Study of Rare B-Meson Decays Related to the CP Observable sin(2beta+gamma) at the BABAR Experiment

Toyoko Jennifer Orimoto

SLAC-R-845

Prepared for the Department of Energy
under contract number DE-AC02-76SF00515

Printed in the United States of America. Available from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, VA 22161.

Study of Rare B-Meson Decays
Related to the CP Observable $\sin(2\beta+\gamma)$
at the *BABAR* Experiment

by

Toyoko Jennifer Orimoto

B.A. (University of California, Berkeley) 2000
M.A. (University of California, Berkeley) 2002

A dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Physics

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Yury Kolomensky, Chair
Professor Marco Battaglia
Professor Chung-Pei Ma

Fall 2006

The dissertation of Toyoko Jennifer Orimoto is approved.

Chair	Date
	Date
	Date

University of California, Berkeley

Fall 2006

Study of Rare B-Meson Decays
Related to the CP Observable $\sin(2\beta + \gamma)$
at the *BABAR* Experiment

by

Toyoko Jennifer Orimoto

Abstract

Study of Rare B-Meson Decays
Related to the CP Observable $\sin(2\beta+\gamma)$
at the *BABAR* Experiment

by

Toyoko Jennifer Orimoto

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Yury Kolomensky, Chair

This study reports the observation of the decays $B^0 \to D_s^{(*)+}\pi^-$ and $B^0 \to D_s^{(*)-}K^+$ in a sample of 230×10^6 $\Upsilon(4S) \to B\overline{B}$ events collected with the *BABAR* detector at the PEP-II asymmetric-energy e^+e^- storage ring, located at the Stanford Linear Accelerator Center. The branching fractions $\mathcal{B}(B^0 \to D_s^+\pi^-) = (1.3 \pm 0.3\ (\text{stat}) \pm 0.2\ (\text{syst})) \times 10^{-5}$, $\mathcal{B}(B^0 \to D_s^-K^+) = (2.5 \pm 0.4\ (\text{stat}) \pm 0.4\ (\text{syst})) \times 10^{-5}$, $\mathcal{B}(B^0 \to D_s^{*+}\pi^-) = (2.8 \pm 0.6\ (\text{stat}) \pm 0.5\ (\text{syst})) \times 10^{-5}$, and $\mathcal{B}(B^0 \to D_s^{*-}K^+) = (2.0 \pm 0.5\ (\text{stat}) \pm 0.4\ (\text{syst})) \times 10^{-5}$ are measured. The significance of the measurements to differ from zero are 5, 9, 6, and 5 standard deviations, respectively. This is a first observation of the decays $B^0 \to D_s^{*+}\pi^-$ and $B^0 \to D_s^{*-}K^+$. These results may potentially be useful in determining the CP asymmetry parameter $\sin(2\beta+\gamma)$ in the decays $B^0 \to D^{(*)+}\pi^-$.

Professor Yury Kolomensky
Dissertation Committee Chair

Contents

Acknowledgements

I would like to thank the following people:

Yury Kolomensky, my thesis advisor, for all of his support and encouragement. His incredible understanding of physics and enduring enthusiasm for scientific research has never ceased to amaze and inspire me. It was his belief in me that really motivated me to finish, and I will never be able to thank him enough for all that he has taught me and everything he has done for me.

The LBNL *BABAR* group, especially Dave Brown and Bob Cahn, for all their support. Mandeep Gill, Gena Kukartsev, and Kerstin Tackmann for being great office-mates and friends. And Aritoki Suzuki and Daniel Troost for their invaluable help with the analysis.

The many other *BABAR*-ians who have helped me along the way, in particular: Cecilia Voena, Anders Borgland, and Paul Jackson... and all the other people on *BABAR* from around the world, whom I've had the pleasure to meet and work with.

The members of the NanoBPM collaboration. Being able to work with such an amazing group of people made all the deadlines on *BABAR* bearable!

The members of my qualifying exam committee and thesis committee: Marco Battaglia, Alex Filippenko, Chung-Pei Ma, and Hitoshi Murayama.

Elisabeth Cutler, who was the first to expose me to the world of physics.

All of my friends and family for their continued support and friendship, especially: My mom and two brothers, Taka and Kazu. Dorothy Kim, for being the awesomest best friend and freakish wondertwin a person could ever imagine having. Athena Gordon, for being a great friend and mentor all these years, and also to her family. And especially Brock Tweedie, whose support, love and devotion made the last few years all worthwhile.

And all the friends and collaborators who I haven't named individually, but who are in my thoughts.

Chapter 1

Introduction

1.1 The Standard Model of Particle Physics

Particle physics attempts to answer the question, "what are the most fundamental constituents of matter and how do they interact?" The particles that we encounter day to day, such as electrons, photons, and protons, and the forces we experience, such as electromagnetism, only make up a small part of a much larger picture. The Standard Model of particle physics is the theory of fundamental particles, and how they interact through the electromagnetic, weak, and strong forces. The fundamental particles consist of three generations of fermions, a set of force-mediating vector bosons, and a scalar boson.

The fundamental fermions, which make up all matter, consist of three generations of quarks and leptons that occur in pairs. All everyday matter is made up of first generation particles. The second and third generation particles, with the possible exception of the neutrinos, live for only a brief time, and are typically only observed by particle detectors. Quarks possess electric charge as well as "color" charge, and interact via the strong, electromagnetic, and weak forces. The leptons, on the other hand, do not possess "color" and only interact through the electromagnetic and weak forces. A striking feature of quark

interactions is that the strong force between quarks and anti-quarks grows stronger with distance, such that they are confined to color-neutral combinations.

In the Standard Model, photons are the mediators of the electromagnetic force, the force which acts between electrically charged particles. Gluons mediate the strong force, which binds nucleons together. The massive $W^{\pm}$ and Z^0 mediate the weak force, which governs such phenomena as radioactive decays and nuclear fusion. Thus far, gravity, the weakest of all forces, has yet to be merged with the Standard Model in a coherent model of elementary particles.

In addition, the Standard Model contains a scalar boson, the Higgs, which arises from the spontaneous symmetry breaking of the electroweak gauge group,the root of inertial mass. Consequently, the interactions of the Higgs with other particles are mass dependent. The Higgs boson is the last remaining piece of the Standard Model that has yet to be discovered experimentally. Direct searches at the Large Electron Positron collider at CERN have placed a lower limit on the Standard Model Higgs mass at $m_h > 114.4\,\mathrm{GeV}/c^2$ at the 90% confidence level.

The fundamental particles are described by their masses and quantum numbers, which include spin, electric charge, "color", and "flavor". The characteristics of the fundamental particles—quarks, leptons, the gauge-mediators, and the Higgs boson—are shown in Table 1.1. In addition to the particles listed here, the anti-particle partner for each also exists, though in some cases particle and anti-particle are the same.

The Standard Model is a quantum field theory, consistent with both quantum mechanics and special relativity. The Lagrangian is invariant under a set of gauge transformations associated with each of the vector bosons. For this reason, the force-mediating particles are also called gauge bosons. The gauge transformations can be described mathematically using unitary groups, and within the Standard Model the gauge group is $SU(3)_C \times SU(2)_L \times U(1)_Y$. The strong interactions are described by $SU(3)_C$, where C stands for color charge. The electromagnetic and weak interactions are described together as $SU(2)_L \times U(1)_Y$, where

Table 1.1. The fundamental particles of the Standard Model—fermions (quarks, leptons) and bosons (gauge bosons, Higgs)—and some of their properties. Charge here corresponds to electric charge. Because of neutrino mixing, masses listed for neutrinos are for the corresponding mass eigenstate. Upper limits on the neutrino masses are from direct searches, while neutrino oscillation experiments have measured the mass differences. The error on the electron and muon masses are both on the order of 10^{-9} and 10^{-6} MeV/c^2, respectively. All masses are quoted from the Particle Data Group (PDG) [12].

Particle	Symbol	Charge (e)	Mass (MeV/c^2)
Quarks, spin = 1/2			
up	u	+2/3	1.5 - 4.0
down	d	-1/3	4 - 8
charm	c	+2/3	1.15 - 1.35 $\cdot 10^{-3}$
strange	s	-1/3	80 - 130
bottom	b	+2/3	4.1 - 4.4 $\cdot 10^{-3}$
top	t	-1/3	$178.1^{+10.4}_{-8.3} \cdot 10^3$
Leptons, spin = 1/2			
electron	e	-1	0.511
electron neutrino	ν_e	0	$< 3 \cdot 10^{-6}$
muon	μ	-1	105.66
muon neutrino	ν_e	0	< 0.19 (90% C.L.)
tau	τ	-1	$1776.99^{+0.29}_{-0.26}$
tau neutrino	ν_τ	0	< 18.2 (95% C.L.)
Gauge Bosons, spin = 1			
photon	γ	0	0
$W^\pm$	$W^\pm$	± 1	$80.425 \pm 0.038 \cdot 10^3$
Z^0	Z^0	0	$91.1876 \pm 0.0021 \cdot 10^3$
gluon	g	0	0
Higgs Boson, spin = 0			
Higgs	h^0	0	$> 114.4 \cdot 10^3$

L indicates that the $SU(2)$ interaction only applies to left-handed fermions and Y stands for hypercharge. The latter symmetry is spontaneously broken in nature, causing all of the fermions and a subset of the $SU(2) \times U(1)$ gauge bosons to acquire masses. The $W^\pm$ and Z^0 are the resulting massive gauge boson states, the photon is the remaining massless state, and the Higgs boson is a byproduct of the symmetry-breaking process.

As mentioned earlier, quarks are confined into colorless bound states. These states are collectively called hadrons. Mesons are bosonic hadrons made of quark-antiquark pairs,

while baryons are fermionic hadrons made of three quarks or three antiquarks. Protons and neutrons are baryons made up of u-u-d quarks and d-d-u quarks, respectively. Hadrons are typically unstable and can decay through the fundamental interactions into other particles, and they can typically decay through a variety of different channels at different decay rates. The decay rate of a particle into a certain decay channel, normalized by the sum of the rates of all channels, is called the branching ratio ($\mathcal{B}$), or branching fraction.

The successes of the Standard Model have been numerous. The Standard Model predicted the existence of the $W^{\pm}$ and Z^0 gauge bosons, as well as the top quark, before they were observed experimentally. Many experiments have tested and successfully confirmed the predictions of the Standard Model. Despite these accomplishments, the theory has some conspicuous failures. For instance, there are a large number of free parameters which cannot be calculated and must be determined experimentally. Also, as mentioned earlier, gravity is not included in the Standard Model, though there have been many attempts to unify all four forces. In addition, the Standard Model cannot account for the large baryon asymmetry observed in the universe. Many experiments, such as the *BABAR* Experiment, continue to test predictions of the Standard Model, in search of evidence of new physics.

1.1.1 C, P, and T

For any field theory Lagrangian, charge conjugation, parity, and time reversal are three discrete operations that may define a symmetry of the theory. Charge conjugation (C) interchanges particles and antiparticles. Parity (P) and time reversal (T) are space-time operations. Parity reverses the handedness of space by sending $(t, \mathbf{x}) \rightarrow (t, -\mathbf{x})$. Time reversal similarly changes $(t, \mathbf{x}) \rightarrow (-t, \mathbf{x})$. The combined operation CP interchanges particles to antiparticles and reverses momentum and helicity. CPT is an exact symmetry for any local Lagrangian field theory, and thus far all observations confirm exact CPT symmetry for the Standard Model.

It was taken for granted that the laws of physics are symmetric under parity until

1956, when T.D. Lee and C.N. Yang discovered that there was no experimental evidence at the time for parity invariance in weak interactions [1]. The first observation of a discrete symmetry violation in weak interactions was made by C.S. Wu in 1957, when she discovered the existence of parity violation in the beta decay of Cobalt-60 [2]. The electromagnetic and strong interactions are C, P, and T symmetric, but the weak interactions violate C and P separately. Moreover, the weak interactions preserve CP and T to very good approximation, but CP violation has been observed in the interactions of neutral K and B mesons. CP violation was first observed in rare kaon decays in 1964 [3] and has subsequently been seen in neutral B meson interactions by B factory experiments such as the *BABAR* Experiment at the PEP-II collider [4].

