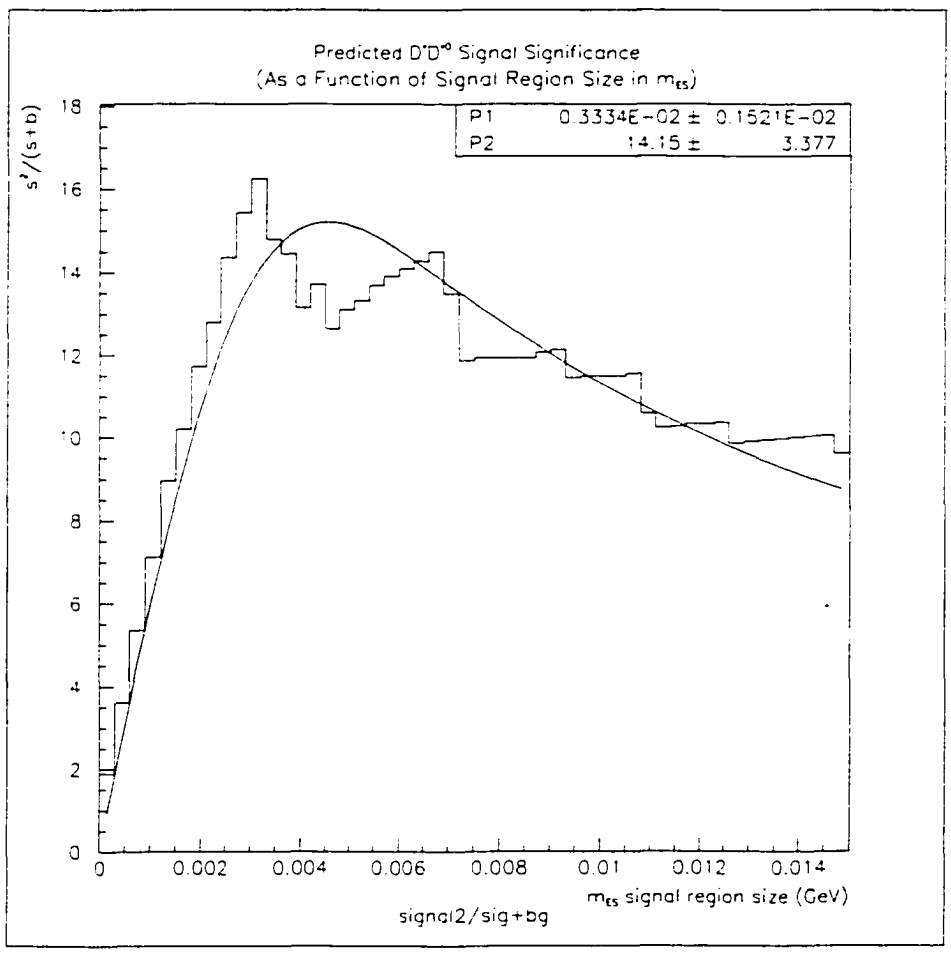


Figure A.16: Optimization of the cut on $m_{\rm ES}$. Plot shows $S^2/(S+B)$ vs. cut on $|m_{\rm ES} - m_B|$ (includes appropriate weights for subdecay mode branching ratios and scaled to the luminosity of the data sample). The fit is to a $S^2/(S+B)$ function if signal were gaussian and background were flat in $m_{\rm ES}$. Although the maximum is at slightly less than 6 MeV, the cut was placed there as there is not much difference in overall $S^2/(S+B)$ in this region, for consistency with the $B^0 \to D^{*+}D^{*-}$ analysis, and because we wished to remain equal to or above 2 σ in $m_{\rm ES}$.

- *Tracking efficiency correction*

  No correction is applied for tracks from `GoodTracksVeryLoose`, but tracks from `GoodTracksLoose` are corrected using a table parameterized as a function of event multiplicity, p_T, ϕ and theta of the track. Two different efficiency correction tables are provided corresponding to the DCH voltage of 1900V and 1960V. The average efficiency correction is in the order of 0.97 to 0.99 per track. This follows the procedure proposed by the tracking efficiency group [30].

- *K_S^0 efficiency*

  The efficiency correction for K_S^0 is parameterized as a function of the 2d flight length of the K_S^0. Two different correction factors are applied according to the DCH voltage of 1900V and 1960V. This follows the procedure proposed by the tracking efficiency group [31].

- *π^0 efficiency*

  For each π^0 a correction factor of 0.95 (2.5% per photon) is applied to the event. This is done for π^0 originating from D decays as well as the slow π^0 from the D^{*+} and D^{*0} decays. This follows the procedure proposed by the neutral working group [32].

- *γ efficiency*

  For each (slow) photon from $D^{*0} \rightarrow D^0\gamma$, a correction factor of 0.975 (2.5% per photon, as above for π^0) is applied. This follows the procedure proposed by the neutral working group [32].

A weight is calculated for each Monte Carlo signal event using the corrections for tracking, K_S^0, π^0, and γ efficiency. Figures A.17, A.18, and A.19

show the weights for 1900V and 1960V determined for signal Monte Carlo. The efficiency is calculated for each combination of sub-decay modes using the ratio of selected events weighted with the correction factor, divided by the number of generated events. The weighted number of selected events, the number of generated Monte Carlo events and the corresponding efficiency for all the modes are summarized in tables A.12, A.13, A.14, and A.15 for 1900V and 1960V. Only slight differences in efficiency between 1900V and 1960V are expected, since mostly `GoodTracksVeryLoose` tracks are used for the reconstruction of B candidates.

For the efficiency determination, the cocktail and exclusive Monte Carlo are both used. The error on the generated cocktail events corresponds to the statistical uncertainty of events generated in a certain mode (no truth mapping used) and is calculated assuming a binomial distribution. In the exclusive Monte Carlo collections the K_S^0 is allowed to decay to $\pi^0\ \pi^0$ and $\pi\pi$. However, we reconstruct only $K_S^0 \to \pi\pi$ and a binomial error is assigned to the events produced. The overall efficiency per mode is the average calculated at 1900V and 1960V and weighted according to the luminosity taken at the different voltages, 1900V and 1960V.

Because of the requirements for D vertex fits (convergence and χ^2 probability > 0.001), and the differences in tracking resolution between data and Monte Carlo, a correction is applied to the overall efficiency calculation. The ratio

$$r = \frac{\epsilon(DATA_{vtx\ cut})/\epsilon(DATA_{no\ vtx\ cut})}{\epsilon(MC_{vtx\ cut})/\epsilon(MC_{no\ vtx\ cut})}$$

is calculated and measures the relative difference between data and Monte Carlo in applying the vertex cut on D candidates. Using the large samples of

$B^0 \to D^*\pi$ data and Monte Carlo, the ratio r was found to be $(98.21 \pm 1.07)\%$ for 2-charged-track vertices and $(90.91 \pm 2.24)\%$ for 4-charged-track vertices.[1] To correct for differences between the vertex efficiency in $B^0 \to D^*\pi$ and $B^0 \to D^{*+}D^-$, the corrected ratio

$$r' = r \times \frac{\sqrt{\epsilon(MC_{D^{*+}D^{*-},\ vtx\ cut})/\epsilon(MC_{D^{*+}D^{*-},\ no\ vtx\ cut})}}{\epsilon(MC_{D^*\pi,\ vtx\ cut})/\epsilon(MC_{D^*\pi,\ no\ vtx\ cut})}$$

is used. This correction, although taken from $B^0 \to D^{*+}D^{*-}$, is thought to be valid for all three decays due to the fact that it corrects for the p^* of the vertices, which are of similar magnitude in each of the three decays. The square root is necessary to account for the two D vertices in each of the three decay modes. r' was calculated to be $(98.96 \pm 1.28)\%$ for 2-charged-track vertices and $(92.53 \pm 3.52)\%$ for 4-charged-track vertices. The efficiency was corrected for each subdecay mode using these values. Since this analysis has two Ds in the final state, the square of r' is used to correct reconstruction efficiencies.

The efficiency and the efficiency $\times$ branching fraction, $\epsilon\mathcal{B}$, for all modes used can be found in Tables A.16, A.17, and A.18. These tables do not include the correction due to the above described vertexing differences. The overall $\epsilon\mathcal{B}$ for all modes, after including the correction from the ratio r', are

$$\epsilon\mathcal{B} = (16.75 \pm 0.47(\text{MC stat}) \pm 0.94(\mathcal{B}\ \text{error}) \pm 0.44(r'\ \text{error})) \times 10^{-4} \qquad (\mathrm{B}^0 \to \mathrm{D}^{*+}\mathrm{D}^{*-})$$

$$\epsilon\mathcal{B} = (15.04 \pm 0.35(\text{MC stat}) \pm 1.12(\mathcal{B}\ \text{error}) \pm 0.40(r\ \text{error}))^{-4} \qquad (\mathrm{B}^0 \to \mathrm{D}^{*\pm}\mathrm{D}^{\mp})$$

$$\epsilon\mathcal{B} = (10.49 \pm 0.60(\text{MC stat}) \pm 0.95(\mathcal{B}\ \text{error}) \pm 0.29(r\ \text{error}))^{-4} \qquad (\mathrm{B}^{\pm} \to \mathrm{D}^{*\pm}\mathrm{D}^{*0})$$

[1] The average of the 2-charged-track and 4-charged-track numbers was used for 3-track vertices.

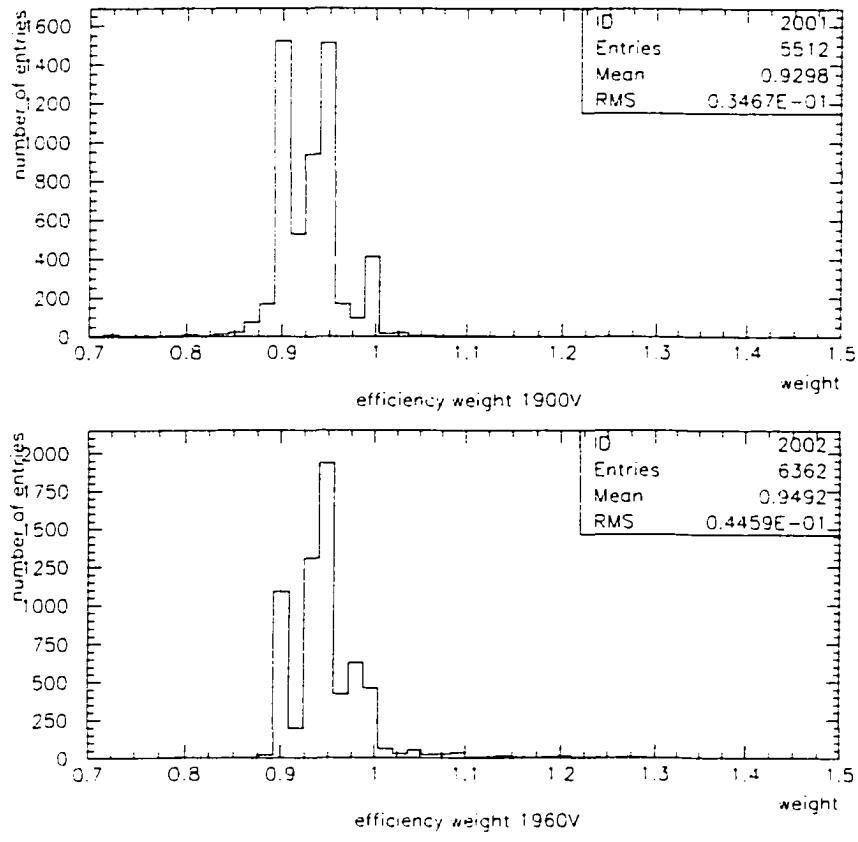


Figure A.17: The plots show the efficiency weight for selected $B^0 \to D^{*+}D^{*-}$ signal events. The upper plot corresponds to the weights determined for Monte Carlo produced with 1900V, the plot below to 1960V. The distributions are not weighted according to individual decay mode branching fractions.

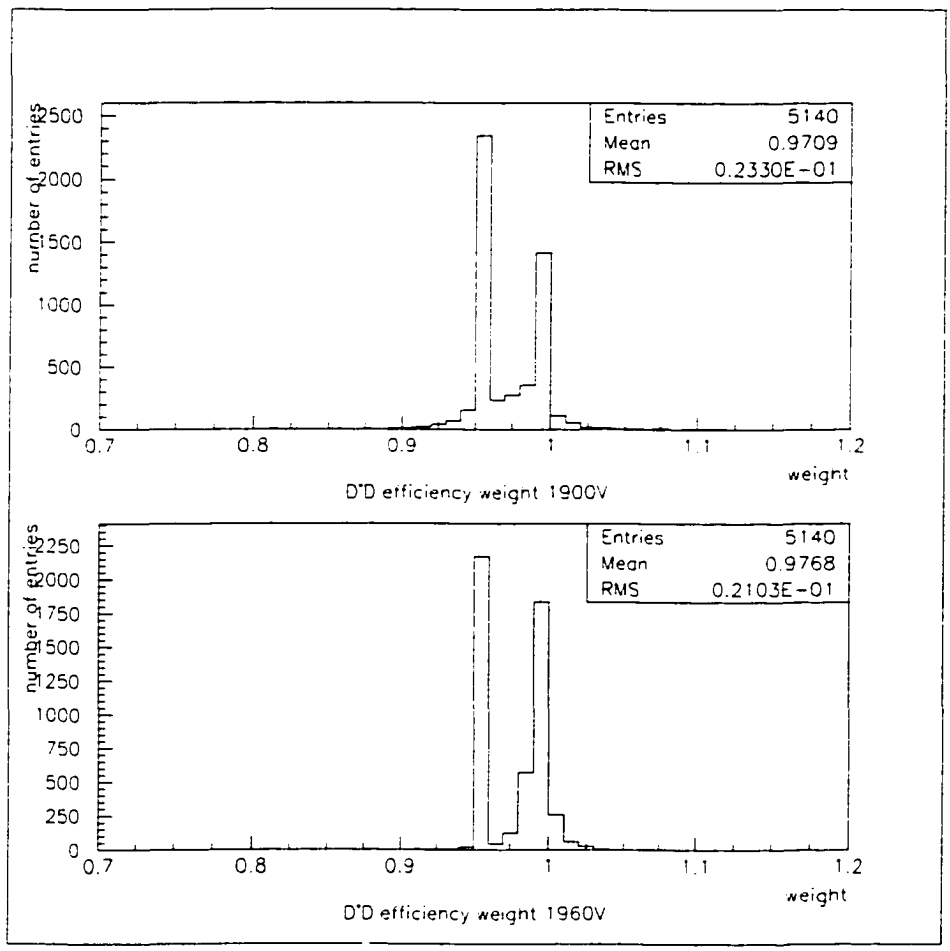


Figure A.18: The plots show the efficiency weight for selected $B^0 \rightarrow D^{*\pm}D^{\mp}$ signal events. The upper plot corresponds to the weights determined for Monte Carlo produced with 1900V, the plot below to 1960V. The distributions are not weighted according to individual decay mode branching fractions. The peaks at 95% correspond to the efficiency correction of 5% per π^0.

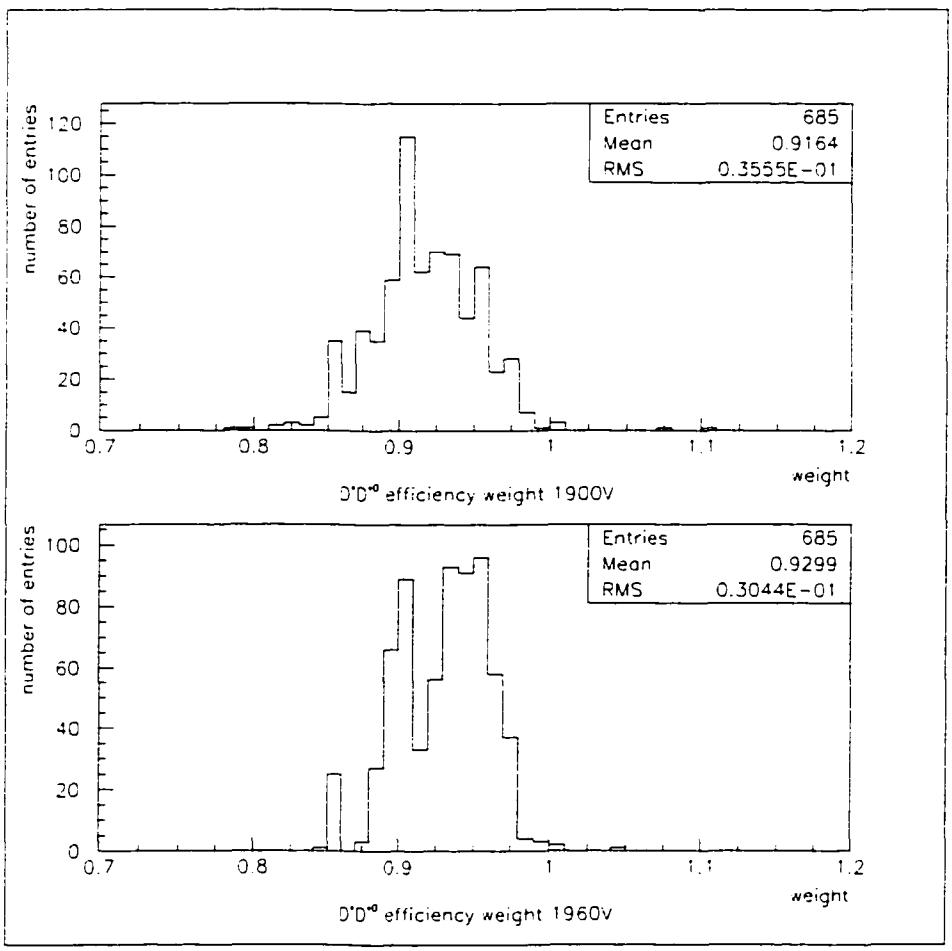


Figure A.19: The plots show the efficiency weight for selected $B^{\pm} \rightarrow D^{*\pm}D^{*0}$ signal events. The upper plot corresponds to the weights determined for Monte Carlo produced with 1900V, the plot below to 1960V. The distributions are not weighted according to individual decay mode branching fractions.

Table A.12: Number of selected and generated $B^0 \rightarrow D^{*+}D^{*-}$ events and the corresponding efficiency for 1900V. The first column represents the mode, the second the weighted number of selected events and the the third and fourth column the number of events produced in cocktail and exclusive Monte Carlo. The last column shows the efficiency for each mode individually, scaled by the weighting factor. The error corresponds to the binomial error and is statistical only. The error on N_{MC} exclusive comes from the fact that the K_S^0 was not required to exclusively decay to $\pi^+ \pi^-$. Zeros in the table indicate that only one type of Monte Carlo (cocktail or exclusive) was generated.

mode	N_{MC} selected	N_{MC} generated cocktail	N_{MC} generated exclusiv	efficiency in %
$(K\pi)$ $(K\pi)$	96.4 ± 8.9	550.3 ± 23.1	0 ± 0	17.52 ± 1.78
$(K\pi)$ $(K\pi\pi^0)$	83.8 ± 8.8	1100.6 ± 32.2	0 ± 0	7.61 ± 0.83
$(K\pi)$ $(K3\pi)$	69.6 ± 8.1	1100.6 ± 32.2	0 ± 0	6.32 ± 0.76
$(K\pi)$ $(K_S^0\pi\pi)$	51.3 ± 6.9	755 ± 26.9	0 ± 0	6.79 ± 0.95
$(K\pi)$ $(K\pi\pi)$	652.7 ± 24.6	0 ± 0	9104 ± 0	7.17 ± 0.27
$(K\pi)$ $(K_S^0\pi)$	480.2 ± 21.1	0 ± 0	6861 ± 46.4	7 ± 0.31
$(K\pi)$ $(KK\pi)$	638 ± 24.4	0 ± 0	10000 ± 0	6.38 ± 0.24
$(K\pi\pi^0)$ $(K\pi\pi^0)$	280.8 ± 16.5	550.3 ± 23.1	10000 ± 0	2.66 ± 0.16
$(K\pi\pi^0)$ $(K3\pi)$	30.3 ± 5.4	1100.6 ± 32.2	0 ± 0	2.75 ± 0.5
$(K\pi\pi^0)$ $(K_S^0\pi\pi)$	22.6 ± 4.7	755.6 ± 26.9	0 ± 0	3 ± 0.63
$(K\pi\pi^0)$ $(K\pi\pi)$	293.6 ± 16.9	0 ± 0	10000 ± 0	2.94 ± 0.17
$(K\pi\pi^0)$ $(K_S^0\pi)$	166.2 ± 12.7	0 ± 0	6861 ± 46.1	2.42 ± 0.19
$(K\pi\pi^0)$ $(KK\pi)$	263.6 ± 16	0 ± 0	10000 ± 0	2.64 ± 0.16
$(K3\pi)$ $(K3\pi)$	324.6 ± 17.7	550.3 ± 23.1	10000 ± 0	3.08 ± 0.17
$(K3\pi)$ $(K_S^0\pi\pi)$	18.3 ± 4.2	755.1 ± 27	0 ± 0	2.42 ± 0.57
$(K3\pi)$ $(K\pi\pi)$	281 ± 16.5	0 ± 0	10000 ± 0	2.81 ± 0.17
$(K3\pi)$ $(K_S^0\pi)$	190.7 ± 13.6	0 ± 0	6861 ± 46.1	2.78 ± 0.2
$(K3\pi)$ $(KK\pi)$	180.4 ± 13.3	0 ± 0	6746 ± 0	2.67 ± 0.2
$(K_S^0\pi\pi)$ $(K_S^0\pi\pi)$	141.2 ± 11.7	259 ± 16	4707.3 ± 49.9	2.84 ± 0.24
$(K_S^0\pi\pi)$ $(K\pi\pi)$	220.1 ± 14.6	0 ± 0	6861 ± 46.1	3.21 ± 0.21
$(K_S^0\pi\pi)$ $(K_S^0\pi)$	137.8 ± 11.6	0 ± 0	4707.3 ± 49.9	2.93 ± 0.25
$(K_S^0\pi\pi)$ $(KK\pi)$	185.1 ± 13.4	0 ± 0	6861 ± 46.4	2.7 ± 0.2

Table A.13: Number of selected and generated $B^0 \to D^{*+}D^{*-}$ events and the corresponding efficiency for 1960V. The first column represents the mode, the second the weighted number of selected events and the the third and fourth column the number of events produced in cocktail and exclusive Monte Carlo. The last column shows the efficiency for each mode individually, scaled by the weighting factor. The error corresponds to the binomial error and is statistical only. The error on N_{MC} exclusive comes from the fact that the K_S^0 was not required to exclusively decay to $\pi^+ \pi^-$. Zeros in the table indicate that only one type of Monte Carlo (cocktail or exclusive) was generated.

mode	N_{MC} selected	N_{MC} generated cocktail	N_{MC} generated exclusiv	efficiency in %
$(K\pi)$ $(K\pi)$	983.9 ± 28.5	555.6 ± 23.4	5000 ± 0	17.71 ± 0.52
$(K\pi)$ $(K\pi\pi^0)$	66.5 ± 7.9	1111.1 ± 32.4	0 ± 0	5.99 ± 0.75
$(K\pi)$ $(K3\pi)$	85 ± 8.9	1111.1 ± 32.4	0 ± 0	7.65 ± 0.82
$(K\pi)$ $(K_S^0\pi\pi)$	55.6 ± 7.2	762.3 ± 27.1	0 ± 0	7.29 ± 0.97
$(K\pi)$ $(K\pi\pi)$	682.9 ± 25.2	0 ± 0	10000 ± 0	6.83 ± 0.25
$(K\pi)$ $(K_S^0\pi)$	528.3 ± 22.1	0 ± 0	6861 ± 46.4	7.7 ± 0.33
$(K\pi)$ $(KK\pi)$	641.6 ± 24.5	0 ± 0	10000 ± 0	6.42 ± 0.25
$(K\pi\pi^0)$ $(K\pi\pi^0)$	220.6 ± 14.7	555.6 ± 23.4	8166 ± 0	2.53 ± 0.17
$(K\pi\pi^0)$ $(K3\pi)$	27.6 ± 5.2	1111.1 ± 32.4	0 ± 0	2.48 ± 0.47
$(K\pi\pi^0)$ $(K_S^0\pi\pi)$	21.4 ± 4.6	762.3 ± 27.1	0 ± 0	2.81 ± 0.61
$(K\pi\pi^0)$ $(K\pi\pi)$	291.9 ± 16.8	0 ± 0	10000 ± 0	2.92 ± 0.17
$(K\pi\pi^0)$ $(K_S^0\pi)$	202.4 ± 14	0 ± 0	6861 ± 46.4	2.95 ± 0.2
$(K\pi\pi^0)$ $(KK\pi)$	249.9 ± 15.6	0 ± 0	10000 ± 0	2.5 ± 0.16
$(K3\pi)$ $(K3\pi)$	250.9 ± 15.6	555.6 ± 23.4	8080 ± 0	2.91 ± 0.18
$(K3\pi)$ $(K_S^0\pi\pi)$	18.8 ± 4.3	762.3 ± 27.1	0 ± 0	2.47 ± 0.57
$(K3\pi)$ $(K\pi\pi)$	286.5 ± 16.7	0 ± 0	10000 ± 0	2.87 ± 0.17
$(K3\pi)$ $(K_S^0\pi)$	211.6 ± 14.3	0 ± 0	6861 ± 46.4	3.08 ± 0.21
$(K3\pi)$ $(KK\pi)$	230.8 ± 15	0 ± 0	10000 ± 0	2.31 ± 0.15
$(K_S^0\pi\pi)$ $(K_S^0\pi\pi)$	167.9 ± 12.7	261.5 ± 16.1	4707.3 ± 49.9	3.38 ± 0.26
$(K_S^0\pi\pi)$ $(K\pi\pi)$	221.8 ± 14.6	0 ± 0	6527.5 ± 45.3	3.4 ± 0.23
$(K_S^0\pi\pi)$ $(K_S^0\pi)$	178.3 ± 13.1	0 ± 0	4707.3 ± 49.9	3.79 ± 0.28
$(K_S^0\pi\pi)$ $(KK\pi)$	206.4 ± 14.2	0 ± 0	6861 ± 46.4	3.01 ± 0.21

Table A.14: Number of selected and generated $B^0 \to D^{*\pm}D^{\mp}$ events and the corresponding efficiency for 1900V and 1960V. The first two columns represent the mode, the second and third the weighted number of (truth-tagged) selected events and the fourth column the number of events produced in Monte Carlo. The last two columns show the efficiency for each mode individually, scaled by the weighting factor. The error corresponds to the binomial error and is purely statistical.

mode		N_{MC} selected	N_{MC} selected	N_{MC}	efficiency	efficiency
D^*	$D^{\pm}$	1900V	1960V	generated	(1900V, in %)	(1960V, in %)
$(K\pi)\pi^{\pm}$	$K\pi\pi$	640.66	650.77	4000	16.02 ± 0.58	16.27 ± 0.58
$(K\pi)\pi^{\pm}$	$K^0_S\pi$	580.48	589.04	4000	14.51 ± 0.56	14.73 ± 0.56
$(K\pi)\pi^{\pm}$	$KK\pi$	449.25	456.39	4000	11.23 ± 0.50	11.41 ± 0.50
$(K\pi\pi^0)\pi^{\pm}$	$K\pi\pi$	342.95	342.95	6000	5.72 ± 0.30	5.72 ± 0.30
$(K3\pi)\pi^{\pm}$	$K\pi\pi$	199.35	199.35	4000	4.98 ± 0.34	4.98 ± 0.34
$(K\pi\pi)\pi^0$	$K\pi\pi$	253.45	253.45	4000	6.34 ± 0.39	6.34 ± 0.39
$(K\pi\pi)\pi^0$	$K^0_S\pi$	253.79	253.79	4000	6.34 ± 0.39	6.34 ± 0.39
$(K^0_S\pi)\pi^0$	$K\pi\pi$	222.86	222.86	4000	5.57 ± 0.36	5.57 ± 0.36
$(K^0_S\pi)\pi^0$	$K^0_S\pi$	215.65	215.65	4000	5.39 ± 0.36	5.39 ± 0.36

Table A.15: Number of selected and generated $B^{\pm} \to D^{*\pm}D^{*0}$ events and the corresponding efficiency for 1900V and 1960V. The first two columns represent the mode, the second and third the weighted number of (truth-tagged) selected events and the fourth column the number of events produced in Monte Carlo. The last two columns show the efficiency for each mode individually, scaled by the weighting factor. The error corresponds to the binomial error and is purely statistical.

mode		N_{MC} selected	N_{MC} selected	N_{MC} generated	efficiency	efficiency
D^*	D^{*0}	1900V	1960V	(cocktail)	(1960V, in %)	(1900V, in %)
$(K\pi)\pi^{\pm}$	$(K\pi)\pi^0$	110.48	103.57	1036.81 ± 31.98	10.65 ± 1.01	9.99 ± 0.98
$(K\pi)\pi^{\pm}$	$(K\pi\pi^0)\pi^0$	42.89	42.37	1036.81 ± 31.98	4.14 ± 0.63	4.09 ± 0.63
$(K\pi)\pi^{\pm}$	$(K3\pi)\pi^0$	46.47	45.52	1036.81 ± 31.98	4.48 ± 0.66	4.39 ± 0.65
$(K\pi)\pi^{\pm}$	$(K\pi)\gamma$	43.07	41.96	518.33 ± 22.69	8.31 ± 1.27	8.10 ± 1.25
$(K\pi)\pi^{\pm}$	$(K\pi\pi^0)\gamma$	19.84	19.67	518.33 ± 22.69	3.83 ± 0.86	3.79 ± 0.86
$(K\pi)\pi^{\pm}$	$(K3\pi)\gamma$	12.43	12.11	518.33 ± 22.69	2.40 ± 0.68	2.34 ± 0.67
$(K\pi)\pi^{\pm}$	$(K_S^0\pi\pi)\gamma$	6.54	6.53	355.57 ± 18.81	1.84 ± 0.72	1.83 ± 0.72
$(K\pi\pi^0)\pi^{\pm}$	$(K\pi)\pi^0$	30.80	30.35	1036.81 ± 31.98	2.97 ± 0.54	2.93 ± 0.53
$(K\pi\pi^0)\pi^{\pm}$	$(K\pi\pi^0)\pi^0$	14.30	14.30	1036.81 ± 31.98	1.38 ± 0.36	1.38 ± 0.36
$(K\pi\pi^0)\pi^{\pm}$	$(K3\pi)\pi^0$	13.24	13.24	1036.81 ± 31.98	1.28 ± 0.35	1.28 ± 0.35
$(K\pi\pi^0)\pi^{\pm}$	$(K\pi)\gamma$	12.65	12.56	518.33 ± 22.69	2.44 ± 0.69	2.42 ± 0.68
$(K\pi\pi^0)\pi^{\pm}$	$(K\pi\pi^0)\gamma$	6.90	6.90	518.33 ± 22.69	1.33 ± 0.51	1.33 ± 0.50
$(K3\pi)\pi^{\pm}$	$(K\pi)\pi^0$	37.11	36.24	1036.81 ± 31.98	3.58 ± 0.59	3.50 ± 0.58
$(K3\pi)\pi^{\pm}$	$(K\pi\pi^0)\pi^0$	17.38	17.38	1036.81 ± 31.98	1.68 ± 0.40	1.68 ± 0.40
$(K3\pi)\pi^{\pm}$	$(K3\pi)\pi^0$	16.28	16.28	1036.81 ± 31.98	1.57 ± 0.39	1.57 ± 0.39
$(K3\pi)\pi^{\pm}$	$(K_S^0\pi\pi)\pi^0$	5.22	5.22	711.25 ± 26.54	0.73 ± 0.32	0.73 ± 0.32
$(K3\pi)\pi^{\pm}$	$(K\pi)\gamma$	30.72	30.44	518.33 ± 22.69	5.93 ± 1.07	5.87 ± 1.06
$(K3\pi)\pi^{\pm}$	$(K3\pi)\gamma$	8.33	8.33	518.33 ± 22.69	1.61 ± 0.56	1.61 ± 0.56
$(K_S^0\pi\pi)\pi^{\pm}$	$(K\pi)\pi^0$	32.47	31.91	711.25 ± 26.54	4.56 ± 0.80	4.49 ± 0.79
$(K_S^0\pi\pi)\pi^{\pm}$	$(K3\pi)\pi^0$	10.45	10.45	711.25 ± 26.54	1.47 ± 0.45	1.47 ± 0.45
$(K_S^0\pi\pi)\pi^{\pm}$	$(K\pi)\gamma$	15.05	14.86	355.57 ± 18.81	4.23 ± 1.09	4.18 ± 1.08
$(K\pi\pi)\pi^0$	$(K\pi)\pi^0$	15.32	15.01	345.60 ± 18.55	4.43 ± 1.13	4.34 ± 1.12

Table A.16: Efficiency and efficiency $\times\mathcal{B}$ for the different $B^0 \to D^{*+}D^{*-}$ modes. The error on the efficiency reflects the Monte Carlo statistical uncertainty only.

mode	efficiency in %	efficiency $\times\mathcal{B}$ ($\times 10^{-4}$)
$(K\pi)$ $(K\pi)$	17.61 ± 0.99	1.184
$(K\pi)$ $(K\pi\pi^0)$	6.87 ± 0.56	3.353
$(K\pi)$ $(K3\pi)$	6.93 ± 0.56	1.822
$(K\pi)$ $(K_S^0\pi\pi)$	7.02 ± 0.68	0.456
$(K\pi)$ $(K\pi\pi)$	7.01 ± 0.19	1.004
$(K\pi)$ $(K_S^0\pi)$	7.32 ± 0.23	0.115
$(K\pi)$ $(KK\pi)$	6.4 ± 0.17	0.089
$(K\pi\pi^0)$ $(K\pi\pi^0)$	2.6 ± 0.11	2.302
$(K\pi\pi^0)$ $(K3\pi)$	2.63 ± 0.35	2.51
$(K\pi\pi^0)$ $(K_S^0\pi\pi)$	2.91 ± 0.44	0.686
$(K\pi\pi^0)$ $(K\pi\pi)$	2.93 ± 0.12	1.524
$(K\pi\pi^0)$ $(K_S^0\pi)$	2.66 ± 0.14	0.152
$(K\pi\pi^0)$ $(KK\pi)$	2.57 ± 0.11	0.129
$(K3\pi)$ $(K3\pi)$	3 ± 0.12	0.771
$(K3\pi)$ $(K_S^0\pi\pi)$	2.44 ± 0.4	0.31
$(K3\pi)$ $(K\pi\pi)$	2.84 ± 0.12	0.796
$(K3\pi)$ $(K_S^0\pi)$	2.92 ± 0.14	0.09
$(K3\pi)$ $(KK\pi)$	2.51 ± 0.13	0.068
$(K_S^0\pi\pi)$ $(K_S^0\pi\pi)$	3.09 ± 0.17	0.048
$(K_S^0\pi\pi)$ $(K\pi\pi)$	3.3 ± 0.16	0.228
$(K_S^0\pi\pi)$ $(K_S^0\pi)$	3.32 ± 0.19	0.025
$(K_S^0\pi\pi)$ $(KK\pi)$	2.84 ± 0.14	0.019
total		17.683

Table A.17: Efficiency and efficiency $\times\mathcal{B}$ for the different $B^0 \to D^{*\pm}D^{\mp}$ modes. The error on the efficiency reflects the Monte Carlo statistical uncertainty only.

mode		efficiency	efficiency $\times\mathcal{B}$
D^*	$D^{\pm}$	(in %)	($\times 10^{-5}$)
$(K\pi)\pi^{\pm}$	$K\pi\pi$	16.13 ± 0.58	38.18 ± 1.38
$(K\pi)\pi^{\pm}$	$K_S^0\pi$	14.61 ± 0.56	4.25 ± 0.16
$(K\pi)\pi^{\pm}$	$KK\pi$	11.31 ± 0.50	2.62 ± 0.12
$(K\pi\pi^0)\pi^{\pm}$	$K\pi\pi$	5.72 ± 0.30	47.78 ± 2.51
$(K3\pi)\pi^{\pm}$	$K\pi\pi$	4.98 ± 0.34	23.28 ± 1.61
$(K\pi\pi)\pi^0$	$K\pi\pi$	6.34 ± 0.39	15.70 ± 0.95
$(K\pi\pi)\pi^0$	$K_S^0\pi$	6.34 ± 0.39	1.93 ± 0.12
$(K_S^0\pi)\pi^0$	$K\pi\pi$	5.57 ± 0.36	1.70 ± 0.11
$(K_S^0\pi)\pi^0$	$K_S^0\pi$	5.39 ± 0.36	0.20 ± 0.01
total			135.67

Table A.18: Efficiency and efficiency $\times\mathcal{B}$ for the different $B^{\pm} \to D^{*\pm}D^{*0}$ modes. The error on the efficiency reflects the Monte Carlo statistical uncertainty only.

mode		efficiency	efficiency $\times\mathcal{B}$
D^*	D^{*0}	(in %)	($\times 10^{-5}$)
$(K\pi)\pi^{\pm}$	$(K\pi)\pi^0$	10.29 ± 1.00	6.33 ± 0.61
$(K\pi)\pi^{\pm}$	$(K\pi\pi^0)\pi^0$	4.11 ± 0.63	9.17 ± 1.40
$(K\pi)\pi^{\pm}$	$(K3\pi)\pi^0$	4.43 ± 0.65	5.33 ± 0.79
$(K\pi)\pi^{\pm}$	$(K\pi)\gamma$	8.19 ± 1.26	3.10 ± 0.48
$(K\pi)\pi^{\pm}$	$(K\pi\pi^0)\gamma$	3.81 ± 0.86	5.23 ± 1.18
$(K\pi)\pi^{\pm}$	$(K3\pi)\gamma$	2.37 ± 0.68	1.75 ± 0.50
$(K\pi)\pi^{\pm}$	$(K_S^0\pi\pi)\gamma$	1.84 ± 0.72	0.34 ± 0.13
$(K\pi\pi^0)\pi^{\pm}$	$(K\pi)\pi^0$	2.95 ± 0.53	6.58 ± 1.19
$(K\pi\pi^0)\pi^{\pm}$	$(K\pi\pi^0)\pi^0$	1.38 ± 0.36	11.17 ± 2.95
$(K\pi\pi^0)\pi^{\pm}$	$(K3\pi)\pi^0$	1.27 ± 0.35	5.57 ± 1.53
$(K\pi\pi^0)\pi^{\pm}$	$(K\pi)\gamma$	2.43 ± 0.68	3.34 ± 0.94
$(K\pi\pi^0)\pi^{\pm}$	$(K\pi\pi^0)\gamma$	1.33 ± 0.51	6.64 ± 2.53
$(K3\pi)\pi^{\pm}$	$(K\pi)\pi^0$	3.53 ± 0.58	4.25 ± 0.70
$(K3\pi)\pi^{\pm}$	$(K\pi\pi^0)\pi^0$	1.68 ± 0.40	7.31 ± 1.75
$(K3\pi)\pi^{\pm}$	$(K3\pi)\pi^0$	1.57 ± 0.39	3.69 ± 0.91
$(K3\pi)\pi^{\pm}$	$(K_S^0\pi\pi)\pi^0$	0.73 ± 0.32	0.43 ± 0.19
$(K3\pi)\pi^{\pm}$	$(K\pi)\gamma$	5.90 ± 1.07	4.36 ± 0.79
$(K3\pi)\pi^{\pm}$	$(K3\pi)\gamma$	1.61 ± 0.56	2.32 ± 0.81
$(K_S^0\pi\pi)\pi^{\pm}$	$(K\pi)\pi^0$	4.52 ± 0.78	1.34 ± 0.24
$(K_S^0\pi\pi)\pi^{\pm}$	$(K3\pi)\pi^0$	1.47 ± 0.45	0.85 ± 0.26
$(K_S^0\pi\pi)\pi^{\pm}$	$(K\pi)\gamma$	4.20 ± 1.09	0.77 ± 0.20
$(K\pi\pi)\pi^0$	$(K\pi)\pi^0$	4.38 ± 1.13	2.87 ± 0.74
total			92.73

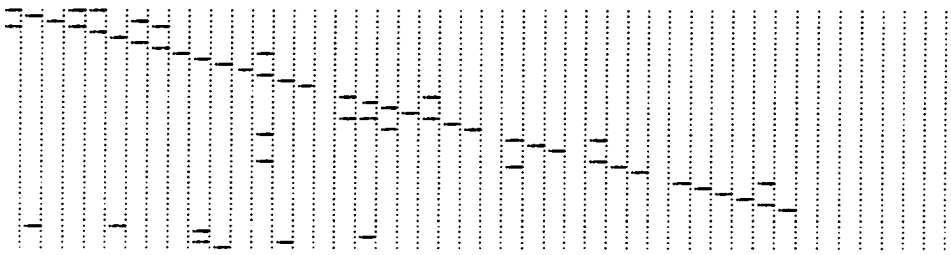


Figure A.20: The $B^{\pm} \to D^{*\pm}D^{*0}$ efficiency crossfeed matrix for all modes used. The total off-diagonal crossfeed is 4.7% of events (weighted by $\epsilon\mathcal{B}$).

where the last error comes from propagating the uncertainties in the measured values of r' into $\epsilon\mathcal{B}$. Therefore we calculate systematic uncertainties due to the vertexing correction of 2.6% for $B^0 \to D^{*+}D^{*-}$, 2.6% for $B^0 \to D^{*\pm}D^{\mp}$, and 2.8% for $B^{\pm} \to D^{*\pm}D^{*0}$.

A.8 Signal Determination for $B^0 \to D^{*+}D^{*-}$

A.8.1 Background Estimation

In order to ascertain the contribution to these candidate events from the actual signal in question, we must estimate the expected contribution from background. This is accomplished by using the previously defined sidebands in the $\Delta E'$ vs. $m_{\rm ES}$ plane.

The number of background events expected in the signal region is taken to be the number of events reconstructed in the sideband region times a scaling factor, f_{Side}, that is a measure of the relative areas of the signal region to the sideband region:

$$N_{Bkg} = f_{Side} \times N_{Sideband}$$

The determination of f_{Side} is done under the assumption that the density of events in the $\Delta E'$ vs. $m_{\rm ES}$ plane can be parameterized as the product of separate functions, $f(\Delta E')$ and $g(m_{\rm ES})$, *i.e.* that $\Delta E'$ and $m_{\rm ES}$ are uncorrelated. Then we have

$$f_{Side} = \frac{\int\int_{SigRegion} f(\Delta E')g(m_{\rm ES})d\Delta E' dm_{\rm ES}}{\int\int_{Sideband} f(\Delta E')g(m_{\rm ES})d\Delta E' dm_{\rm ES}}$$

We determine f_{Side} and its associated systematic error by fitting projections of $\Delta E'$ and $m_{\rm ES}$ from Monte Carlo $B\overline{B}$ and $c\overline{c}$ events to several different parameterized functions. For the $\Delta E'$ distributions our trial functions are a first order polynomial, a second order polynomial, and an exponential. The results of these fits are shown in Figure A.21 . For the $m_{\rm ES}$ projections

Table A.19: Summary of values for f_{Side} determined using different parameterizations of the background shape. Numbers are $\times 10^{-2}$.

	$f(\Delta E')$		
$g(m_{\rm ES})$	Linear	Quadratic	Exponential
Argus	1.72	1.77	1.71
Constant	1.82	1.88	1.81
Linear	1.83	1.89	1.82

our trial functions are a first order polynomial and the Argus background function. The results of these fits are shown in Figure A.22.

Combining the different parameterizations of $\Delta E'$ and $m_{\rm ES}$ yield the results for f_{Side} summarized in Table A.19. The values of f_{Side} range from 0.0171 to 0.0189. We take as our baseline parameterization the combination $f(\Delta E')$ = first order polynomial and $g(m_{\rm ES})$ = Argus function and assign a systematic error due to variation in parameterization of 0.001, giving

$$f_{Side} = (1.72 \pm 0.10) \times 10^{-2}$$

Based on our samples of Monte Carlo events, we see no evidence of peaking of the background in the signal region.

We observe a total of 363 events in the sideband region in data. This is shown in Figure A.24. Combined with the value of f_{Side} determined above, we estimate the number of background events appearing in the signal region to be

$$N_{Bkg} = 6.24 \pm 0.33(stat) \pm 0.36(syst)$$

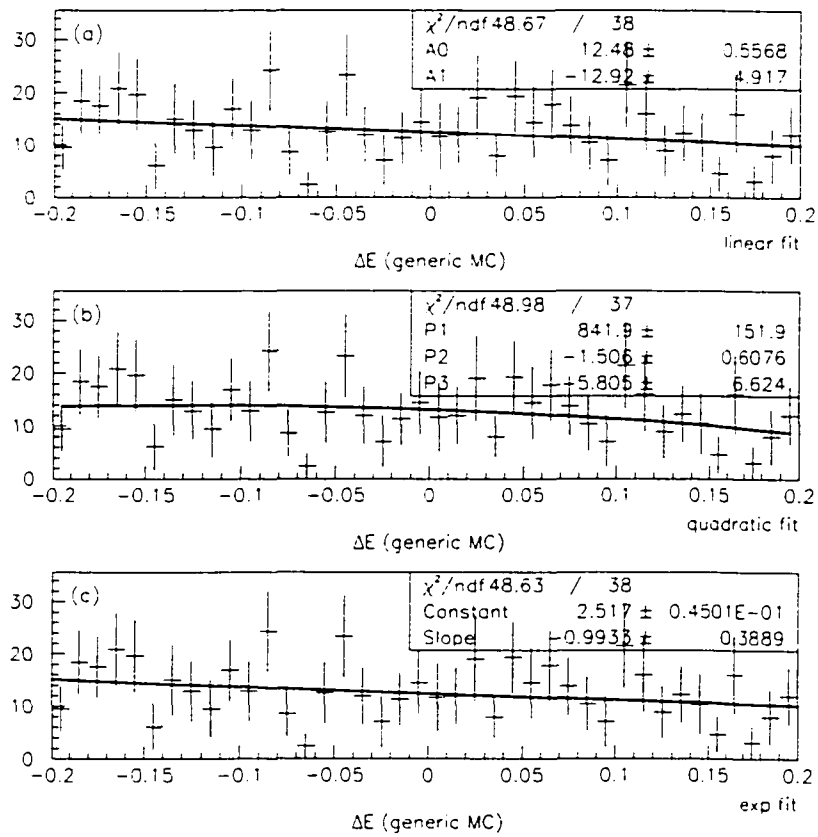


Figure A.21: Fits of trial functions to projections of $\Delta E'$ from the background Monte Carlo sample. Figure (a) shows the fit to a straight line. Figure (b) shows the fit to a second-order polynomial. Figure (c) shows the fit to an exponential. The units on the x-axis are GeV.

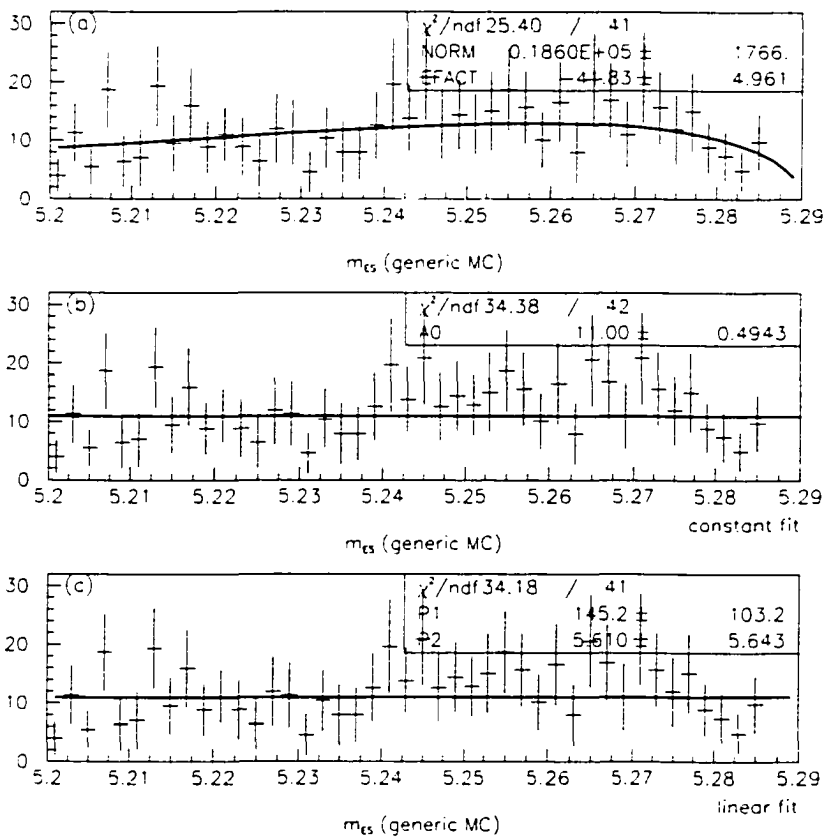


Figure A.22: Fits of trial functions to projections of m_{ES} from the background Monte Carlo sample. Figure (a) shows the fit to the Argus background function. Figure (b) shows the fit to a constant. Figure (c) shows the fit to a linear function. The units on the x-axis are GeV.

where the statistical error comes from the number of events in the sideband region and the systematic error comes from our determination of f_{Side}.

Because of the number of submodes that are used, the possibility of misreconstructing one mode as another (crossfeed) was investigated. Using individual signal Monte Carlo samples, 1000 events per submode, we show in Figure A.23 the number of events after passing all selection criteria reconstructed in the different submodes. The horizontal axis represents each submode signal Monte Carlo collection, and the vertical axis shows the number of events reconstructed in each submode. The number of events reconstructed on the off-diagonal (indicating crossfeed between modes) was two out of a total of 21000 events generated; the effect was considered negligible.

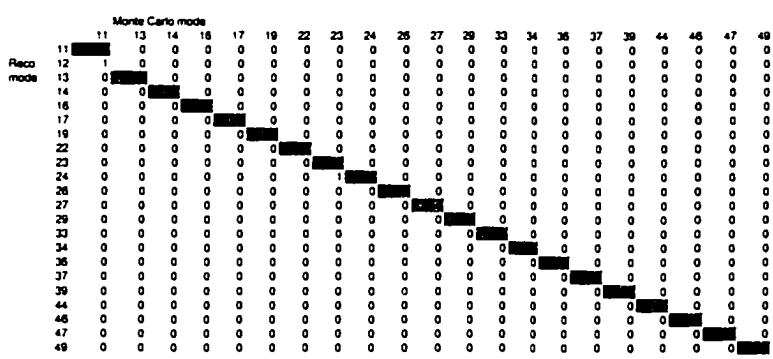


Figure A.23: The number of events out of one thousand generated that pass all selection criteria as a function of the individual submodes. The mode generated in signal Monte Carlo is represented on the horizontal axis and the mode reconstructed is shown on the vertical axis.

A.8.2 $B^0 \to D^{*+}D^{*-}$ Signal

After all analysis selection criteria are applied, as described in previous sections, we are left with 38 candidates in the signal region of the $\Delta E'$ vs. $m_{\rm ES}$ plane. This is shown in Figure A.24. Based on the 38 candidate events in the signal region and the background estimation from above, the number of signal events is

$$N_{Sig} = 31.8 \pm 6.2(stat) \pm 0.4(syst),$$

where the statistical error includes the statistical error on the number of background events and the systematic error is the uncertainty quoted previously on f_{Side}.

As a consistency check, we break up the data and Monte Carlo samples by final state. Table A.20 shows a breakdown of the signal region by D decay mode. Also, Figure A.25 shows the χ^2_{Mass} distribution for the signal candidates.

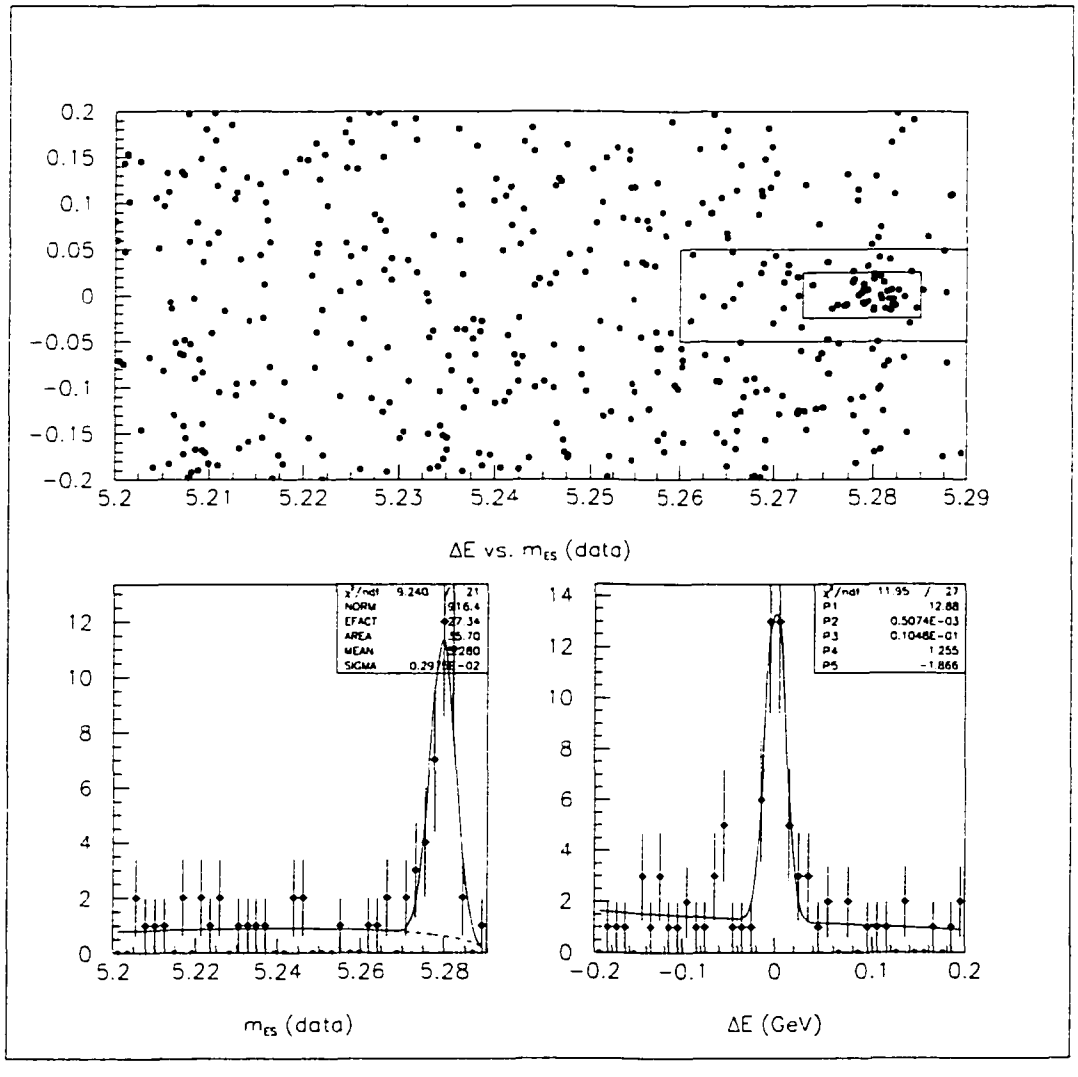


Figure A.24: (a)$\Delta E'$ vs. m_{ES} for data. (b) Projection on to the m_{ES} axis after applying the signal region cut on $\Delta E'$. (c) Projection on to the $\Delta E'$ axis after applying the signal region cut on m_{ES}.

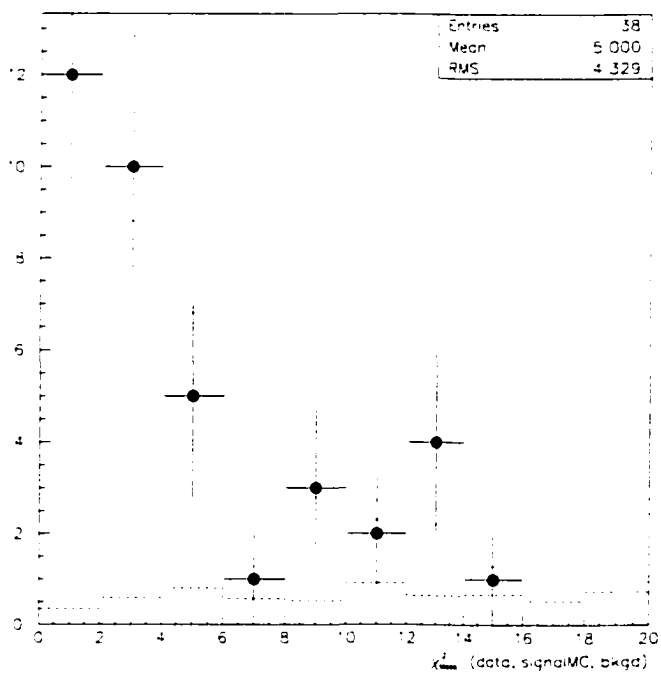


Figure A.25: χ^2_{Mass} for events in the signal region. Points with error bars are data. The dashed histogram is the χ^2_{Mass} distribution from the data sideband region, scaled by f_{Side}. The dotted histogram is the sum of the dashed histogram and the χ^2_{Mass} distribution from signal Monte Carlo, normalized to the number of events in the data signal region.

Table A.20: Comparison of events seen in signal region with expectation from signal Monte Carlo, broken down by D decay mode combination. The Monte Carlo expectation for signal is the $\epsilon\mathcal{B}$ for the given mode combination divided by the total $\epsilon\mathcal{B}$, all multiplied by the total number of signal events seen in data (for normalization purposes). The background expectation is the number of events in the data sidebands for that mode combination multiplied by f_{Side}. The sources of uncertainty in the expected values include the Monte Carlo statistics in determining the efficiency, the uncertainties on the D^* branching fractions, the uncertainty on r (see Section A.7), the uncertainty on f_{Side}, and the statistical uncertainty from the number of events in the sideband for that mode combination.

Final State	N_{Sig} (MC Expected)	N_{Bkg} (Sidebands)	N_{Total}	N_{Data}
$(K\pi)(K\pi)$	2.07 ± 0.16	0.12 ± 0.05	2.19 ± 0.17	1
$(K\pi)(K\pi\pi^0)$	5.98 ± 0.67	0.17 ± 0.06	6.15 ± 0.67	8
$(K\pi)(K3\pi)$	3.12 ± 0.32	0.17 ± 0.06	3.29 ± 0.32	1
$(K\pi)(K_S^0\pi\pi)$	0.80 ± 0.10	0.09 ± 0.04	0.88 ± 0.11	0
$(K\pi)(K\pi\pi)$	1.91 ± 0.16	0.21 ± 0.06	2.11 ± 0.17	0
$(K\pi)(K_S^0\pi)$	0.21 ± 0.02	0.12 ± 0.05	0.33 ± 0.05	0
$(K\pi)(KK\pi)$	0.16 ± 0.02	0.10 ± 0.04	0.26 ± 0.05	0
$(K\pi\pi^0)(K\pi\pi^0)$	4.19 ± 0.60	0.79 ± 0.13	4.98 ± 0.62	8
$(K\pi\pi^0)(K3\pi)$	4.43 ± 0.70	0.60 ± 0.11	5.03 ± 0.71	4
$(K\pi\pi^0)(K_S^0\pi\pi)$	1.23 ± 0.23	0.26 ± 0.07	1.49 ± 0.24	1
$(K\pi\pi^0)(K\pi\pi)$	2.95 ± 0.32	1.03 ± 0.15	3.98 ± 0.36	6
$(K\pi\pi^0)(K_S^0\pi)$	0.28 ± 0.03	0.26 ± 0.07	0.53 ± 0.08	1
$(K\pi\pi^0)(KK\pi)$	0.23 ± 0.03	0.69 ± 0.12	0.92 ± 0.12	0
$(K3\pi)(K3\pi)$	1.31 ± 0.16	0.05 ± 0.03	1.36 ± 0.16	2
$(K3\pi)(K_S^0\pi\pi)$	0.53 ± 0.10	0.17 ± 0.06	0.70 ± 0.12	3
$(K3\pi)(K\pi\pi)$	1.50 ± 0.16	0.55 ± 0.10	2.05 ± 0.19	1
$(K3\pi)(K_S^0\pi)$	0.16 ± 0.02	0.15 ± 0.05	0.31 ± 0.06	0
$(K3\pi)(KK\pi)$	0.12 ± 0.01	0.27 ± 0.07	0.39 ± 0.07	1
$(K_S^0\pi\pi)(K_S^0\pi\pi)$	0.08 ± 0.01	0.02 ± 0.02	0.10 ± 0.02	0
$(K_S^0\pi\pi)(K\pi\pi)$	0.44 ± 0.05	0.29 ± 0.07	0.73 ± 0.09	0
$(K_S^0\pi\pi)(K_S^0\pi)$	0.04 ± 0.01	0.02 ± 0.02	0.06 ± 0.02	0
$(K_S^0\pi\pi)(KK\pi)$	0.03 ± 0.01	0.10 ± 0.04	0.14 ± 0.04	0

A.9 Signal Determination for $B^0 \to D^{*\pm}D^{\mp}$ and $B^{\pm} \to D^{*\pm}D^{*0}$

A.9.1 Background Estimation

In order to ascertain the contribution to these candidate events from the actual signal in question, we must estimate the expected contribution from background. This is accomplished by fitting the data sideband in the ΔE vs. $m_{\rm ES}$ plane.

An unbinned maximum likelihood fit (using the RooFitTools package [33]) is used to fit the full 2-dimensional data ΔE vs. $m_{\rm ES}$ sideband (see Figure A.12) to a variety of functions. ΔE , rather than $\Delta E'$, is used for the background fit. This is due to the fact that, although $\Delta E'$ removes the correlations in the signal between ΔE and $m_{\rm ES}$, $\Delta E'$ adds such correlations in the background. Thus, in order to fit using uncoupled functions for background in ΔE and $m_{\rm ES}$, ΔE must be used instead of $\Delta E'$. The standard ARGUS parametrization for $m_{\rm ES}$ and a linear polynomial for ΔE are used for the background parametrization. We also fit a large variety of other functions to the background and use the results of these fits to establish the systematic error on the fit from the parametrization. A summary of the fit functions used for $B^0 \to D^{*\pm}D^{\mp}$ are in the table below:

m_{ES} bkgd.	ΔE bkgd.
ARGUS	**linear**
linear	quadratic
	constant

This gives a total of 6 ($= 2 \times 3$) possible $B^0 \to D^{*\pm}D^{\mp}$ fits. Each of these 6 fit combinations is done, and the results for the amount of background in the signal region are summarized in Table A.21.

Table A.21: Amount of background in the signal region for each of the $B^0 \to D^{*\pm}D^{\mp}$ fits. Each of the numbers is $\pm 3.7\%$ due to the sideband statistics.

Background parametrization					
ARGUS m_{ES}			linear m_{ES}		
ΔE linear	ΔE quad.	ΔE const.	ΔE linear	ΔE quad.	ΔE const.
8.89	8.96	8.57	10.06	10.14	9.71

A summary of the fit functions used for $B^{\pm} \to D^{*\pm}D^{*0}$ are in the table below:

m_{ES} bkgd.	$\Delta E'$ bkgd.
ARGUS	**linear**
linear	quadratic
constant	constant

This gives a total of 9 ($= 3 \times 3$) possible $B^{\pm} \rightarrow D^{*\pm}D^{*0}$ fits. Each of these 9 fit combinations is done, and the results for the amount of background in the signal region are summarized in Table A.22.

For the central value of the background measurement, we choose the fit where m_{ES} background is ARGUS and ΔE background is linear. Other fits are used to establish a systematic error for the combinatoric background measurement. This gives measurements of

$$N_{comb.bkgd} = 8.89 \pm 0.33(stat) \pm 1.25(syst) \qquad (B^0 \rightarrow D^{*\pm}D^{\mp})$$

$$N_{comb.bkgd} = 17.25 \pm 0.54(stat) \pm 1.63(syst) \qquad (B^{\pm} \rightarrow D^{*\pm}D^{*0})$$

taking the systematic error to cover the full range of all the above fits.

The parameters and their errors of the fit used for the central value are listed in Tables A.23 and A.24. Plots of the 2D fits are in Figures A.26 and A.28 and m_{ES} and ΔE projections are in Figures A.27 and A.30.

A.9.2 $B^0 \rightarrow D^{*\pm}D^{\mp}$ Signal

After all analysis selection criteria are applied, as described in previous sections, we are left with 31 $B^0 \rightarrow D^{*\pm}D^{\mp}$ candidates in the signal region of the ΔE vs. m_{ES} plane.

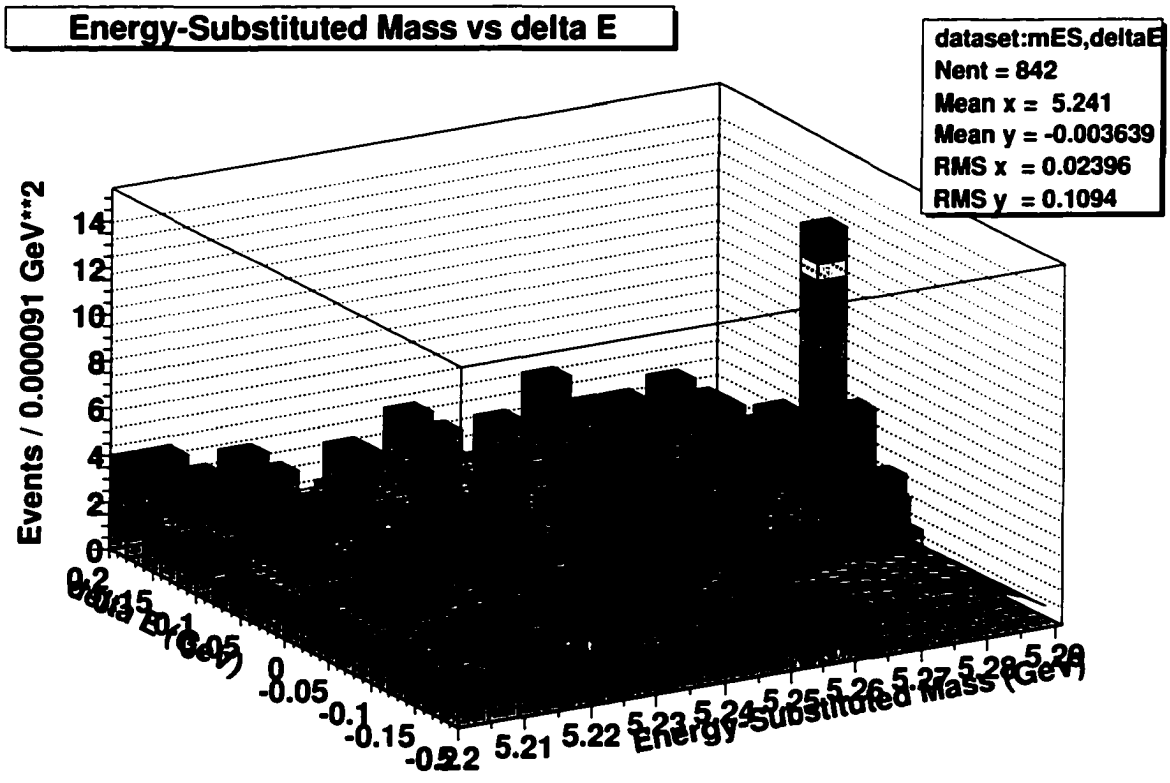


Figure A.26: 2-dimensional unbinned maximum likelihood fit of the $B^0 \rightarrow D^{*\pm}D^{\mp}$ ΔE vs. $m_{\rm ES}$ sideband in data (after all analysis cuts). The fit is to the product of an ARGUS function (in $m_{\rm ES}$) and a linear polynomial (in ΔE) for background. Only the sideband (see Figure A.12) is fit; the feed-down region and area around the signal region are excluded from the background fit (as explained in the previous pages). This fit is used for the central value of the background estimating, with 5 other such fits to different parametrizations to establish the systematic error on the background measurement from the parametrization. **NOTE**: the feed-down region ($m_{\rm ES} > 5.26\ GeV, \Delta E < -0.05\ GeV$) has its *data bins removed* in the above plot and, as explained above, is *not part of the fit.*

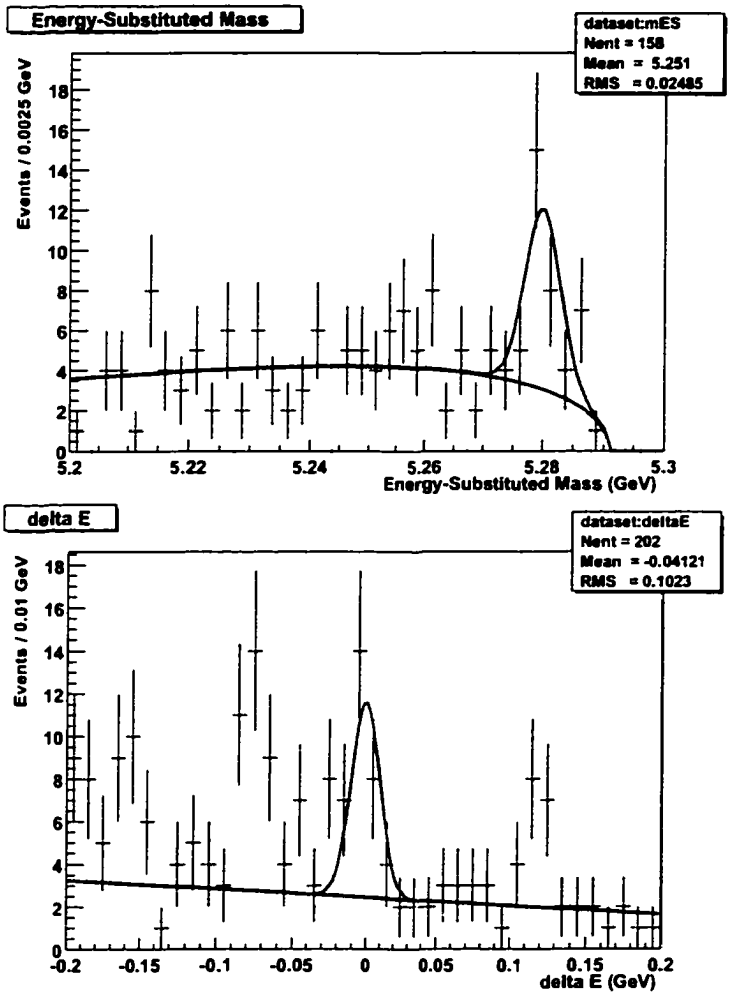


Figure A.27: (top) Projection in $m_{\rm ES}$ (with $-0.025 < \Delta E < 0.025$) of the 2D fit to the ΔE vs. $m_{\rm ES}$ plane. Points with error bars are data and blue line is background fit projection. Red line is the background fit projection plus a gaussian fit to the residual in the region $5.26 < m_{\rm ES} < 5.291$. (bottom) Projection in ΔE (with $5.273 < m_{\rm ES} < 5.285$) of the 2D $B^0 \to D^{*\pm}D^{\mp}$ fit to the ΔE vs. $m_{\rm ES}$ plane. Points with error bars are data and blue line is background fit projection. **NOTE**: points below ΔE = -0.05 GeV are in the feed-down region and are thus *not part of the fit.* Red line is the background fit projection plus a gaussian fit (centered at 0.) to the residual in the region $-0.05 < \Delta E < 0.05$.

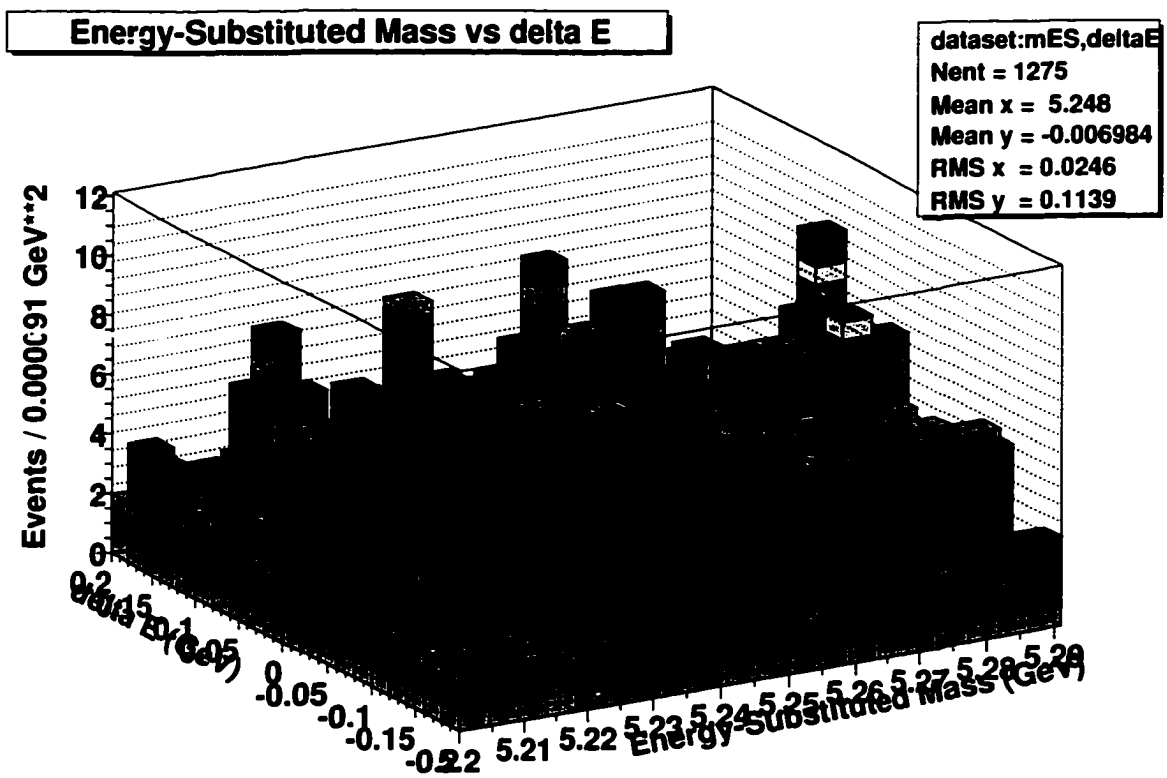


Figure A.28: 2-dimensional unbinned maximum likelihood fit of the $B^{\pm} \rightarrow D^{*\pm}D^{*0}$ $\Delta E'$ vs. $m_{\rm ES}$ sideband in data (after all analysis cuts). The fit is to the product of an ARGUS function (in $m_{\rm ES}$) and a linear polynomial (in $\Delta E'$) for background. Only the sideband (see Figure A.12) is fit; the area around the signal region is excluded from the background fit (as explained in previous pages). This fit is used for the central value of the background estimating, with 8 other such fits to different parametrizations to establish the systematic error on the background measurement from the parametrization.

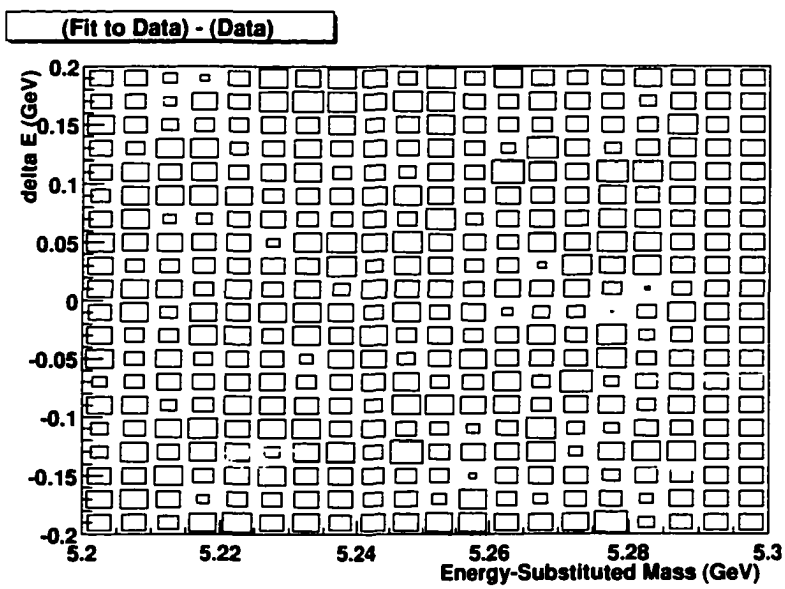


Figure A.29: Difference of the 2-dimensional $B^{\pm} \rightarrow D^{*\pm}D^{*0}$ fit (see previous page) and the data. This plot indicates that there are no signs of any significant clumping in data (other than, of course, the signal itself) not modelled by the fit.