The CP transformation properties of the fields, Lorentz invariance, and the hermiticity of the Lagrangian require that CP transforms any combination of fields and derivatives in the Lagrangian to its hermitian conjugate. If there are any complex coefficients in front of these terms, then the coefficients of CP-related terms are complex conjugates of each other. In this case, CP may not be a preserved symmetry. Of course, not all phases that appear in the Lagrangian must be physical. Any complex field of the Lagrangian may be transformed by a phase rotation, which does not change any physical quantities but redefines the phases of the Lagrangian. CP violation arises from nonzero phases which remain after all such phase rotations have been performed to remove any unphysical phases. In the Standard Model, CP violation arises from a nonzero phase in the quark mixing matrix, which is described in the following section.

1.1.2 The Cabibbo-Kobayashi-Maskawa Matrix

In the Standard Model, quarks can change flavor through interactions with the $W^{\pm}$ bosons. For instance, interactions in which an up-type quark changes to a down-type quark with the emission of a W ($u \to d + W^{+}$) can occur. Moreover, cross-generational couplings such as $u \to s + W^{+}$ can also occur.

In 1963 N. Cabibbo suggested that there was a factor of $\cos\theta_C$ for vertices like $u \to d + W^+$ and a factor of $\sin\theta_C$ for cross-generational vertices like $u \to s + W^+$, where θ_C is a small angle [5]. In matrix form, this can be summarized as:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}, \tag{1.1}$$

where the d' and s' are the weak eigenstates which couple to the $W^{\pm}$. Such "rotation" of quark fields successfully explained the ratio of kaon and pion leptonic decay rates.

M. Kobayashi and T. Maskawa expanded Cabibbo's mixing matrix to include a third generation, before the charm quark was even discovered [6]. The Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix, as seen in Eq.(1.3), embodies the fact that the quark mass eigenstates and the quark flavor eigenstates are not the same, and the mixing matrix describes how these states are related. Eq. (1.2) shows how the weak eigenstates (on the left hand side) are mixtures of the mass eigenstates (on the right hand side), with the mixing governed by the CKM matrix.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.2}$$

The Standard Model lends no theoretical predictions for the CKM matrix, and it therefore must be measured experimentally. The 90% confidence limits on the magnitudes of the CKM matrix elements, determined from level tree-level constraints together with unitarity and the assumption of only three quark generations, are shown below [12]:

$$V_{CKM} = \begin{pmatrix} 0.9739 \text{ to } 0.9751 & 0.221 \text{ to } 0.227 & 0.0029 \text{ to } 0.0045 \\ 0.221 \text{ to } 0.227 & 0.9730 \text{ to } 0.9744 & 0.039 \text{ to } 0.044 \\ 0.0048 \text{ to } 0.014 & 0.037 \text{ to } 0.043 & 0.9990 \text{ to } 0.9992 \end{pmatrix} \tag{1.3}$$

Different processes are described as relatively "CKM-suppressed" or "CKM-favored" depending on what element of the CKM matrix accompanies the process. For instance, we

can see from the above matrix that a $c \to d$ transition, which would have a factor of V_{cd}, is CKM-suppressed with respect to a $u \to d$ transition, which is governed by a factor of V_{ud}.

1.2 *CP* Violation in the Standard Model

The standard parametrization of the CKM matrix is:

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{1.4}$$

where $c_{ij} \equiv \cos\theta_{ij}$, $s_{ij} \equiv \sin\theta_{ij}$, and δ is the *CP* violating phase.

The condition of unitarity of the CKM matrix gives rise to three relationships that are relevant for understanding *CP* violation:

$$V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0 \tag{1.5}$$

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \tag{1.6}$$

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{1.7}$$

Each of these equations can be represented geometrically as a unitarity triangle in the complex plane. *CP* violation manifests itself as a nonzero area of these unitarity triangles [7]. For the first two triangles, one of the sides is much smaller than the other two, resulting in a highly flattened triangle. This suggests that there are small *CP* asymmetries in the K system, which is associated with the first triangle, and in B_s decays, which are associated with the second triangle. The openness of the last triangle, referred to as "the Unitarity Triangle", however, suggests that there are large *CP* asymmetries in B_d meson decays. From henceforth, "B meson" will refer to B_d mesons, unless otherwise specified.

The Unitarity Triangle is depicted in Figure 1.1a. It can be rescaled and rotated, as shown in Figure 1.1b, after choosing a phase convention such that $V_{cd}V_{cb}^*$ is real, dividing all the sides by $|V_{cd}V_{cb}^*|$. This aligns one side with the real axis and makes the length of this

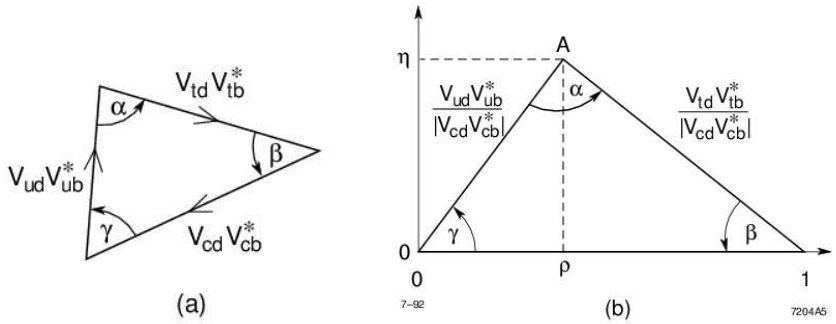


Figure 1.1. The Unitarity Triangle (a) before scaling and (b) after scaling and rotation.

side 1. Then, two of the vertices are fixed at $(0,0)$ and $(0,1)$, and the apex of the triangle is denoted as (ρ,η). The three angles of the Unitarity Triangle, α, β, and γ are physical quantities which can be measured by studying *CP* asymmetries in B decays. They can be written in terms of the CKM matrix elements as:

$$\alpha \equiv \arg\left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right],\ \beta \equiv \arg\left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right],\ \gamma \equiv \arg\left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right]. \tag{1.8}$$

Furthermore, if the Standard Model poses the correct explanation for *CP* violation, then the angles as extracted from the *CP* observables will obey the relationship $\gamma \equiv \pi - \alpha - \beta$.

The Unitarity Triangle can also be expressed in terms of the Wolfenstein parameters [8]:

$$V_{CKM} = \begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \tag{1.9}$$

where $\lambda = |V_{us}| = 0.22$ and η is the imaginary component which gives rise to *CP* violation. This parametrization is particularly attractive since λ is a relatively small quantity, serving as an expansion parameter. The relationship between the former parametrization in Eq. (1.4) and this one is:

$$s_{12} \equiv \lambda,\ s_{23} \equiv A\lambda^2,\ s_{13}e^{-i\delta} \equiv A\lambda^3(\rho - i\eta), \tag{1.10}$$

to $\mathcal{O}(\lambda^4)$. From henceforth, λ from the Wolfenstein parametrization will be referred to as $\lambda_{Cabibbo}$ to prevent confusion with notation in other sections.

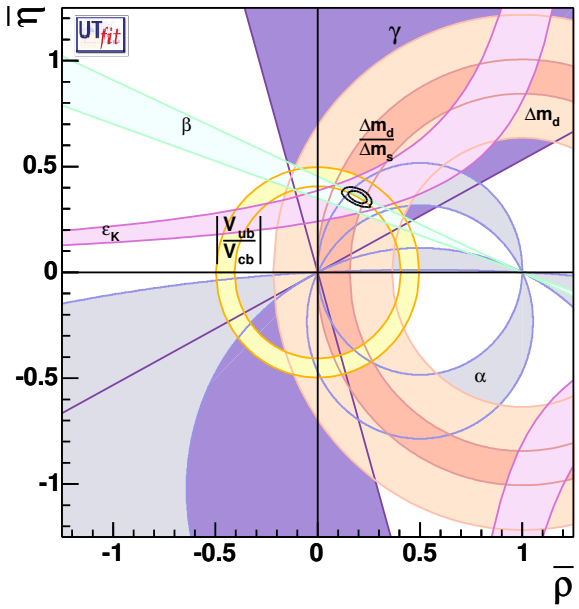


Figure 1.2. Allowed regions of for $(\rho - \eta)$, as determined by the UT_{fit} group. Closed contours for 68% and 95% probability are shown. Full lines correspond to 95% probability regions for the constraints, given by measurements of $|V_{ub}|/|V_{cb}|, \epsilon_K, \Delta m_d, \Delta m_d/\Delta_m s, \sin 2\beta, \cos 2\beta, \alpha, \gamma, \beta,$ and $\sin(2\beta + \gamma)$ [9].

Although it is sufficient to measure one of the angles to demonstrate the existence of *CP* violation, the Unitarity Triangle must be over-constrained to demonstrate that the CKM mechanism is the correct explanation of this phenomenon. Figure 1.2 shows the constraints on the Unitarity Triangle in the (ρ - η) plane, from a variety of measurements [9].

1.3 *CP* Violation in the $B^0 - \bar{B}^0$ System

Neutral B_d mesons are made from one b-type and one d-type quark or antiquark. For these mesons, there are two different types of states: the flavor eigenstates, which have definite quark content and are related to particle production and decay processes, and the eigenstates of the Hamiltonian, which have definite mass and lifetime and which propagate through space in a definite way. If *CP* were conserved, then the mass eigenstates would also be *CP* eigenstates. Since *CP* is not conserved in weak interactions, the mass eigenstates

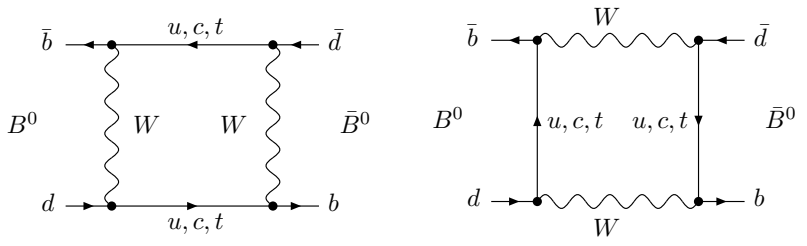


Figure 1.3. Box diagrams for $B^0 - \bar{B}^0$ mixing in the Standard Model.

are not necessarily the same as the *CP* eigenstates, and the flavor eigenstates can mix with one another as they propagate through space.

The flavor eigenstates of the neutral B_d system are $B^0 = \bar{b}d$ and $\overline{B}^0 = \bar{d}b$. A linear combination of neutral B flavor eigenstates can be written as:

$$a\langle B^0| + b\langle \overline{B}^0|. \tag{1.11}$$

These states, B^0 and $\overline{B}^0$, are related by the *CP* transformation:

$$CP\langle B^0| = e^{2i\xi_B}\langle \overline{B}^0|,\ CP\langle \overline{B}^0| = e^{-2i\xi_B}\langle B^0|. \tag{1.12}$$

The phase ξ_B is an arbitrary *CP* phase, and the freedom in defining it comes from the fact that the strong interactions are flavor conserving.

The evolution of such a state is governed by the time-dependent Schrödinger equation:

$$i\frac{d}{dt}\begin{pmatrix} a \\ b \end{pmatrix} = H\begin{pmatrix} a \\ b \end{pmatrix} \equiv (M - \frac{i}{2}\Gamma)\begin{pmatrix} a \\ b \end{pmatrix}, \tag{1.13}$$

where H is the Hamiltonian of the system, and M and Γ are 2×2 Hermitian matrices. CPT invariance dictates that $H_{11} = H_{22}$. M_{12} and Γ_{12} are the dispersive and absorptive parts of the transition amplitude from B^0 to $\overline{B}^0$ and hence, particularly important for *CP* violation. Within the Standard Model, these elements come from box diagrams with two W exchanges. The Feynman diagrams for these box diagrams are shown in Figure 1.3.