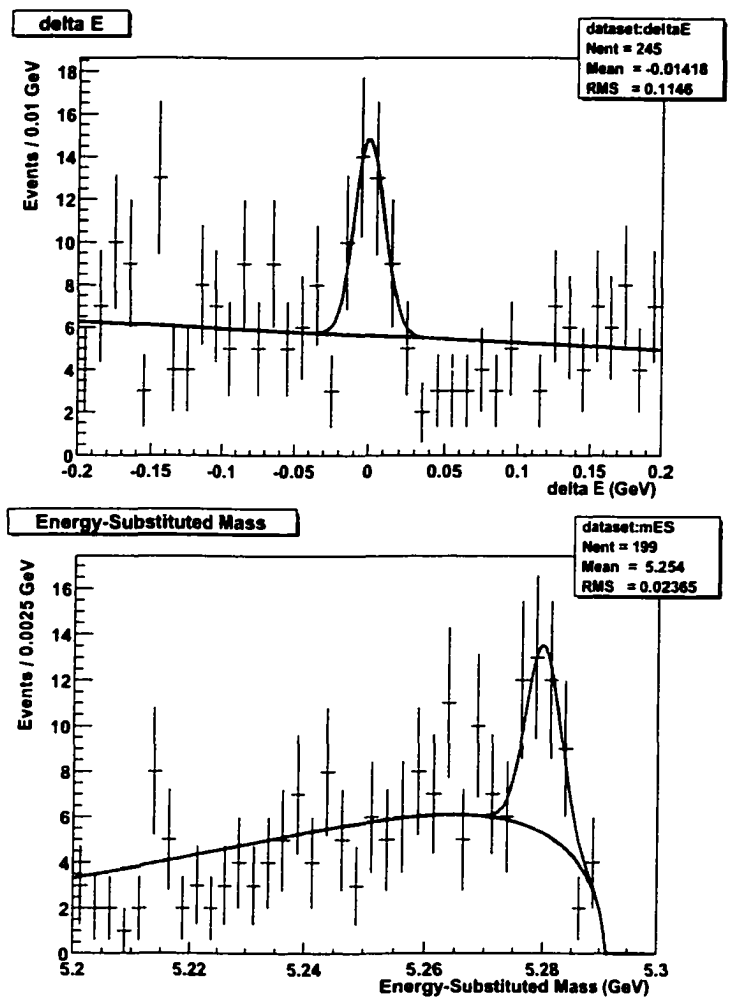


Figure A.30: (top) Projection in $\Delta E'$ (with $5.273 < m_{\rm ES} < 5.285$) of the 2D $B^{\pm} \to D^{*\pm}D^{*0}$ fit to the $\Delta E'$ vs. $m_{\rm ES}$ plane. Points with error bars are data and blue line is fit projection. Red line is the background fit projection plus a gaussian fit (centered at 0.) to the residual in the region $-0.05 < \Delta E' < 0.05$. (bottom) Projection in $m_{\rm ES}$ (with $-0.025 < \Delta E' < 0.025$) of the 2D fit to the $\Delta E'$ vs. $m_{\rm ES}$ plane. Points with error bars are data and blue line is fit projection. Red line is the background fit projection plus a gaussian fit to the residual in the region $5.26 < m_{\rm ES} < 5.291$.

Table A.22: Amount of background in the signal region for each of the $B^{\pm} \rightarrow D^{*\pm}D^{*0}$ fits. Each of the numbers is $\pm 3.0\%$ due to the sideband statistics.

Bkgd. parametrization		
$m_{\rm ES}$ param.	$\Delta E'$ param.	$N_{bkgd.}$
ARGUS	linear	**17.25**
	quad.	16.99
	const.	17.25
linear	linear	15.98
	quad.	15.73
	const.	15.98
constant	linear	15.87
	quad.	15.62
	const.	15.87

Based on the 31 candidate events in the signal region and the background estimation from above, the number of signal events before peaking background subtraction (described below) is $N_{Sig+PeakBkg} = 22.1^{+6.8}_{-5.7}(stat) \pm 1.3(syst)$. The statistical error includes the both the statistical error on the number of events in the signal region (calculated using the Gehrels approximation to Poisson [34], where the upper error is $1 + \sqrt{N + 0.75}$ and the lower error is $\sqrt{N - 0.25}$) which predominates, and the statistical error on the number of background events. The systematic error is the uncertainty quoted previously on N_{bkgd}.

Table A.23: Fitted parameters in the data 2D signal+background ΔE vs. $m_{\rm ES}$ fit.

Parameter	Value	
ARGUS parameter	-27.11±4.55	
ΔE slope	-1.72±0.29	$(N_{evts}/90\ \mathrm{MeV}^2)/\mathrm{GeV}$
	With fixed:	
ARGUS kinematic limit	5291.00	MeV

Studies of potential peaking background in the signal region have been done and are described in the following section on systematic errors. The studies described there allow us to determine that the fraction of of (combinatoric background-subtracted) signal events that are in fact peaking background (primarily from D^*KX events) is $7.3 \pm 7.6\%$, or $+1.6 \pm 1.7$ events. This is subtracted from the measured number of signal events, giving

$$N_{Sig} = 20.5^{+6.8}_{-5.7}(\mathrm{stat}) \pm 1.3(\mathrm{comb.\ bkg.\ syst}) \pm 1.7(\mathrm{peak\ bkg.\ syst})$$

As a consistency check, we break up the data and Monte Carlo samples by final state. Table A.25 shows a breakdown of the signal region by D decay mode.

A.9.3 $B^\pm \to D^{*\pm}D^{*0}$ Signal

After all analysis selection criteria are applied, as described in previous sections, we are left with 39 $B^\pm \to D^{*\pm}D^{*0}$ candidates in the signal region of

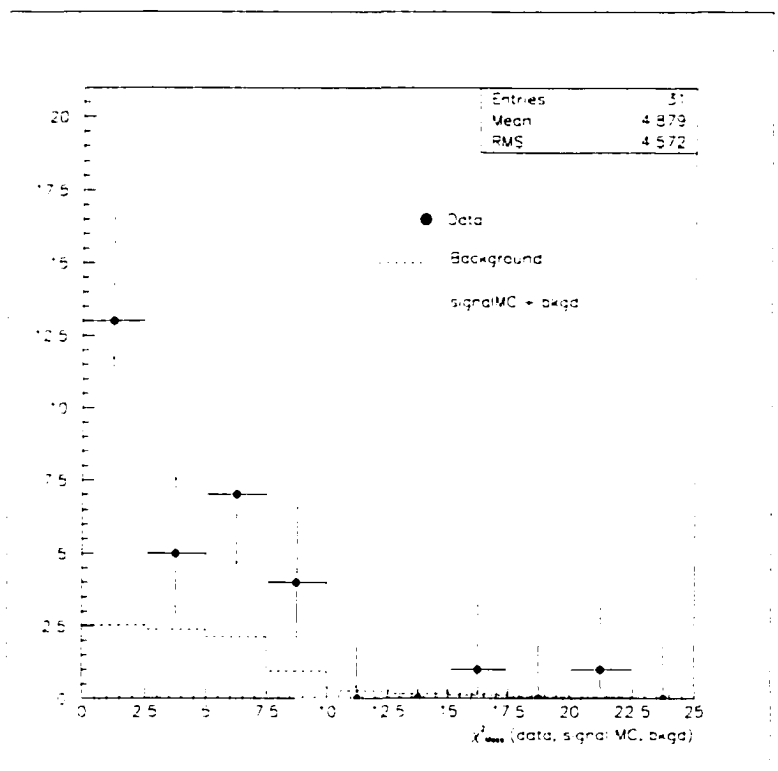


Figure A.31: χ^2_{Mass} for events in the $B^0 \rightarrow D^{*\pm}D^{\mp}$ signal region. Points with error bars are data. The error bars are the asymmetric Poisson errors (using the Gehrels approximation [34]). The dashed histogram is the χ^2_{Mass} distribution from the data sideband region, scaled by the ratio of the fitted background in the signal region over the total number of events in the sideband. The dotted histogram is the sum of the dashed histogram and the χ^2_{Mass} distribution from signal Monte Carlo, normalized to the number of signal events in the data signal region.

Table A.24: Fitted parameters in the data 2D signal+background $\Delta E'$ vs. $m_{\rm ES}$ fit.

Parameter	Value	
ARGUS parameter	-50.07±3.33	
$\Delta E'$ slope	-0.50±0.25	$(N_{evts}/90\ \mathrm{MeV}^2)/\mathrm{GeV}$
	With fixed:	
ARGUS kinematic limit	5291.00	MeV

the $\Delta E'$ vs. $m_{\rm ES}$ plane.

Based on the 39 candidate events in the signal region and the background estimation from above, the number of signal events is

$$N_{Sig} = 21.7^{+7.3}_{-6.2}(stat) \pm 1.6(syst)$$

The statistical error includes the both the statistical error on the number of events in the signal region (calculated using the Gehrels approximation to Poisson [34], where the upper error is $1 + \sqrt{N + 0.75}$ and the lower error is $\sqrt{N - 0.25}$) which predominates, and the statistical error on the number of background events. The systematic error is the uncertainty quoted previously on N_{bkgd}.

As a consistency check, we break up the data and Monte Carlo samples by final state. Table A.26 shows a breakdown of the signal region by D decay mode.

Table A.25: Comparison of $B^0 \to D^{*\pm}D^{\mp}$ events seen in signal region with expectation from signal Monte Carlo, broken down by D decay mode combination. The Monte Carlo expectation for signal is the $\epsilon\mathcal{B}$ for the given mode combination divided by the total $\epsilon\mathcal{B}$, all multiplied by the central value of the branching fraction seen in data (for normalization purposes). The background expectation is the number of events in the data sidebands for that mode combination multiplied by the ratio of the integral of the fitted background under the signal region to that in the sidebands. The sources of uncertainty in the expected values include the Monte Carlo statistics in determining the efficiency, the uncertainties on the D^* and D branching fractions, the uncertainty on r' (see Section A.7), and the statistical uncertainty from the number of events in the sideband for that mode combination.

Final State		N_{Sig}	N_{Bkg}	N_{Total}	N_{Data}		
D^*	$D^{\pm}$	(MC Expected)	(Sideband)		$D^{*+}D^-$	$D^{*-}D^+$	Total
$(K\pi)\pi^{\pm}$	$K\pi\pi$	5.83 ± 0.23	1.43 ± 0.13	7.26 ± 0.26	3	5	8
$(K\pi)\pi^{\pm}$	$K^0_S\pi$	0.65 ± 0.03	0.18 ± 0.05	0.83 ± 0.06	1	1	2
$(K\pi)\pi^{\pm}$	$KK\pi$	0.40 ± 0.02	0.34 ± 0.07	0.74 ± 0.07	1	0	1
$(K\pi\pi^0)\pi^{\pm}$	$K\pi\pi$	7.29 ± 0.40	3.14 ± 0.20	10.43 ± 0.45	6	7	13
$(K\pi\pi^0)\pi^{\pm}$	$KK\pi$	0.05 ± 0.01	0.02 ± 0.02	0.07 ± 0.02	1	0	1
$(K3\pi)\pi^{\pm}$	$K\pi\pi$	3.55 ± 0.25	1.14 ± 0.12	4.69 ± 0.28	1	1	2
$(K\pi\pi)\pi^0$	$K\pi\pi$	2.40 ± 0.15	1.61 ± 0.14	4.01 ± 0.21	0	1	1
$(K\pi\pi)\pi^0$	$K^0_S\pi$	0.29 ± 0.02	0.42 ± 0.07	0.71 ± 0.07	1	2	3
$(K^0_S\pi)\pi^0$	$K\pi\pi$	0.26 ± 0.02	0.25 ± 0.06	0.51 ± 0.06	1	0	1
$(K^0_S\pi)\pi^0$	$K^0_S\pi$	0.03 ± 0.00	0.04 ± 0.02	0.07 ± 0.02	0	0	0

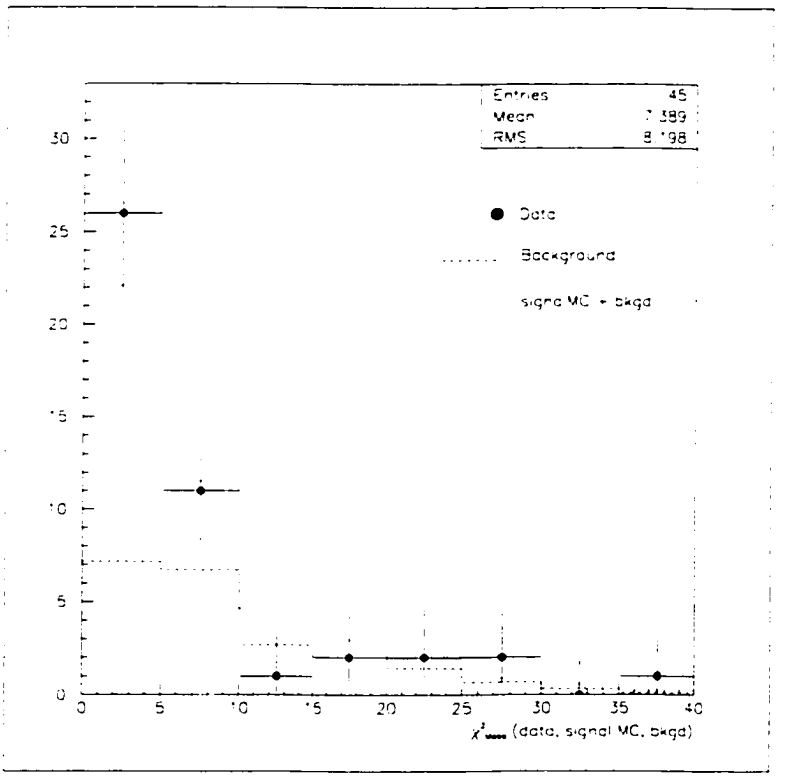


Figure A.32: χ^2_{Mass} for events in the $B^{\pm} \rightarrow D^{*\pm}D^{*0}$ signal region. Points with error bars are data. The error bars are the asymmetric Poisson errors (using the Gehrels approximation [34]). The dashed histogram is the χ^2_{Mass} distribution from the data sideband region, scaled by the ratio of the fitted background in the signal region over the total number of events in the sideband. The dotted histogram is the sum of the dashed histogram and the χ^2_{Mass} distribution from signal Monte Carlo, normalized to the number of signal events in the data signal region.

Table A.26: Comparison of events seen in the $B^{\pm} \to D^{*\pm}D^{*0}$ signal region with expectation from signal Monte Carlo, broken down by D decay mode combination. The Monte Carlo expectation for signal is the $\epsilon\mathcal{B}$ for the given mode combination divided by the total $\epsilon\mathcal{B}$, all multiplied by the central value of the branching fraction seen in data (for normalization purposes). The background expectation is the number of events in the data sidebands for that mode combination multiplied by the ratio of the integral of the fitted background under the signal region to that in the sidebands. The sources of uncertainty in the expected values include the Monte Carlo statistics in determining the efficiency, the uncertainties on the D^* and D^{*0} branching fractions, the uncertainty on r' (see Section A.7), and the statistical uncertainty from the number of events in the sideband for that mode combination.

Final State		N_{Sig}	N_{Bkg}	N_{Total}	N_{Data}
D^*	D^{*0}	(MC Expected)	(Sideband)		
$(K\pi)\pi^{\pm}$	$(K\pi)\pi^0$	1.44 ± 0.55	0.08 ± 0.03	1.52 ± 0.55	4
$(K\pi)\pi^{\pm}$	$(K\pi\pi^0)\pi^0$	2.09 ± 0.83	0.82 ± 0.11	2.91 ± 0.84	3
$(K\pi)\pi^{\pm}$	$(K3\pi)\pi^0$	1.21 ± 0.48	0.37 ± 0.08	1.58 ± 0.49	0
$(K\pi)\pi^{\pm}$	$(K\pi)\gamma$	0.70 ± 0.28	0.25 ± 0.06	0.95 ± 0.29	0
$(K\pi)\pi^{\pm}$	$(K\pi\pi^0)\gamma$	1.19 ± 0.51	1.60 ± 0.16	2.79 ± 0.53	3
$(K\pi)\pi^{\pm}$	$(K3\pi)\gamma$	0.40 ± 0.18	0.73 ± 0.11	1.13 ± 0.21	1
$(K\pi)\pi^{\pm}$	$(K_S^0\pi\pi)\gamma$	0.08 ± 0.04	0.15 ± 0.05	0.23 ± 0.06	0
$(K\pi\pi^0)\pi^{\pm}$	$(K\pi)\pi^0$	1.50 ± 0.61	0.19 ± 0.05	1.69 ± 0.61	3
$(K\pi\pi^0)\pi^{\pm}$	$(K\pi\pi^0)\pi^0$	2.54 ± 0.73	4.89 ± 0.28	7.43 ± 0.78	5
$(K\pi\pi^0)\pi^{\pm}$	$(K3\pi)\pi^0$	1.27 ± 0.37	1.13 ± 0.13	2.40 ± 0.39	2
$(K\pi\pi^0)\pi^{\pm}$	$(K\pi)\gamma$	0.76 ± 0.41	0.59 ± 0.10	1.35 ± 0.42	3
$(K\pi\pi^0)\pi^{\pm}$	$(K\pi\pi^0)\gamma$	1.51 ± 0.80	2.09 ± 0.18	3.60 ± 0.82	2
$(K3\pi)\pi^{\pm}$	$(K\pi)\pi^0$	0.97 ± 0.59	0.12 ± 0.04	1.09 ± 0.59	0
$(K3\pi)\pi^{\pm}$	$(K\pi\pi^0)\pi^0$	1.66 ± 0.72	1.75 ± 0.16	3.41 ± 0.74	2
$(K3\pi)\pi^{\pm}$	$(K3\pi)\pi^0$	0.84 ± 0.37	0.67 ± 0.10	1.51 ± 0.38	2
$(K3\pi)\pi^{\pm}$	$(K_S^0\pi\pi)\pi^0$	0.10 ± 0.06	0.08 ± 0.03	0.18 ± 0.07	1
$(K3\pi)\pi^{\pm}$	$(K\pi)\gamma$	0.99 ± 0.41	0.33 ± 0.07	1.40 ± 0.42	1
$(K3\pi)\pi^{\pm}$	$(K3\pi)\gamma$	0.53 ± 0.27	0.64 ± 0.10	1.17 ± 0.29	1
$(K_S^0\pi\pi)\pi^{\pm}$	$(K\pi)\pi^0$	0.30 ± 0.12	0.15 ± 0.04	0.45 ± 0.13	1
$(K_S^0\pi\pi)\pi^{\pm}$	$(K3\pi)\pi^0$	0.19 ± 0.09	0.29 ± 0.07	0.48 ± 0.11	1
$(K_S^0\pi\pi)\pi^{\pm}$	$(K\pi)\gamma$	0.18 ± 0.08	0.06 ± 0.03	0.24 ± 0.09	1
$(K\pi\pi)\pi^0$	$(K\pi)\pi^0$	0.65 ± 0.29	0.46 ± 0.08	1.11 ± 0.29	3

A.10 Systematic Uncertainties

A.10.1 Tracking Efficiency

The systematic error on `GoodTracksLoose` is 1.2% and for `GoodTracksVeryLoose` 1.3% per track. For slow pions a systematic error of 2.6% is assigned. The tracks from K_S^0 decays come from the `ChargedTracks` list and are each assigned a 1.0% systematic uncertainty, or 2.0% per K_S^0. Tables A.27, A.28, and A.29 show the average number of tracks in each category, weighted by $\epsilon\mathcal{B}$, for each of the decay modes. Each of these tracking efficiencies is treated as correlated, so the total systematic is the linear sum as opposed to the quadrature sum. Tracks from charged kaons are not included here as they are accounted for separately, except in the decay $D^0 \to K^-\pi^+$ where no kaon identification is used.

Table A.27: Average number of tracks falling in each of the various track categories for $B^0 \to D^{*+}D^{*-}$ decays, and their associated systematic uncertainty due to tracking efficiency.

Track Type	Syst. per Track (%)	Avg. Number	Net Syst. (%)
`GoodTracksLoose`	1.2	1.873	2.25
`GoodTracksVeryLoose`	1.3	3.087	4.01
K_S^0	2.0	0.127	0.25
Slow Charge π	2.6	1.763	4.58
Total			9.94

Table A.28: Average number of tracks falling in each of the various track categories for $B^0 \to D^{*\pm}D^{\mp}$ decays, and their associated systematic uncertainty due to tracking efficiency.

Track Type	Syst. per Track (%)	Avg. Number	Net Syst. (%)
`GoodTracksLoose`	1.2	0.299	0.36
`GoodTracksVeryLoose`	1.3	4.981	6.48
K_S^0	2.0	0.146	0.29
Slow Charge π	2.6	0.868	2.26
Total			9.39

Table A.29: Average number of tracks falling in each of the various track categories for $B^{\pm} \to D^{*\pm}D^{*0}$ decays, and their associated systematic uncertainty due to tracking efficiency.

Track Type	Syst. per Track (%)	Avg. Number	Net Syst. (%)
`GoodTracksLoose`	1.2	0.605	0.73
`GoodTracksVeryLoose`	1.3	4.327	5.63
K_S^0	2.0	0.072	0.14
Slow Charge π	2.6	0.908	2.36
Total			8.86

A.10.2 π^0 Efficiency

The systematic uncertainty on the π^0 efficiency is taken as 2.5% per π^0, including the soft π^0 from the decay $D^{*+} \to D^+\pi^0$ [32]. The $\epsilon\mathcal{B}$ weighted average number of π^0s is 0.99 for $B^0 \to D^{*+}D^{*-}$, 0.480 for $B^0 \to D^{*\pm}D^{\mp}$, and 1.54 for $B^\pm \to D^{*\pm}D^{*0}$. We therefore assign a systematic uncertainty due to the π^0 efficiency of 2.5

A.10.3 γ Efficiency

In $B^\pm \to D^{*\pm}D^{*0}$, the systematic uncertainty on the (slow) γ (from $D^{*0} \to D^0\gamma$) efficiency is taken as 1.25% per γ [32]. The $\epsilon\mathcal{B}$ weighted average number of γs is 0.32. We therefore assign a systematic uncertainty of 0.4% due to the γ efficiency.

The π^0 and γ efficiencies as correlated, so the systematic from these two sources in $B^\pm \to D^{*\pm}D^{*0}$ is the linear sum (4.3%).

A.10.4 K_S^0 Efficiency

The systematic uncertainty on the K_S^0 efficiency is already included in the charged track efficiency. Since the total systematic assigned to K_S^0 efficiency comes from the tracking efficiency for `ChargedTracks`, an additional systematic is not assigned here [31].

A.10.5 Charged Kaon ID

To determine the systematic uncertainty from Kaon ID Monte Carlo events without PID correction are compared with Monte Carlo with PID correc-

tion [29]. The $\epsilon\mathcal{B}$ for Monte Carlo events without PID correction is enhanced by 2.8% in each of $B^0 \to D^{*+}D^{*-}$, $B^0 \to D^{*\pm}D^{\mp}$, and $B^{\pm} \to D^{*\pm}D^{*0}$, and therefore a systematic uncertainty of 2.8% to charged Kaon ID is assigned for each of the three modes.

A.10.6 Resolution Differences

To understand the sensitivity to unaccounted-for differences between the data and Monte Carlo resolutions on reconstructed masses, $\Delta E'$, and $m_{\rm ES}$, the cuts on χ^2_{Mass}, $\Delta E'$, and $m_{\rm ES}$ are each varied and a systematic uncertainty on the total $\epsilon\mathcal{B}$ equal to the fractional change in $\epsilon\mathcal{B}$ is assigned.

For the variation of χ^2_{Mass}, we are guided by the pull distribution plots for the D^0 and D^+ shown in [26]. It is from these plots that the factor of 1.2 that is used to scale the error on the D mass is extracted when constructing χ^2_{Mass} in data. This factor of 1.2 has a 4% relative error. Because the Ds in our sample are not of the exact same momentum spectrum, and we include some additional D^+ decay modes, we take a conservative uncertainty on the scaling factor of 10%. The χ^2_{Mass} is inversely proportional to the square of this scale factor. This translates to a change in the χ^2_{Mass} scale of $\pm 20\%$. As we change the cut on χ^2_{Mass} from 16 to 24, the total $\epsilon\mathcal{B}$ changes from 18.00×10^{-4} to 18.88×10^{-4}, leading to a systematic error of 2.7%. It should be noted, however, that from Figures A.25, A.31, and A.32 the mean of the χ^2_{Mass} in data and the mean of the sum of background and signal Monte Carlo agree with each other within the errors of each.

For the variation of $\Delta E'$ and $m_{\rm ES}$, we also look at a 10% variation on the cut values. Comparing Figure A.27 for data and Figure A.3, one sees that

although the errors on the resolutions for the data fits are rather large, both the $\Delta E'$ and $m_{\rm ES}$ resolutions are within 1 sigma of the Monte Carlo values. Changing the cut on $\Delta E'$ by $\pm 10\%$ produces a systematic uncertainty on $\epsilon\mathcal{B}$ of 2.3%. Variation of the $m_{\rm ES}$ cut produces an uncertainty of 2.0%.

Adding these three resolution uncertainties in quadrature produces a net uncertainty of 4.1%.

A.10.7 Vertex Fitting

As described in Section A.7, the ratio r' was calculated to determine a correction factor for the efficiency measurement due to the requirement of converging vertex fits on D candidates and the cut of χ^2 probability > 0.001. The systematic uncertainty on $\epsilon\mathcal{B}$ determined in Section A.7 was 2.6% for both $B^0 \to D^{*+}D^{*-}$ and $B^0 \to D^{*\pm}D^{\mp}$, and 2.8% for $B^{\pm} \to D^{*\pm}D^{*0}$.

A.10.8 Background Estimation

The systematic uncertainty on the (combinatoric) background estimation is described in Section A.9. The statistical error on the background estimation coming from the number of events in the sideband region will be incorporated in the statistical error on the branching fraction. The systematic uncertainty propagates into the branching fraction as

$$\frac{\sigma_{Bkg}(syst)}{N_{Sig}} = 1.1\% \qquad (B^0 \to D^{*+}D^{*-})$$

$$\frac{\sigma_{Bkg}(syst)}{N_{Sig}} = 7.1\% \qquad (B^0 \to D^{*\pm}D^{\mp})$$

$$\frac{\sigma_{Bkg}(syst)}{N_{Sig}} = 7.5\% \qquad (B^{\pm} \to D^{*\pm}D^{*0})$$

relative systematics on the branching fractions.

A.10.9 Peaking Backgrounds

Possible sources of peaking backgrounds for $B^0 \to D^{*+}D^{*-}$ potentially not well modeled in the Monte Carlo sample were explored. Possible sources include charged $B \to D^{*0}D^*$ (which is currently not included in generic MC) and $B^0 \to D_s^{(*)}D^*$ (which is included in generic MC but for which the current published branching ratios have a relative error of over 25%[24]).

Figure A.33 shows the $\Delta E'$ vs m_{ES} behavior for events in these channels reconstructed as $B^0 \to D^{*+}D^{*-}$. From the expected relation $BR(B^+ \to D^{*0}D^*) \approx BR(B^0 \to D^{*+}D^{*-})$ and the error on the branching fraction for $BR(B^0 \to D_s^{(*)}D^*)$ we determine the relative systematic uncertainty on the branching ratio due to peaking backgrounds not well modeled in Monte Carlo to be 1.1%.

Examination of background events within the $B^0 \to D^{*\pm}D^{\mp}$ signal region in generic $B^0\overline{B}^0$ MC indicted that peaking background could occur from events of the type $B^0 \to D^*KX$, where the KX system has the same final state as a D submode and happens to be reconstructed within the D mass resolution range. Four such events were found in 8 fb^{-1} of generic $B^0\overline{B}^0$ MC, indicating that it could potentially be a significant effect.

To determine the amount of peaking background within the observed data signal peak, events were reconstructed with a shifted D mass window. Since, as described above, the background is primarily from events with a false D, shifting the D mass window and determining the resulting "signal" will give an estimate of the amount of false D background included in the

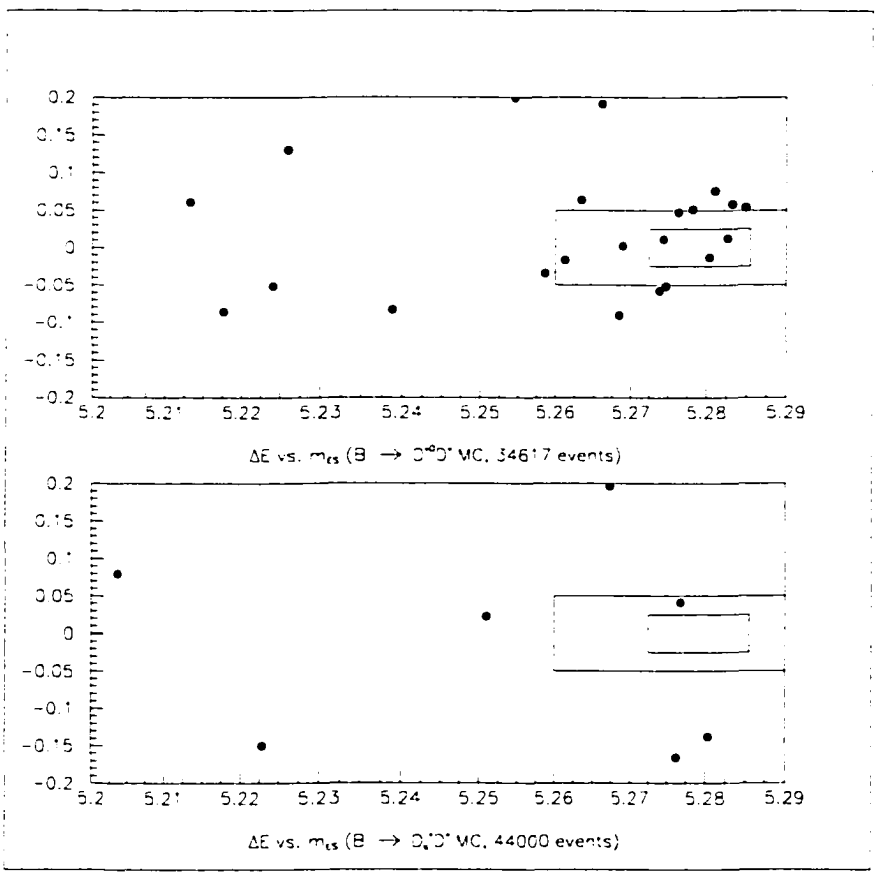


Figure A.33: The upper plot shows the background on the $\Delta E'$ vs m_{ES} plane from 34617 generated charged $B \longrightarrow D^{*0}D^{*}$ events (which corresponds to an effective luminosity of 550 fb^{-1}). The lower plot shows the background on the $\Delta E'$ vs m_{ES} plane from 44000 generated $B^0 \longrightarrow D_s^{*}D^{*}$ events (which corresponds to an effective luminosity of 20 fb^{-1}).

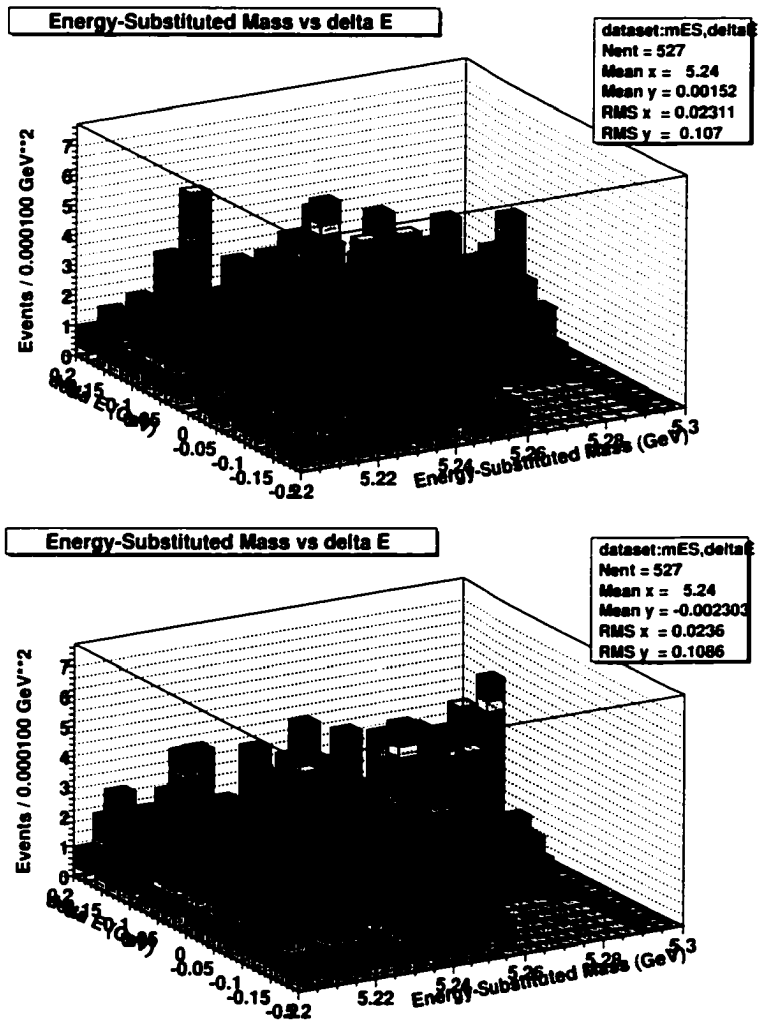


Figure A.34: 2-dimensional ΔE vs. $m_{\rm ES}$ plots of shifted D mass window events, used for estimating peaking background, for (top) D mass 20 MeV above nominal and (bottom) D mass 20 MeV below nominal. After combinatoric background subtraction (via the two-dimensional unbinned maximum likelihood fit to the sideband, as shown in both plots), the signals for these D mass shifted events were used to estimate the contribution from peaking background (from false D events) to $B^0 \to D^{*\pm}D^{\mp}$ signal.

$B^0 \rightarrow D^{*\pm}D^{\mp}$ signal itself. Data events were reconstructed with D mass shifted ± 20 MeV (the limits set in ntuple production) with respect to the nominal D mass (see Figure A.34). The resulting false signals produced an estimate of $7.3 \pm 7.6\%$ of the $B^0 \rightarrow D^{*\pm}D^{\mp}$ signal to be from peaking background. As described in the previous section, this value was subtracted from the number of signal events reconstructed. The 7.6% error on the measurement is taken to be the systematic error from peaking background estimation.

A.10.10 Transversity Amplitudes

For $B^0 \rightarrow D^{*+}D^{*-}$ and $B^{\pm} \rightarrow D^{*\pm}D^{*0}$, to obtain the systematic due to lack of knowledge of true transversity amplitudes in data (which result in differing slow pion kinematics and thus a transversity-dependent efficiency), 5000 Monte Carlo events in each of the two extremes of transversity amplitudes: 1) $|G_{1+}| = 1 \quad |G_{0+}| = 0 \quad |G_{1-}| = 0$ and 2) $|G_{1+}| = 0 \quad |G_{0+}| = 1 \quad |G_{1-}| = 0$, and each of the two modes: a) $D^0\pi + D^0\pi \rightarrow (K\pi)\pi + (K\pi)\pi$ and b) $D^0\pi + D^{\pm}\pi^0 \rightarrow (K\pi)\pi + (K\pi\pi)\pi^0$, were generated and reconstructed (4*5000 = a total of 20000 events). Figure A.35 shows the p_T distributions at the generator level for different values of the transversity amplitudes, including the set of amplitudes used for signal Monte Carlo generation.

Efficiencies for reconstructing signal events were calculated from each of the 4 samples and the differences in resulting efficiencies between samples of differing transversity were used to calculate the systematic error from our lack of knowledge of the true transversity amplitudes. The efficiency differences were weighted by the relative $\epsilon\mathcal{B}$ for $D^0\pi + D^0\pi$ and $D^0\pi + D^{\pm}\pi^0$ events.

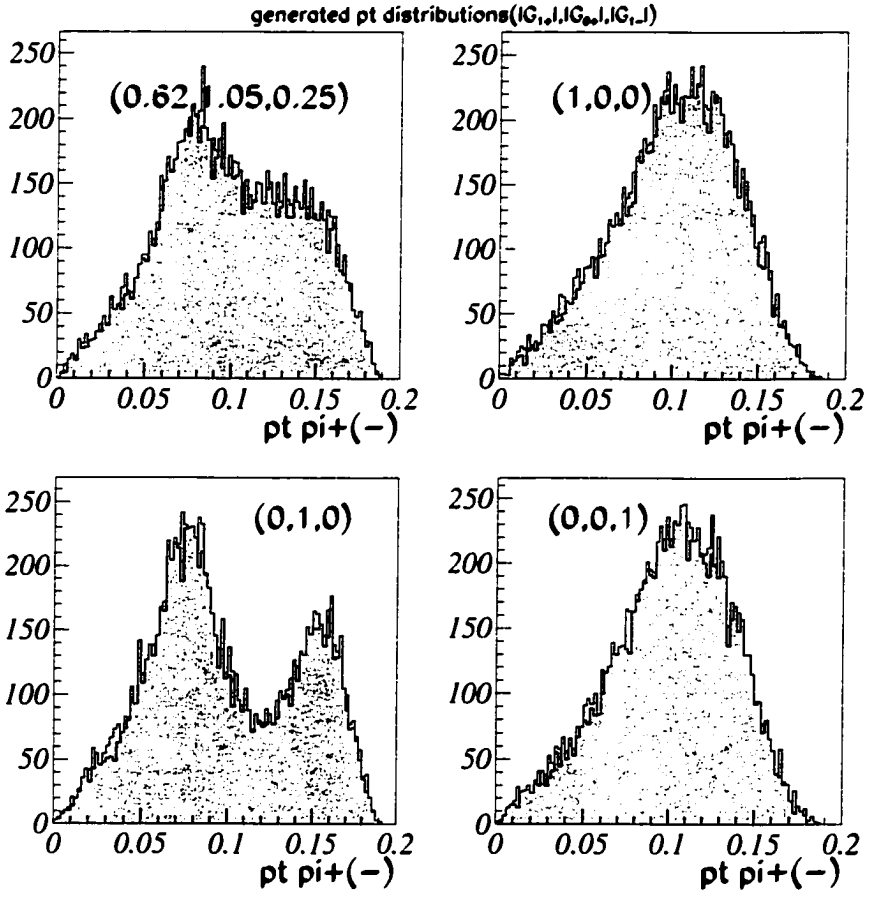


Figure A.35: Generator-level p_T distributions for the slow pion for various values of the transversity amplitudes. The upper-left plot shows the set of amplitudes used for signal Monte Carlo generation. The upper-right and lower-left show the same distribution for the sets of amplitudes used for the systematic uncertainty study.

The resulting relative systematic uncertainty due to transversity amplitudes of is 6.6%.

A.10.11 Summary

In Tables A.30, A.31, and A.32 the systematic uncertainties are summarized as their fractional error on the branching fraction.

Table A.30: Summary of systematic uncertainties for $B^0 \to D^{*+}D^{*-}$.

Source	Uncertainty (%)
Monte Carlo Statistics	2.7
Sub-mode Branching Fractions	5.6
Charged Tracking Efficiency	9.9
π^0 Efficiency	2.5
K_S^0 Efficiency	0.3
Kaon Identification	3.4
Vertex Fitting	2.6
Transversity Amplitudes	6.6
Background Estimation	1.1
Peaking Backgrounds	1.1
Resolution Differences	4.1
Number of B^0	1.6
Total	15.0

Table A.31: Summary of systematic uncertainties for $B^0 \to D^{*\pm}D^{\mp}$.

Source	Uncertainty (%)
Monte Carlo Statistics	2.3
Sub-mode Branching Fractions	7.4
Charged Tracking Efficiency	9.4
π^0 Efficiency	1.2
Kaon Identification	2.8
Vertex Fitting	2.6
Background Estimation	7.1
Peaking Backgrounds	7.6
Resolution Differences	4.1
Number of B^0	1.6
Total	17.1

Table A.32: Summary of systematic uncertainties for $B^{\pm} \to D^{*\pm}D^{*0}$.

Source	Uncertainty (%)
Monte Carlo Statistics	5.7
Sub-mode Branching Fractions	9.1
Charged Tracking Efficiency	8.9
π^0 and γ Efficiency	4.3
Kaon Identification	2.8
Vertex Fitting	2.8
Transversity Amplitudes	6.6
Background Estimation	7.5
Peaking Backgrounds	1.1
Resolution Differences	4.1
Number of B^0	1.6
Total	18.7

Appendix B

Details of $B \to D^{(*)}\overline{D}^{(*)}$ sin2β Measurement

B.1 Event Yields

B.1.1 *CP* Data Samples

Here we summarize the yields in 56 fb^{-1} of on-resonance data for the event selection described in Chapter 3. For the branching ratio analyses, the extraction of signal events is done via a cut-and-count procedure. For the *CP* fit the $m_{\rm ES}$ distribution is used to assign a signal probability to each event. We therefore quote below the signal yields determined by fitting to the $m_{\rm ES}$ distribution to an ARGUS+gaussian function. The resolution of the gaussian is fixed to that measured in the exclusive B data sample (2.63 MeV) and we obtain the yield from the area of the gaussian. Tables B.1 and B.2 contain the yields for $B^0 \to D^{*+}D^{*-}$ and $B^0 \to D^{*\pm}D^{\mp}$ in the 56 fb^{-1} datasample. In

Table B.3 we see the mode by mode break down of events in $B^0 \to D^{*+}D^-$, as well as the number of events reconstructed as $B^0 \to D^{*+}D^-$ versus those reconstructed as $B^0 \to D^{*-}D^+$.

Table B.1: Below are the results of the fits to the $m_{\rm ES}$ and ΔE projections shown in Figures B.2-B.3 and Figures B.5-B.8. These yields neglect any possible peaking background component. The value for $\sigma_{m_{\rm ES}} = 2.63\,\mathrm{MeV}$ was determined from the Breco data sample and was fixed in the fit to the $m_{\rm ES}$ projection.

mode	total area $m_{\rm ES}$ fit	$m_{\rm ES}$ yield with cuts	total area from ΔE fit	ΔE yield with cuts	purity
$B^0 \to D^{*+}D^{*-}$	94.1 ± 11.0	92.3	93.8 ± 11.6	91.5	77.9%
$B^0 \to D^{*+}D^-$	65.0 ± 11.0	60.4	60.1 ± 11.7	55.6	51.1%

Table B.2: Below are the details of the fits to the ΔE projections shown in Figures B.3 and B.8.

		$\sigma_{\Delta E}$ from ΔE fit	mean of ΔE fit
$B^0 \to D^{*+}D^{*-}$	Run1	16.0 ± 3.4	$+3.6 \pm 3.6$
	Run2	10.6 ± 1.6	-7.9 ± 1.8
	Total	12.6 ± 1.6	-4.6 ± 1.7
$B^0 \to D^{*+}D^{-}$	Run1	5.5 ± 1.1	-5.3 ± 1.9
	Run2	11.7 ± 2.4	$+0.7 \pm 2.7$
	Total	10.3 ± 2.0	-2.3 ± 2.2

B.1.2 Peaking backgrounds

Peaking background studies have already been performed for the BR determinations of $B^0 \to D^{*+}D^{*-}$ (PRL draft and BAD 166) and $B^0 \to D^{*+}D^{-}$ (BAD 274). These studies, based on 20.4 fb^{-1}, were done by using the generic MC (corresponding to 8 fb^{-1} of $B^0\bar{B}^0$, 12.8 fb^{-1} of B^+B^-, and 6.8 fb^{-1} of $c\bar{c}$ of MC SP3), simulated events for some specific potentially dangerous modes and finally data sidebands. More studies have now been done and the results are summarized in the following two sections.

Peaking backgrounds for $B^0 \to D^{*+}D^{*-}$

Based on the described samples of generic SP3 Monte Carlo events, we see no evidence of peaking of the background in the signal region.

Possible sources of peaking backgrounds potentially not well modeled in

Table B.3: Below is the mode by mode breakdown of the events seen in the $m_{\rm ES}$ projection for $B^0 \to D^{*+}D^-$. Given are the number of events in each reconstruction mode seen in $m_{\rm ES} > 5.27\,{\rm MeV}$. The mode name is denoted by the D^* decay mode first and the D mode second.

mode	number of events in signal region	no. of evts. $D^{*-}D^+$	no. of evts. $D^{*+}D^-$
$(K\pi)(K\pi\pi)$	22	18	4
$(K\pi\pi^0)(K\pi\pi)$	56	33	23
$(K3\pi)(K\pi\pi)$	31	10	21
$(K^0_S\pi\pi)(K\pi\pi)$	6	1	5
$(K\pi)(K^0_S\pi)$	3	1	2
$(K\pi\pi^0)(K^0_S\pi)$	8	5	3
$(K3\pi)(K^0_S\pi)$	6	1	5
$(K^0_S\pi\pi)(K^0_S\pi)$	0	0	0
Total	132	69	63

the Monte Carlo sample were also explored. We studied in detail the decays of charged $B \to D^{*0}D^*$ (which is currently not included in generic MC) and $B^0 \to D_s^{(*)}D^*$ (which is included in generic MC but for which the current published branching ratios have a relative error of over 25%[37]).

The reconstruction of $B^0 \to D^{*+}D^{*-}$ candidates was performed on a sample of 44000 generated $B^0 \to D_s^*D^*$ events, which corresponds to an effective luminosity of 20 fb^{-1}, and no events pass our selection criteria.

We repeated the same procedure on the channel $B \to D^{*0}D^*$, where we

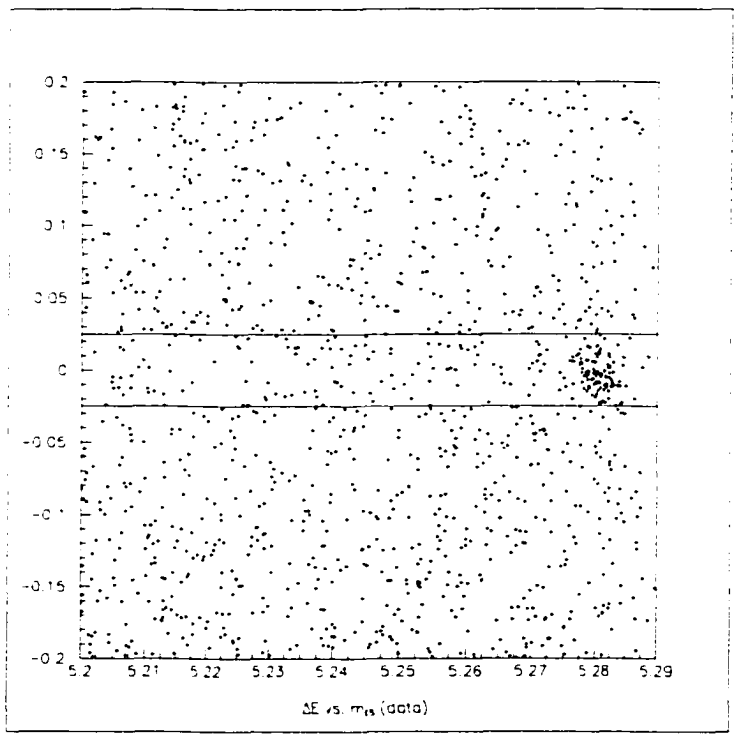


Figure B.1: Displayed is the ΔE vs. $m_{\rm ES}$ projection for $B^0 \to D^{*+}D^{*-}$. The lines show where the ΔE cut is applied for the $m_{\rm ES}$ projection.

expect events to peak both in m_{ES} and ΔE. We analyzed 34617 generated $D^{*0}D^*$ events, corresponding to an effective luminosity of 550 fb^{-1}, assuming $BR(B^+ \to D^{*0}D^*) \approx BR(B^0 \to D^{*+}D^{*-})$ and we found 3 events in the signal region.

Other possible sources of background that might peak in m_{ES} buth not in ΔE can be investigated by using data sidebands.

Figure B.9 shows the $m_{\rm ES}$ distribution in the ΔE sideband ($50 < |\Delta E| < 200 MeV$) fitted with an ARGUS function plus a Gaussian distribution, where all the parameters except normalizations were fixed at the fitted values ob-

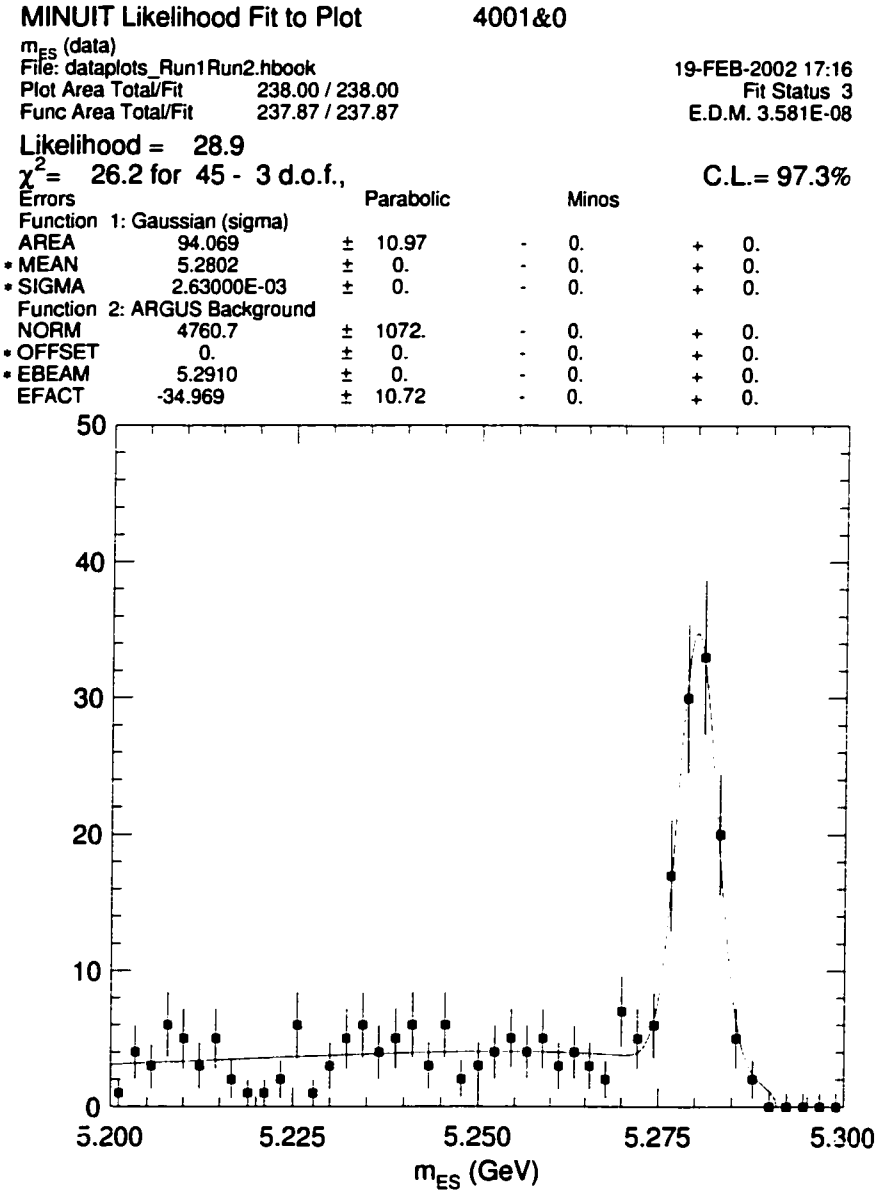


Figure B.2: Displayed is the $m_{\rm ES}$ projection for $B^0 \to D^{*+}D^{*-}$ after applying the signal region cut on ΔE . The line is a fit to an ARGUS+gaussian shape where the gaussian resolution is fixed to the value seen in Monte Carlo.

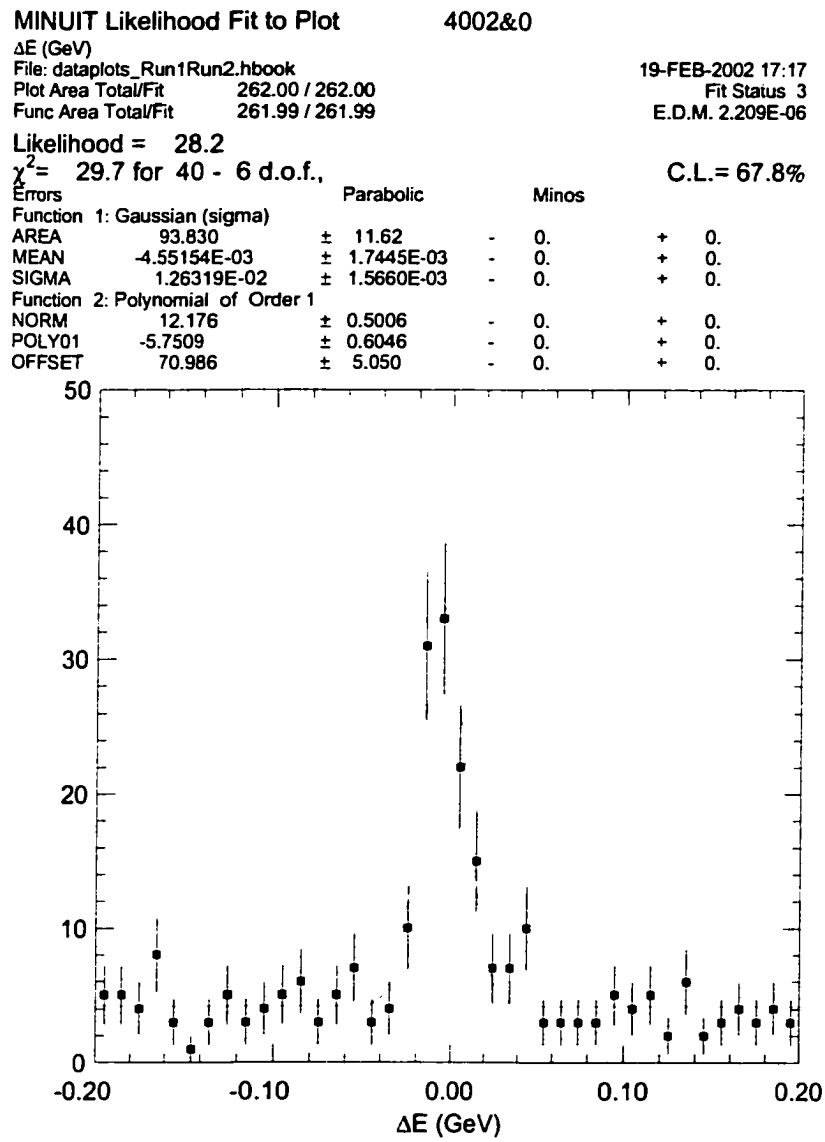


Figure B.3: Displayed is the ΔE projection for $B^0 \to D^{*+}D^{*-}$ after applying the signal region cut on $m_{\rm ES}$. The line is a fit to two gaussians+exponential shape. The second gaussian is fitted to the feeddown region where $B^0 \to D_S D^*$ is misreconstructed. While this fit is not used in the final CP analysis, a comparison of the yields extracted from the $m_{\rm ES}$ projection can be made from this figure.

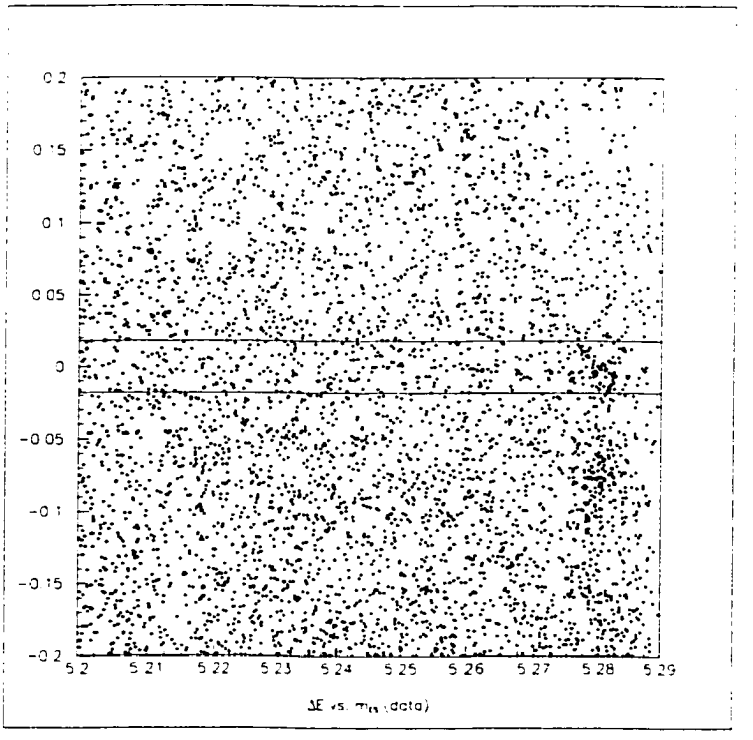


Figure B.4: Displayed is the ΔE vs. $m_{\rm ES}$ projection for $B^0 \to D^{*+}D^-$. The lines show where the ΔE cut is applied for the $m_{\rm ES}$ projection.

MINUIT Likelihood Fit to Plot 4001&0

m_{ES} (data)

File: dstard_dataplots_Run1Run2.hbook 22-FEB-2002 08:03

Plot Area Total/Fit 432.00 / 432.00 Fit Status 3

Func Area Total/Fit 431.69 / 431.69 E.D.M. 1.338E-06

Likelihood = 32.5

χ^2= 30.0 for 45 - 3 d.o.f., C.L.= 91.7%

Errors			Parabolic		Minos		
Function 1: Gaussian (sigma)							
AREA	64.965	±	11.04	-	0.	+	0.
* MEAN	5.2802	±	0.	-	0.	+	0.
* SIGMA	2.63000E-03	±	0.	-	0.	+	0.
Function 2: ARGUS Background							
NORM	10750.	±	1526.	-	0.	+	0.
* OFFSET	0.	±	0.	-	0.	+	0.
* EBEAM	5.2910	±	0.	-	0.	+	0.
EFACT	-28.158	±	6.635	-	0.	+	0.

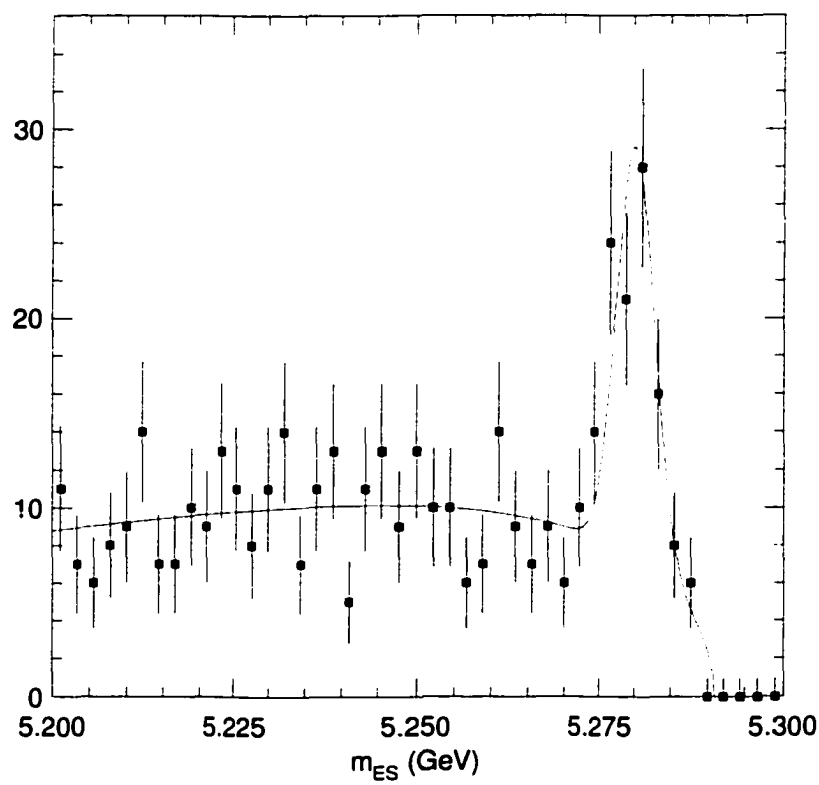


Figure B.5: Displayed is the m_{ES} projection for $B^0 \to D^{*+}D^-$ after applying the signal region cut on ΔE . The line is a fit to an ARGUS+gaussian shape where the gaussian resolution is fixed to the value seen in Monte Carlo.

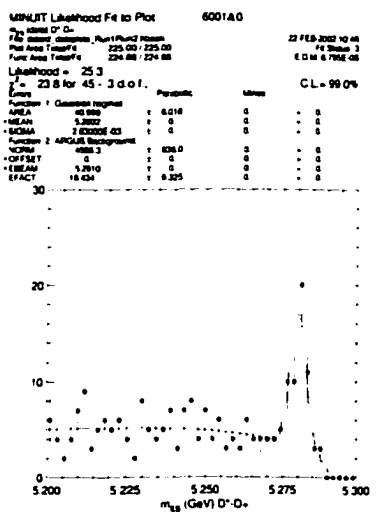


Figure B.6: Displayed is the $m_{\rm ES}$ projection for $B^0 \to D^{*-}D^+$ after applying the signal region cut on ΔE . This excludes all events where $B^0 \to D^{*+}D^-$. The line is a fit to an ARGUS+gaussian shape where the gaussian resolution is fixed to the value seen in Monte Carlo.

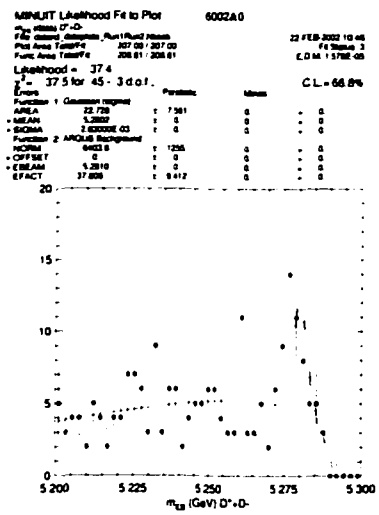


Figure B.7: Displayed is the $m_{\rm ES}$ projection for $B^0 \to D^{*+}D^-$ after applying the signal region cut on ΔE . This excludes all events where $B^0 \to D^{*-}D^+$. The line is a fit to an ARGUS+gaussian shape where the gaussian resolution is fixed to the value seen in Monte Carlo.

MINUIT Likelihood Fit to Plot 4002&0

ΔE (GeV)

File: dstard_dataplots_Run1Run2.hbook 27-FEB-2002 10:09

Plot Area Total/Fit 610.00 / 610.00 Fit Status 3

Func Area Total/Fit 610.02 / 610.02 E.D.M. 1.918E-06

Likelihood = 33.6

χ^2= 33.0 for 40 - 9 d.o.f., C.L.= 36.8%

Errors		Parabolic		Minos		
Function 3: Gaussian (sigma)						
AREA	60.946	± 11.73	-	0.	+	0.
MEAN	-2.27534E-03	± 2.1985E-03	-	0.	+	0.
SIGMA	1.02837E-02	± 1.9902E-03	-	0.	+	0.
Function 4: Gaussian (sigma)						
AREA	90.504	± 16.70	-	0.	+	0.
MEAN	-7.54041E-02	± 3.2364E-03	-	0.	+	0.
SIGMA	1.59558E-02	± 2.9532E-03	-	0.	+	0.
Function 5: Exponential						
NORM	1009.7	± 49.24	-	0.	+	0.
SLOPE	3.7775	± 0.4442	-	0.	+	0.
OFFSET	8.90687E-03	± 2.0839E-02	-	0.	+	0.

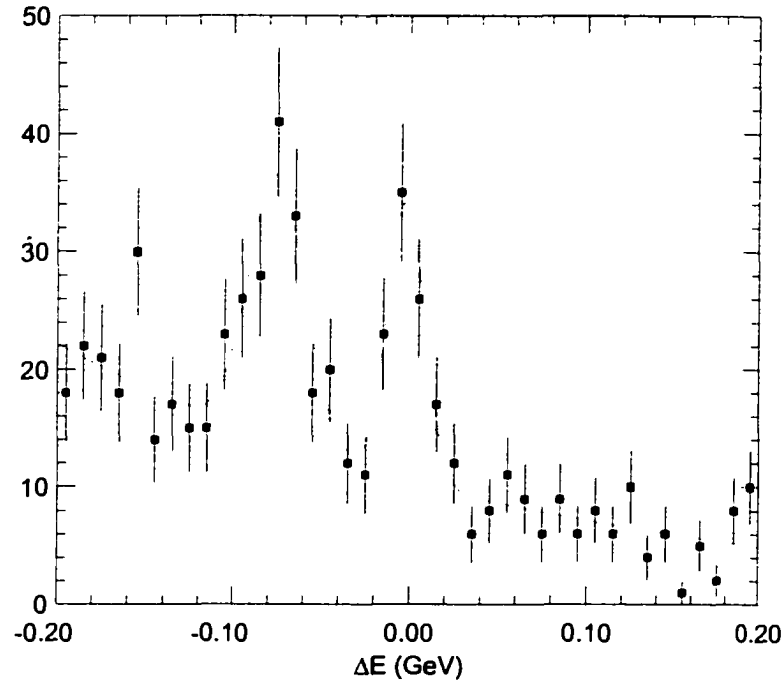


Figure B.8: Displayed is the ΔE projection for $B^0 \to D^{*+}D^-$ after applying the signal region cut on m_{ES}. The fit is to two gaussians + exponential. The region $\Delta E < 50\,\text{MeV}$ is fitted with a gaussian show where other B^0 decay modes can appear as misreconstructed D^*D decays. While this fit is not used in the final CP analysis, a comparison of the yields extracted from the m_{ES} projection can be made from this figure.

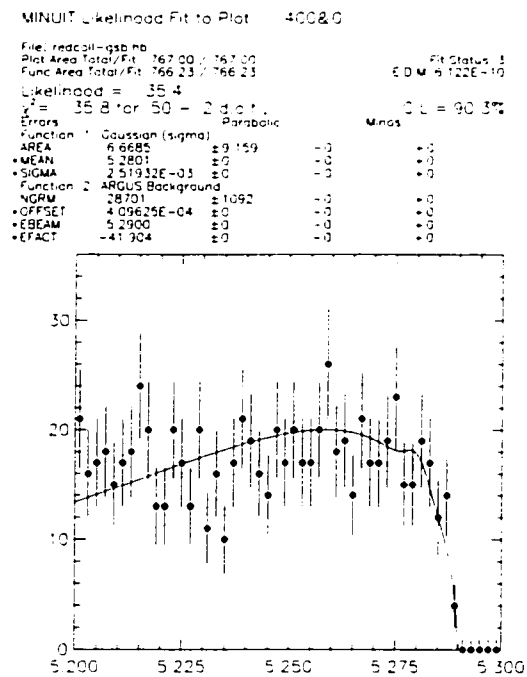


Figure B.9: m_{ES} distribution of $B^0 \to D^{*+}D^{*-}$ in the ΔE sideband fitted with an Argus function plus a Gaussian distribution to evaluate peaking background.

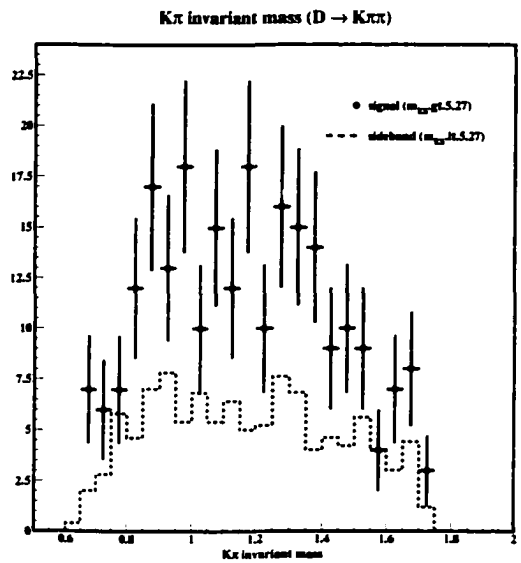


Figure B.10: Invariant mass of the $K\pi$ pairs in $B^0 \to D^{*\pm}D^{\mp}$ data events where the $D^{\mp}$ decays to $K\pi\pi$. Both pion choices per event are used for the plot. No unusual peaks are seen in signal.

tained from a fit to the $m_{\rm ES}$ distribution in the ΔE signal region. From this fit we estimate an upper limit of 1.1 ± 1.5 events of peaking background in the signal region from source having linear distribution on ΔE.

Peaking backgrounds for $B^0 \to D^{*+}D^-$

Examination of background events within the $B^0 \to D^{*\pm}D^{\mp}$ signal region in generic $B^0\overline{B}^0$ MC indicated that peaking background could occur from events of the type $B^0 \to D^*KX$, where the KX system has the same final state as a D submode and happens to be reconstructed within the D mass resolution range. Four such events were found in 8 fb^{-1} of generic $B^0\overline{B}^0$ MC,

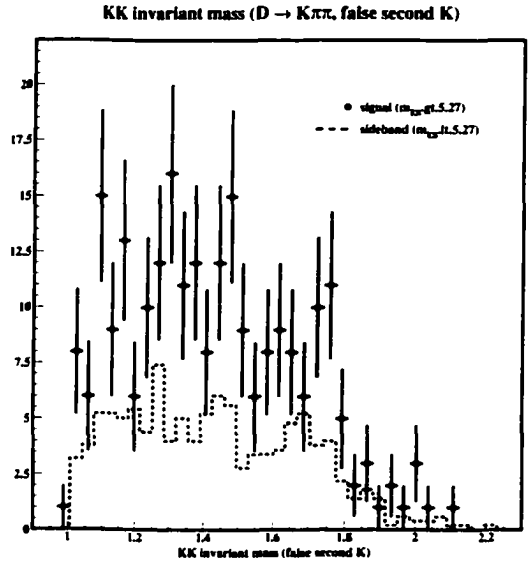


Figure B.11: Invariant mass of the $K\pi$ pairs in the K^+K^- hypothesis for $B^0 \longrightarrow D^{*\pm}D^{\mp}$ data events where the $D^{\mp}$ decays to $K\pi\pi$. Both pion choices per event are used for the plot. No unusual peaks are seen in signal.

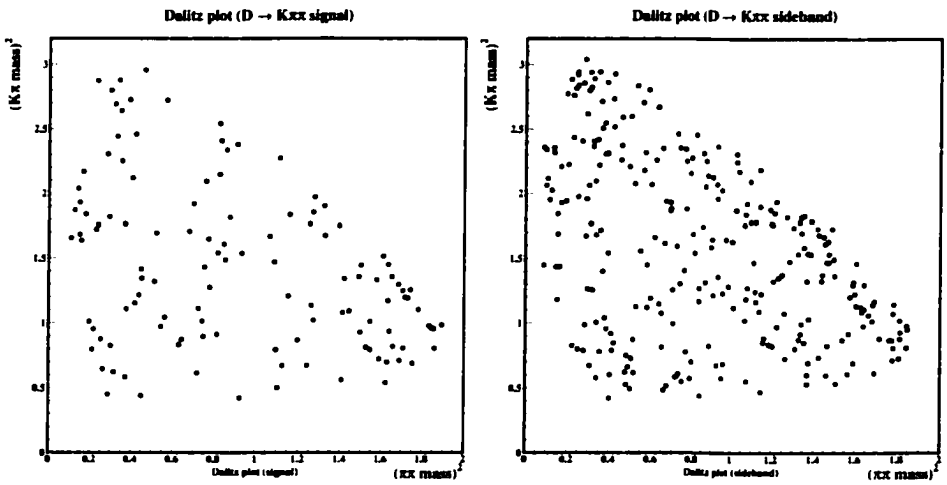


Figure B.12: Dalitz plots for signal region and $m_{\rm ES}$ sideband for $B^0 \to D^{*\pm}D^{\mp}$ data events where the $D^{\mp}$ decays to $K\pi\pi$.

indicating that it could potentially be a significant effect.

The possible peaking backgrounds in $B^0 \to D^{*\pm}D^{\mp}$ can be split into three types: 1) events with a false D^*, 2) events with a true D^* but false D reconstructed from a different final state (these can peak in $m_{\rm ES}$ but should not peak in ΔE), and 3) events with a true D^* and false D but which have exactly the same final state as the signal (these can peak in ΔE as well as $m_{\rm ES}$).

Peaking background from 1) has been checked by looking at the $\Delta M = M(D^*) - M(D^0)$ distribution for the signal box and the $m_{\rm ES}$ sideband, as well as looking at the $m_{\rm ES}$ distribution in windows around the nominal ΔM. Given the narrowness of acceptable ΔM, it is reasonable to predict that there is no significant peaking background with a fake D^*. This prediction is borne out in the data. Figure B.13 shows that there is no significant excess

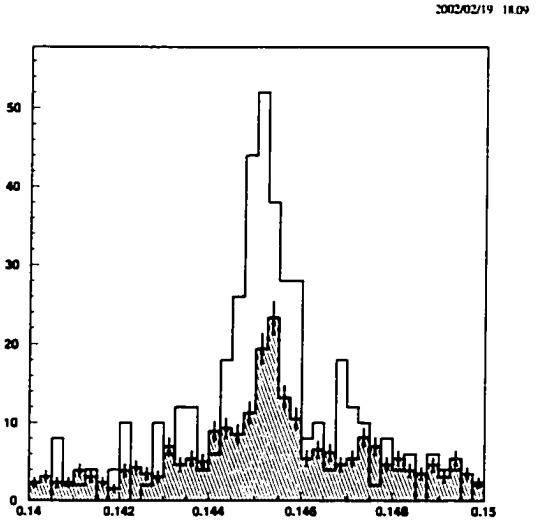


Figure B.13: ΔM for signal and m_{ES} sideband normalized to the ARGUS+gaussian fit. The lack of significant excess of ΔM sidebands within the signal is evidence for lack of peaking backgrounds with a false D^*.