The eigenstates of the Hamiltonian, which are mass states distinguished as the light B_L and heavy B_H states, can be written in terms of the flavor eigenstates:

$$\langle B_L| = p\langle B^0| + q\langle \overline{B}^0|, \tag{1.14}$$

$$\langle B_H| = p\langle B^0| - q\langle \overline{B}^0|, \tag{1.15}$$

such that the normalization condition

$$|q|^2 + |p|^2 = 1 \tag{1.16}$$

is satisfied.

The mass difference and the width difference between B_H and B_L are defined as:

$$\Delta m_B \equiv M_H - M_L,\ \Delta\Gamma_B \equiv \Gamma_H - \Gamma_L. \tag{1.17}$$

By diagonalizing the Hamiltonian, we can relate Δm_B and $\Delta\Gamma_B$ to the matrix elements M_{12} and Γ_{12}.

$$(\Delta m_B)^2 - \frac{1}{4}(\Delta\Gamma_B)^2 = 4(|M_{12}|^2 - \frac{1}{4}|\Gamma_{12}|^2), \tag{1.18}$$

$$\Delta m_B \Delta\Gamma_B = 4\mathcal{R}e(M_{12}\Gamma_{12}^*). \tag{1.19}$$

The ratio q/p can then be written as:

$$\frac{q}{p} = \frac{\Delta m_B - \frac{i}{2}\Delta\Gamma_B}{2(M_{12} - \frac{i}{2}\Gamma_{12})} = \frac{2(M_{12}^* - \frac{i}{2}\Gamma_{12}^*)}{\Delta m_B - \frac{I}{2}\Delta\Gamma_B}. \tag{1.20}$$

The difference in widths $\Delta\Gamma_B$, though not yet measured, is expected to be negligibly small:

$$\Delta\Gamma_B/\Gamma_B = \mathcal{O}(10^{-2}). \tag{1.21}$$

This difference in width arises from decay modes that are common to the B^0 and $\overline{B}^0$, and the branching ratios for such modes are on the order of 10^{-3}. Since these channels contribute with differing signs, we assume that the sum of these contributions do not exceed the individual values, and thus $\Delta\Gamma_B \ll \Gamma_B$.

On the other hand, a number of measurements for Δm_B exist, particularly from the B factories [12]:

$$x \equiv \Delta m_B/\Gamma_B = 0.774 \pm 0.009. \tag{1.22}$$

From this quantity we can see that

$$\Delta\Gamma_B \ll \Delta m_B. \tag{1.23}$$

Using Eq. (1.21) and (1.23), we can simplify Eq. (1.18), (1.19), and (1.56):

$$\Delta m_B = 2|M_{12}|,\ \Delta\Gamma_B = 2\mathcal{R}e(M_{12}\Gamma_{12}^*)/|M_{12}|, \tag{1.24}$$

$$q/p = -|M_{12}|/M_{12}. \tag{1.25}$$

Any B state can be written in terms of the mass eigenstates B_H and B_L. The amplitudes of such an admixture evolve in time according to the Schrödinger equation:

$$a_H(t) = a_H(0)e^{-iM_Ht}e^{-\frac{1}{2}\Gamma_Ht},\ a_L(t) = a_L(0)e^{-iM_Lt}e^{-\frac{1}{2}\Gamma_Lt}. \tag{1.26}$$

A state that is initially a pure B^0 state at $t = 0$ is indicated by $\langle B^0_{phys}|$ and has $a_L(0) = a_H(0) = 1/(2p)$. A state that is initially a pure $\overline{B}^0$ state, $\langle \bar{B}^0_{phys}|$, has $a_L(0) = -a_H(0) = 1/(2q)$. The time evolution of these initially pure states can be written in terms of the flavor eigenstates:

$$\langle B^0_{phys}| = g_+(t)\langle B^0| + (q/p)g_-(t)\langle \overline{B}^0|, \tag{1.27}$$

$$\langle \bar{B}^0_{phys}| = (p/q)g_-(t)\langle B^0| + g_+(t)\langle \overline{B}^0|, \tag{1.28}$$

where

$$g_+(t) = e^{-iMt}e^{-i\Gamma t}\cos(\Delta m_B t/2), \tag{1.29}$$

$$g_-(t) = e^{-iMt}e^{-i\Gamma t}i\sin(\Delta m_B t/2), \tag{1.30}$$

and $M = \frac{1}{2}(M_H + M_L)$ and $\Gamma = \frac{1}{2}(\Gamma_H + \Gamma_L)$. We also define $\tau = 1/\Gamma$ as the average lifetime of the neutral B mesons.

1.3.1 Time Evolution at the $\Upsilon(4S)$

At PEP-II, electrons and positrons collide to produce $\Upsilon(4S)$ mesons, which can subsequently decay into pairs of $B\bar{B}$ mesons. Since the $\Upsilon(4S)$ has spin-1 and the B mesons have zero spin, the $B\bar{B}$ state has a definite orbital angular momentum quantum number, and the $B\bar{B}$ pair is produced in a coherent $L = 1$ state. In other words, the two B mesons are in a state of quantum entanglement. Each particle evolves in time, as described above

for a single B meson, but they evolve coherently, such that at any time until one of the particles decays, there is exactly one B^0 and one $\overline{B}^0$. Once one of the B mesons decays, however, the other B continues to evolve as a single particle. Hence, events with two B^0 or two $\overline{B}^0$ decays may be observed. The probability of such an occurrence depends on the time between the two decays.

In the $\Upsilon(4S)$ rest frame, the two-B state can be written in terms of the angle θ they make with the e^- beam direction and the angle ϕ in the plane transverse to the beam direction:

$$S(t_f, t_b) = \frac{1}{\sqrt{2}}(B^0_{phys}(t_f, \theta, \phi)\bar{B}^0_{phys}(t_f, \pi - \theta, \phi + \pi) \\ - \bar{B}^0_{phys}(t_f, \theta, \phi)B^0_{phys}(t_f, \pi - \theta, \phi + \pi))\sin(\theta), \quad (1.31)$$

where t_f is the proper time of the B_f, the B meson in the forward half-space at angle $(\theta_f < \pi/2, \phi_f)$, and t_b is the proper time for the backward-moving B_b, at $(\pi - \theta_f, \phi_f + \pi)$.

In terms of the quantities that we defined in the previous section, we can write this as:

$$S(t_f, t_b) = \frac{1}{\sqrt{2}}e^{-((\Gamma/2)+iM)(t_f+t_b)}(\cos[\Delta m_B(t_f - t_b)/2](B^0_f\bar{B}^0_b - \bar{B}^0_f B^0_b) \\ - i\sin[\Delta m_B(t_f - t_b)/2](\frac{p}{q}B^0_f B^0_b - \frac{q}{p}\bar{B}^0_f\bar{B}^0_b))\sin(\theta_f). \quad (1.32)$$

Since the B's are produced with equal momenta back-to-back in the $\Upsilon(4S)$ rest frame, this state contains one B^0 and one $\overline{B}^0$ until one of the particles decays. Before this happens, $t_f = t_b$. Once a particle decays, its clock is stopped and the second term that depends on $\sin[\Delta m_B(t_f - t_b)/2]$ will play a role.

We can use the above equation to derive the amplitude for one of the two B mesons to decay to any state f_1 at time t_1 and the other B to decay to f_2 at time t_2:

$$A(t_1, t_2) = \frac{1}{\sqrt{2}}e^{-((\Gamma/2)+iM)(t_1+t_2)}\zeta(t_1, t_2)(\cos[\Delta m_B(t_1 - t_2)/2](A_1\bar{A}_2 - \bar{A}_1 A_2) \\ - i\sin[\Delta m_B(t_1 - t_2)/2](\frac{p}{q}A_1 A_2 - \frac{q}{p}\bar{A}_1\bar{A}_2)), \quad (1.33)$$

where A_i is the amplitude for a B^0 to decay to final state f_i and $\bar{A}_i$ is the amplitude for a $\overline{B}^0$ to decay to the same final state f_i. Any state that determines the flavor of the parent B has A_f or $\bar{A}_f = 0$. We also define

$$\zeta(t_1, t_2) = \begin{cases} +1, & t_1 = t_f, t_2 = t_b, \\ -1, & t_1 = t_b, t_2 = t_f \end{cases} \tag{1.34}$$

The time-dependent rate for producing the final states f_1, f_2 is:

$$\begin{aligned} R(t_1, t_2) = Ce^{-\Gamma(t_1+t_2)}[(|A_1|^2 + |\bar{A}_1|^2)(|A_2|^2 + |\bar{A}_2|^2) - 4\mathcal{R}e(\frac{q}{p}A_1^*\bar{A}_1)\mathcal{R}e(\frac{q}{p}A_2^*\bar{A}_2) \\ - \cos(\Delta m_B(t_1 - t_2))[(|A_1|^2 - |\bar{A}_1|^2)(|A_2|^2 - |\bar{A}_2|^2)] - 4\mathcal{I}m(\frac{q}{p}A_1^*\bar{A}_1)\mathcal{I}m(\frac{q}{p}A_2^*\bar{A}_2)] \\ + 2\sin(\Delta m_B(t_1 - t_2))[\mathcal{I}m(\frac{q}{p}A_1^*\bar{A}_1)(|A_2|^2 - |\bar{A}_2|^2) - (|A_1|^2 - |\bar{A}_1|^2)\mathcal{I}m(\frac{q}{p}A_2^*\bar{A}_2)]], \end{aligned} \tag{1.35}$$

where we have integrated over all directions, and an overall normalization factor C has been introduced. We have also approximated $|q/p| = 1$.

CP studies distinguish whether a B is a B^0 or $\overline{B}^0$ by looking at its decay products; this process is called *flavor tagging*, or simply *tagging*. These studies often look at events in which one B decays to a *CP* eigenstate f_{CP} at time t_f and the other B decays to a tagging mode at time t_{tag}. For instance, if we find a tagging mode with $A_2 = 0, \bar{A}_2 = \bar{A}_{tag}$, this then tells us that the other B is a B^0 at time $t_2 = t_{tag}$. Even when $t_{tag} > t_f$, the other B at any time after t_{tag} must be in a mixture such that if it had not decayed, it would have evolved to become B^0 at $t_f = t_{tag}$. Rewriting the time-dependent rate for this scenario:

$$\begin{aligned} R(t_{tag}, t_{f_{CP}}) = Ce^{-\Gamma(t_{tag}+t_{f_{CP}})}|\bar{A}_{tag}|^2|A_{f_{CP}}|^2[1 + |\lambda_{f_{CP}}|^2 \\ + \cos[\Delta m_B(t_{f_{CP}} - t_{tag})](1 - |\lambda_{f_{CP}}|^2) - 2\sin[\Delta m_B(t_{f_{CP}} - t_{tag})]\mathcal{I}m(\lambda_{f_{CP}})], \end{aligned} \tag{1.36}$$

where

$$\lambda_{f_{CP}} \equiv \frac{q}{p}\frac{\bar{A}_{f_{CP}}}{A_{f_{CP}}} = \eta_{f_{CP}}\frac{q}{p}\frac{\bar{A}_{\bar{f}_{CP}}}{A_{f_{CP}}}, \tag{1.37}$$

since

$$A_{\bar{f}_{CP}} = \eta_{f_{CP}}\bar{A}_{\bar{f}_{CP}}, \tag{1.38}$$

where $\eta_{f_{CP}}$ is the CP eigenvalue for state f_{CP}.

If the tagged state has $\bar{A}_2 = 0, A_2 = A_{tag}$, then the second B was a $\overline{B}^0$ at time t_{tag} and an equation like Eq. (1.36) applies, except that the signs of the cosine and sine terms are reversed. Since $|q/p| = 1$, the amplitudes for these opposite tagging scenarios are the same.