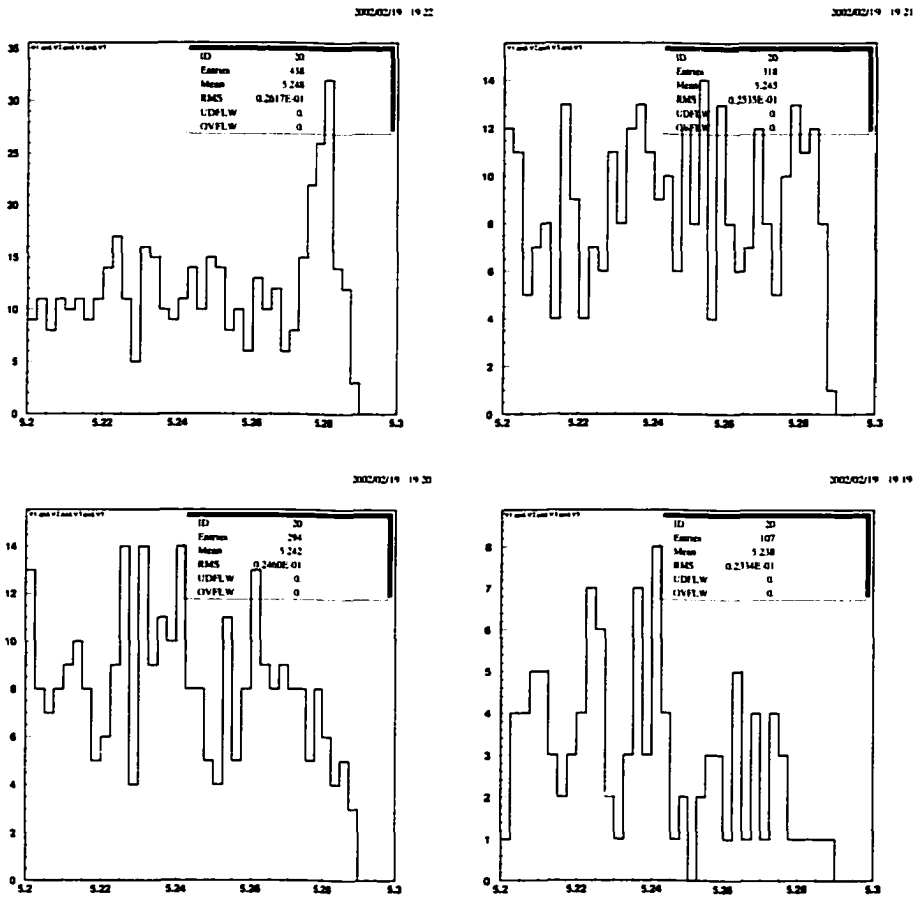


Figure B.14: m_{ES} distribution in data in windows of ΔM. The first plot is within 1 MeV, the second between 1 and 2.5, the third between 2.5 and 5, and the last beyond 2.5. No signal is seen in the outer bins, indicating lack of false D^* contribution to the signal.

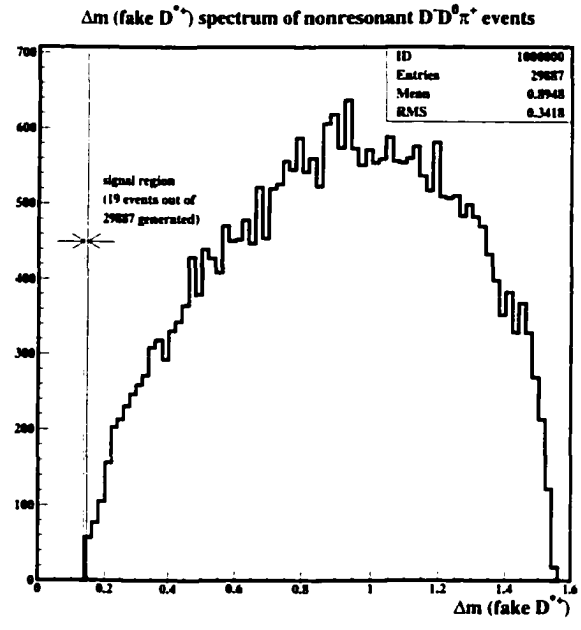


Figure B.15: Generator-level distribution of ΔM in $D^+D^0\pi^-$. The small size of the signal region effectively precludes any significant peaking background from this source.

of ΔM sidebands within the signal. To check this result slices in ΔM have been done and the Mes plots have been looked at for the following windows around the nominal 0.1454; these are shown in Figure B.14. We do not believe there should be any background contribution from this source, because even if it existed it would be accounted for in the non-ΔE peaking background described below (B1). The only fake D^* that would peak in ΔE is $D^+D^0\pi^-$, which, assuming a production rate similar to the signal, is innocuous: the ΔM mass window is narrow enough that nonresonant $D^0\pi$ have a very low efficiency. Generator level studies (see Figure B.15) show that the efficiency for the selection on $D^+D^0\pi^-$ is 0.06%.

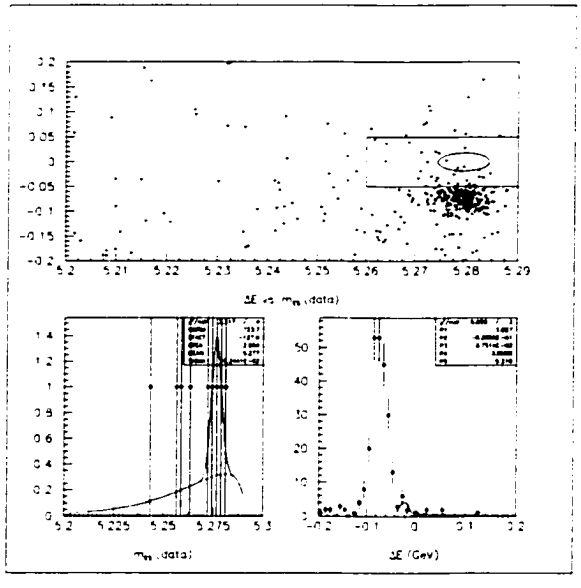


Figure B.16: m_{ES} and ΔE plots in 130 fb^{-1} of $D_s D^*$ SP4 signal MC.

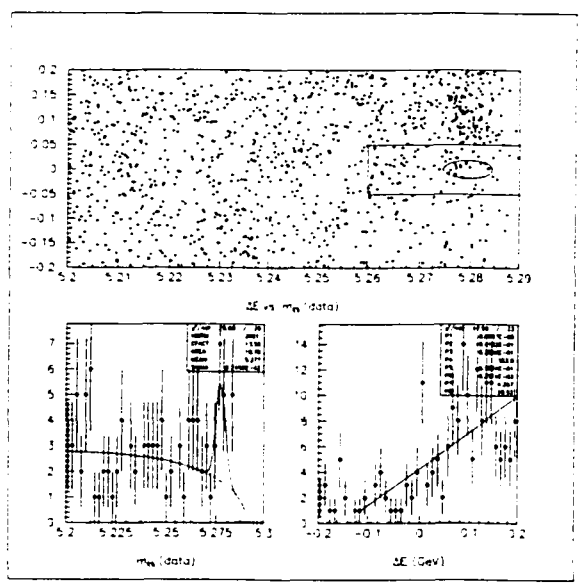


Figure B.17: m_{ES} and ΔE plots in 240 fb^{-1} of B^0 cocktail $(D^{(*)}\pi, \rho, a_1)$ SP4 signal MC.

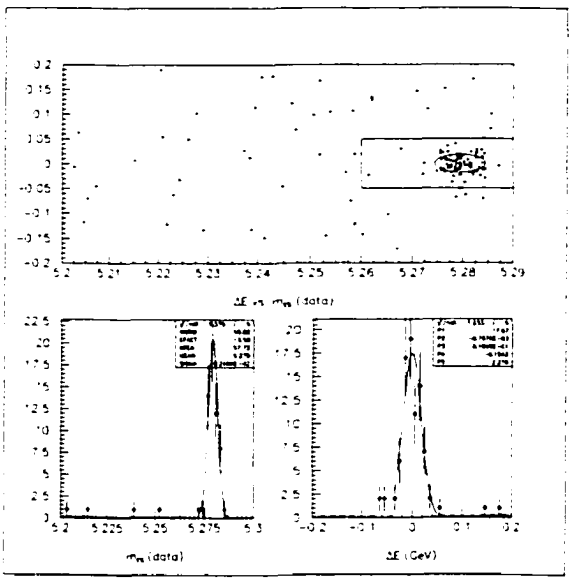


Figure B.18: m_{ES} and ΔE plots in 1.1 ab^{-1} of $D^* K_s \pi$ SP4 signal MC.

Peaking background from 2) may be checked via the yields as measured on m_{ES} and ΔE , that are 65.0 ± 11.0 and 60.1 ± 11.7 respectively. In the case that there were a large amount of peaking background that did not peak in ΔE , the ΔE fit would yield a significantly smaller number of events. That is not the case, so the amount is not large. We can improve on the quantitative measurement by checking individual sources:

- One of the candidates for such a source is $D_s D^*$, that appears as a peak at -70 MeV in the ΔE plot. This background has been checked on MC (see Figure B.16): from the ΔE and m_{ES} distributions on 130 fb^{-1} equivalent MC one infers that there is 1.0 ± 0.5 events expected from this source.

- $D^* a_1 (a_1 \rightarrow 3\pi)$ can be a concern. It has been checked running on B^0

cocktail MC ($D^{(*)}\pi, \rho, a_1$ with $D^* \to D^0\pi^+$ or $D^+\pi^0$). The $m_{\rm ES}$ plot on 240 fb^{-1} equivalent MC is in Figure B.17: the peaking background from this source has been estimated to be 2.5 ± 1.2.

- The $D^*3\pi\pi^0$ can also be a concern if the π^0 is lost and one of the pions is called a kaon. The kinematics has been checked at generator level. The distribution on 250,000 generated events (If one assumes a BF of 0.2% and an $\epsilon * \mathcal{B}(D^* \to X)$ 10% this is $1.2\ ab^{-1}$), yields 7 events in the signal region, consistent with the expectation that this background should be small. In order to test the possibility of a pion $\leftrightarrow$ kaon mismatch causing peaking background, we checked the Kπ and K"K" invariant mass (calling one of the two pions a kaon), see Figures B.10, B.11, and B.12, and no unusual peaks were seen.

Adding these three contributions (1.0 ± 0.5, 2.5 ± 1.2, and 0.5 ± 0.4), one then obtains the estimate of the peaking background from different final states of 4.0 ± 1.4.

Peaking background from 3), *i.e.* with a true D^* and the same final state as the signal, can potentially come from several sources:

- $D^{**}D^0$ ($D^0 \to K\pi$, $D^{**} \to D^*\pi$). This background should have BF similar to the one of the signal, however the $K\pi\pi$ system would have to coincidentally fall within the $D^{\mp}$ mass window, so it should not be a concern.

- $D^*K\pi\pi$ – the charge of the K must be opposite to the D^*, while in the signal the two charges are the same. Thus these modes cannot constitute a significant background.

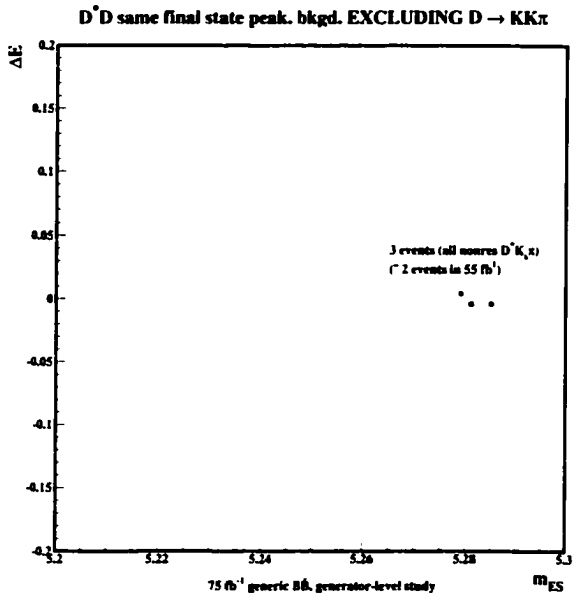


Figure B.19: Generator-level m_{ES} vs. ΔE plot of same-final-state peaking background in 75 fb^{-1} MC.

- $D^* K_s \pi$ - this mode is expected to have a branching fraction of about $5x10^{-4}$, thus some could feed in but not a very large amount. To quantitatively estimate the amount of $D^* K_s \pi$ in the signal region, a sample equivalent to 1.1 ab^{-1} (with the above branching fraction hypothesis) has been generated (see Figure B.18). Thus this mode would constitute a peaking background in 55 fb^{-1} of 2 ± 2 events (error mainly due to the lack of knowledge of the exact branching fraction of this mode, as it is unmeasured).

Two additional studies to estimate the peaking background from the above sources have been performed. The first is a generator-level study on generic $B^0\overline{B}^0$ to observe real D^* background with the same final state as

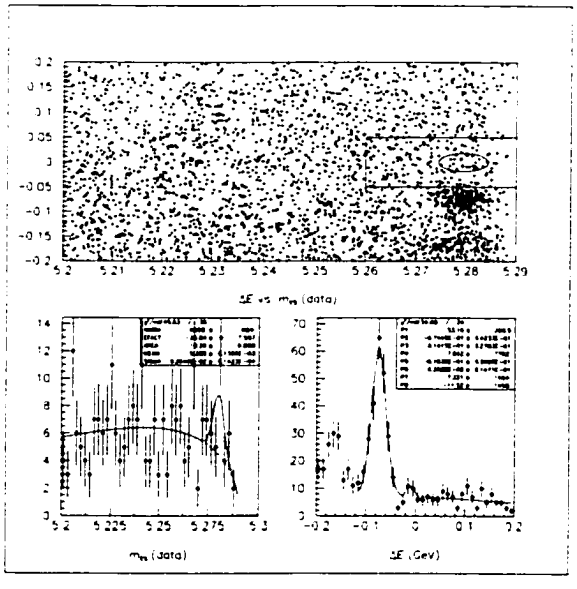


Figure B.20: m_{ES} and ΔE plot of peaking background in 150 fb^{-1} generic filtered MC.

the signal. The equivalent of 75 fb^{-1} yields just 2 events (as mentioned and expected above) in $D^* K_s \pi$, see Figure B.19. The second used filtered SP4 MC which is both generated and reconstructed. The equivalent of 150 fb^{-1} of MC reconstructed using the kinematic peaking backgrounds filter is seen in Figure B.20. No significant source other than $D^* K_s \pi$ (which is within expectation) is seen.

Thus we find a peaking background contribution of 6 ± 4 events to the $B^0 \to D^{*\pm} D^{\mp}$ signal peak. The error is a very conservative addition of the errors on different final state peaking background (as above, 4.0 ± 1.4) and same final state peaking background (as above, 2 ± 2).

B.1.3 Background Studies

The following section describes in more detail the background composition and the feed down region for $B^0 \to D^{*+} D^-$. Also included is a preliminary discussion of self cross feed within the sub-decay modes for both $B^0 \to D^{*+} D^{*-}$ and $B^0 \to D^{*+} D^-$.

Feed Down in $B^0 \to D^{*+} D^-$

As seen in Figures B.4 and B.8, a clear clustering of candidates is seen near $\Delta E = -75\,\text{MeV}$. This peak is primarily due to the Cabbibo favored decay $B^0 \to D_s D^*$. Here the highly favored decay $D_s \to K K \pi$ is misreconstructed as $D \to K \pi \pi$ because a kaon is misidentified as a pion. The incorrect mass assumption is the cause of the shift in ΔE .

A study of the decay $B^0 \to D_s D^*$ in Monte Carlo gives an indication of how much this decay is spilling over into the signal region for $B^0 \to D^{*+} D^-$.

Looking at 98k sp4 Monte Carlo events, when all selection criteria is applied excluding the ΔE cut, we see a total of 338 events peaking near $\Delta E = -75\,\mathrm{MeV}$, as expected. In the region $|\Delta E| < 18\,\mathrm{MeV}$, the cut applied to the m_{ES} projection, we see 3 events; this gives an approximate spillover into the signal region from this decay of 0.9%.

Lastly, in Figures B.4 and B.8, a second smaller clustering of events is seen near $\Delta E = -170\,\mathrm{MeV}$. This is attributed to the misreconstruction of $B^0 \to D^{*+}D^{*-}$; here, if a soft pion is not included, then a shift in ΔE would be seen in the reconstruction of $B^0 \to D^{*+}D^{-}$. Looking at 96k sp4 exclusive D^*D^* Monte Carlo events, when all selection criteria is applied we see a total of 4034 events peaking near $\Delta E = -170\,\mathrm{MeV}$, as expected. In the signal region for D^*D, we see 23 events; this gives an approximate spillover from $B^0 \to D^{*+}D^{*-}$ of 0.2%.

Feed Across

Because of the number of submodes that are used, the possibility of misreconstructing one mode as another (crossfeed) was investigated. For both $B^0 \to D^{*+}D^{*-}$ and $B^0 \to D^{*+}D^{-}$, this study was done previously on sp3 Monte Carlo, and with a slightly looser selection. Though the event selection is not the same as described above, it is shown here that the magnitude of any cross feed effect is negligable.

In $B^0 \to D^{*+}D^{*-}$, using individual signal Monte Carlo samples, 1000 events per submode, we show in Figure B.21 the number of events after passing the selection criteria reconstructed in the different submodes. The horizontal axis represents each submode signal Monte Carlo collection, and

the vertical axis shows the number of events reconstructed in each submode. The number of events reconstructed on the off-diagonal (indicating crossfeed between modes) was two out of a total of 21000 events generated; hence, the effect was considered negligible.

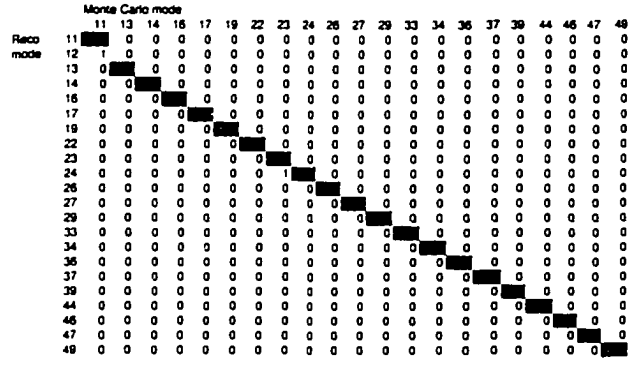


Figure B.21: The number of events out of one thousand generated $B^0 \rightarrow D^{*+}D^{*-}$ events per submode that pass selection criteria, as a function of the individual modes. The mode generated in signal Monte Carlo is represented on the horizontal axis and the mode reconstructed is shown on the vertical axis.

In $B^0 \rightarrow D^{*+}D^{-}$, again using individual signal MC samples, approx. 4000 events per submode, we show in Figure B.22 the number of events after passing the selection criteria reconstructed in the different submodes. For a few sub-decay mode the off-diagonal entries are > 5 events; however, the overall effect is still small when the percent rates are considered. The largest feed across occurs when the decay $(D^*)(D^+) \rightarrow (K_S^0\pi)(K\pi\pi)$ (mode 7,6) is misreconstructed as $(D^*)(D^+) \rightarrow (K\pi\pi)(K_S^0\pi)$ (mode 6,7) or vice versa. However, the overall effect of self cross feed is seen to be small.

Reco \ Monte Carlo mode	16	17	19	26	27	29	36	37	39	46	47	49	66	67	69	76
16		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0
26	1	0	0		0	0	0	0	0	0	0	0	0	0	1	0
27	0	1	0	0		0	0	0	0	0	0	0	0	0	0	0
29	0	0	2	0	0		0	0	0	0	0	0	0	0	0	0
36	0	0	0	0	0	0		0	1	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
39	0	0	0	0	0	1	0	1		0	0	1	0	0	0	0
46	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0
47	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0
49	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0
66	0	0	0	2	0	0	0	0	0	0	0	0		1	0	0
67	0	0	0	0	0	0	0	0	0	0	0	0	0		0	9
69	0	0	0	0	0	0	0	0	0	0	0	0	1	0		0
76	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	

Figure B.22: The number of events out of 4000 generated $B^0 \to D^{*+}D^-$ events per submode that pass selection criteria as a function of the individual modes. The mode generated in signal Monte Carlo is represented on the horizontal axis and the mode reconstructed is shown on the vertical axis.

Kaon ID in $D^+ \to K\pi\pi$

In $B^0 \to D^{*+}D^-$, the bachelor D^+ decays are potentially a major source of background. In order to understand the magnitude of this we study the effect of adding a tighter kaon selection to the kaon in $D^+ \to K\pi\pi$. In all the decay modes included, those containing $K\pi\pi$ contribute approximately 80% to the total signal. Thus, adding the tighter selection could have a significant effect.

Using the data sample, and adding the KTight selection to the bachelor D (only), the resulting yield from the $m_{\rm ES}$ fit is 51.3 ± 9.4, *purity* $= 56.9$, and the total $S^2/(S+B) = 29.2$. While the purity increases by a small ammount, the total $S^2/(S+B)$ decreases and indicates that the cut is not optimal. The signal decreases by approximately 21%.

To verify the effect, we looked at 11000 sp4 Monte Carlo events for the mode $B^0 \to D^* D \to (K\pi\pi)(K\pi\pi)$. With current kaon selection (notAPion), the approximate efficiency for this mode is 7.3%. Adding tight kaon ID to the bachelor D reduces the efficiency to approx. 6.0% – this corresponds to a change in efficiency of about 18%. Thus the drop in signal seen in data is believable and we conclude that adding tight kaon ID is not optimal for this analysis.

Pion Veto in $D^+ \to K\pi\pi$

Another possible means of lowering background in $B^0 \to D^{*+} D^-$ is to add a kaon veto selection to the pions in $D^+ \to K\pi\pi$. Given that our current selection for kaons is notAPion, requiring the pions in this mode to explicitly NOT have this selection may aid in reducing background.

In the data sample, adding the pion veto to the bachelor D (only) gives a resulting yield from the $m_{\rm ES}$ fit of 55.9 ± 9.7, *purity* = 58.9%, and the total $S^2/(S+B) = 32.9$. Again, the purity increases by a small ammount, but the overall $S^2/(S+B)$ decreases and indicates the cut is not optimal. The signal decreases by approximately 14%.

To validate the effect, we looked at 11000 sp4 Monte Carlo events for the mode $B^0 \to D^* D \to (K\pi\pi)(K\pi\pi)$. With current selection (no pion veto), the approximate efficiency for this mode is 7.3%. Adding the pion veto to the bachelor D reduces the efficiency to approx. 5.7% – this corresponds to a change in efficiency of about 22%. This would indicate that a pion veto in these modes is not only reducing background but also significantly reducing the signal – this is an undesirable effect and hence the tighter selection for

pions will not be included.

It is also interesting to note the change in the feeddown region in ΔE . As explained above the dominate cause of the feeddown in $B^0 \to D^{*+}D^-$ is the misreconstruction of $B^0 \to D_s D^*$ where a pion is misidentified as a kaon. With the pion veto this peak disappears as expected (see Figure B.23).

Finally, another study on pion veto was also done to see the effects of looser selection. Instead of requiring the pions in $D^+ \to K\pi\pi$ to be NOT notAPion, here we only require that they NOT have KTight kaon selection. This is a much looser selection and resulted in a signal yield of 64.5 ± 10.7, purity of 54.2% and $S^2/(S+B) = 35.0$. Clearly this is a slightly more optimial selection for this analysis; however, because of the lack of validation studies and time constraints, it will not be included in the current analysis.

B.2 Validation Analyses

B.2.1 Tagging and Vertexing in $D^{(*)+}D^{(*)-}$

Dedicated studies of tagging and vertexing in $D^{(*)+}D^{(*)-}$ are necessary due to the fact that one of the measurements of interest is the *difference* in measured $\sin 2\beta$ between $D^{(*)+}D^{(*)-}$ and charmonium. Any differences in tagging and vertexing could cause a systematic difference between those measurements, which would obfuscate investigation of Standard Model consistency.

Large differences between $D^{(*)+}D^{(*)-}$ and charmonium are not expected, due to the fact that vertexing resolution is dominated by the resolution on the tag vertex, not the reconstructed vertex, and there is no indication that major differences in tag side reconstruction occur between these two types

of decays. However, even though effects are expected to be quite small in comparison with our statistical errors, it is important to check that this is the case and to assign any systematics or corrections if differences exist.

One method of checking for differences would be to obtain tagging dilutions and resolution function parameters using $D_s^{(*)}D^{(*)}$ events and compare those results with the standard ($D^{(*)}\{\pi, \rho, a_1\}$) mixing sample in data. This would provide a comparison of differences in tagging and vertexing from double charm decays vs. charmonium and the mixing sample. However, that analysis is not feasible on the timescale of this first measurement and the present checks are based on simulated data.

$D^{*+}D^{*-}$ comparison of tagging dilutions between CP and B_{flav} MC

We have performed tests to measure possible bias introduced while fitting for sine and cosine coefficients in the $D^{*+}D^{*-}$ CP sample, due to the use of tagging dilutions from the fit of B_{flav} sample.

By using a large statistics sample of $D^{*+}D^{*-}$ and B_{flav} Montecarlo we evaluated the average dilution $\mathcal{D}$ and difference $\Delta\mathcal{D}$ between B^0 and $\bar{B}^0$ for the four tagging categories, at truth level for both Montecarlo sample and by fitting B_{flav}. The values of average dilutions $\mathcal{D}$ and difference $\Delta\mathcal{D}$ together with the fitted value of sine and cosine coefficients, obtained with a fit where dilutions were fixed at the truth or fitted values, are shown in table B.4.

$D^{*\pm}D^{\mp}$ comparison of tagging dilutions between CP and B_{flav} MC

A similar study of tagging dilutions to that done above for $D^{*+}D^{*-}$ is done for $D^{*\pm}D^{*\mp}$. The equivalent table for $D^{*\pm}D^{*\mp}$ fitted $\mathcal{D}$ and $\Delta\mathcal{D}$ is below:

Parameter	$D^{*+}D^{*-}$ MC Truth	B_{flav} MC Truth	B_{flav} MC fitted
$\mathcal{D}$(Lepton)	0.882 ± 0.008	0.869 ± 0.006	0.868 ± 0.010
$\mathcal{D}$(Kaon)	0.673 ± 0.008	0.658 ± 0.006	0.665 ± 0.008
$\mathcal{D}$(NT1)	0.608 ± 0.018	0.574 ± 0.013	0.603 ± 0.018
$\mathcal{D}$(NT2)	0.305 ± 0.016	0.298 ± 0.011	0.289 ± 0.015
$\Delta\mathcal{D}$(Lepton)	0.056 ± 0.017	0.014 ± 0.013	0.0077 ± 0.0170
$\Delta\mathcal{D}$(Kaon)	0.009 ± 0.016	0.027 ± 0.011	0.0382 ± 0.0126
$\Delta\mathcal{D}$(NT1)	-0.046 ± 0.036	-0.091 ± 0.026	-0.0784 ± 0.0282
$\Delta\mathcal{D}$(NT2)	0.064 ± 0.032	0.058 ± 0.022	0.0675 ± 0.0229
$sinCoef$	-0.231 ± 0.022	-0.236 ± 0.023	-0.235 ± 0.022
$cosCoef$	-0.002 ± 0.015	-0.009 ± 0.015	-0.001 ± 0.0147

Table B.4: Value of mistag dilution parameters in $D^{*+}D^{*-}$ and B_{flav} MC sample and corresponding value of CP fitted parameters.

More detailed tables of tagging efficiency and mistag for $B^0 \to D^{*\pm}D^{\mp}$ signal MC are summarized in the tables below:

```
B0 events only:

 Results from tagperf tag=MCB0 fmt=text:

 Category        TOTAL    CAT   WRONG EFF(%) SEFF(%)  W(%)    SW(%)  Q(%)   SQ(%)
 ---------------------------------------------------------------------------------
 NOTElKaon        7117    137     1.0    1.9 +- 0.2   0.7 +- 0.7   1.9 +- 0.2
 NOTMuKaon        7117    122     1.0    1.7 +- 0.2   0.8 +- 0.8   1.7 +- 0.2
 NOTElectron      7117    335    24.0    4.7 +- 0.3   7.2 +- 1.4   3.5 +- 0.3
 NOTMuon          7117    303    26.0    4.3 +- 0.2   8.6 +- 1.6   2.9 +- 0.3
 NOTKaon          7117   2497   380.0   35.1 +- 0.6  15.2 +- 0.7  17.0 +- 0.8
 NT1              7117    516   111.0    7.3 +- 0.3  21.5 +- 1.8   2.4 +- 0.3
```

Parameter	$D^{*\pm}D^{\mp}$ MC Truth	B_{flav} MC Truth	B_{flav} MC fitted
$\mathcal{D}$(Lepton)	0.882 ± 0.011	0.869 ± 0.006	0.868 ± 0.0100
$\mathcal{D}$(Kaon)	0.679 ± 0.010	0.658 ± 0.006	0.665 ± 0.0079
$\mathcal{D}$(NT1)	0.611 ± 0.024	0.574 ± 0.013	0.603 ± 0.0175
$\mathcal{D}$(NT2)	0.312 ± 0.021	0.298 ± 0.011	0.289 ± 0.0147
$\Delta\mathcal{D}$(Lepton)	0.002 ± 0.022	0.014 ± 0.013	0.008 ± 0.0170
$\Delta\mathcal{D}$(Kaon)	0.034 ± 0.020	0.027 ± 0.011	0.038 ± 0.0126
$\Delta\mathcal{D}$(NT1)	-0.079 ± 0.047	-0.091 ± 0.026	-0.078 ± 0.0282
$\Delta\mathcal{D}$(NT2)	0.114 ± 0.042	0.058 ± 0.022	0.068 ± 0.0229
$C_{D^{*+}D^{-}}$	0.042 ± 0.027	0.035 ± 0.028	0.042 ± 0.028
$S_{D^{*+}D^{-}}$	0.694 ± 0.039	0.773 ± 0.040	0.765 ± 0.040
$C_{D^{*-}D^{+}}$	-0.048 ± 0.027	-0.057 ± 0.028	-0.049 ± 0.028
$S_{D^{*-}D^{+}}$	0.694 ± 0.039	0.714 ± 0.040	0.707 ± 0.040
$abs(\lambda)$	1.046 ± 0.020	1.047 ± 0.020	1.047 ± 0.020
$\sin(2\beta + \delta)$	0.750 ± 0.039	0.774 ± 0.040	0.765 ± 0.040
$\sin(2\beta - \delta)$	0.695 ± 0.039	0.715 ± 0.040	0.708 ± 0.040

Table B.5: Value of mistag dilution parameters in $D^{*\pm}D^{\mp}$ and B_{flav} MC sample and corresponding value of CP fitted parameters.

```
NT2            7117    995   312.0    14.0 +-  0.4  31.4 +-  1.5    1.9 +-  0.3

Total Q = 31.2% +/- 1.0%   (eff = 68.9% +/- 1.0%)

B0bar events only:

Results from tagperf tag=MCB0bar fmt=text:
```

Category	TOTAL	CAT	WRONG	EFF(%)	SEFF(%)	W(%)	SW(%)	Q(%)	SQ(%)
NOTElKaon	7059	132	1.0	1.9 +-	0.2	0.8 +-	0.8	1.8 +-	0.2
NOTMuKaon	7059	92	1.0	1.3 +-	0.1	1.1 +-	1.1	1.2 +-	0.1
NOTElectron	7059	339	26.0	4.8 +-	0.3	7.7 +-	1.4	3.4 +-	0.3
NOTMuon	7059	304	24.0	4.3 +-	0.2	7.9 +-	1.5	3.1 +-	0.3
NOTKaon	7059	2468	417.0	35.0 +-	0.6	16.9 +-	0.8	15.3 +-	0.7
NT1	7059	569	100.0	8.1 +-	0.3	17.6 +-	1.6	3.4 +-	0.4
NT2	7059	986	369.0	14.0 +-	0.4	37.4 +-	1.5	0.9 +-	0.2

Total Q = 29.2% +/- 1.0% (eff = 69.3% +/- 1.0%)

B0 and B0bar together:

Results from tagperf tag=MC fmt=text:

Category	TOTAL	CAT	WRONG	EFF(%)	SEFF(%)	W(%)	SW(%)	Q(%)	SQ(%)
NOTElKaon	14176	269	2.0	1.9 +-	0.1	0.7 +-	0.5	1.8 +-	0.1
NOTMuKaon	14176	214	2.0	1.5 +-	0.1	0.9 +-	0.7	1.5 +-	0.1
NOTElectron	14176	674	50.0	4.8 +-	0.2	7.4 +-	1.0	3.4 +-	0.2
NOTMuon	14176	607	50.0	4.3 +-	0.2	8.2 +-	1.1	3.0 +-	0.2
NOTKaon	14176	4965	797.0	35.0 +-	0.4	16.1 +-	0.5	16.1 +-	0.5
NT1	14176	1085	211.0	7.7 +-	0.2	19.4 +-	1.2	2.9 +-	0.2
NT2	14176	1981	681.0	14.0 +-	0.3	34.4 +-	1.1	1.4 +-	0.2

Total Q = 30.1% +/- 0.7% (eff = 69.1% +/- 0.7%)

$D^{*+}D^{*-}$ vertex resolution

To check the vertex resolution of the reconstructed B in data we look at the projections in x and y directions of the distribution of B_{reco} vertex and

the Beam Spot positions. Figure B.24 shows a comparison between data and MC, where a realistic svt misalignment was simulated. The following figures B.25, B.26 show the distribution of the same quantity for the two x and y projections respectively in data and MC. Due to the small dimension of the beam spot in y (7μm) the value of the standard deviation of the Gaussian fit to the y distribution can be interpreted as a rough estimation of the $B \to D^{*+}D^{*-}$ vertex resolution in the transverse plane.

$D^{*+}D^{*-}$ comparison of resolution function between CP and B_{flav} MC

A comparison of signal resolution function parameters between $D^{*+}D^{*-}$ signal MC and B_{flav} MC sample is summarized in table B.6.

Figure B.27 (left) shows δt residual in MC signal events (points). The two curves superimposed are obtained by using the fitted parameters on the signal sample (dashed curve) and on B_{flav} sample (plain curve). Figure B.27 (right) shows the pull distribution of δt residual difference between the two MC samples.

$D^{*\pm}D^{\mp}$ comparison of resolution function between CP and B_{flav} MC

A comparison of signal resolution function parameters between $D^{*\pm}D^{\mp}$ signal MC and the B_{flav} MC sample is summarized in table B.7.

Figure B.28 (left) shows the δt residual in $D^{*\pm}D^{\mp}$ signal events (points with errors) and in the B_{flav} MC sample (histogram). Figure B.28 (right) shows the pull distribution of δt residual difference between the two MC samples.

Parameter	$D^{*+}D^{*-}$ MC Truth	$D^{*+}D^{*-}$ MC fitted	B_{flav} MC Truth	B_{flav} MC fitted
Scale (core)	1.266 ± 0.010	1.178 ± 0.069	1.230 ± 0.007	1.192 ± 0.038
Scale (tail)	3.0 (fixed)	3.0 (fixed)	3.0 (fixed)	3.0 (fixed)
$\delta(\Delta t)$ lepton (c)	-0.149 ± 0.024	-0.066 ± 0.070	-0.126 ± 0.018	-0.084 ± 0.045
$\delta(\Delta t)$ kaon (c)	-0.299 ± 0.015	-0.185 ± 0.043	-0.292 ± 0.011	-0.294 ± 0.029
$\delta(\Delta t)$ NT1 (c)	-0.225 ± 0.032	-0.159 ± 0.086	-0.171 ± 0.022	-0.191 ± 0.058
$\delta(\Delta t)$ NT2 (c)	-0.241 ± 0.023	-0.120 ± 0.058	-0.237 ± 0.016	-0.257 ± 0.041
$\delta(\Delta t)$ (tail)	-1.403 ± 0.129	-1.501 ± 0.505	-1.131 ± 0.077	-0.704 ± 0.257
f(tail)	0.075 ± 0.006	0.117 ± 0.035	0.087 ± 0.004	0.090 ± 0.016
f(outlier)	0.003 ± 0.001	0.002 ± 0.001	0.005 ± 0.001	0.003 ± 0.001

Table B.6: Signal parameters from exponential $\otimes$ resolution-function fits for $D^{*+}D^{*-}$ and B_{flav} MC samples. The scale factor for the tail component was fixed to 3.0 in these fits.

Parameter	$D^{*\pm}D^{\mp}$ MC Truth	$D^{*\pm}D^{\mp}$ MC fitted	$B_{\rm flav}$ MC Truth	$B_{\rm flav}$ MC fitted
Scale (core)	1.2848 ± 0.0138	1.2039 ± 0.0700	1.2304 ± 0.0070	1.1921 ± 0.0378
Scale (tail)	3.0 (fixed)	3.0 (fixed)	3.0 (fixed)	3.0 (fixed)
$\delta(\Delta t)$ lepton (c)	-0.1295 ± 0.0321	-0.0801 ± 0.0703	-0.1258 ± 0.0175	-0.0841 ± 0.0452
$\delta(\Delta t)$ kaon (c)	-0.2827 ± 0.0199	-0.1912 ± 0.0435	-0.2920 ± 0.0105	-0.2942 ± 0.0294
$\delta(\Delta t)$ NT1 (c)	-0.1649 ± 0.0407	-0.1775 ± 0.0865	-0.1705 ± 0.0218	-0.1909 ± 0.0584
$\delta(\Delta t)$ NT2 (c)	-0.2398 ± 0.0307	-0.1188 ± 0.0577	-0.2371 ± 0.016	-0.2565 ± 0.0405
$\delta(\Delta t)$ (tail)	-1.5155 ± 0.2155	-1.5663 ± 0.5700	-1.1311 ± 0.0770	-0.7044 ± 0.2536
f(tail)	0.0641 ± 0.0079	0.1098 ± 0.0367	0.0868 ± 0.0040	0.0898 ± 0.0164
f(outlier)	0.0056 ± 0.0010	0.0020 ± 0.0013	0.0048 ± 0.0005	0.0033 ± 0.0009

Table B.7: Signal parameters from exponential $\otimes$ resolution-function fits for $D^{*\pm}D^{\mp}$ and B_{flav} MC samples. The scale factor for the tail component was fixed to 3.0 in these fits.

A comparison of Δz vertexing resolution between $B^0 \to D^{*\pm}D^{\mp}$ events and $J/\psi K_s$ events can be seen in Figure B.30. As can be seen, any differences are quite small.

B.2.2 MC studies: $B^0 \to D^{*+}D^{*-}$

Toy MC studies of fitting properties

The performance of the fit to sine and cosine coefficients was checked using a toy Monte Carlo. Experiments were generated with 1000 times the statistic and the same properties of data as far as tagging dilutions, resolution functions and background fractions are concerned. At generation level several sets of experiments with different values of $|\lambda|$, $Im\lambda$ were created. A fit was then performed floating the sine and cosine coefficients.

Figure B.31 shows the linearity of the fit of the sine coefficient while $Im\lambda$ was varied at generation level in the range $[-1., 1.]$ and $|\lambda|$ were fixed at 1. Figure B.32 shows the cosine coefficient linearity plots obtained fitting the distributions generated with $Im\lambda$=0.59, and $|\lambda|$ set in order to give the values between ± 0.45 at step of 0.15 for the cosine coefficient.

All those results show the good linearity of the fit of the sine and cosine coefficients. To check the convergence properties at low statistics we repeated the fit for one hundred experiments of 134 events each. In data and in this toy study we have 59 tagged signal events and 75 tagged background events distributed following the Argus distribution in the $m_{\rm ES}$ range from 5.2 to 5.29. None of the fits failed to converge.

Full MC studies

We used a large sample of signal $D^{*+}D^{*-}$ MC to study the distributions of mean fitted values and errors of CP parameters for several experiments with the same number of events as the data sample (90 untagged events each). The input values of the sine coefficient in the Monte Carlo is -0.231 corresponding to $Im\lambda = -0.7$ and flat angular distribution giving $K = 0.33$. The input cosine coefficient is zero corresponding to $|\lambda| = 1$.

The fit was performed fixing the tagging dilutions and resolution function paramerets at the fitted value of Montecarlo B_{flav} sample and floating the sine and cosine coefficients. The results are shown in figure B.33.

A fit was also done using the whole MC sample (26853 events), again fixing the tagging dilutions and resolution function parameters at the fitted value of the B_{flav} MC sample and floating the sine and cosine coefficients. The result obtained is: sinCoef $=-0.232 \pm 0.022$, cosCoef$=-0.004 \pm 0.015$.

The large MC sample we used was generated as signal $D^{*+}D^{*-}$, decaying in the two chains: $D^* \to D^0\pi$, $D^0 \to K\pi$ for both D^* mesons or one D^* meson decaying in $D^0\pi$, $D^0 \to K\pi$ and the other $D^* \to D^{\pm}\pi 0$, $D^{\pm} \to K\pi\pi$. In the future this sample will also be used to study the acceptance corrections for the angular analysis.

B.2.3 MC studies: $B^0 \to D^{*\pm}D^{\mp}$

Toy MC studies

A CP fit for $B^0 \to D^{*\pm}D^{\mp}$ is implemented in the tFit package, which performs an unbinned maximum likelihood fit according to the likelihoods spec-

ified in section 4.2.2. In addition, a toy Monte Carlo event generator for $B^0 \rightarrow D^{*\pm}D^{\mp}$ events has been implemented within the same package, providing a fast and simple way of generating events for testing the fit behavior. All known resolution and dilution effects, as well as adjustable background levels, are incorporated within the toy MC generator, allowing for an accurate statistical representation of the data.

The toy Monte Carlo is used to check for fit biases and to predict the range of statistical error and log likelihood expected from the data. Samples of events with exactly the same statistics as the data, *i.e.* same number of events in every flavor tag category and flavor tag value as the data, with the same dilutions and resolution as the data, are generated.

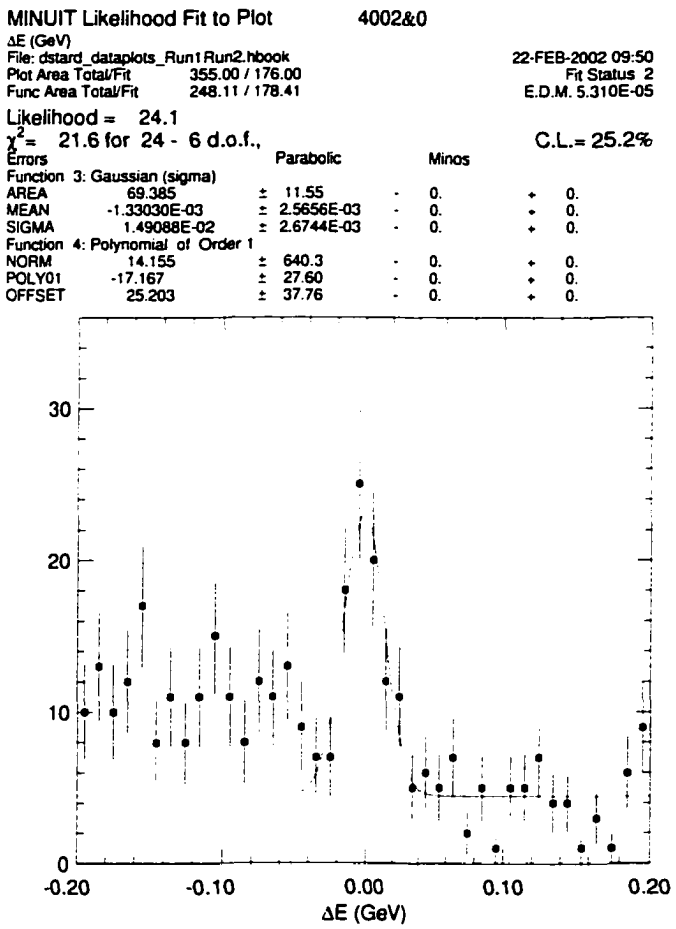


Figure B.23: Displayed is the ΔE projection for $B^0 \to D^{*+}D^-$ after applying the signal region cut on $m_{\rm ES}$ and the additional requirement that the pions in the bachelor D decay veto the kaon selection. The fit is a line+gaussian and is truncated to avoid the region $\Delta E < 50\,\text{MeV}$ where other B^0 decay modes can appear as misreconstructed D^*D decays. The dotted line shows the extention of the fit into the excluded region.

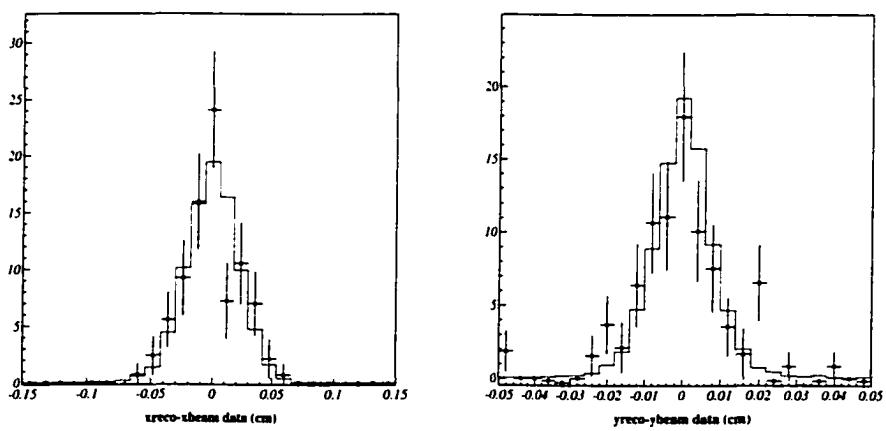


Figure B.24: Position difference distributions of the B_{reco} vertex and the Beam Spot in the x (left) and y (right) projection. The dots shows the distribution on data, the red histogram superimposed correspond to the distribution of the svt misaligned MC.

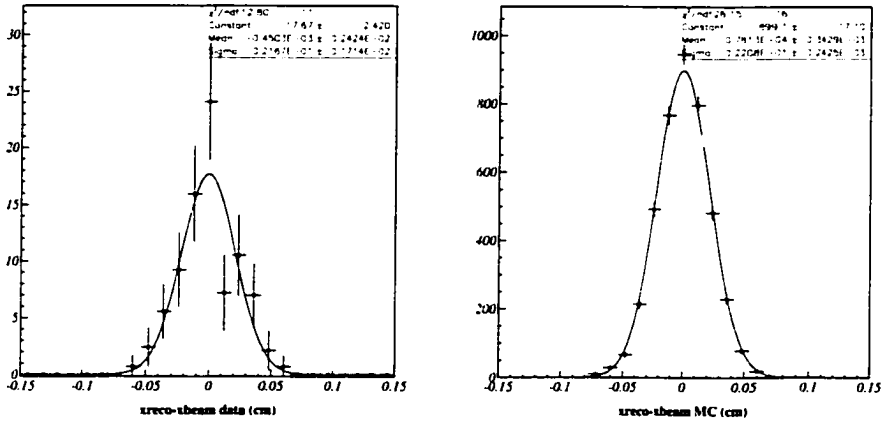


Figure B.25: Position difference distributions of the B_{reco} vertex and the Beam Spot in the x projection in data and MC wit SVT simulated misalignment.

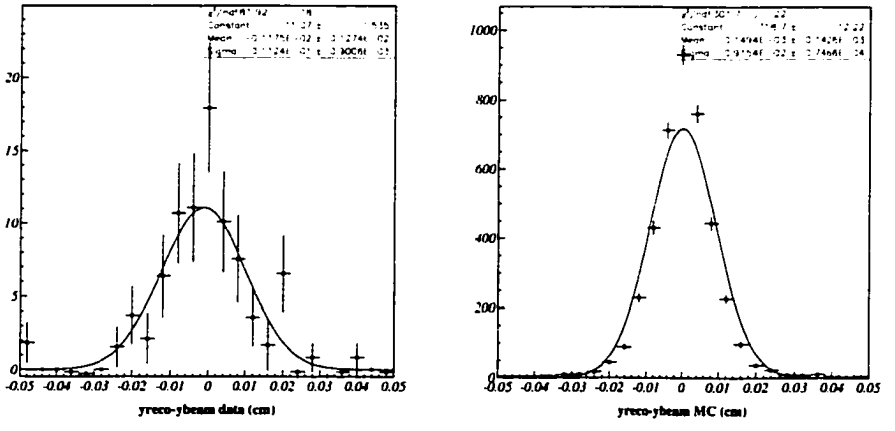


Figure B.26: Position difference distributions of the B_{reco} vertex and the Beam Spot in the y projection in data and MC wit SVT simulated misalignment.

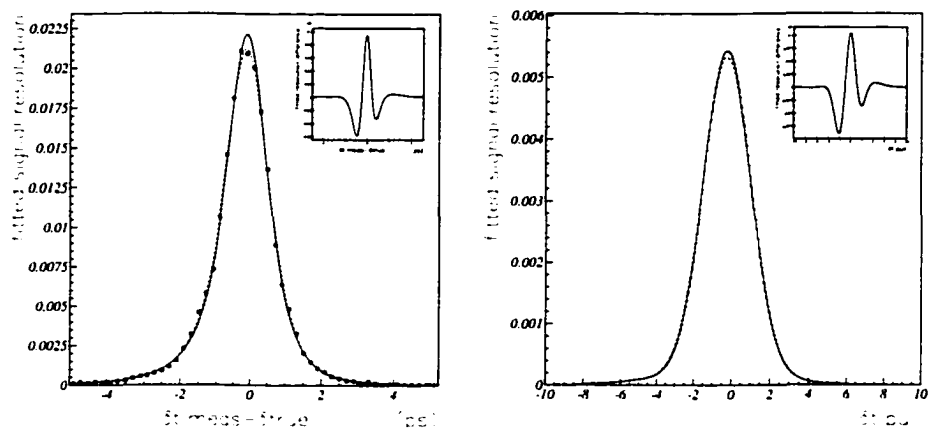


Figure B.27: Comparison of the fitted δt residual (left) and δt pull (right) between the B_{flav} and $D^{*+}D^{*-}$ MC samples. The tail scale factor for the signal component was fixed to 3.0 for these fits.

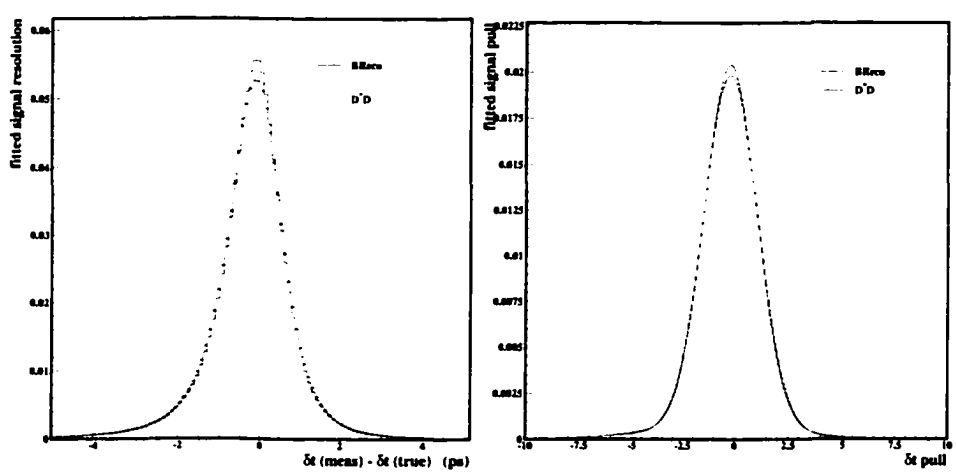


Figure B.28: Comparison of the fitted δt residual (left) and δt pull (right) between the B_{flav} and $D^{*\pm}D^{\mp}$ MC samples. The tail scale factor for the signal component was fixed to 3.0 for these fits.

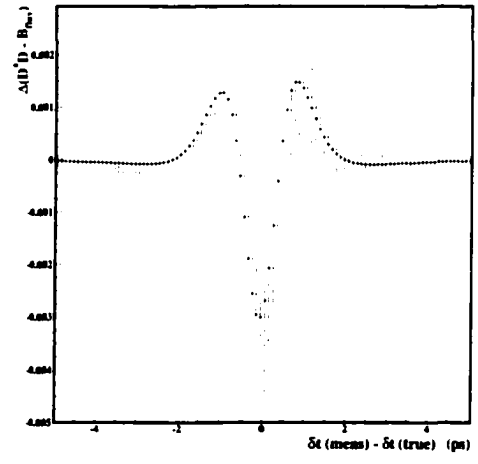


Figure B.29: The fitted δt residual in $D^{*\pm}D^{\mp}$ MC minus that from the B_{flav} sample (equivalent to the difference of the two histograms in figure B.28 (left)). As can be seen, the difference is a decrease in the core and corresponding increase in the tail. The differences are small, as can be seen in figure B.28 above.

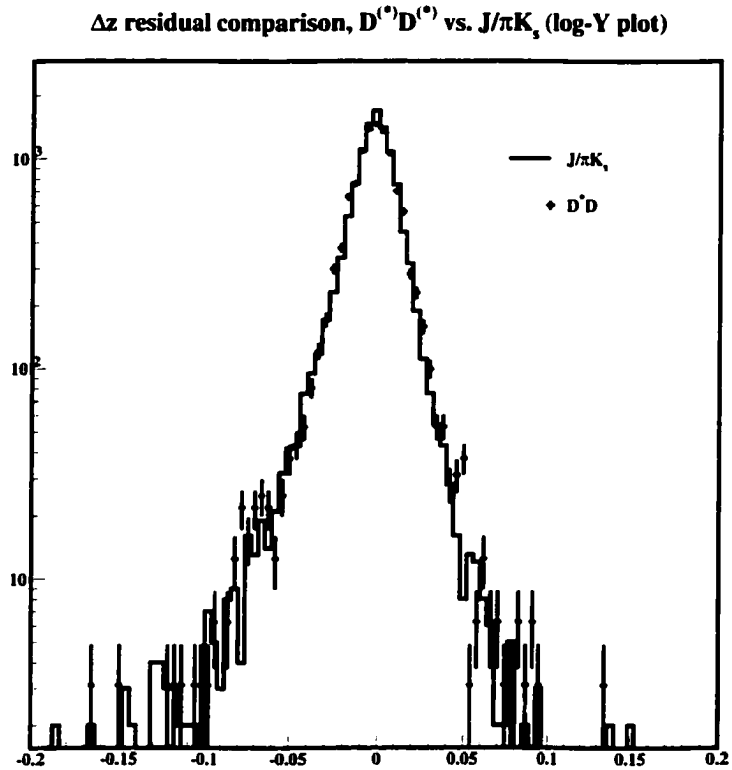


Figure B.30: Δz residuals ($= \Delta z_{rec} - \Delta z_{true}$) for $D^{*\pm}D^{\mp}$ and $J/\psi K_s$ events. As may be seen, differences in vertexing resolution are small. (NOTE: despite the title, these are just $D^{*\pm}D^{\mp}$ and not general $D^{(*)}\bar{D}^{(*)}$ events. This plot will need to be updated with SP4 results (this is SP3) and separate distribitions for $D^{*+}D^{*-}$ and $D^{*\pm}D^{\mp}$ events.)

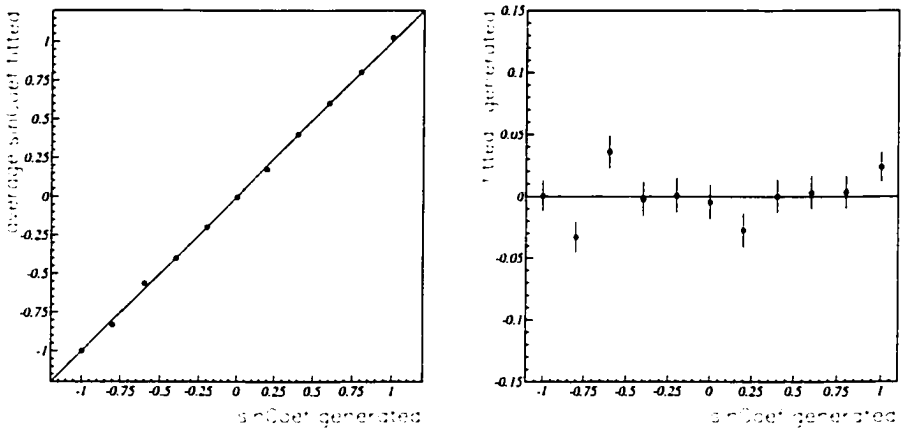


Figure B.31: Fitted sine coefficient parameter versus generated one (left) and difference between Fitted and generated value versus generated (right) while changing $Im\lambda$ value at generation time.

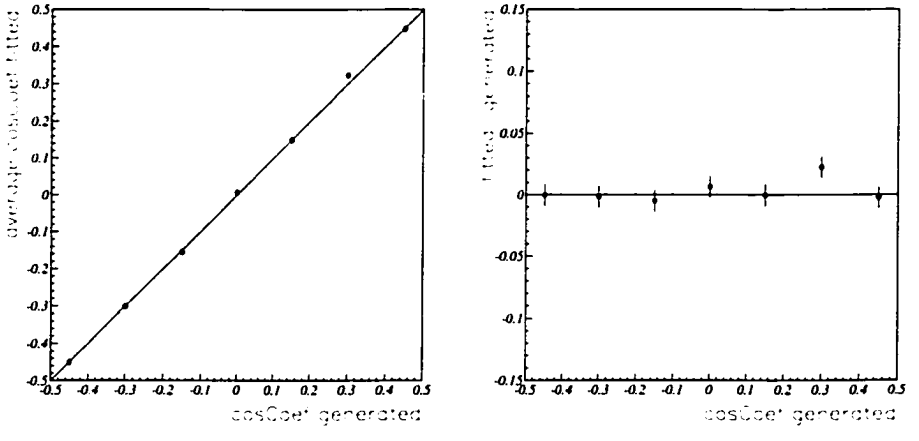


Figure B.32: Fitted cosine coefficient parameter versus generated one (left) and difference between fitted and generated value versus generated (right) while changing $|\lambda|$ value at generation time.

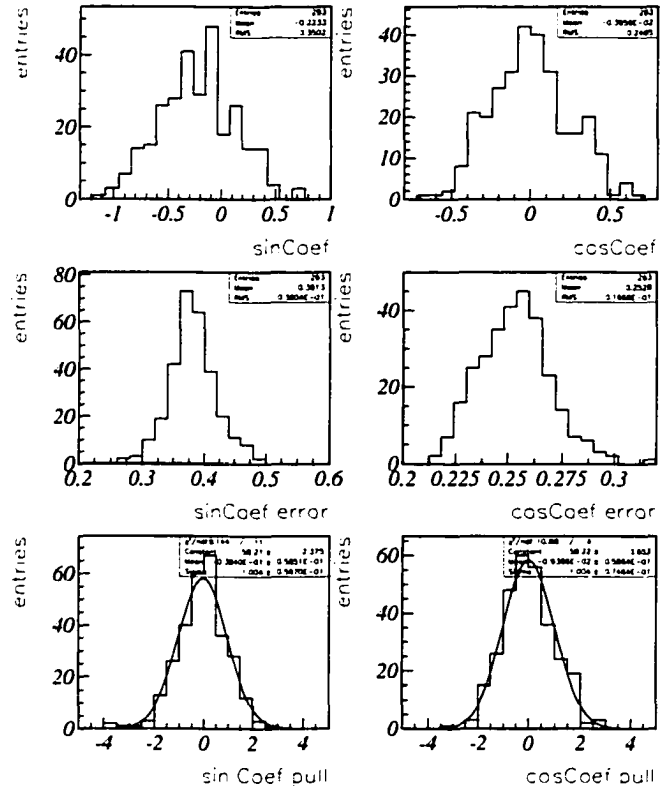


Figure B.33: Distribution of the sine (left) and cosine (right) coefficients in $D^{*+}D^{*-}$ MC (top); distribution of the errors (middle) and pull distribution (bottom).

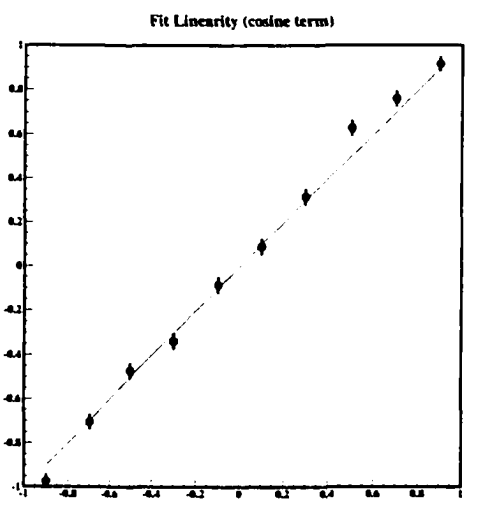


Figure B.34: Fitted vs. generated cosine coefficient parameter for $D^{*+}D^{-}$. Each point is 100 70-event experiments; the errors are the errors on the weighted mean.

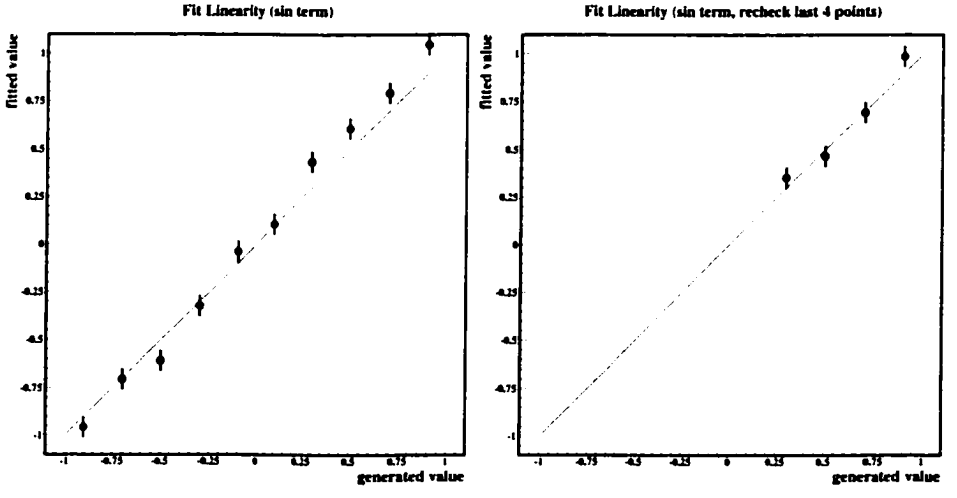


Figure B.35: Fitted vs. generated sine coefficient parameter for $D^{*+}D^{-}$ (left). Each point is 100 70-event experiments; the errors are the errors on the weighted mean. The last 4 points on the left plot are above the line. Rethrowing new experiments for these points yields the plot on the right, indicating that those points were a large statistical fluctuation.

Plots showing the distributions of expected errors and fitted values are in Figure B.36 for the "3-parameter" CP fit, and in Figures B.37 and B.38 for the "4-parameter" fit The parameters of the 4-parameter fit are very nearly uncorrelated; this is shown in Figure B.39.

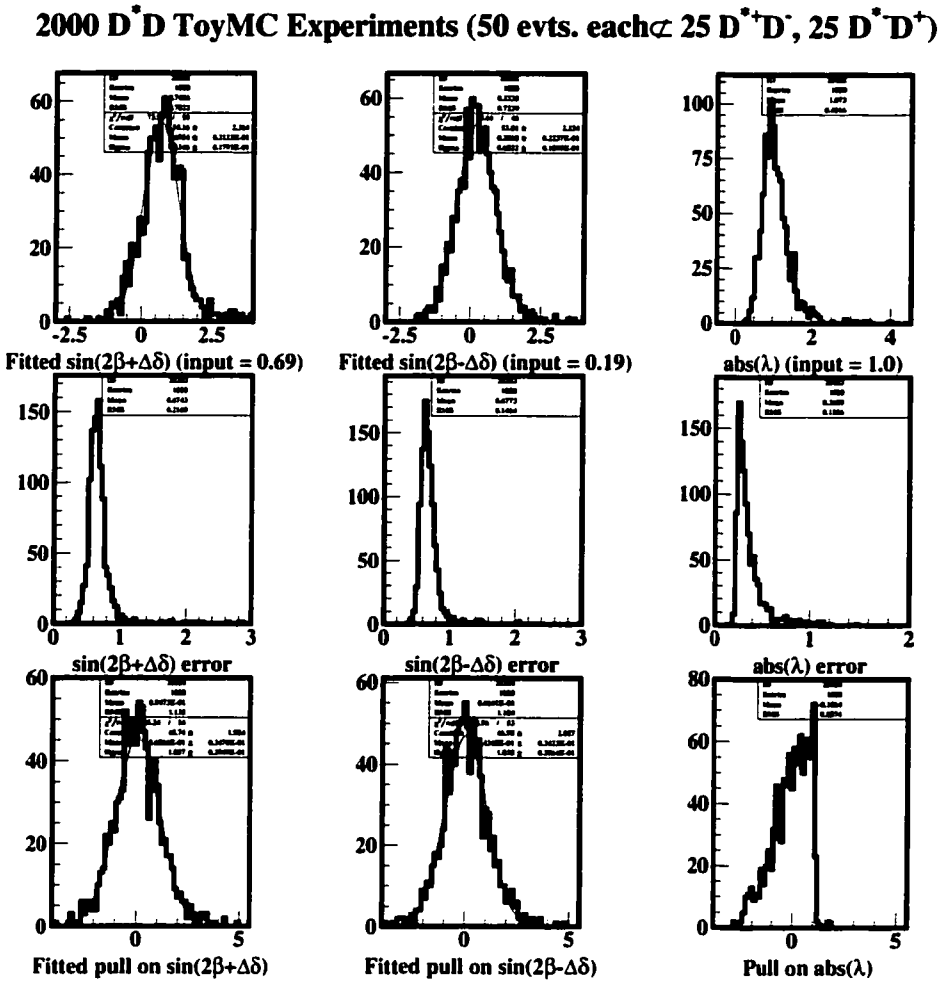


Figure B.36: The upper plots show the distributions of fitted values for the physical parameters of the "3-parameter" fit: $sin(2\beta + \delta)$, $sin(2\beta - \delta)$, and $abs(\lambda)$ for 2000 50-event toy Monte Carlo $B^0 \rightarrow D^{*\pm}D^{\mp}$ experiments. The middle plots show the expected (symmetric) errors on these quantities, and the lower plots show the pull distributions $(x_{true} - x_{fit})/\sigma_x$. Note the non-gaussian pull for $abs(\lambda)$ — this is due to a highly non-parabolic error for this parameter, which is displayed in figures following this.

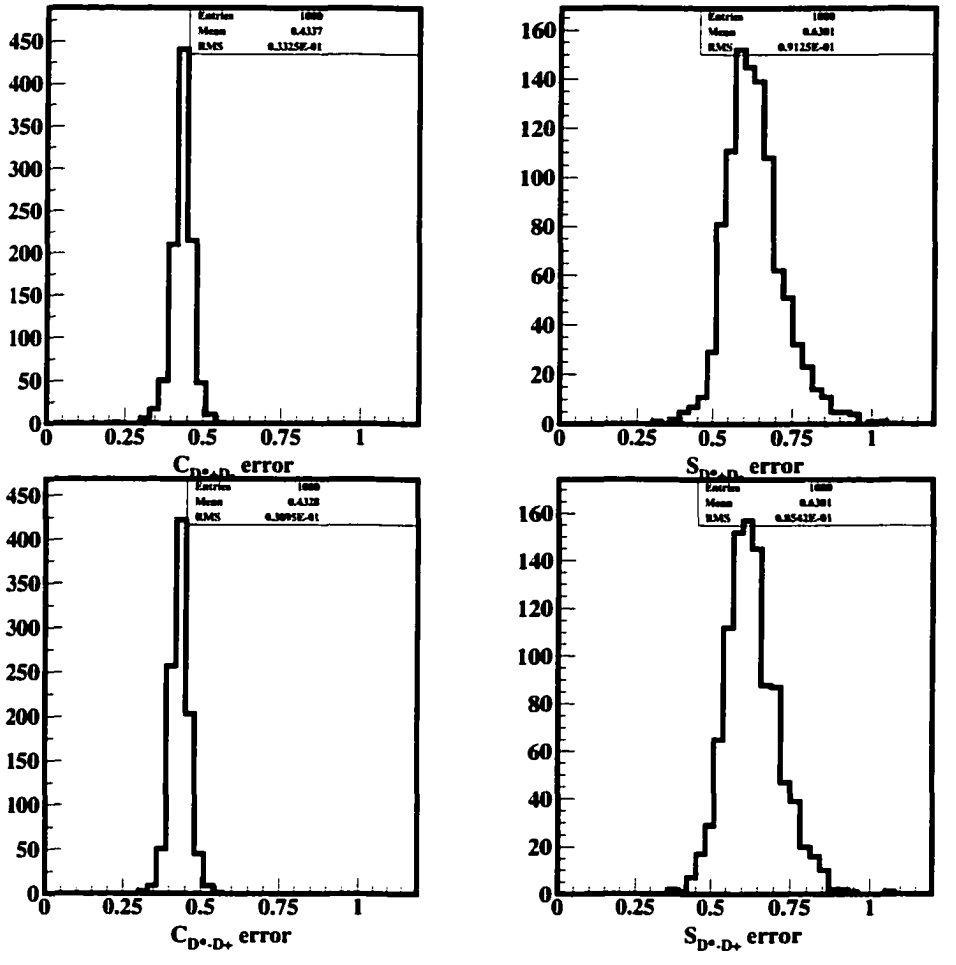


Figure B.37: The plots show the distributions of expected (symmetric) errors on (upper left) $C_{D^{*+}D^-}$, (upper right) $S_{D^{*+}D^-}$, (lower left) $C_{D^{*-}D^+}$, and (lower right) $S_{D^{*-}D^+}$ from 1000 50-event toy Monte Carlo $B^0 \to D^{*\pm}D^{\mp}$ experiments.

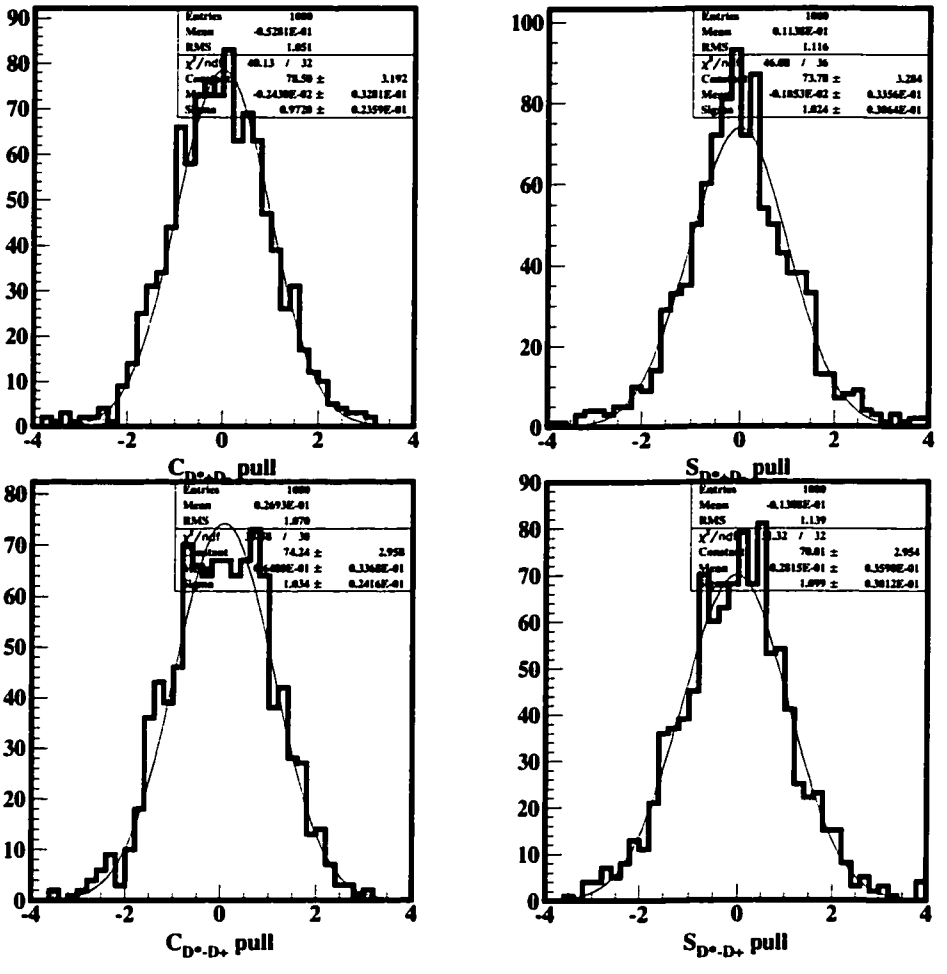


Figure B.38: The plots show the pull distributions $(x_{true} - x_{fit})/\sigma_x$ on (upper left) $C_{D^{*+}D^-}$, (upper right) $S_{D^{*+}D^-}$, (lower left) $C_{D^{*-}D^+}$, and (lower right) $S_{D^{*-}D^+}$ from 1000 50-event toy Monte Carlo $B^0 \to D^{*\pm}D^{\mp}$ experiments.

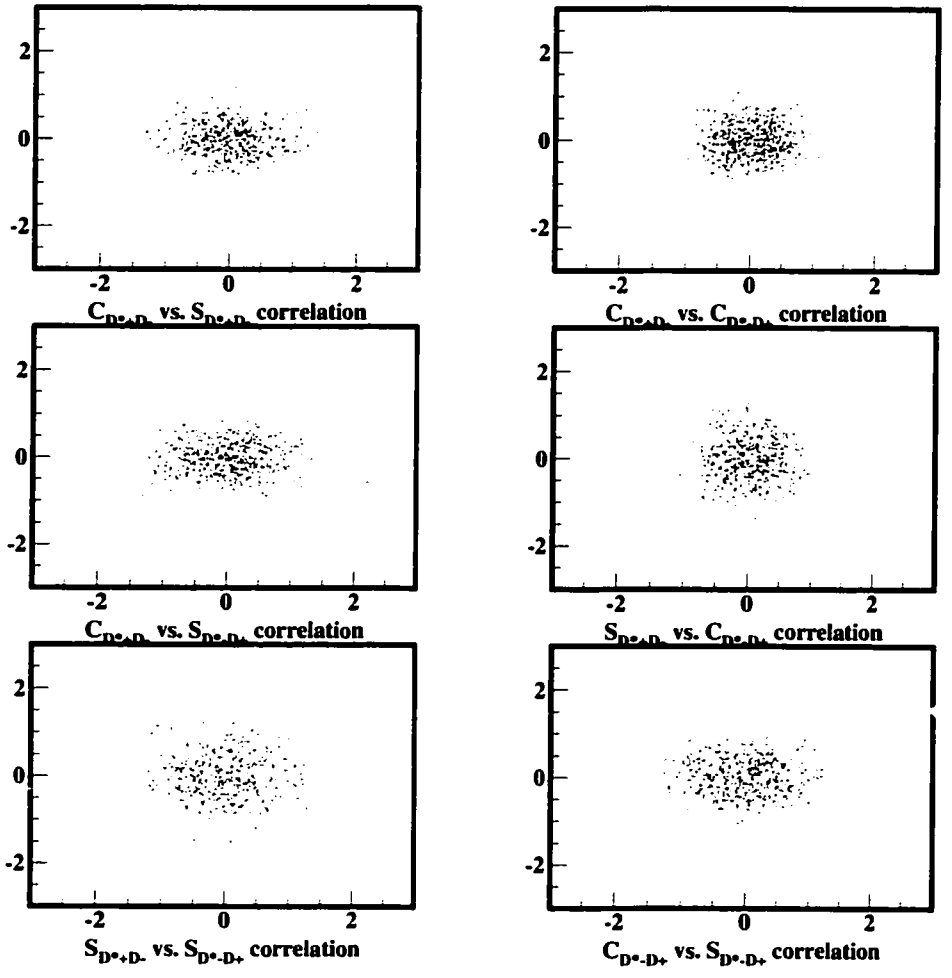


Figure B.39: The plots show the correlation between the 4 fitted parameters $C_{D^{*+}D^-}$, $S_{D^{*+}D^-}$, $C_{D^{*-}D^+}$, and $S_{D^{*-}D^+}$. using 1000 50-event toy Monte Carlo $B^0 \to D^{*\pm}D^{\mp}$ experiments. As can be seen, the correlation is very small.