From the equation above, we can write the decay rate distributions $f_+(f_-)$ for $B \to f$ when B_{tag} is a B^0 ($\overline{B}^0$):

$$f_\pm(\Delta t) = \frac{e^{-|\Delta t|/\tau}}{4\tau}[1 \pm S_f \sin(\Delta m_d \Delta t) \mp C_f \cos(\Delta m_d \Delta t)], \tag{1.39}$$

where

$$S_f = \frac{2\mathcal{I}m\lambda_f}{1+|\lambda_f|^2} \text{ and } C_f = \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}. \tag{1.40}$$

These distributions are normalized such that $f_+ + f_- = 1$. In time-dependent CP studies, $f_\pm$ are the distributions used to fit the data, and C_f and S_f or $|\lambda_f|$ and $\mathcal{I}m\lambda_f$ are the parameters extracted from the analysis.

The difference of the rates divided by their sum gives us a time-dependent CP asymmetry of:

$$a_{f_{CP}} = \frac{(1-|\lambda_{f_{CP}}|^2)\cos(\Delta m_B t) - 2\,\mathcal{I}m\lambda_{f_{CP}} \sin(\Delta m_B t)}{1+|\lambda_{f_{CP}}|^2}, \tag{1.41}$$

where $t = t_{f_{CP}} - t_{tag}$, the time difference between the two B decays. The ability to relate this time difference to the distance between the two decay vertices is the main reason for building an asymmetric collider. If the beams were collided with the same energy, the distance would be unmeasurable due to the resolution limits of current vertex detectors. If we did not have the ability to measure this time difference and had to integrate over this variable, the term in front of the $\sin(\Delta m_B t)$ would be lost and we would only be sensitive to CP violating effects with $|\lambda| \neq 1$. This, in effect, is a consequence of the two B mesons being produced in a coherent state. At a hadronic machine where the B mesons might be produced incoherently, this effect does not occur.

1.4 Classification of CP Violation in B Decays

CP violation in B decays can be classified into three categories: CP violation in decay, CP violation in mixing, and CP violation in the interference between decays with and without mixing. These three categories are discussed in more detail in the following sections.

1.4.1 CP Violation in Decay, or Direct CP Violation

CP violation in decay occurs when there is interference between different terms in the amplitudes for a decay. Such interference depends on the complex phases in A_f and $\bar{A}_{\bar{f}}$, of which there are two types. The first type of phase comes from complex parameters of any Lagrangian term that contributes to the amplitude. Such phases appear as the complex conjugate in the CP conjugate amplitude, and hence these phases appear in A_f and $\bar{A}_{\bar{f}}$ with opposite signs. For the Standard Model these types of phases only exist in the CKM matrix, and they are therefore called "weak phases". The second type of phase occurs in scattering or decay amplitudes, typically due to strong interactions, even when the Lagrangian is real. These phases appear in A_f and $\bar{A}_{\bar{f}}$ with the same sign. For both weak and strong phases, only relative phases between different terms have physical relevance.

We can deconstruct A in terms of its magnitude A_i, weak phase ϕ_i, and strong phase δ_i. Summing over all amplitude contributions:

$$A_f = \sum_i A_i e^{i(\delta_i+\phi_i)}, \bar{A}_{\bar{f}} = e^{2i(\xi_f-\xi_B)} \sum_i A_i e^{i(\delta_i-\phi_i)}, \tag{1.42}$$

where ξ_f and ξ_B were defined in Eq. (1.12).

For any final state f, $|\frac{\bar{A}_{\bar{f}}}{A_f}|$ is independent of phase conventions and is a physically meaningful quantity. Using Eq. (1.42), we can write this quantity as:

$$\left|\frac{\bar{A}_{\bar{f}}}{A_f}\right| = \left|\frac{\sum_i A_i e^{i(\delta_i-\phi_i)}}{\sum_i A_i e^{i(\delta_i+\phi_i)}}\right|. \tag{1.43}$$

If CP is conserved, then the weak phases are all equal and $|\frac{\bar{A}_{\bar{f}}}{A_f}| = 1$. Therefore, CP violation

occurs when:

$$\text{Direct } CP \text{ violation}: \ |\frac{\bar{A}_{\bar{f}}}{A_f}| \neq 1. \tag{1.44}$$

This type of *CP* violation is called *CP* violation in decay, or direct *CP* violation. Direct *CP* violation can only occur when at least two terms with different weak phases also have different strong phases. This is the only type of *CP* violation accessible to charged B mesons, since they do not exhibit mixing. The *CP* asymmetry for direct *CP* violation can be written as:

$$a_{CP} = \frac{1 - |\bar{A}/A|^2}{1 + |\bar{A}/A|^2}. \tag{1.45}$$

1.4.2 *CP* Violation in Mixing, or Indirect *CP* Violation

CP violation in mixing, also referred to as indirect *CP* violation, is related to the fact that when *CP* is conserved, the mass eigenstates of a Hamiltonian are also *CP* eigenstates.

The relevant physically meaningful quantity, independent of phase conventions is:

$$\left|\frac{q}{p}\right|^2 = \left|\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}\right|. \tag{1.46}$$

When the mass eigenstates and *CP* eigenstates are the same, the relative phase between M_{12} and Γ_{12} is zero. Indirect *CP* violation therefore occurs when:

$$\text{Indirect } CP \text{ violation}: \ |q/p| \neq 1 \tag{1.47}$$

This type of *CP* violation has been observed for the neutral kaon system, but has yet to be seen in neutral B mesons. The effect of indirect *CP* violation in neutral B_d decays is expected to be small, on the order of $\mathcal{O}(10^{-2})$. In addition, calculating q/p involves calculating M_{12} and Γ_{12}, which involves large hadronic uncertainties. Even if indirect *CP* violation is observed, it would be difficult to relate the asymmetry rates to CKM parameters.

1.4.3 *CP* Violation in the Interference Between Mixing and Decay

The last type of *CP* violation concerns neutral B decays into *CP* eigenstates which are accessible to both B^0 and $\overline{B}^0$ decays. If *CP* is conserved, $|q/p| = 1$ and $|\bar{A}_{\bar{f}_{CP}}/A_{f_{CP}}| =$

1, as discussed for the other two types of *CP* violation. In addition, the relative phase between (q/p) and $(\bar{A}_{\bar{f}_{CP}}/A_{f_{CP}})$ disappears. The physical meaningful quantity in this case is $\lambda = \eta_{f_{CP}} \frac{q}{p} \frac{\bar{A}_{\bar{f}_{CP}}}{A_{f_{CP}}}$. *CP* violation in the interference between mixing and decay, sometimes abbreviated as "interference between mixing and decay" occurs when:

$$CP \text{ violation in interference between mixing and decay : } \lambda \neq 1. \tag{1.48}$$

To a good approximation, $|q/p| = 1$ and $|\bar{A}/A| = 1$, and yet we can still have *CP* violation when:

$$|\lambda| = 1,\ \mathcal{I}m\lambda \neq 0. \tag{1.49}$$

This type of *CP* violation has been observed in the neutral kaon system, as well as in neutral B mesons. *CP* violation in the interference between mixing and decay can be studied by comparing decays into final *CP* eigenstates of time-evolving B mesons that start at time zero as B^0 to those that start as $\overline{B}^0$. The *CP* asymmetry in this case can be written as:

$$a_{f_{CP}} = \frac{\Gamma(B^0_{phys}(t) \to f_{CP}) - \Gamma(\bar{B}^0_{phys}(t) \to f_{CP})}{\Gamma(B^0_{phys}(t) \to f_{CP}) + \Gamma(\bar{B}^0_{phys}(t) \to f_{CP})}. \tag{1.50}$$

Using Eq. (1.39), the asymmetry can be written as:

$$a_{f_{CP}} = C_{f_{CP}} \cos(\Delta m_B t) - S_{f_{CP}} \sin(\Delta m_B t). \tag{1.51}$$

CP violation in the interference between mixing and decay occurs when $S_{f_{CP}} \neq 0$. Direct *CP* violation occurs when $C_{f_{CP}} \neq 0$.

When there is no *CP* violation in decay ($C_{f_{CP}} = 0$), the *CP* violating parameters can be related to the parameters of the Lagrangian in a theoretically clean way. For instance, the *CP* asymmetry for B decays that are dominated by a single *CP* violating phase, and therefore, have a negligible effect from *CP* violation in decay, can be cleanly related to electroweak Lagrangian parameters with little hadronic uncertainty. An example of such a "golden" mode is $B^0 \to J/\psi K^0_S$. However, aside from these cases, it is nontrivial to relate *CP* asymmetries to Standard Model parameters.

CP violation in the interference between mixing and decay can also be observed in decays that are not *CP* eigenstates, as long as the final state is accessible to both the B^0

and $\overline{B}^0$. This is the type of *CP* violation that motivates this analysis and it is discussed in further detail in the next section.

1.5 $\sin(2\beta+\gamma)$ in $B \to D^{(*)}\pi$ Decays

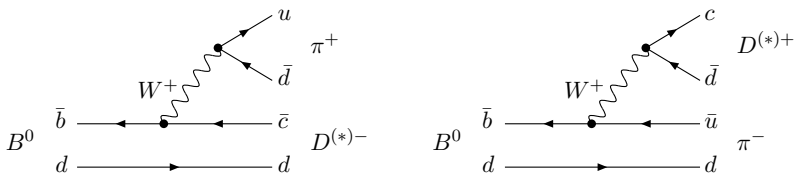


Figure 1.4. Dominant Feynman diagrams for the CKM-favored decay $B^0 \to D^-\pi^+$ (left) and the doubly CKM-suppressed decay $B^0 \to D^+\pi^-$ (right).

An important experimental test of the model described in the previous sections is the determination of the angle $\gamma = \arg(-V_{ud}V_{ub}^*/V_{cd}V_{cb}^*)$ of the unitarity triangle. A theoretically-clean measurement of $\sin(2\beta+\gamma)$ can be obtained from the study of the time dependent decay rate distributions of $B^0, \overline{B}^0 \to D^{(*)-}\pi^+$. Henceforth, charge conjugation will be implied, unless otherwise stated. The time evolution of these decays involves the interference between CKM-favored decays $B^0 \to D^{(*)-}\pi^+$ and CKM-suppressed decay $B^0 \to D^{(*)+}\pi^-$ [10], the Feynman diagrams for which can be seen in Figure 1.4. Although the states $D^{(*)\pm}\pi^\mp$ are not *CP* eigenstates, *CP* eigenstates can be constructed from a linear combination of $D^{(*)+}\pi^-$ and $D^{(*)-}\pi^+$. *CP* violation can be observed since both final states are accessible to both the B^0 and $\overline{B}^0$.

The time dependent decay rate distributions for these decays, following the discussion in Section 1.3, is:

$$f^{\pm}_{D^{(*)-}\pi^+} = \frac{e^{-|\Delta t|/\tau}}{4\tau}[1 \pm S_{D^{(*)-}\pi^+}\sin(\Delta m_d \Delta t) \mp C_{D^{(*)-}\pi^+}\cos(\Delta m_d \Delta t)], \tag{1.52}$$

$$f^{\pm}_{D^{(*)+}\pi^-} = \frac{e^{-|\Delta t|/\tau}}{4\tau}[1 \pm S_{D^{(*)+}\pi^-}\sin(\Delta m_d \Delta t) \mp C_{D^{(*)+}\pi^-}\cos(\Delta m_d \Delta t)]. \tag{1.53}$$

The upper (lower) sign refers to the flavor of the tagged B as B^0 ($\overline{B}^0$). The sine and cosine

coefficients can be expressed as:

$$S_{D^{(*)\pm}\pi^{\mp}} = \frac{2\ Im(\lambda_{D^{(*)\pm}\pi^{\mp}})}{1+|\lambda_{D^{(*)\pm}\pi^{\mp}}|^2},\ C_{D^{(*)\pm}\pi^{\mp}} = \frac{1-|\lambda_{D^{(*)\pm}\pi^{\mp}}|^2}{1+|\lambda_{D^{(*)\pm}\pi^{\mp}}|^2}, \tag{1.54}$$

where

$$\lambda_{D^{(*)\pm}\pi^{\mp}} \equiv \frac{q}{p}\frac{\bar{A}_{D^{(*)\pm}\pi^{\mp}}}{A_{D^{(*)\pm}\pi^{\mp}}} \tag{1.55}$$

is proportional to the ratio of the amplitudes for the CKM-suppressed and CKM-favored decays.