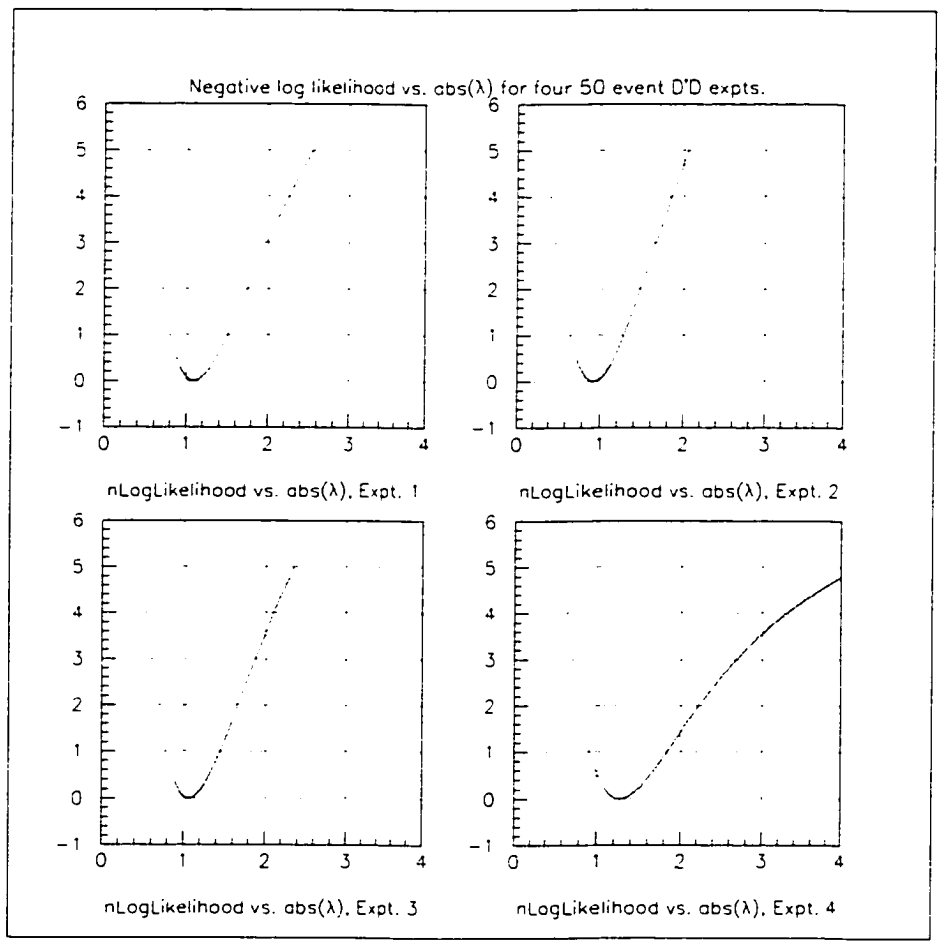


Figure B.40: Negative log likelihood $(-\ln(\frac{likelihood}{max(likelihood)}))$ vs. $abs(\lambda)$ for 4 $B^0 \rightarrow D^{*\pm}D^{\mp}$ experiments. Note the asymmetric errors and the non-parabolicity for this parameter.

Note that the parameter $abs(\lambda)$ appears to behave in a non-gaussian manner, as seen from its pull plot in Figure B.36. This is confirmed by viewing likelihood curves of $abs(\lambda)$, which display the non-parabolic errors for this quantity. Figure B.40 shows some example likelihood curves for $abs(\lambda)$.

Full MC studies

$B^0 \to D^{*\pm} D^{\mp}$ CP fitting studies using the full *BABAR* Monte Carlo simulation have also been done. A combined "4-parameter" fit to 20000 Breco and 15000 SP4 signal MC events, floating signal resolution function and dilution parameters, gives the following fit output:

```
===== Final value for Scale (core) - sig                1.1282 +/-    0.0535
===== Final value for d(dt) (core) [ps] - lepton       -0.0247 +/-    0.0386
===== Final value for d(dt) (core) [ps] - kaon         -0.1894 +/-    0.0298
===== Final value for d(dt) (core) [ps] - NT1          -0.0836 +/-    0.0502
===== Final value for d(dt) (core) [ps] - NT2          -0.1186 +/-    0.0415
===== Final value for d(dt) (tail) [ps] - sig          -0.3089 +/-    0.1527
===== Final value for fraction in tail - sig            0.1543 +/-    0.0231
===== Final value for outlier fraction - sig            0.0043 +/-    0.0014
===== Final value for dilu ave - sig, lepton            0.8733 +/-    0.0181
===== Final value for dilu ave - sig, kaon              0.6714 +/-    0.0147
===== Final value for dilu ave - sig, NT1               0.6154 +/-    0.0299
===== Final value for dilu ave - sig, NT2               0.3411 +/-    0.0256
===== Final value for dilu diff - sig, lepton           0.0013 +/-    0.0271
===== Final value for dilu diff - sig, kaon             0.0426 +/-    0.0197
===== Final value for dilu diff - sig, NT1             -0.0937 +/-    0.0395
===== Final value for dilu diff - sig, NT2              0.0722 +/-    0.0308
===== Final value for cos coeff D*-D+ signal           -0.0652 +/-    0.0294 (+    0.0294    -0.0294)
===== Final value for sin coeff D*-D+ signal            0.6987 +/-    0.0399 (+    0.0399    -0.0400)
===== Final value for cos coeff D*+D- signal            0.0247 +/-    0.0294 (+    0.0294    -0.0294)
===== Final value for sin coeff D*+D- signal            0.7540 +/-    0.0404 (+    0.0404    -0.0405)
```

showing that the physical parameters of the resulting fit are unbiased up to the statistical significance of the fit. An equivalent "3-parameter" fit also demonstrates a lack of visible bias:

```
===== Final value for abs(lambda)                    1.0460 +/-    0.0196 (+    0.0198    -0.0194)
===== Final value for Scale (core) - sig             1.1289 +/-    0.0535
===== Final value for d(dt) (core) [ps] - lepton    -0.0200 +/-    0.0382
===== Final value for d(dt) (core) [ps] - kaon      -0.1865 +/-    0.0296
===== Final value for d(dt) (core) [ps] - NT1       -0.0808 +/-    0.0501
===== Final value for d(dt) (core) [ps] - NT2       -0.1174 +/-    0.0414
===== Final value for d(dt) (tail) [ps] - sig       -0.3136 +/-    0.1532
===== Final value for fraction in tail - sig         0.1539 +/-    0.0231
===== Final value for outlier fraction - sig         0.0043 +/-    0.0014
===== Final value for dilu ave - sig, lepton         0.8730 +/-    0.0181
===== Final value for dilu ave - sig, kaon           0.6712 +/-    0.0147
===== Final value for dilu ave - sig, NT1            0.6164 +/-    0.0299
===== Final value for dilu ave - sig, NT2            0.3411 +/-    0.0256
===== Final value for dilu diff - sig, lepton        0.0087 +/-    0.0258
===== Final value for dilu diff - sig, kaon          0.0501 +/-    0.0178
===== Final value for dilu diff - sig, NT1          -0.0864 +/-    0.0387
===== Final value for dilu diff - sig, NT2           0.0767 +/-    0.0303
===== Final value for sin(2b+d) signal               0.7544 +/-    0.0405 (+    0.0405    -0.0405)
===== Final value for sin(2b-d) signal               0.7002 +/-    0.0400 (+    0.0400    -0.0400)
```

Similar checks were performed on ensembles of SP4 signal MC samples with more realistic statistics (70 $B^0 \to D^{*\pm}D^{\mp}$ events, equally divided between $D^{*+}D^-$ and $D^{*-}D^+$) as shown in Figure B.41 and Figure B.42. In a separate check, Breco signal MC was treated as $B^0 \to D^{*\pm}D^{\mp}$ signal (assigned randomly as 50% $D^{*+}D^-$ and 50% $D^{*-}D^+$) and equivalent checks were performed to make sure that the resulting fitted CP violation is zero. Results are shown in Figure B.43 and Figure B.44.

A plot showing the raw CP asymmetry for 376 $B^0 \to D^{*\pm}D^{\mp}$ signal events is shown in Figure B.45.

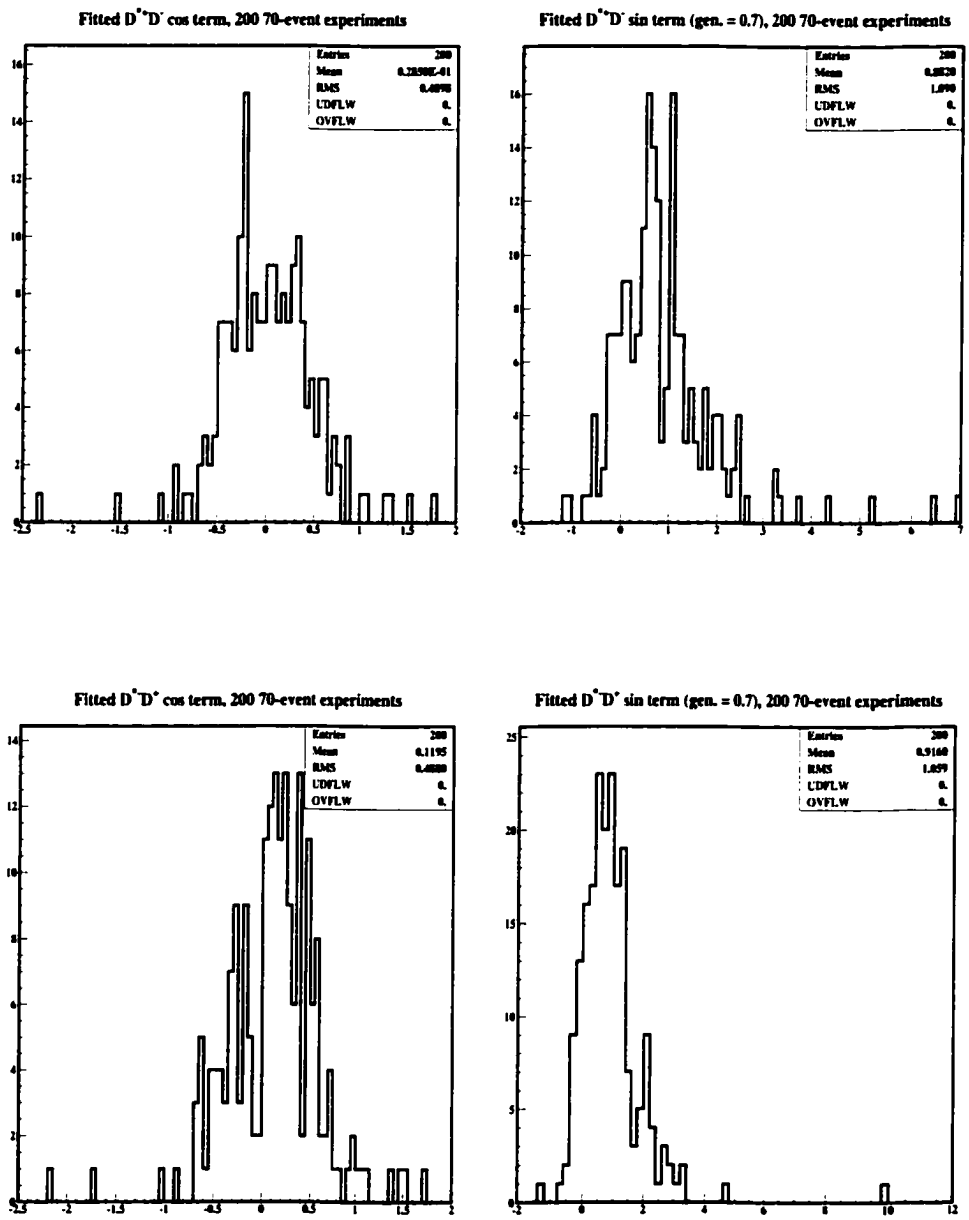


Figure B.41: Ensemble bias check of the "4-parameter" fit for 200 $B^0 \rightarrow D^{*\pm}D^{\mp}$ SP4 signal MC samples. Generated sin2β is 0.7 and abs(λ) is 1, thus the cos terms (left) should center around 0., whereas the sin terms (right) should center around 0.7. Pulls are shown on the folowing page.

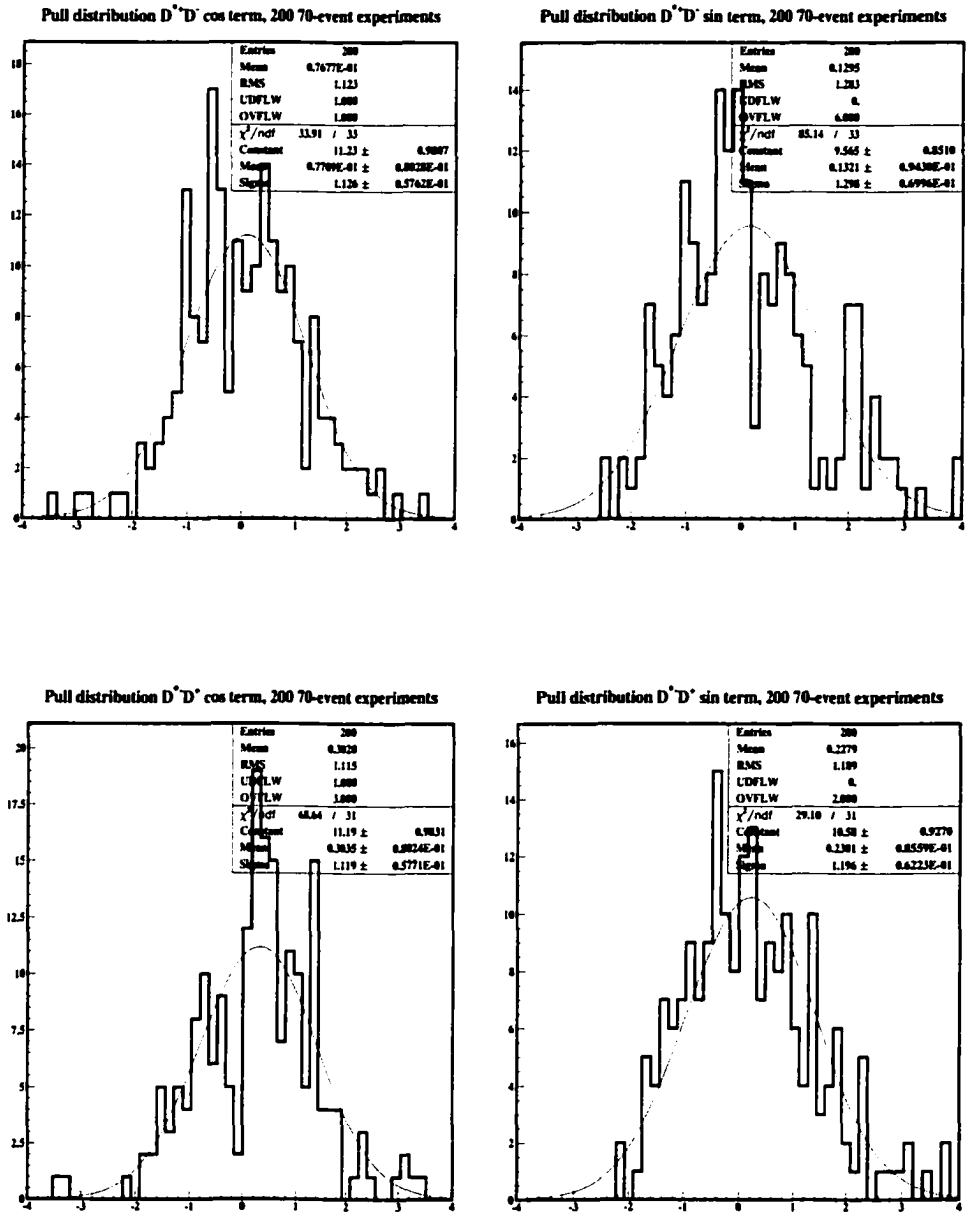


Figure B.42: Fitted pulls $(x_{fit} - x_{gen}/\sigma_{x_{fit}})$ of the 4 physical parameters for 200 $B^0 \rightarrow D^{*\pm}D^{\mp}$ SP4 signal MC samples. No statistically significant deviations are seen.

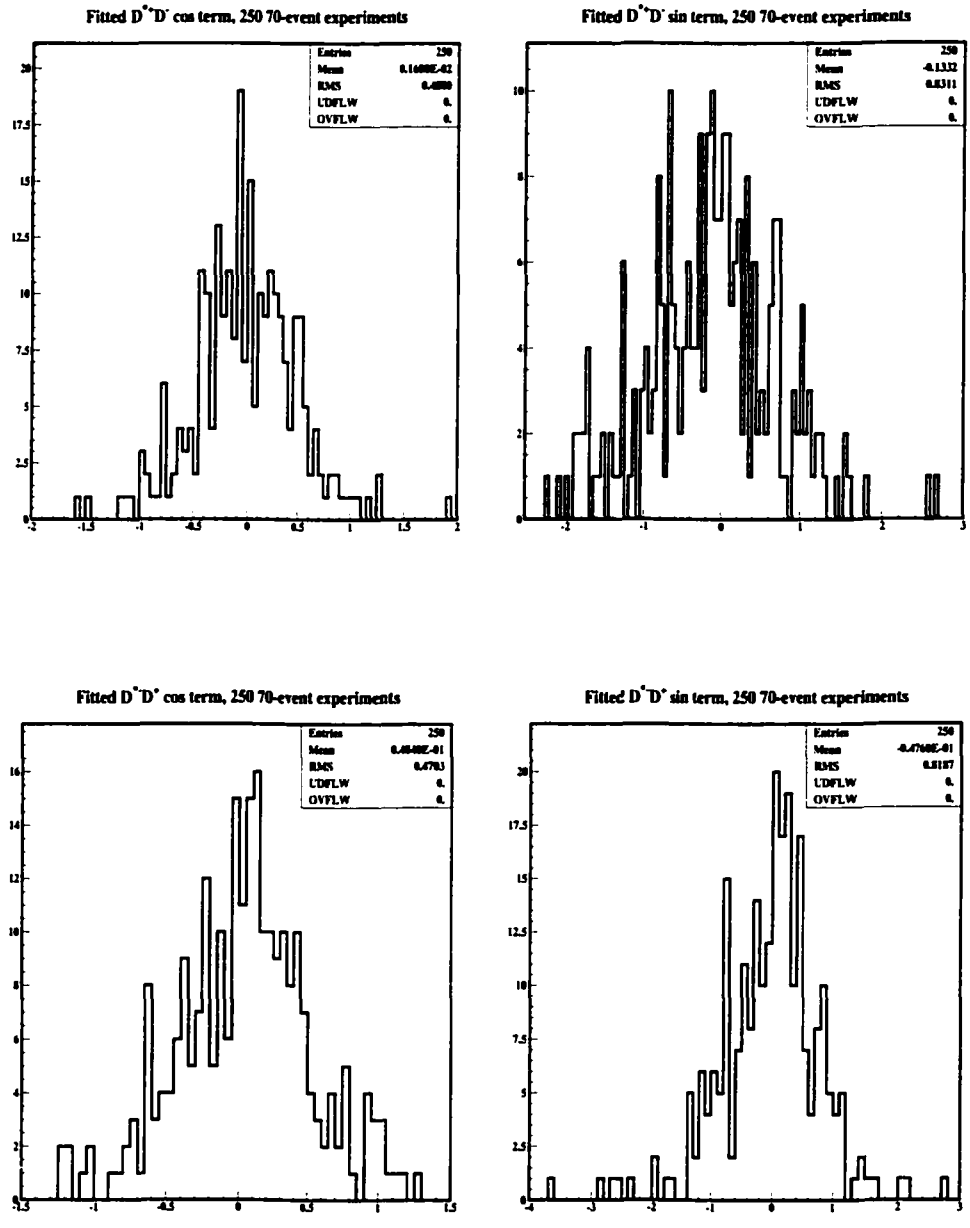


Figure B.43: Ensemble bias check of the "4-parameter" fit for 250 experiments where Breco MC samples are treated as signal (assigned randomly as 50% $D^{*+}D^{-}$ and 50% $D^{*-}D^{+}$). The cos and sin terms shouls each center around 0. Pulls are shown on the following page.

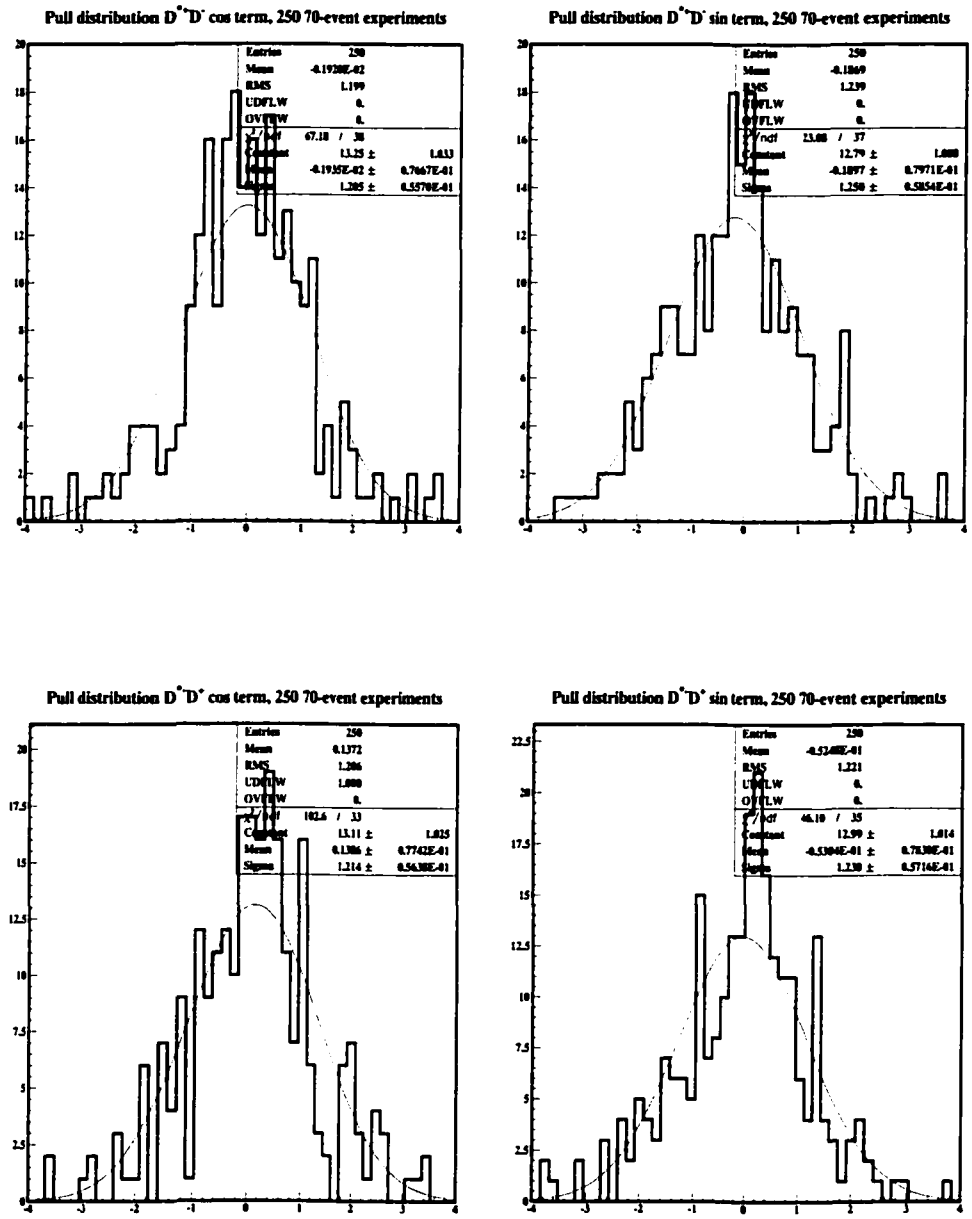


Figure B.44: Fitted pulls $(x_{fit} - x_{gen}/\sigma_{x_{fit}})$ of the 4 physical parameters for 250 experiments where Breco MC samples are treated as signal. No statistically significant deviations are seen.

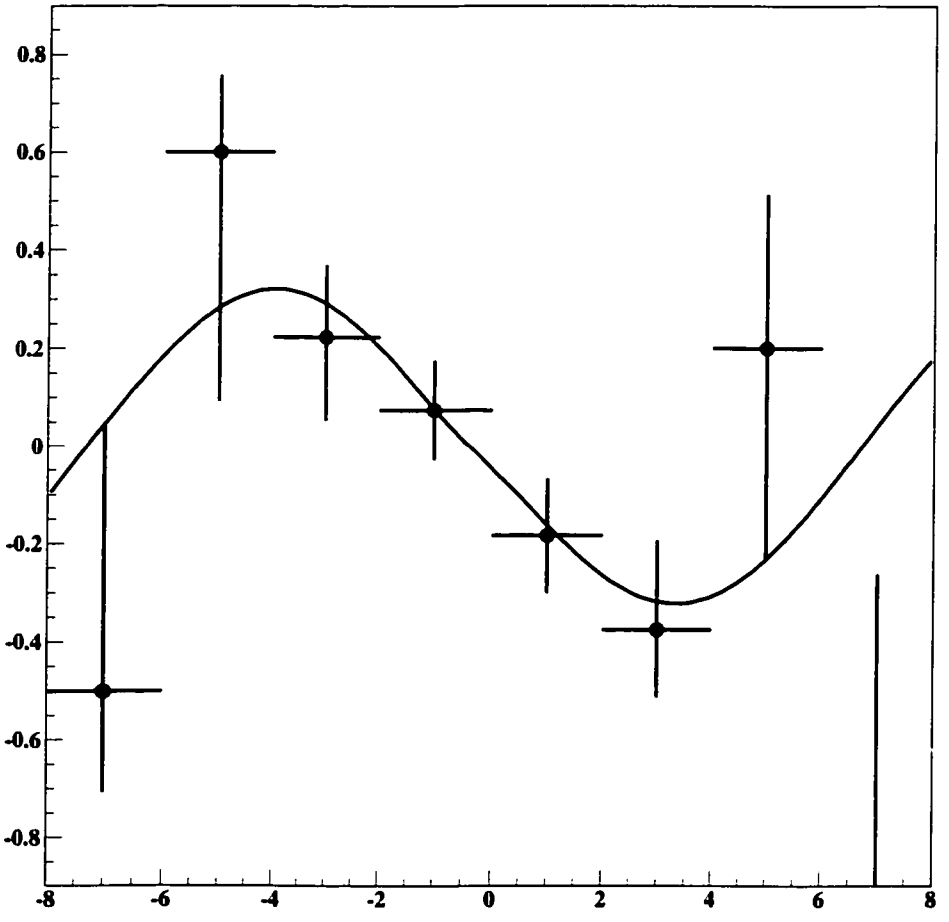


Figure B.45: Raw CP asymmetry for 376 $B^0 \rightarrow D^{*\pm}D^{\mp}$ signal MC events using the full Monte Carlo simulation (sp3). The fitted value of sin2β, 0.664$\pm$ 0.187 (by combining the fitted values of $sin(2\beta + \delta)$ and $sin(2\beta - \delta)$ as described in Section 4.2.1), is consistent with the input value of 0.7.

Appendix C

Kink-Finding Software

C.1 Overview

Kink-finding is necessary at BaBar for two primary reasons: 1) as a fake rejection criterion for muons found in the outer subdetectors and 2) as a means of increasing yield of kaons that decay in the tracking volume, for kaon tagging purposes. These two separate purposes also divide the strategies involved in finding kink decays in the experiment.

Pion decays within the tracking volume, especially decays at the relatively high momentum (> approximately 500 MeV) needed to reach the muon detection subsystems (IFR and EMC), tend to be small angle and are predominantly reconstructed as a single track. Using Babar's Kalman filter track fit, trajectory pieces before and after a hypothesized decay in a track can be reconstructed. Obtaining these trajectory pieces allows both consistency variables developed in previous experiments and novel consistency variables to be used. As this involves searching for kinks within a *single* reconstructed

track, this is referred to as **"1-track" kink-finding**.

For decaying kaons, a slightly different strategy is required. The kaon tagging spectrum is dominated by low-momentum (< 1GeV) particles (a large percentage of which [10-11%] decay in the tracking volume). Charged kaon decays, unlike pion decays, have a large p (205/236 MeV) associated with the decay. The resulting large angle between the parent and child tracks tends to result in 2 *separate* tracks being reconstructed by initial track pattern recognition. As this involves searching through *pairs* of reconstructed tracks for kinks, this is referred to as **"2-track" kink-finding**.

As the output of the kink-finding algorithms must be carefully observed to watch for unusual detector or reconstruction behavior, **kink-finding monitoring** must be employed as an automatic method for producing diagnostics. As the results of the algorithms must be stored after reconstruction, **kink persistance** using BaBar's Objectivity database must take place. And for users wishing to obtain kink-finding information in analyses, a clear and logical **kink-finding interface** is explained.

C.2 "1-Track" Kink Finding

C.2.1 Previous Methodologies

As explained above, "1-track" kink-finding involves searching for kinks within a single reconstructed track. Several methods for this form of kink detection have been used in previous experiments. Of these, methods using the output of Kalman filter track fitting algorithms have been particlarly successful. The Kalman filter algorithm is able to produce a piecewise trajectory of optimal

parameters at any point along the particle path [62]; one can exploit the output of the Kalman filter by using the differences between these sets of parameters as a function of path length in a search for kinks. The DELPHI collaboration was the first to use such information in a kink-finding algorithm [63]. On can simply compare the track parameters before and after a given point on the trajectory and form a χ^2 for the matching of the two parameter sets.

$$\chi^2_{match} = (\boldsymbol{p}^B - \boldsymbol{p}^F)^T [C^B + C^F]^{-1} (\boldsymbol{p}^B - \boldsymbol{p}^F) \tag{1}$$

where p are the forward and backward parameter vectors at a given point and C are the associated covariance matrices. Small subtleties include the fact that one must be sure to use the unsmoothed Kalman parameters, ie those obtained directly from the forward and backward filters, and also that one must not double-count the information from any Kalman sites along the trajectory when separately computing the forward and backward parameter sets. The BaBar Kalman algorithm stores the full filter outputs internally when the track fit is made, so no fit recomputation is necessary to compute this variable, making computation of this χ^2 simple and fast.

Cousins *et al* describes an enhancement to this method which was used by the NOMAD collaboration [64]. Using the fact that kinks are divided into three physical categories:

I. *Bremsstrahlung*, in which only the momentum *magnitude* is changed at the kink site.

II. *Hard elastic scattering*, in which the momentum magnitude is unchanged,

but the directional angles in theta and phi change.

III. *Decays* (most often $\pi \rightarrow \mu + \nu_\mu$) may change both momentum magnitude and direction.[1]

one can determine separate χ^2 consistencies for each of these hypotheses. To achieve this using the conventional BaBar track fit parameters[65], one must transform the parameters and covariance to use the hypothesized kink site as a reference point, rather than the nominal use of the origin (so that the parameters d_0 and z_0 will simply describe continuity at the kink site and thus can be eliminated). This is somewhat of a speed penalty; therefore this information is only calculated at a given site if the basic mismatch χ^2 described above provides an indication that a kink may be located there.

C.2.2 1-Track Kink Finding and the BaBar Kalman Fit

The BaBar Kalman filter fit supports kink finding by providing the optimal track fit parameters (and their covariance) on either side of a kink candidate. This pair of parameters (one inside the kink, one outside) are equivalent to the result that would have been obtained had a full Kalman fit been performed on the subset of hits found inside or outside the kink candidate flight length. Since these parameters include all material (dE/dx and multiple scattering) effects and magnetic field inhomogeneitiy corrections, they can be

[1]Inelastic scatters with a single charged daughter particle would also fall under this category.

used to create an optimal descriminant for selecting a particular kind of kink according to the physics model of a specific kind of kink.

A trivial model which exploits the interior fit parameters is the model that the track was produced by a single particle (no kink). The statistic produced by this model is the so-called matching chisquared, which is a much more sensitive test of kinking than the fit chisquared itself, since it is sensitive to the correlated displacement of hits produced by a kink. The matching chisquared is an appropriate measure to use when initially selecting kink candidates from the general track list. Discriminants for specific kink models can be formed by allowing appropriate degrees of freedom (change in track direction and/or momentum) to occur at the kink location without a chisquared penalty.

Both the matching chisquared and the interior fit parameters are provided as part of the KalRep interface, which is the transient representation of the Kalman track fit result. In order to facilitate the process of finding a kink, whose location is *a priori* unknown, the KalRep interface to the matching chisquared and the interior fit parameters takes as input the flight length of the candidate kink. This allows the implementation of kink finding which actually searches for the location (flight length) of the kink by testing various flight lengths, and choosing the one which minimizes the chisquared.

The ability to implement such powerful kink-finding tools in an efficient way is a unique feature of the BaBar Kalman filter track fit implementation. This implementation caches the intermediate results of processing the KalSites when a track is initially fit. Subsequent computation of the interior fit parameters or the matching chisquared can implemented extremely

efficiently using these intermediate results, requiring at most one five-by-five matrix inversion and matrix multiply.

C.2.3 1-Track Kink Finding Strategy

The 1-track kink finding module, Trk1TrkKinkFind,[2] runs at the very end of the tracking reconstruction sequences, after all DCH and SVT pattern recognition and fitting has taken place. Thus the kink finder can take advantage of the full DCH-SVT tracking volume and not be restricted to either subsystem.[3]

Several cuts can be made to reduce the quantity of events and tracks on which a kink search is performed; such cuts increase the speed of kink finding. The first such cut that can be made is the number of tracks in the event — as it is not important for physics to search for kinks in 2-track events (which constitute a large percentage of the events passing through reconstruction [dimuons and prescaled bhabhas]), such events are skipped. The parameter representing the minimum number of tracks in an event for a kink search to proceed can be varied in tcl (see Appendix A). In order for there to be sufficient lever arm for finding a kink, a minimum number of active hits are required. At bare minimum this number must be at least 10, to satisfy the minimum number of degrees of freedom for obtaining the 5 forward and 5 backward filter parameters at a point along the track. The tcl parameter (minNActive) is currently set by default to 12 — at least 12

[2]The prefix "Trk" marks the module as part of the tracking reconstruction.

[3]A disadvantage of this is that it enhances the dependence of kink-finding on the accuracy of tracking alignment calibrations, however alignment is universally critical for physics measurements.

active hits (SVT or DCH) are required before kink finding takes place (this cut does not affect efficiency, as one needs sufficient lever arm for a kink to be found via a search anyway). Another simple cut that may be used to reduce the number of tracks that unecessarily pass through a kink search is pT. As the major physics goal of 1-track kink finding is to reduce the number of muon fakes found in the IFR, finding kinks in tracks that do not have sufficient momentum to reach the IFR is not an important goal. Therefore tracks which fall below the minPt tcl value, which is set by default to 500 MeV (as tracks below that pT will not be reconstructed as muons in the IFR). Thus *one will never see a 1-track kink in a track that has* $pT < 500$ *MeV*. The last such simple cut that is made uses the χ^2 consistency of the reconstructed track. Studies showed that the vast majority of kinked tracks have χ^2 consistency < 0.2 (see below), which is the default setting for the tcl parameter maxChiSqProbab. These simple cuts remove the majority of unkinked tracks, providing a reduced sample on which a kink search can take place quickly and efficiently.

After these initial cuts are made to eliminate tracks that are unlikely to contain kinks, a full search can proceed on the remaining tracks. Mismatch χ^2 is evaluated at positions along the track from the 5th to the 5th-to-last active hit (the 5 is set by the tcl parameter trkEndMargin, however, as tracking does not plan to change the fact that it uses 5 parameters, this is unlikely to ever need modification). The distance between the positions at which the mismatch χ^2 is evaluated is set by the tcl parameter rangeStep — this is set by default to 1 cm.[4] The flightlength with the maximum mismatch χ^2

[4]The fact that there will sometimes be no intervening KalSites between neighboring

is found and the χ^2/nDof is evaluated at that site. A momentum-dependent cut is then applied on the χ^2/nDof. Momentum-dependence is expected for the mismatch χ^2 according to the dependence of the mean kink angle and δp on momentum (as shown in figure C.1). Tcl parameters specifying three terms of the Laurent series of momentum dependence are provided — at present only one of these parameters is used (the others are set to zero) and the mismatch χ^2 cut is simply considered to be inversely proportional to momentum. From the plots in figure C.1, this appears to be close to the true dependence (although fitting for these parameters in MC may be a useful tune for the future). If the track passes this cut, the maxiumum mismatch χ^2 is again used in a cut that attempts to separate kinks from tracks that are simply poorly fitted. As the relationship

$$\chi^2_{trk} = \chi^2_{forw} + \chi^2_{back} + \chi^2_{match} \tag{2}$$

holds at any point along the track, one can determine the percentage of total track χ^2 which is due to the mismatch at any given point. This percentage is evaluated and required to be greater than the tcl parameter minMismatch-Percent (set by default to 0.6) in order to reject tracks that are simply poorly fitted. After this cut is made, the Cousins χ^2 for the decay hypothesis is evaluated at the flightlength with the maximum mismatch χ^2. The Cousins χ^2/nDof is given a cut with the same momentum dependence as that given for the mismatch χ^2 (controlled by the tcl parameter cousConsistencyCut)

locations is not a problem (having an ambiguity in the position with the maximum χ^2 is not a problem — the first one is simply chosen), in fact it is advantageous to throughly search the full track.

and this is the final cut.

C.2.4 1-Track Kink Finding Performance

The performance of 1-track kink finding was measured on data and both single pion and generic $B\bar{B}$ monte carlo in 8.6.2c. The pions in the single pion MC had their lifetimes reduced by a factor of 100 (to 0.2603 ps) to increase the percentage of tracks that decay in the tracking volume and were generated with momenta between 0.75 and 2 GeV. The following additional tags were used (all of these tags, or more recent ones, are in all the 9.X series releases):

`KalPhysProcess`	`V01-01-07`
`TrkKink`	`V01-01-07`
`TrkKinkTests`	`V00-00-07`
`TrkSequence`	`jea-04-16-00`
`BbrGeom`	`V00-07-08`

Efficiency

Efficiency was measured in generic $B\bar{B}$ and in the single pion MC mentioned above. Efficiency is defined as the chance of reconstructing a kink given that there is a single decay (represented by a single change in GTrack associated with the GHits that correspond to the reconstructed HOTs) along the track between the 1st and last hit (unless otherwise mentioned), also given that the pT is greater than 500 MeV (unless otherwise mentioned). See table 1

for efficiencies.

Purity

Purity was similarly measured in generic $B\bar{B}$ and in the single pion MC. See table 2 for purity information.