Neglecting *CP* violation in mixing, the ratio $\frac{q}{p}$ can be estimated from box diagrams, such as those shown in Figure 1.3, for $B^0 - \overline{B}^0$ mixing to a very good approximation:

$$\frac{q}{p} = -\frac{M_{12}^*}{|M_{12}|} = \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} e^{2i\xi_B}. \tag{1.56}$$

where M_{12} was defined in Eq. (1.13) and ξ_B was defined in Eq. (1.12).

For $B^0 \to D^+\pi^-$, the ratio $\frac{\bar{A}_f}{A_f}$ can be expressed as:

$$\frac{\bar{A}_{D^+\pi^-}}{A_{D^+\pi^-}} = \frac{V_{cb}V_{ud}^*}{V_{cd}V_{ub}^*} e^{-2i\xi_B} \frac{\bar{M}_{D^+\pi^-}}{M_{D^+\pi^-}}, \tag{1.57}$$

where $\bar{M}_{D^+\pi^-}$ and $M_{D^+\pi^-}$ are the hadronic matrix elements which take into account the fact that the quarks are bound into color-neutral hadrons. Combining Eq (1.56) and (1.62), $\lambda_{D^+\pi^-}$ can be written as:

$$\lambda_{D^+\pi^-} = \left(\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}\right)\left(\frac{V_{cb}V_{ud}^*}{V_{cd}V_{ub}^*}\right)\left(\frac{\bar{M}_{D^+\pi^-}}{M_{D^+\pi^-}}\right). \tag{1.58}$$

Extracting the weak phase $2\beta+\gamma$, this becomes:

$$\lambda_{D^+\pi^-} = e^{-i(2\beta+\gamma)}\left(\frac{|V_{tb}^* V_{td}|}{|V_{tb} V_{td}^*|}\right)\left(\frac{|V_{cb}V_{ud}^*|}{|V_{cd}V_{ub}^*|}\right)\left(\frac{\bar{M}_{D^+\pi^-}}{M_{D^+\pi^-}}\right). \tag{1.59}$$

Extracting the relative strong phase δ between the hadronic matrix elements, we have:

$$\lambda_{D^+\pi^-} = e^{-i(2\beta+\gamma+\delta)}\left(\frac{|V_{tb}^* V_{td}|}{|V_{tb} V_{td}^*|}\right)\left(\frac{|V_{cb}V_{ud}^*|}{|V_{cd}V_{ub}^*|}\right)\left(\frac{|\bar{M}_{D^+\pi^-}|}{|M_{D^+\pi^-}|}\right). \tag{1.60}$$

Assuming there is no *CP* violation in mixing, $\lambda_{D^+\pi^-}$ can then be written as:

$$\lambda_{D^+\pi^-} = e^{-i(2\beta+\gamma+\delta)}\left(\frac{|V_{cb}V_{ud}^*|}{|V_{cd}V_{ub}^*|}\right)\left(\frac{|\bar{M}_{D^+\pi^-}|}{|M_{D^+\pi^-}|}\right). \tag{1.61}$$

For $B^0 \to D^-\pi^+$, the ratio $\frac{\bar{A}_f}{A_f}$ is:

$$\frac{\bar{A}_{D^-\pi^+}}{A_{D^-\pi^+}} = \frac{V_{ub}V_{cd}^*}{V_{ud}V_{cb}^*} e^{-2i\xi_B} \frac{\bar{M}_{D^+\pi^-}}{M_{D^+\pi^-}}. \tag{1.62}$$

Since CP is conserved in strong interactions within the Standard Model,

$$\bar{M}_f = M_{\bar{f}} \text{ and } \bar{M}_{\bar{f}} = M_f. \tag{1.63}$$

We can write the corresponding expression for $\lambda_{D^-\pi^+}$ as:

$$\lambda_{D^-\pi^+} = e^{-i(2\beta+\gamma-\delta)} \left(\frac{V_{ub}V_{cd}^*}{V_{ud}V_{cb}^*}\right) \left(\frac{|M_{D^+\pi^-}|}{|\bar{M}_{D^+\pi^-}|}\right). \tag{1.64}$$

From Eq. (1.61) and (1.64), we can see that the final states $D^+\pi^-$ and $D^-\pi^+$ have the same strong phase with opposite signs and identical weak phases. In addition, we can see that $|\lambda_{D^+\pi^-}| = 1/|\lambda_{D^-\pi^+}|$. Similar expressions for $\lambda_{D^{*\pm}\pi^\mp}$ can be written, although the hadronic matrix elements, and thus the strong phase, may differ. We can thus express $\lambda_{D^{(*)\pm}\pi^\mp}$ as:

$$\lambda_{D^{(*)-}\pi^+} = e^{-i(2\beta+\gamma-\delta^{(*)})}|\lambda_{D^{(*)}\pi}|, \tag{1.65}$$

$$\lambda_{D^{(*)+}\pi^-} = e^{-i(2\beta+\gamma+\delta^{(*)})}\frac{1}{|\lambda_{D^{(*)}\pi}|}. \tag{1.66}$$

Rewriting Eq. (1.52) and (1.53) in terms of $2\beta+\gamma$, $\delta^{(*)}$, and $|\lambda_{D^{(*)}\pi}|$:

$$\begin{aligned} f^\pm = \frac{e^{-|\Delta t|/\tau}}{4\tau(1+|\lambda_{D^{(*)}\pi}|^2)}[1+|\lambda_{D^{(*)}\pi}|^2 \\ \mp 2|\lambda_{D^{(*)}\pi}|\sin(2\beta+\gamma-\xi\delta^{(*)})\sin(\Delta m_d\Delta t) \mp \xi(1-|\lambda_{D^{(*)}\pi}|^2)\cos(\Delta m_d\Delta t)], \end{aligned} \tag{1.67}$$

where the upper (lower) sign is for a B^0 ($\overline{B}^0$) tagged decay and $\xi = 1(-1)$ for $D^{(*)-}\pi^+(D^{(*)+}\pi^-)$. The goal for the study of CP violation in $B \to D^{(*)}\pi$ decays is to extract the quantity $\sin(2\beta+\gamma)$ from the measurement of the time dependent rate expressed in Eq. (1.67). From henceforth, we will use the notation $r_{D^{(*)}\pi} = |\lambda_{D^{(*)}\pi}|$.

We can estimate the expectation for $r_{D^{(*)}\pi}$ by inputting the values of the CKM-matrix elements into Eq. (1.61), neglecting the hadronic matrix elements. This results in an estimate for $r_{D^{(*)}\pi}$ of ≈ 0.02. The smallness of $r_{D^{(*)}\pi}$ indicates the smallness of the amplitude

for the doubly CKM-suppressed decay with respect to the amplitude for the CKM-favored decay. The CKM-suppressed mode receives a factor of $\lambda^3_{Cabibbo}$ from V_{ub} and an additional factor of $\lambda_{Cabibbo}$ from V_{cd}, whereas the CKM-favored mode only has a factor of $\sim \lambda^2_{Cabibbo}$ from V_{cb}.

As can be seen in Eq. (1.67), the value of $r_{D^{(*)}\pi}$ indicates the sensitivity to measuring *CP* violation using this method, since the term containing the weak phase is necessarily weighed by the factor of $r_{D^{(*)}\pi}$. Because $r_{D^{(*)}\pi}$ is expected to be so small, the experimental sensitivity is limited, and it is not possible to extract both $\sin(2\beta+\gamma-\xi\delta^{(*)})$ and $r_{D^{(*)}\pi}$ from a fit to the available dataset. An external measurement or calculation of $r_{D^{(*)}\pi}$ is therefore needed. A measurement of the branching fractions for $B^0 \to D^{(*)+}\pi^-$ would give us this required external input. Unfortunately, the direct measurement of the branching fractions for the doubly CKM-suppressed mode $B^0 \to D^{(*)+}\pi^-$ is not possible with the currently available data sample due to the presence of the overwhelming background from $\overline{B}^0 \to D^{(*)+}\pi^-$. The branching fraction for the CKM-favored mode has been measured and is about 10^{-3}, whereas the branching fraction for the CKM-suppressed mode is expected to be on the order of 10^{-6} [12].

1.6 $B^0 \to D_s^{(*)+}\pi^-$ Decays

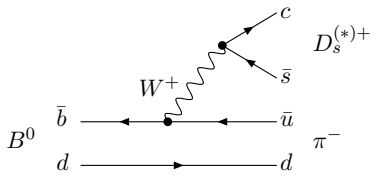


Figure 1.5. Dominant Feynman diagram for the decays $B^0 \to D_s^{(*)+}\pi^-$

Although the branching fraction for the doubly CKM-suppressed decay, $B^0 \to D^{(*)+}\pi^-$, cannot be measured experimentally with our current dataset, it can be related to the branching fraction for $B^0 \to D_s^{(*)+}\pi^-$ (Fig. 1.5). The measurement of $\mathcal{B}(B^0 \to D_s^{(*)+}\pi^-)$ is more feasible than the measurement of the $\mathcal{B}(B^0 \to D^{(*)+}\pi^-)$ since the amplitude is $\sim \lambda^3_{Cabibbo}$

Table 1.2. Previous measurements of the branching fractions and upper limits of $B^0 \to D_s^{(*)+}\pi^-$and $B^0 \to D_s^{(*)-}K^+$, from the *BABAR* and Belle experiments, as well as the world averages from the PDG [11, 12].

	BABAR		Belle	World Averages
Mode	$\mathcal{B}(10^{-5})$	90% CL (10^{-5})	$\mathcal{B}(10^{-5})$	$\mathcal{B}(10^{-5})$
$B^0 \to D_s^+\pi^-$	$3.2 \pm 0.9 \pm 1.0$ (3.3σ)	-	$2.4^{+1.1}_{-0.9} \pm 0.6$	2.7 ± 1.0
$B^0 \to D_s^{*+}\pi^-$	$1.9^{+1.2}_{-1.3} \pm 0.5$ (2.3σ)	< 4.1	-	-
$B^0 \to D_s^-K^+$	$3.2 \pm 1.0 \pm 1.0$ (3.5σ)	-	$4.5^{+1.4}_{-1.2} \pm 1.1$	3.8 ± 1.3
$B^0 \to D_s^{*-}K^+$	-	< 2.5	-	-

rather than $\sim \lambda^4_{Cabibbo}$ and, there is no corresponding CKM-favored mode to overwhelm the signal. The branching fractions can be related as [10]:

$$\mathcal{B}(B^0 \to D_s^{(*)+}\pi^-) = \frac{\mathcal{B}(B^0 \to D^{(*)+}\pi^-)}{\tan^2\theta_C} \cdot \left(\frac{f_{D_s^{(*)}}}{f_{D^{(*)}}}\right)^2 \quad (1.68)$$

$$= \frac{\mathcal{B}(B^0 \to D^{(*)-}\pi^+)}{\tan^2\theta_C} \cdot r^2_{D^{(*)}\pi} \cdot \left(\frac{f_{D_s^{(*)}}}{f_{D^{(*)}}}\right)^2, \quad (1.69)$$

where θ_C is the Cabibbo angle and $\frac{f_{D_s^{(*)}}}{f_{D^{(*)}}}$ is the ratio of the decay constants, which takes into account the effects from the factorization approximation and $SU(3)$ flavor symmetry breaking. Lattice QCD calculations estimate this ratio to be [13]:

$$\frac{f_{D_s^{(*)}}}{f_{D^{(*)}}} = 1.22 \pm 0.04. \quad (1.70)$$

Other $SU(3)$-breaking effects are believed to affect $r_{D^{(*)}\pi}$ by less than 30% [14]. Estimating $r_{D^{(*)}\pi}$ from the CKM matrix elements gives us an estimate of $\mathcal{B}(B^0 \to D_s^{(*)+}\pi^-)$ of $2 \cdot 10^{-5}$.

The results of our previous measurements for $B^0 \to D_s^{(*)+}\pi^-$ and $B^0 \to D_s^{(*)-}K^+$ are reported in Table 1.2, in addition to results from the Belle Experiment and the world averages from the PDG.

1.7 $B^0 \to D_s^{(*)-} K^+$ Decays

$B^0 \to D_s^{(*)+}\pi^-$ has a single amplitude contributing to the decay since it has four distinct quark flavors in the final state (Fig 1.5). On the other hand, there are two diagrams contributing to $B^0 \to D^{(*)-}\pi^+$ and $B^0 \to D^{(*)+}\pi^-$: tree amplitudes (Fig. 1.4) and color-suppressed direct W-exchange amplitudes (Fig. 1.6). Thus the relationship in Eq. (1.69) is only valid if the W-exchange diagram contribution for $B^0 \to D^+\pi^-$ is negligible compared to the tree diagram one.

To probe the size of the W-exchange amplitudes relative to the dominant processes $B^0 \to D^{(*)-}\pi^+$, we also consider the decay $B^0 \to D_s^{(*)-}K^+$ which proceeds through a W-exchange process (Fig. 1.6). Such W-exchange contributions are difficult to compute since the spectator quark plays an important role and factorization cannot be assumed.

As can be seen from the diagrams, both charge combinations, $D_s^{\pm}K^{\mp}$, are allowed for the B^0 decay: one is CKM-favored ($D_s^- K^+$) while the other is CKM-suppressed ($D_s^+ K^-$). The amplitude for the CKM-favored decay, taking into account only the W-exchange diagram, is expected to be $\lambda^2 \times \frac{f_B}{m_B}$, where f_B (~ 200 MeV) is the B meson decay constant and m_B is the B mass (5.28 GeV) [15]. This results in an estimate of roughly λ^4 for the amplitude of the decay.

In addition, the decay rate of $B^0 \to D_s^{(*)-}K^+$ could be enhanced by final state rescattering from other B decays, such as $B^0 \to D^-\pi^+$. It is difficult to accurately calculate these effects, since they involve large hadronic uncertainties. Such rescattering could enhance the

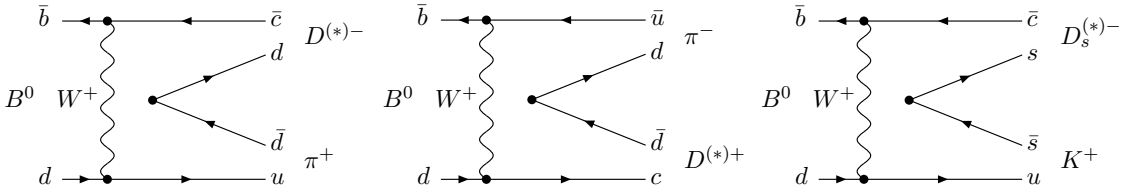


Figure 1.6. Dominant Feynman diagrams for the color-suppressed W-exchange contributions to $B^0 \to D^{(*)-}\pi^+$ (left), $B^0 \to D^{(*)+}\pi^-$ (middle), and the decay $B^0 \to D_s^{(*)-}K^+$ (right).

amplitude of $\mathcal{B}(B^0 \to D_s^- K^+)$ to λ^3, similar in magnitude to $\mathcal{B}(B^0 \to D_s^+ \pi^-)$. In this case, the W-exchange contribution would be overwhelmed by the rescattering process.

By measuring the rate of the rescattering process, we can set an upper limit on the contribution of the W-exchange diagrams. If rescattering processes do not contribute significantly, the ratio of $\mathcal{B}(B^0 \to D_s^{(*)-} K^+)$ over $\mathcal{B}(B^0 \to D^{(*)-} \pi^+)$ would give us an estimate of the ratio of W-exchange contribution over emission diagram contributions, apart from effects due to $SU(3)$ symmetry breaking.

Predictions of $\mathcal{B}(B^0 \to D_s^{(*)-} K^+)$ give $\mathcal{O}(10^{-4})$ when using the naive f_B/m_B suppression factor and on the order of $6 \cdot 10^{-4}$ when including rescattering contributions [16]. Estimates using perturbative QCD also exist [17].

Chapter 2

The BABAR Detector at PEP-II

2.1 The PEP-II B-Factory

The *BABAR* detector collects data at the Positron Electron Project II (PEP-II) storage ring at the Stanford Linear Accelerator Center (SLAC). A schematic drawing of the linear accelerator (Linac) and PEP-II can be seen in Figure 2.1. PEP-II is a storage ring designed for studying *CP* violation in the B meson system, and it is often called a B factory because it is optimized for producing copious amounts of B mesons.

PEP-II was designed for a luminosity of $3 \times 10^{33}\,\mathrm{cm}^{-2}s^{-1}$ and has exceeded this with a

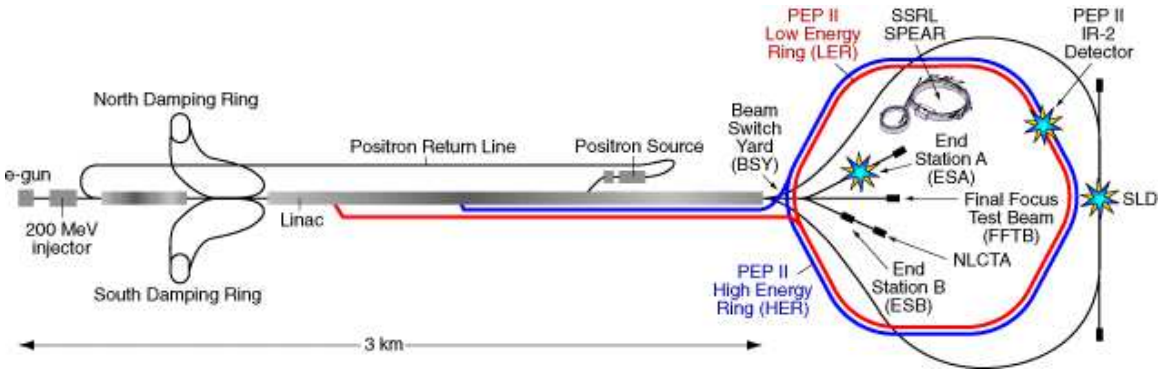


Figure 2.1. The SLAC Linac and PEP-II Storage Ring.

Table 2.1. PEP-II beam parameters, for the design and for typical colliding beam operation in the first year. 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 horizontal, vertical, and longitudinal RMS size of the luminous region.

Parameters	Design	Typical
Energy HER/LER (GeV)	9.0/3.1	9.0/3.1
Current HER/LER (A)	0.75/2.15	1.7/2.9
# of bunches	1658	1732
Bunch spacing (ns)	4.2	4.2
σ_{Lx} (μm)	110	157
σ_{Ly} (μm)	3.3	4.7
σ_{Lz} (cm)	0.9	1.0
Luminosity ($10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$)	3	10.0
Luminosity (pb^{-1}/day)	135	728

peak luminosity of $1 \times 10^{34}\,\mathrm{cm}^{-2}s^{-1}$. Table 2.1 summarizes the beam parameters at PEP-II, and Figure 2.2 shows a plot of the integrated luminosity delivered by PEP-II, and the corresponding luminosity recorded by the *BABAR* detector.

PEP-II is an asymmetric energy collider, primarily operating at the $\Upsilon(4S)$ resonance of $\sqrt{s} = 10.58\,\mathrm{GeV}$. 9.0 GeV electrons collide with 3.1 GeV positrons to create an optimal environment for *CP* studies. This asymmetric mode of operation results in B^0 mesons with a boost in the lab frame of $\beta\gamma = 0.56$, making it possible to measure the B^0 decay times which are crucial for time dependent *CP* analyses.

In addition, there are a number of advantages to operating at the $\Upsilon(4S)$ resonance. First and foremost, data-taking at a resonance enhances the production cross-section, and the $\Upsilon(4S)$ resonance is particularly well-suited for studying B mesons since it predominantly decays as $\Upsilon(4S) \to B\bar{B}$ at a rate $> 96\%$ [12]. Also, the lack of fragmentation products from the $\Upsilon(4S)$ decay results in lower combinatorial backgrounds. Moreover, the knowledge of the exact 4-momentum of the $B\bar{B}$ system and the magnitudes of the momenta for the individual B mesons in the center-of-mass (CM) frame can be used as kinematic constraints for suppressing backgrounds. Although the majority of data is recorded at the $\Upsilon(4S)$

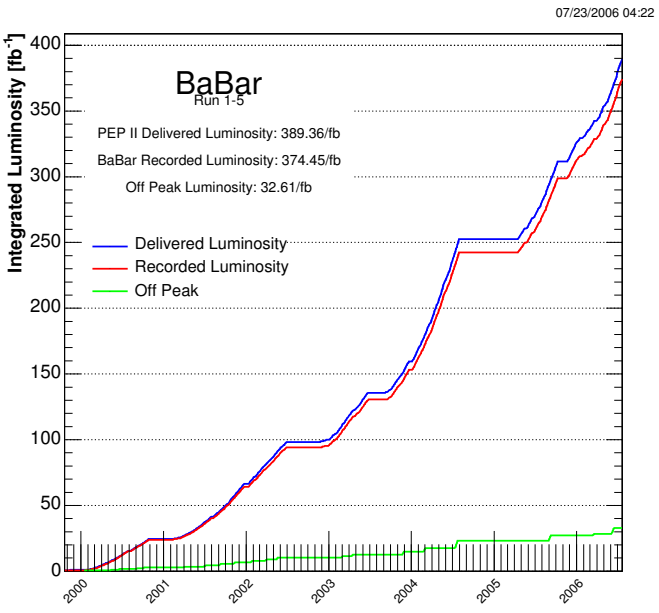


Figure 2.2. PEP-II delivered luminosity and *BABAR* recorded luminosity as a function of time.

resonance, about 12% is recorded 40 MeV lower for studying non-resonant backgrounds, such as light quark pair production. Table 2.2 shows the cross-sections for the predominant production processes at the $\Upsilon(4S)$ resonance. Figure 2.3 shows the ratio of the cross sections of $e^+e^- \to$ hadrons and $e^+e^- \to \mu^+\mu^-$ as a function of $\sqrt{s}$, near the $\Upsilon(4S)$ resonance.

Before the particles reach the 2.2 km circumference PEP-II storage ring, they are ac-

Table 2.2. Production cross-sections at $\sqrt{s} = M_{\Upsilon(4S)}$. The e^+e^- cross-section is the effective cross-section, expected within the experimental acceptance.