Timing

CPU timing was measured on the "tersk" machines at SLAC, on code compiled optimized for SunOS5. Note that the Trk1TrkKinkFind module is after the background filter, so it does not run on every event. The following table, obtained from the AppAST module timing utility, summarizes the information:

real data, 1000 events from run 12225

AppAST: 0.00 || 311 || 1.59 || 5.11254 || 0.00 ||Trk1TrkKinkFind

real data isMultiHadron events (run 12225)

AppAST: 0.00 || 500 || 3.49 || 6.98000 || 0.00 ||Trk1TrkKinkFind

SP3 BBbar w/bkg MC (run 200104)

AppAST: 0.00 || 500 || 2.75 || 5.50000 || 0.00 ||Trk1TrkKinkFind

At 5-7 ms/event, kink finding is one of the faster reconstruction modules and timing is not a major concern.

[5]Note low statistics: > 1GeV pion kinks are not such common occurrances in $B\bar{B}$.

[6]This is the major background.

[7]Note that these are single track events.

Table C.1: **Efficiencies**

	SP3 BBbar + digi bkg, 2000 evts (runs 200104 and 200105)	Single particle pion decays, 100x reduced ct, 0.75-2.0 GeV, 1000 evts
a) Pion decays occurring in tracking volume, given that a single track is reconstructed, pt > 500 MeV	21/46 = 46%	289/557 = 52%
b) Same as a), if the decay occurs between the 5th and 5th-to-last hit on track	19/29 = 66%	276/394 = 70%
c) Same as a), pt > 1 GeV	3/6 = 50%[5]	117/221 = 53%
d) Same as b), pt > 1 GeV	2/3 = 66%[5]	108/146 = 74%

Miscellaneous

Information such as the number of 1-track kinks one can expect to find in data and in MC is often of use. The following table summarizes this:

real data, 1000 events from run 12225

# of reconstructed 1-track kinks: 17 ($\approx$ **1 per 59 data events**)

real data, 1000 isMultiHadron events from run 12225

# of reconstructed 1-track kinks: 90 ($\approx$ **1 per 11 isMH events**)

SP3 B$\bar{B}$ w/bkg MC, 1000 events from run 200104

# of reconstructed 1-track kinks: 85 ($\approx$ **1 per 12 $B\bar{B}$ MC events**)

Locations of found 1-track kinks in data and in MC can be found in figures C.3 and C.4 respectively. The difference in the distributions can be attributed to the loss of efficiency towards the center of the chamber for data in which the DCH was at 1900V.

C.3 "2-Track" Kink Finding

C.3.1 2-Track Kink Finding Strategy

2-track kink finding, as mentioned in the overview, involves searching for kinks among pairs of reconstructed tracks. The 2-track kink finding module, Vtx2TrkKinks, runs in the physics path, unlike 1-track kink finding (which, as mentioned in the previous section, runs in reconstruction). This decision was made in order to increase the flexibility of running 2-track kink finding.

1-track kink finding is restricted to running in the reconstruction by necessity, due to the fact that it requires detailed information about piecewise track trajectories, which is not available to users running physics modules from the micro database. 2-track kink finding, however, does not require such detailed information, and by making 2-track kink finding a physics module, users do not have to rerun the reconstruction in order to run it at their leisure (and with whatever tcl settings they prefer). There does exist the ability to run 2-track kink finding in the reconstruction, though, if this is ever necessary.

Vtx2TrkKinks takes a list of BtaCandidates as input, which is ChargedTracks by default (this is controlled by the tcl parameter inputBtCandidateList). It then looks for pairs of tracks within this list that pass kink cuts. Alternatively, if the tcl parameter selectTracks is set to false, one can instead provide the module with 2 separate BtaCandidate lists in which to search for kink "mothers" and "daughters" respectively (such as within lists of kaons and muons) — these lists are specified by the tcl parameters mothersList and daughtersList respectively. Events with just 2 or less ChargedTracks are eliminated (this is controlled by the tcl parameter minNumTracks). If selectTracks is true (the default setting), internal "possibleMothers" and "possibleDaughters" lists of BtaCandidates are created. The criteria for a BtaCandidate ending up on possibleMothers is: the DOCA to the beam spot must be within a radius specified by the tcl parameter motherDocaToOriginCut (2 cm by default) and it must not already end in a decay vertex. The criterion for a track ending up on possibleDaughters is: the DOCA to the beam spot must be *outside* a radius specified by the tcl parameter daughtDocaToOriginCut (5 cm by default).

All possible pairs of tracks within these internal "possibleMothers" and "possibleDaughters" lists are then considered. Pairs in which the tracks have different charges are dropped (as long as the tcl parameter checkSameCharge remains at its default value of true). The POCA of each track pair is calculated and a cut on the maximum DOCA is made (set by the tcl parameter maxTrkDist, by default 1 cm [which is fairly tight]) are dropped. A cut is then made on the endFoundRange of the mother track (the flightlength of the last HOT) — this flightlength must be within a distance to the flightlength at the POCA to the daughter track which is specified by the tcl parameters motherPocaBfrEndCut and motherPocaAftEndCut. These parameters are set to -5 cm and 47.5 cm by default, meaning that 47.5 cm > (flightlength at kink) − (flightlength at endFoundRange of mother) > -5. cm. This help eliminate ostensible mother tracks which unphysically extend beyond the kink site itself. Similarly, a cut is made on the startFoundRange of the daughter track — this flightlength must be within a distance to the flightlength at the POCA to the mother tracks which is specified by the tcl parameters daughtPocaBfrStartCut and daughtPocaAftStartCut. These parameters are set to -5 cm and 25 cm by default, ie 25. cm > (flightlength at startFoundRange of daughter) − (flightlength at kink) > -5. cm. If this is satisfied, a mass hypothesis cut is made.

The invariant mass of the kink is calculated in both the decay to $\mu\nu_\mu$ hypothesis and the decay to $\pi\pi^0$ hypothesis. This is achieved by considering the missing 3-momentum vector $\boldsymbol{p}^{mother} - \boldsymbol{p}^{daughter}$, which is considered to be the 3-momentum of the neutral. The hypothesis is then applied to create 4-momenta for both the reconstructed daughter and the hypothesized neutral

daughter. The invariant mass of the mother is then calculated from the sum of these two 4-momenta. A peak is obtained around the kaon mass (see figures C.7 and C.8) which provides the ability to make a cut.

A cut is made around the mass in the $\mu\nu_\mu$ hypothesis which is set by the tcl parameters invMassMotherMuMin and invMassMotherMuMax. The defaults for these values are 400 MeV and 570 MeV (the PDG mass of the charged K is 493.7 MeV). This is wide enough to allow decays to $\pi\pi^0$ to pass as well (such decays tend to form a second peak in $\mu\nu_\mu$ hypothesis mass plots). If this cut is passed, the kink is then vertexed.

Vertexing is done using the LeastChi algorithm in Fast mode by default. These settings can be changed via the tcl parameters vtxFitAlg and vtxFitMode respectively (a change of default to GeoKin may be made in the near future due to the lack of continuing support for the LeastChi algorithm). A VtxKink object (which inherits from a BtaAbsVertex) is created to represent the vertex of the mother $\rightarrow$ daughter decay (the interface is explained in more detail in the Kinkfinding Interface section). A cut is made on the χ^2 of the vertex fit — if the χ^2 is greater than the tcl parameter maxVtxChi2 (default is 8.), the kink is rejected. A composite track representing the neutral daughter is created and added to the VtxKink vertex. A consistency for the decay is obtained by assuming a Lorentzian over flat background mass distribution and a decaying exponential in mother-daughter track DOCA. Parameters representing the Lorentzian shapes, the height of the flat background, and the coefficienct for the decaying exponential in DOCA are given by the tcl parameters alphaMuNumu, alphaPiPi0, flatBkg, and docaCoeff. The consistency is considered to be the height of the Lorenzian at the calculated mass

divided by the total Lorenzian plus flat background height, all multiplied by the value of the decaying exponential in DOCA, and this value is stored in the VtxKink object. Finally, the mother BtaCandidate is put on the output kink mother list (set by the tcl parameter outputKinkMothers, by default **KinkMothers**) and the daughter and neutral daughter are put on the output kink daughter list (set by the tcl parameter outputKinkDaughters, by default **KinkDaughters**).

C.3.2 2-Track Kink Finding Performance

The performance of 2-track kink finding was last measured in detail in release 8.2.8 (the code has not changed much since then). The following table summarizes the efficiency of each of the cuts with 500 generic $B\bar{B}$ events.

```
                                                        475 scatters
                                                         53 albedo scatters
 track pairs in 500 generic bbar events:              +- 1046 non-merges
 ---------------------------------------              |
                                                      |    topological
100 K -> mu numu  kink decays      background         |      kinks(*)
 21 K -> pi pi0   -----------      ----------         |    -----------
      +------------> 122              19713          +------> 1574
```

*topological kinks = elastic scatters, inelastic scatters with not more than 1 charged daughter, albedo scatters, bremsstrahlung, and failure to merge SVT and DCH tracks. These were considered separately from signal and from combinatoric background (since they are neither especially "bad" nor especially "good"–it depends on what the kink-finding is being used for).

mother DOCA to beamspot < 2 cm	110	16546	1163
daught DOCA to beamspot > 5 cm	104	9122	958
r of POCA > 10 cm	91	6119	804
DOCA of track trajectories < 1 cm	50	273	177
47.5 cm > mother POCA flight length - mother endFoundRange > -5 cm	49	118	127
25 cm > daught startFoundRange - daught POCA flight length > -5 cm	47	64	78
vtx chi^2 < 8.	46	52	74

```
0.575 > inv. mass of
mother in mu + numu
hypo. > 0.4 GeV        46                    12                    8
                        |  34 K->mu numu          3 scatters        |
                        +- 12 K->pi pi0           5 non-merges --+
```

Timing infomation on $B\bar{B}$ MC and on data using an executable compiled opimized on Sun Ultra-5 machines is summarized in the following table:

real data, 500 events from "Padova" reprocessing

8.62 ms/event

$B\bar{B}$ SP2 MC, 500 events

14.4 ms/event

A kink is found in approximately 1 out of every 50 $B\bar{B}$ events. Event displays of found 2-track kinks in real data and in MC can be found in figure C.5. The distribution of the found kink vertices in real data can be found in figure C.6.

The reconstructed invariant mass in the $\mu\nu_\mu$ hypothesis for found 2-track kinks can be plotted (see figure C.7) and, as expected, shows a peak at the charged Kaon mass in both MC and real data (invariant mass in the $\pi\pi^0$ hypothesis, as expected, shows a smaller peak in MC and in data,[5] as shown in figure C.8).

[5]The central values for the branching ratios for $\mu\nu_\mu$ and $\pi\pi^0$ are 63.51% and 21.16% respectively (PDG 1998)

C.4 Kink Finding Interface

A uniform, clear, and logical physics interface for kink finding is of course necessary, and this section outlines the interface that is in place currently and details its use. There are 2 lists of BtaCandidates, KinkMothers and KinkDaughters, that are in the event after kinkfinding is run. Each KinkMother BtaCandidate ends in a decay vertex that can be accessed through the decayVtx() member function of the BtaCandidate. The decayVtx() function returns a pointer to a BtaAbsVertex; however, if a given candidate is on the KinkMothers list, the decay vertex may be safely casted to a VtxKink pointer (which inherits from a BtaAbsVertex). VtxKink resides in the package VtxBase and the interface it presents is shown in figure C.10.

C.4.1 Typical Use

Some code fragments providing examples of typical use of kinkfinding in analysis code are presented here:

a) To add the 2-track kinkfinding module (Vtx2TrkKinks) to your AppUserBuild:

```
#include "VertexingTools/Vtx2TrkKinks.hh"
add( new Vtx2TrkKinks( "Vtx2TrkKinks", "2-track kinks" ) );
```

b) To add the 2-track kinkfinding module to your path (in myAnalysis.tcl):

```
sequence append MyAnalysis Vtx2TrkKinks
```

c) To get the full HepAList of KinkMother BtaCandidates (after doing a) and b)):

```
HepAList<BtaCandidate>* kinkMotherList;
IfdStrKey kinkKey("KinkMothers");
getTmpAList (anEvent, kinkMotherList, kinkKey);
if ( 0 == kinkMotherList ) {
    ErrMsg(warning) << "no kinkMotherList found, skipping event" << endmsg;
    return APPMODULERETURNVAL1;
}
```

d) To get the probability that a given BtaCandidate (*cand) is a decay (after doing a), b), and c)):

```
#include "VtxBase/VtxKink.hh"

    double decayProba(0.);
    HepAListIterator<BtaCandidate> kinkMothersIter(*kinkMotherList);
    const BtaCandidate* decay;
    while (decay = kinkMothersIter()) {
        if (cand->uid() ==
            ((VtxKink*) decay->decayVtx())->originalMotherCand()->uid()) {
          decayProba = ((VtxKink*) decay->decayVtx())->kinkConsistency();
        }
    }
```

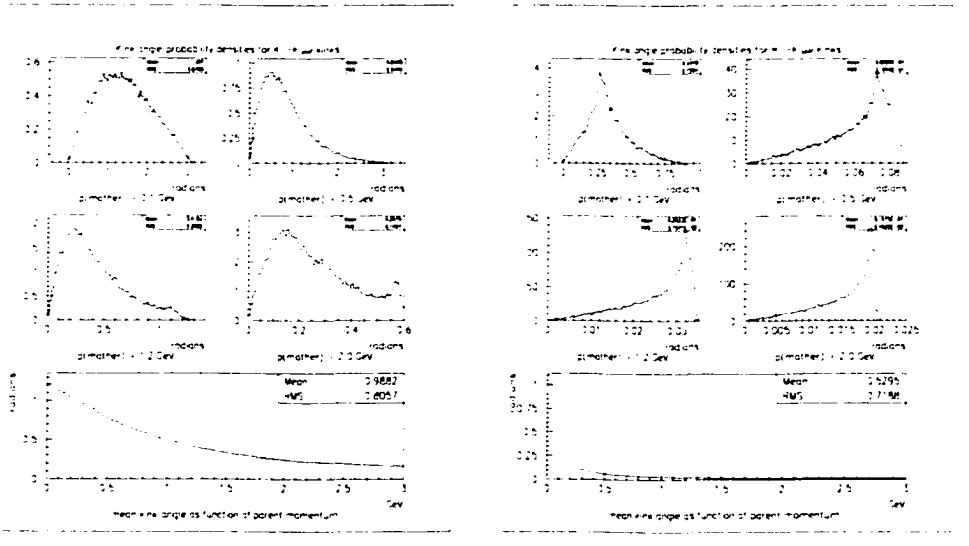


Figure C.1: Distributions of kink angles in the two most common kink decays: $K \rightarrow \mu\nu_\mu$ and $\pi \rightarrow \mu\nu_\mu$. The former has a large average kink angle (> 0.25 radians for most momenta found in BaBar) whereas the latter has a very small average kink angle (usually < 0.1 radians). Therefore the former is usually reconstructed as 2 separate tracks by track pattern recongition, whereas the latter tends to be reconstructed as a single track. This creates the need for 2 separate kink finding software strategies to be run.

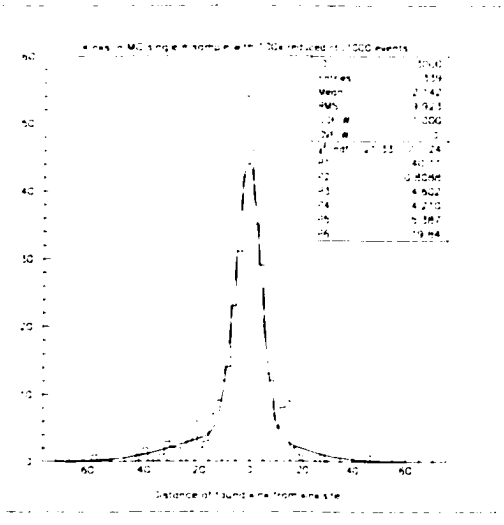


Figure C.2: Resolution of reconstructed flightlength at which kink occurred $foundkinkflightlength - truekinkflightlength$ in single pion MC (lifetime-reduced to increase the percentage of decays in the sample). The distribution is fitted to a double gaussian — the inner gaussian has a width of 4.6 cm.

Table C.2: **Purity**

	SP3 BBbar + digi bkg, 2000 evts (runs 200104 and 200105)	Single particle pion decays, 100x reduced ct, 0.75-2.0 GeV, 1000 evts
1) Given a found kink, what is the chance it is really just a track with no decays, no background hits, and no anomalous hits from other tracks (may be brem or scatter)	49/170 = 29%	4/293 = 1%
2) What is the chance it is not a decay and has no background hits, but has ≥1 anomalous hits from other tracks?	68/170 = 40%[6]	0/293 = 0%[7]
3) What is the chance that it is not a decay but contains ≥1 background hits?	12/170 = 7%	0/293 = 0%
4) Actual decays	41/170 = 24%	289/293 = 99%

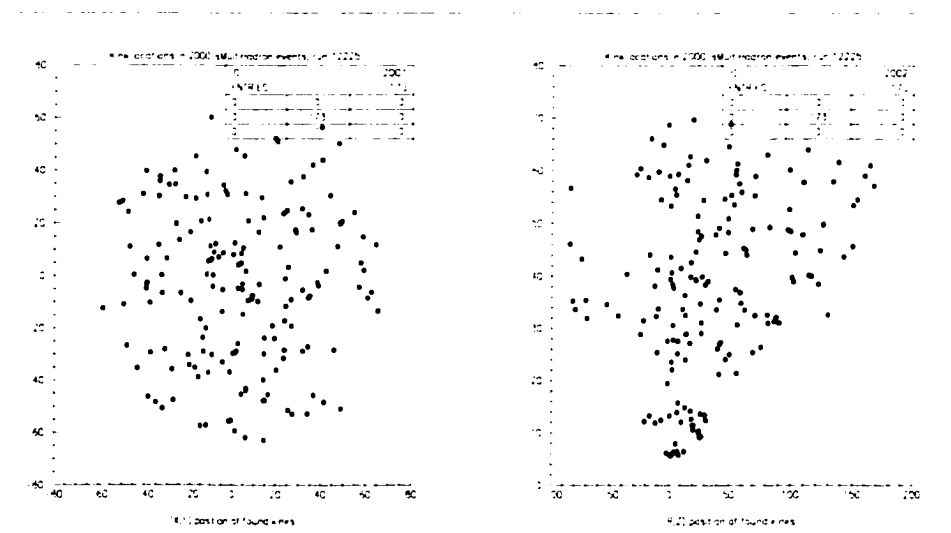


Figure C.3: Locations of reconstructed 1-track kinks in data (2000 isMulti-Hadron events)

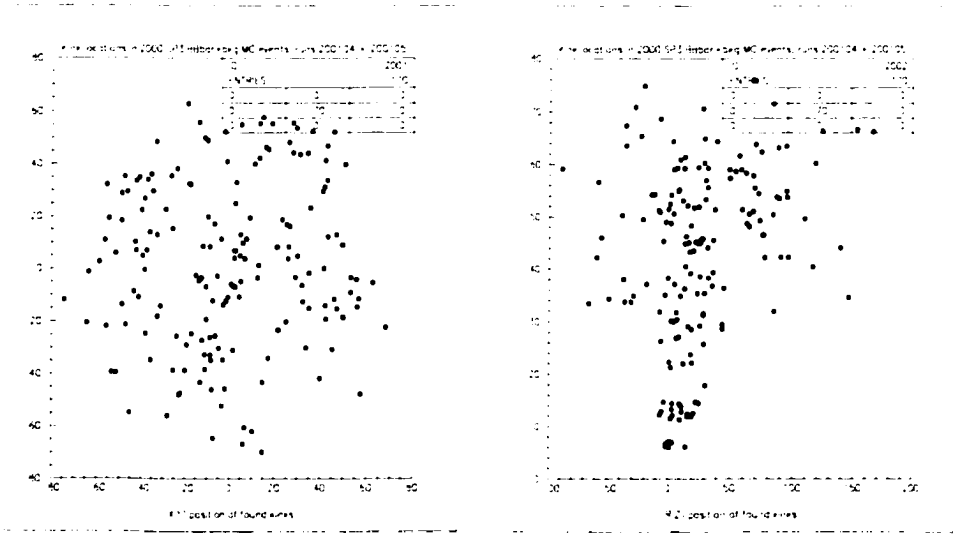


Figure C.4: Locations of reconstructed 1-track kinks in SP3 $B\bar{B}$ MC (2000 events)

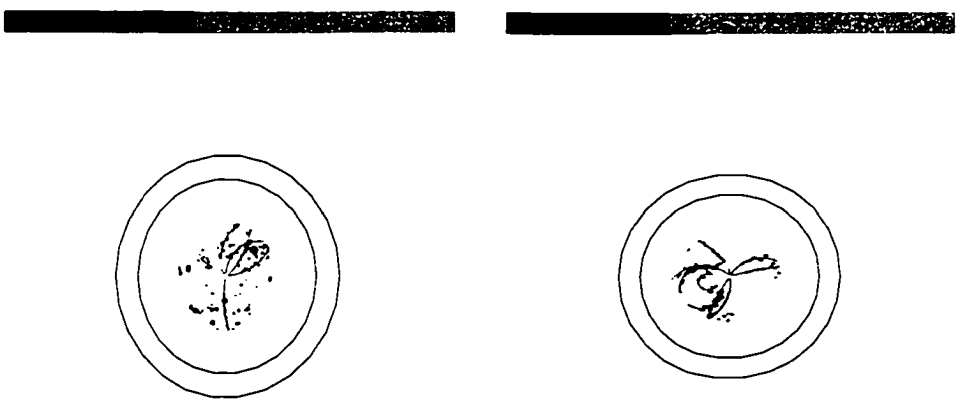


Figure C.5: Event displays of found 2-track kinks in real data and in SP2 $B\bar{B}$ MC

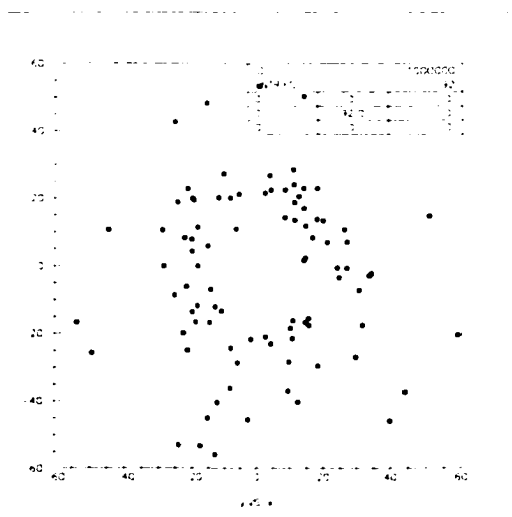


Figure C.6: Distribution of found kink vertices in real data (Padova reprocessing). There is a predominance near the support tube — this can be attributed to both support tube scatters and the fact that this is where the lever arm for creating two separate tracks is largest.

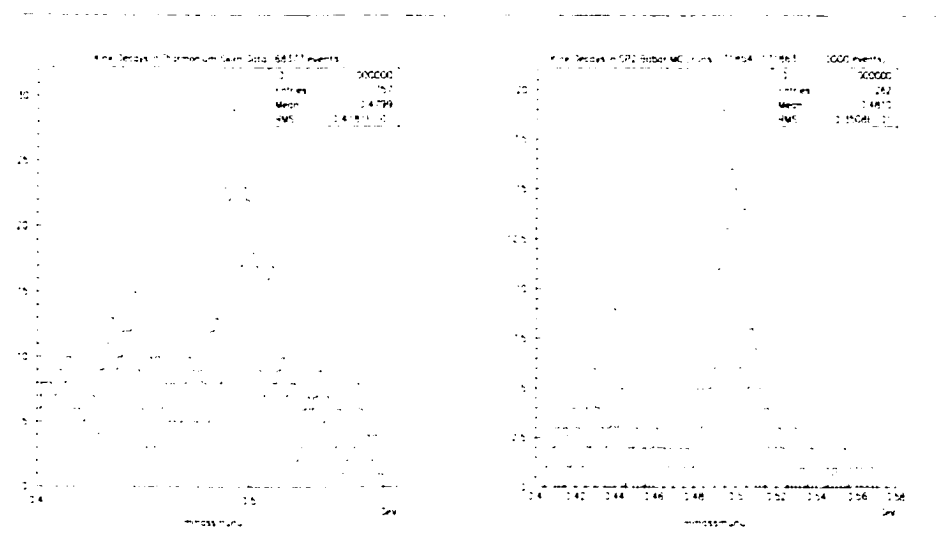


Figure C.7: Reconstructed 2-track kink invariant mass in the $\mu\nu_\mu$ hypothesis for real data (Charmonium skims) and for SP3 $B\bar{B}$ MC.

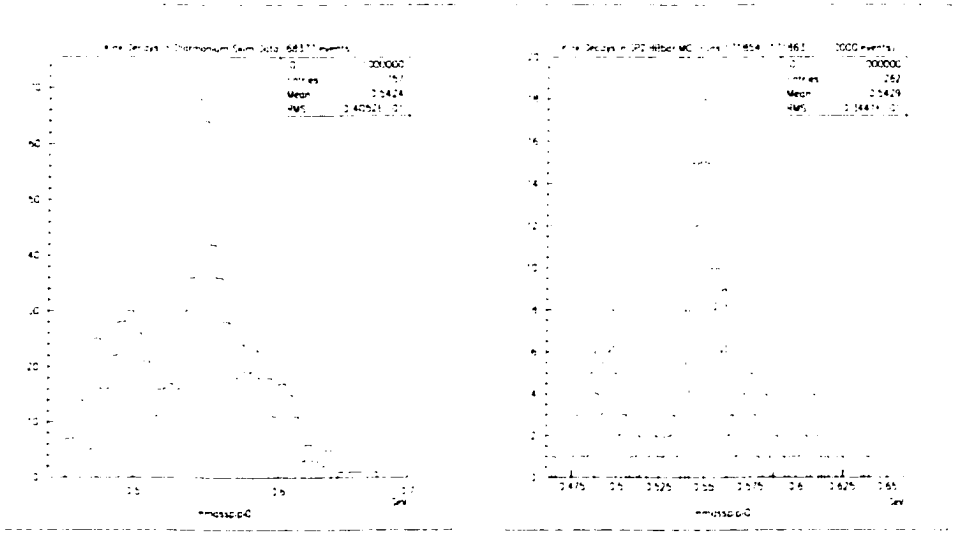


Figure C.8: Reconstructed 2-track kink invariant mass in the $\pi\pi^0$ hypothesis for real data (Charmonium skims) and for SP3 $B\bar{B}$ MC.

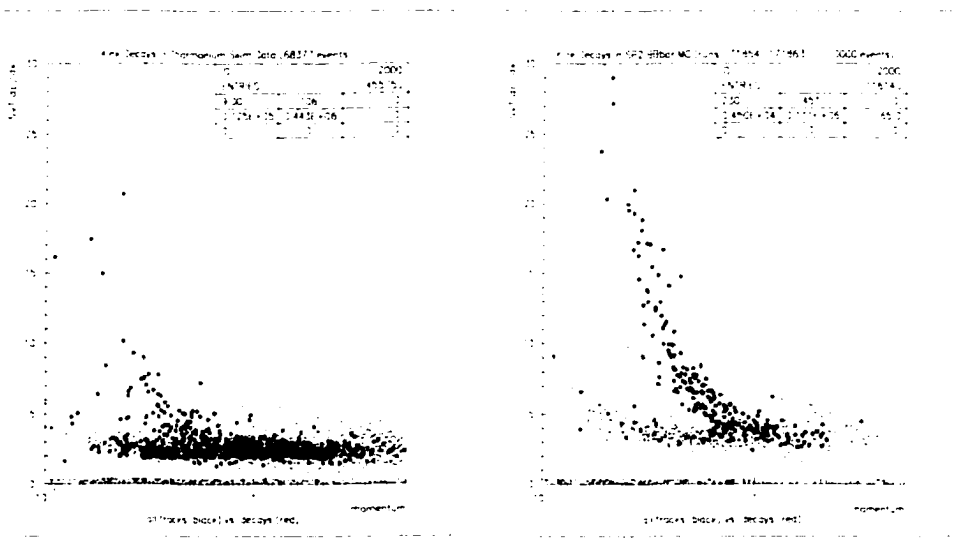


Figure C.9: Reconstructed SVT dE/dx for all charged tracks vs. parent tracks of found 2-track kinks in real data (Charmonium skims) and for SP3 $B\bar{B}$ MC. The kaons band dominates — moreso in MC than in data due to both better tracking resolution and better SVT dE/dx resolution in MC than in data (in release 8.6.3x).

```
class VtxKink : public BtaAbsVertex
{
public:

  //*********************************
  // Kink-specific Information
  //*********************************

// "kinkConsistency" = percentage chance that it really *is* a kink
    virtual double              kinkConsistency()       const  = 0;
// this will have to wait for a calibration:
    virtual double              kinkConsistencyErr()    const  = 0;
// the type of kink
    enum KinkType OneTrack=1, TwoTrack=2 ;
    virtual KinkType            kinkType()              const  = 0;

  //*********************************
  // Links to tracks
  //*********************************

// equivalent to inComingCand() -- here to save confusion
    virtual const BtaCandidate* motherCand()            const  = 0;

    virtual BtaCandidate*       daughtCand()            const  = 0;

// the uncharged daughter (neutrino in the case of pi -> mu numu and
// K -> mu numu or pi0 in the case of K -> pi pi0)
    virtual BtaCandidate*       neutDaughtCand()        const  = 0;

// "original" = before the vertexing refit occurs
    virtual BtaCandidate*       originalMotherCand()    const  = 0;

// this one is null if there originally was only a single track found
    virtual BtaCandidate*       originalDaughtCand()    const  = 0;

  //*******************************************
  // Position and fit-quality information
  //*******************************************

// The kink's ** POSITION ** in space & time lives here:
    virtual HepLorentzVector    v4()                    const  = 0;

    virtual HepSymMatrix        covariance()            const  = 0;
// Cross-covariance of x and p.
    virtual HepMatrix           xpCov()                 const  = 0;

    virtual double              chiSquared()            const  = 0;
    virtual int                 nDof()                  const  = 0;

// Flight length along (fitted) mother track & daughter track at which
// kink occurred
    virtual double              flightLenMother()       const  = 0;
    virtual double              flightLenDaught()       const  = 0;
```

Figure C.10: VtxKink interface for 1-track and 2-track kinks. VtxKink.hh (this file) is located in the package VtxBase and provides the physics interface for kinks in tracks.

Bibliography

[1] T. D. Lee and C. N. Yang, Phys. Rev. **104** (1956) 254.

[2] C. S. Wu *et al.*, Phys. Rev. **105** (1957) 1413.

[3] J.H. Christensen, J.W. Cronin, V.L. Fitch, and R. Turlay, Phys. Rev. Lett. **13** (1964) 138.

[4] M. Kobayashi and T. Maskawa, Prog. Th. Phys. **49** (1973) 652.

[5] V. Fitch, in Nobel Lectures (Physics) World Scientific, Singapore (1991), lecture for 1980 prize.

[6] I. I. Bigi *et al.*, in *CP Violation*, ed. C. Jarlskog, World Scientific, Singapore (1992).

[7] M. Gell-Mann and A. Pais, Phys. Rev. **97** (1955) 1387.

[8] M. Peskin and D. Schroeder, An Introduction to Quantum Field Theory Addison-Wesley, New York (1995).

[9] R. Aleksan, Quark Mixing and *CP* Violation, Proceedings of the NATO-ASI Virgin Islands Summer School (1995).

[10] S. L. Glashow, J. Iliopoulis, and L. Maiani, Phys. Rev. D **2** (1970) 1285.

[11] A. D. Sakharov, *ZhETF Pis. Red.* **5**, 32 (1967); *JETP Lett.* **5**, 24 (1967).

[12] V. A. Kuzmin, V. A. Rubakov and M. Shaposhnikov, *Phys. Lett. D* **155** (1985) 36.

[13] S. Adler, Phys. Rev. **117** (1969) 2426.

[14] J. Bell and R. Jackiw, Nuo. Cim. **A 60** (1969) 47.

[15] G. 't Hooft, Phys. Rev. D **14** (1976) 3432.

[16] M.B. Gavela *et al.*, Nucl. Phys. B **430** (1994) 345.

[17] The *BABAR* Collaboration, Nucl. Instr. and Methods A **479** (2002) 1.

[18] B. Aubert *et al.* (The *BABAR* Collaboration), hep-ex/0203008, accepted for publication in Phys. Rev. Lett. .

[19] Y. Grossman and M. Worah, Phys. Lett. B **395**, 241 (1997), R. Fleischer, Int. Jour. Mod. Phys. **A 12**, 2459 (1997)

[20] E. Lipeles, *et al.*, Phys. Rev. D **62** (2000).

[21] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).

[22] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. **B185**, 218 (1987).

[23] P. H. Harrison and H. R. Quinn, eds. "The *BABAR* Physics Book", SLAC-R-504 (1998).

[24] Particle Data Group. D. E. Groom *et al.*, "Review of Particle Physics", Eur. Phys. Jour. C **15**, 1 (2000).

[25] C. Hearty, "Measurement of the Number of $\Upsilon(4S)$ Mesons Produced in Run 1 (B Counting)", *BABAR* Analysis Document 134, (2001).

[26] BReco AWG, "Exclusive Reconstruction of Hadronic B Decays to Open Charm", *BABAR* Analysis Document 150, (2001).

[27] `http://www.slac.stanford.edu/BFROOT/www/Physics/BaBarData/GoodRuns/dataSets.html`

[28] `http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/BetaTools/MicroKilling.htm`

[29] private communication from Prof. Aaron Roodman

[30] E. Varnes "Measurement of the Tracking Efficiency Using 3+1 tau Events", *BABAR* Analysis Document 87, (2000).
`http://babar-hn.slac.stanford.edu:5090/HyperNews/get/physAnal/1037.html`
`http://babar-hn.slac.stanford.edu:5090/HyperNews/get/physAnal/1024.html`
`http://www.slac.stanford.edu/BFROOT/www/Physics/TrackEfficTaskForce/Recipe/TrackingEfficiencies.html`

[31] `http://babar-hn.slac.stanford.edu:5090/HyperNews/get/physAnal/1042.html`

[32] `http://babar-hn.slac.stanford.edu:5090/HyperNews/get/physAnal/1028.html`

[33] D. Kirkby *et al.* "A User's Guide to the RooFitTools Fitting Package", *BABAR* Analysis Document 12 (2000).

[34] N. Gehrels, ApJ, **303**, 336 (1986).

[35] K. Abe *et al.*, Belle-Conf-0104 (2001).

[36] B.A.D. 351 Measurement of the Branching Ratio for the Decays $B \to D^{(*)}\overline{D}^{(*)}$

[37] PDG 2000.

[38] B.A.D. 223 A Study Of The Transversity Amplitudes In The $B^0 \to D^{*+}D^{*-}$ Decays.

[39] B.A.D. 205 Supporting document for summer 2001 sin2beta analysis.

[40] B.A.D. 356 Supporting document for the Moriond, 2002 sin2beta analysis.

[41] B. Aubert *et al.* (The *BABAR* Collaboration), hep-ex/0201020, accepted for publication in Phys. Rev. D .

[42] R. Aleksan *et al.*, Nucl. Phys. B **361** (1991) 141.

[43] R. Aleksan *et al.*, Phys. Lett. B **317** (1993) 173.

[44] http://babar-hn.slac.stanford.edu:5090/HyperNews/get/sin2beta/198.html (R. N. Cahn).

[45] J. D. Richman, Heavy-Quark Physics and *CP* violation (Les Houches lectures) (1998).

[46] B.A.D. 125 Measurement of B^0 mixing using fully reconstructed hadronic B^0 decays

[47] Defined as $A \sim \sqrt{1-(m_{ES}/m_0)^2} \times exp(\kappa(1-(m_{ES}/m_0)^2)$ for $m_{ES} < m_0$. The default value of m_0 is 5.291 GeV.

[48] B.A.D. 56 (contact D. Wright) Analysis of $B^0 \rightarrow J/\psi K_L^0$

[49] B.A.D. 119 (contact J. Beringer) B Tagging in BaBar: Status for the sin2beta Journal Publication

[50] B.A.D. 102 (contact F.Martinez Vidal) The BaBar Vertexing

[51] B.A.D. 130 (contact F.Martinez Vidal) Performances and control samples of the *BABAR* vertexing

[52] B.A.D. 99 (contact O. Long) Measurement of sin2β with the decay mode $B^0 \rightarrow J/\psi K_L^0$

[53] B.A.D. 151 (S. Plaszczynski, L. Roos, M.H. Schune) Measurement of the mistag fractions with counting methods

[54] J. Beringer, *Changes in tagging and deltaT for events used for Osaka results*, sin2beta HN 192, January 19, 2001

[55] J. Beringer, *Comparison with and Unblinding of the Osaka Subsample*, sin2beta meeting, January 22, 2001

[56] J. Beringer et al., *Cut Based Tagging*, BaBar Analysis Document 118, Version 1, (2000)

[57] F. Porter, Interval Estimation Using the Likelihood Function.

[58] http://babar-hn.slac.stanford.edu:5090/HyperNews/get/physAnal/1023.html (A. Gritsan).

[59] B.A.D. 153 (contact H. Lacker) Constraining the CKM Matrix

[60] B.A.D. 92 (contact G. Eigen) Inputs for the Unitarity Triangle Fits

[61] BaBar CKM Working Group,
http://babar-hn.slac.stanford.edu:5090/HyperNews/get/ckmfit/5.html

[62] D.N. Brown, E. Charles, and D. Roberts, "The BaBar Track Fitting Algorithm", submitted to the CHEP 2000 conference (Berkeley, Jan. 2000)

[63] F. Frühwirth, "Applications of Filter Method to the Reconstruction of Tracks and Vertices in Events of Experimental High Energy Physics", HEPHY-PUB 516/88 (Vienna, Dec. 1988)

[64] R. Cousins *et al.*, "Kalman Filter Track Fits and Track Breakpoint Analysis", submitted to Nucl. Instr. Meth. (Geneva, Dec. 1999)

[65] G. Lynch *et al.*, "BaBar Drift Chamber Tracking Conventions", BaBar Note 488, TNDC Note 99-95 (Stanford, Apr. 1999)