$e^+e^- \to$	Cross-section (nb)
$b\overline{b}$	1.05
$c\overline{c}$	1.30
$s\overline{s}$	0.35
$u\overline{u}$	1.39
$d\overline{d}$	0.35
$\tau^+\tau^-$	0.94
$\mu^+\mu^-$	1.16
e^+e^-	~ 40

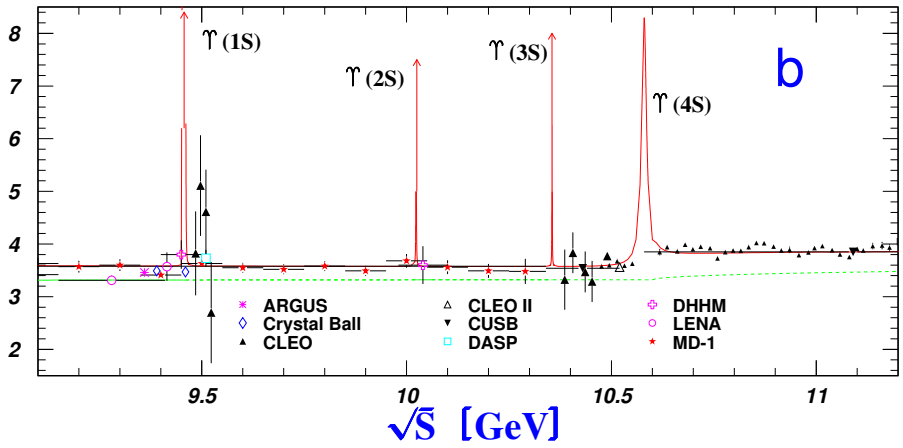


Figure 2.3. R, the ratio of the $\sigma(e^+e^- \to hadrons)$ to $\sigma(e^+e^- \to \mu^+\mu^-)$, near the $\Upsilon(4S)$ resonance, as a function of $\sqrt{s}$ [12].

celerated in the 3 km long Linac. The electrons that go into the accelerator are created by thermal emission from a metal cathode or by laser-induced photo-emission from a GaAs photocathode, and the positrons are created by a firing a high energy electron beam at a tungsten target, producing e^+e^- pairs. The electrons and positrons are partially accelerated in the Linac and then sent to damping rings, where the phase space of the beams is reduced by synchrotron radiation to maximize the luminosity. The beams are next accelerated in the Linac, gaining energy proportional to how far they travel, and then kicked out of the Linac in bunches into the storage ring.

The beams collide head-on at Interaction Region 2 (IR-2, as shown in Figure 2.1), where the *BABAR* detector is located. After the beams collide, they are separated by a series of magnets and then returned to their circular orbits in the storage ring. Figure 2.4 shows a transverse view of the interaction region and the configuration of magnets used to separate the beams. The beams are first separated by a pair of dipole magnets (B1) in the horizontal plane located ±21 cm on either side of the interaction point (IP). The dipoles are followed by a series of offset quadrupoles: the Q1 samarium-cobalt quadrupoles located within the field of the *BABAR* solenoid and the Q2, Q4, and Q5 iron quadrupoles located outside or in the fringe field of the solenoid.

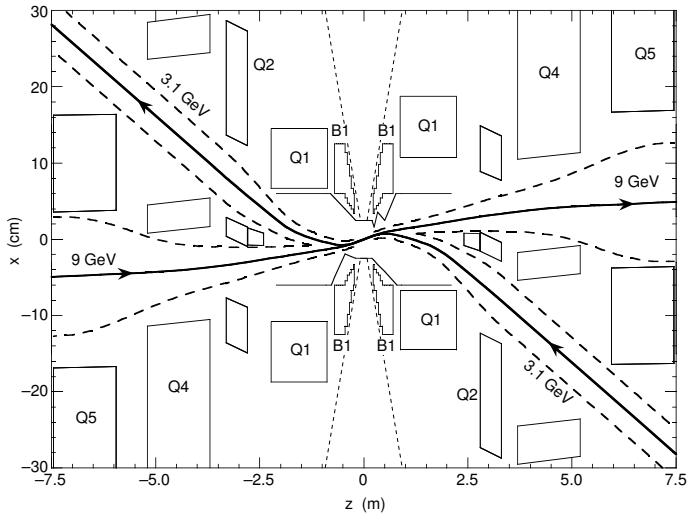


Figure 2.4. Interaction Region 2 (IR-2), transverse view. A series of magnets separates the beams post-collision.

The most important beam parameters for the performance of *BABAR* are the luminosity, beam energies, and position, angle and size of the luminous region. The luminosity is monitored online by PEP-II using radiative Bhabha scattering ($e^+e^- \to e^+e^-\gamma$). Offline, absolute luminosity is monitored using $e^+e^- \to e^+e^-$ and $e^+e^- \to \mu^+\mu^-$ events.

The beam energies are calculated using the total magnetic bending strength and the deviations of the accelerating frequencies from their central values. The systematic error on the beam energies is estimated to be between 5 and 10 MeV. The relative energy setting for the beams are accurate and stable to $\sim$1 MeV. The low energy and high energy beam energy spreads are 2.3 MeV and 5.5 MeV, respectively. In order to stay close to the $\Upsilon(4S)$ peak, the ratio of $B\bar{B}$ enriched hadronic events to leptonic pair production is monitored online.

The size and position of the beam spot, the luminous region of the beam, are important parameters for time-dependent analyses, and their values are monitored online and offline. These quantities are measured relative to the *BABAR* coordinate system, which is right-handed and anchored to the drift chamber, with the z-axis coinciding with the principal axis of the drift chamber. The positive y-axis points upward, and the positive x-axis points

away from the center of the PEP-II ring. The vertical size of the beam spot is too small to be measured directly, and it is inferred from the luminosity, horizontal size, and beam current and varies by 1-2μm. The transverse position, size, and angles of the beam spot are measured from the distribution of the distance of closest approach to the z-axis as a function of the azimuthal angle, ϕ, for tracks from well measured two-track events. The longitudinal parameters are computed from the longitudinal vertex distribution of the two tracks. Beam position uncertainties are of the order of a few μm in the transverse plane and 100μm along the beam axis. Beam position variations on a run-by-run basis are comparable to these uncertainties, which indicates that the beams are typically stable over a run.

2.2 Detector Overview

The *BABAR* detector, like many particle detectors, is comprised of several coaxial layers surrounding the interaction region, with the different detector subsystems making up the consecutive layers. Figure 2.5 depicts the overall layout of the *BABAR* detector, with numbers indicating the different detector subsystems, as itemized below. Figures 2.6 and 2.7 depict a longitudinal cross section and an end view of the detector, respectively, with scales to give a sense of the size.

The detector was designed with a number of goals, including maximum acceptance in the center-of-mass system. Since the collisions are asymmetric in energy, the detector was designed to be asymmetric for maximal acceptance. In addition, since the trajectories of the B mesons are nearly parallel to the z-axis, their decay time difference is measured by measuring the z components of their decay vertices. This requires a very high resolution vertex detector. High efficiency tracking in the range $\sim$60 MeV/$c < p_t < \sim$4 GeV/c is also a requirement, in addition to the ability to discriminate between e, μ, π, K and p over a large kinematic range. The identification of different particles is crucial for CP violation studies which distinguish B^0 from $\overline{B}^0$ mesons by looking at the flavor of the decay products. Moreover, π-K discrimination at high momenta is critical for distinguishing between decay

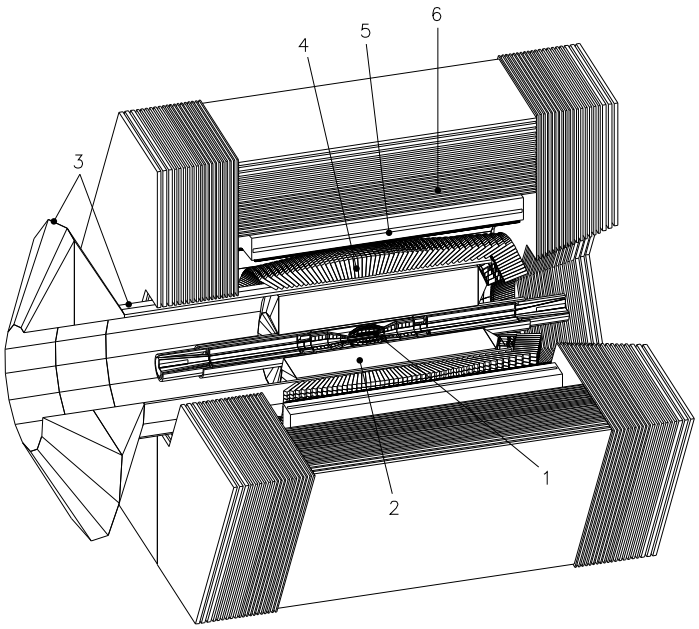


Figure 2.5. Layout of the *BABAR* detector, with numbers corresponding to detector subsystems, as itemized in the text.

channels such as $B^0 \to \pi^+\pi^-$ and $B^0 \to K^\pm\pi^\mp$. Detection of photons and π^0s over a large energy range and the ability to identify neutral hadrons, such as muons and neutrons, were also important factors in the design of the detector.

The five detector subsystems are listed from the inner to outer layers are:

1. Silicon Vertex Tracker
2. Drift Chamber
3. Detector of Internally Reflected Cherenkov Light
4. Electromagnetic Calorimeter
5. Instrumented Flux Return

In addition, a super-conducting 1.5 T solenoidal magnet is located between the drift chamber and electromagnetic calorimeter. These subsystems are highlighted in the following sections. More detail about the *BABAR* detector can be found elsewhere [19].

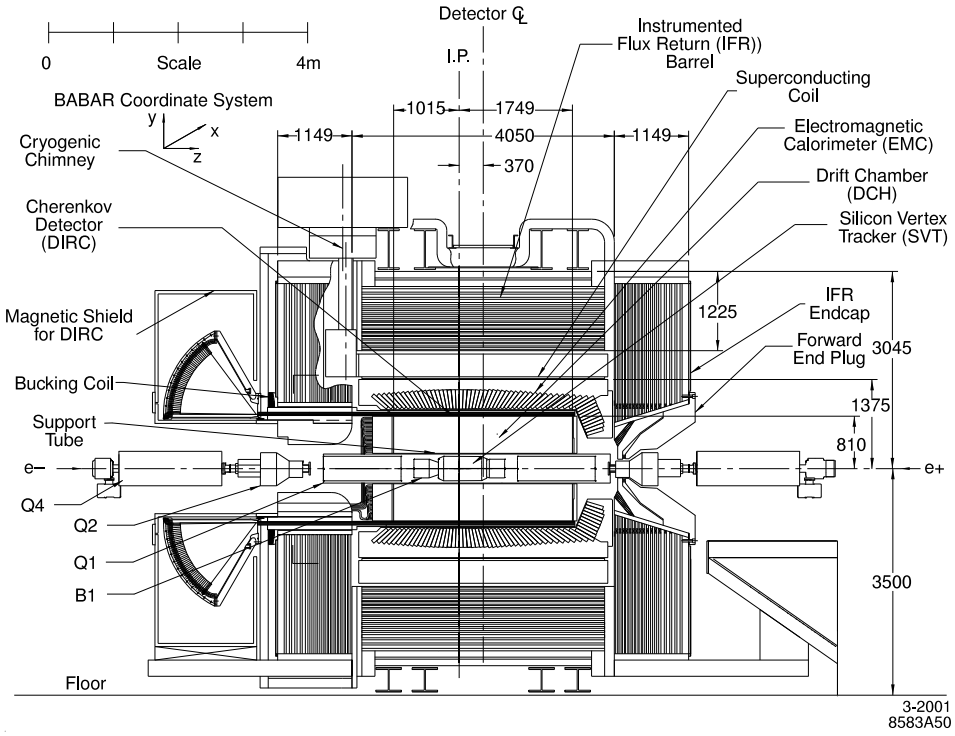


Figure 2.6. The *BABAR* detector, longitudinal cross section.

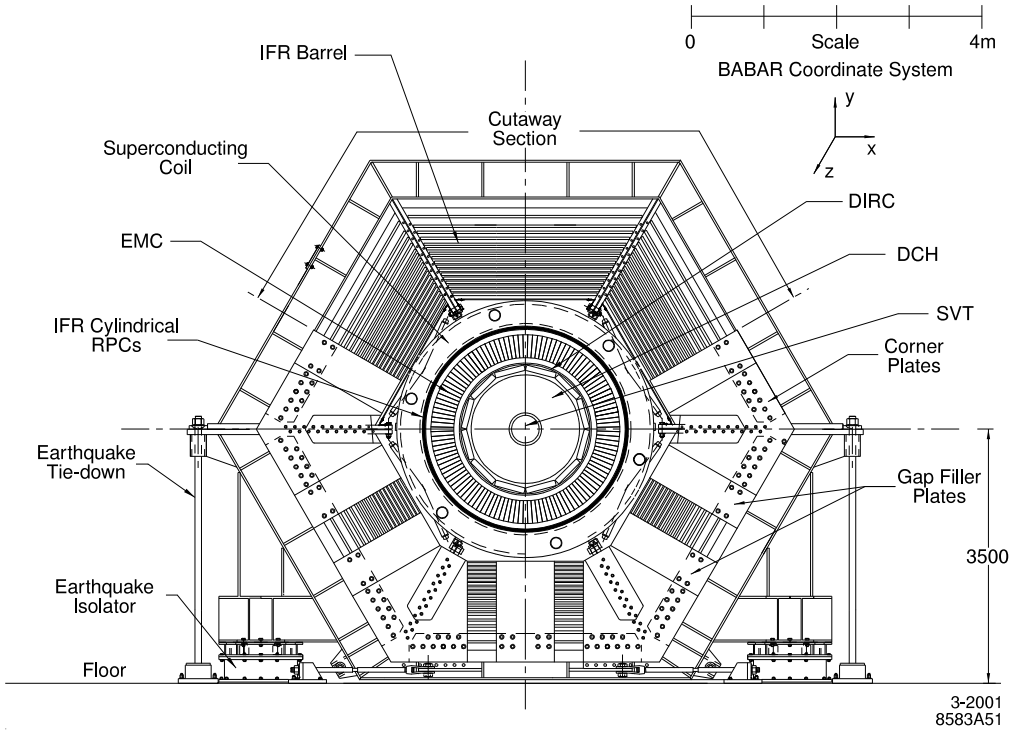


Figure 2.7. The *BABAR* detector, end view.

2.3 Silicon Vertex Tracker

The *BABAR* tracking system is comprised of the silicon vertex tracker (SVT) and the drift chamber (DCH). The purpose of the tracking system is to measure the momenta and angles of charge particles with high precision and efficiency.

The SVT is designed to measure charged particle angles and positions just outside the beam pipe. The primary purpose of the SVT is the determination of the decay vertices of the two B daughters of the $\Upsilon(4S)$. This provides us with the time difference between the two B decays, which is crucial for time-dependent CP asymmetry studies. The SVT has been optimized for high efficiency, good resolution, and the ability to withstand exposure to high levels of radiation.

The SVT is made of five concentric cylindrical layers of double-sided silicon strip detectors with readout at the ends of each module to reduce the inactive material in the fiducial volume of the detector. Figures 2.8 and 2.9 depict schematic cross-sections of the SVT. The layers are divided azimuthally into overlapping modules. The inner three layers, located at a radius of $\sim 3\,\mathrm{cm}$ from the beam pipe, have six detector modules and are traditional barrel-style structures with modules overlapping in a pin-wheel format. The outer two layers are positioned farther from the beam pipe at a radius of $\sim 9\,\mathrm{cm}$ and made up of 16 and 18 modules, respectively, with modules alternating at slightly smaller/larger radii for overlap.

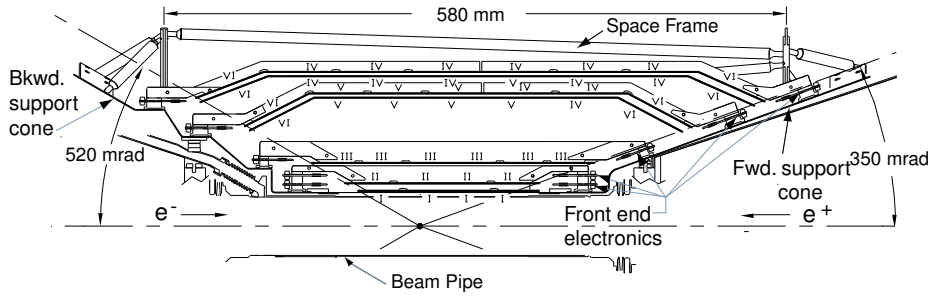


Figure 2.8. SVT, longitudinal section.

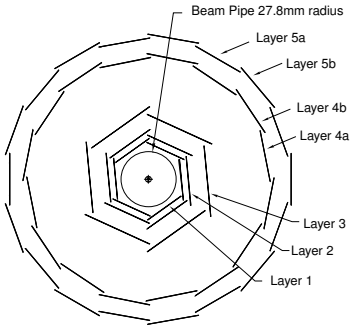


Figure 2.9. SVT, transverse section.

These outer layers use an arch structure, as can be seen in the longitudinal cross-section, in order to increase solid angle coverage with minimal silicon needed and to avoid very large track incidence angles near the edges of the acceptance regions.

The main purpose of the inner layers of the SVT is to provide position and angle information for measuring the vertex. These inner layers are mounted as close to the beam pipe as possible in order to reduce the effect of multiple-scattering in the beam pipe on the vertex determination. The outer layers are placed closer to the DCH in order to provide the measurements needed to link the SVT and DCH tracks. Since charged particles with transverse momenta less than 100 MeV/c will not reach the DCH, the SVT is the only detector subsystem which can provide information for reconstructing such very low-energy tracks. Likewise, the SVT is the sole detector for short-lived particles, such as Λs, that will decay within the SVT. In addition, both the hit time and time-over-threshold, related to the pulse height, are measured in the SVT. This information can be used to compute specific energy loss, dE/dx, of charged particles in the SVT and used in conjunction with measurements from the DCH and DIRC for particle identification.

2.4 Drift Chamber

The drift chamber is the second component of the *BABAR* tracking system. Figure 2.10 shows a schematic longitudinal cross-section of the DCH. The DCH is the main tracking

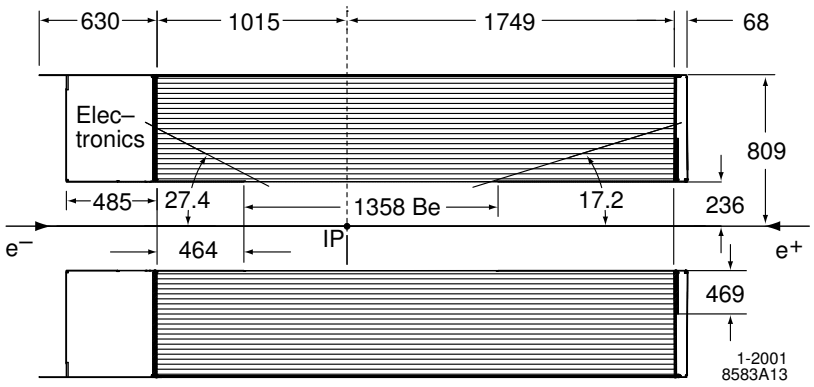


Figure 2.10. DCH, longitudinal section. Dimensions are shown in mm.

device for the experiment and its primary purpose is the precise and efficient measurement of charged particle momenta and angles. The measurements made by the DCH complement those made by the SVT close to the interaction point. Together the DCH and SVT provide the tracking resolution required for *CP* analyses and the study of rare B decays. The DCH provides the only reconstruction information for particles that decay outside of the SVT.

In addition to tracking, the DCH provides ionization loss (dE/dx) information used for particle identification of low momentum particles, complementing the measurements made in the SVT and DIRC barrel region. For particles in the extreme forward or backward regions, the DCH and SVT are the only sources of particle identification. Moreover, the DCH provides the charged particle trigger, one of primary triggers for *BABAR*.

The DCH is 2.8 m long with an inner radius of 23.6 cm and an outer radius of 80.9 cm. Since the events will be boosted in the forward direction, the DCH is optimized to minimize the amount of material in the forward end and is offset from the IP, as can be seen in Figure 2.10. It is comprised of 40 cylindrical layers of small hexagonal cells, thus providing up to 40 measurements of position and dE/dx for charged particles with momentum larger than 180 MeV/c. The layers are grouped as sets of four into ten superlayers, which are sequentially staggered by half a cell and alternate between axial (A) and stereo (U,V) pairs, in the order AUVAUVAUVA, as shown in Figure 2.11. The axial superlayers have sense wires arranged parallel to the z axis, whereas the stereo superlayers have a nonzero stereo angle, ranging

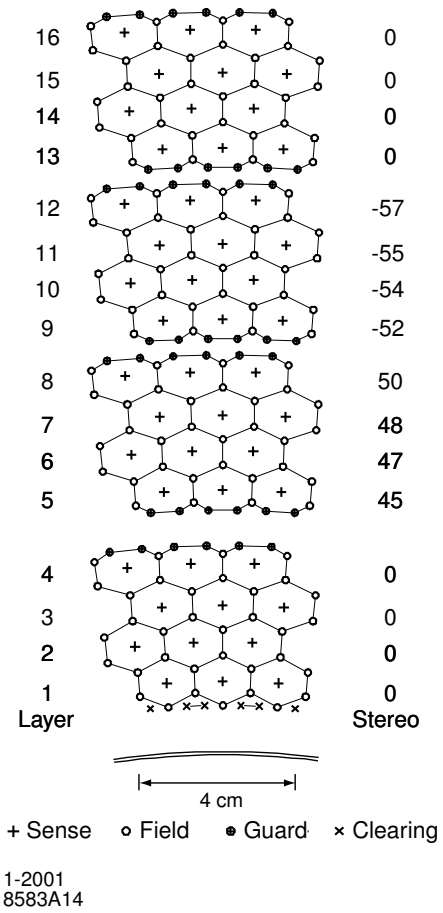


Figure 2.11. Schematic layout of DCH drift cells for the first four superlayers. Numbers on the right denote the stereo angles (mrad) of sense wires for each layer. Lines connecting the field wires have been drawn to help visualize cell boundaries.

between ± 45 mrad and ± 76 mrad, which allows us to measure the radius and azimuthal angle of tracks, in addition to z.

There are a total of 7,104 DCH drift cells, each of which consists of a sense wire surrounded by six field wires. The field wires are held at ground potential, whereas the sense wires have a positive high voltage applied to them. The cells are approximately hexagonal so that a near circular symmetry can be achieved close to the center the cell. The isochrones, contours of constant drift times which are circular near the sense wires, and drift paths for ions in two cells are shown in Figure 2.12.

One of the main limits on tracking resolution comes from multiple-scattering in the

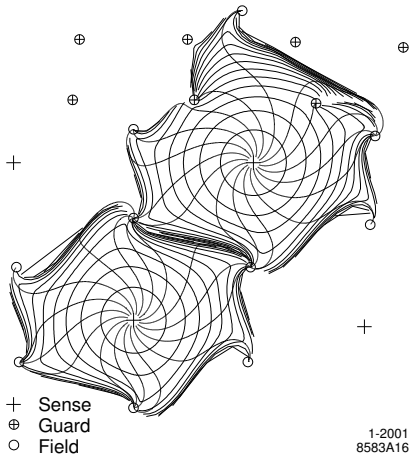


Figure 2.12. DCH drift cell isochrones in adjacent cells of layers 3 and 4 of an axial superlayer. The isochrones shown are separated by 100 ns.

DCH. To minimize such limitations, low-mass aluminum field wires and a helium-based gas mixture is used in the DCH. In addition, the material of the DCH is minimized to prevent reducing the performance of the DIRC and EMC, which are located just outside the DCH.

2.5 Detector of Internally Reflected Cherenkov Light

The detector of internally reflected Cherenkov light (DIRC) is the primary component of the *BABAR* particle identification (PID) system, providing information for discrimination of particles of different mass. A schematic of the longitudinal cross-section of the DIRC is shown in Figure 2.13. The ability to distinguish between kaons and pions, for instance, is crucial to the study of $B^0 \to D_s^{(*)+}\pi^-$ decays, for which $B^0 \to D_s^{(*)-}K^+$ decays are a significant background, and vice versa. Likewise, the ability to flavor tag one of the B mesons through its decay products is crucial for time-dependent *CP* analyses, such as the study of $B^0 \to D^{\pm}\pi^{\mp}$ decays used for measuring $\sin(2\beta+\gamma)$.

Charged particles with momenta that exceed the Cherenkov threshold will emit a cone of light with an angle θ_c with respect to the particle trajectory. This angle is related to the velocity of the particle as $\cos\theta_c = 1/n\beta$, where n is the index of refraction of the material

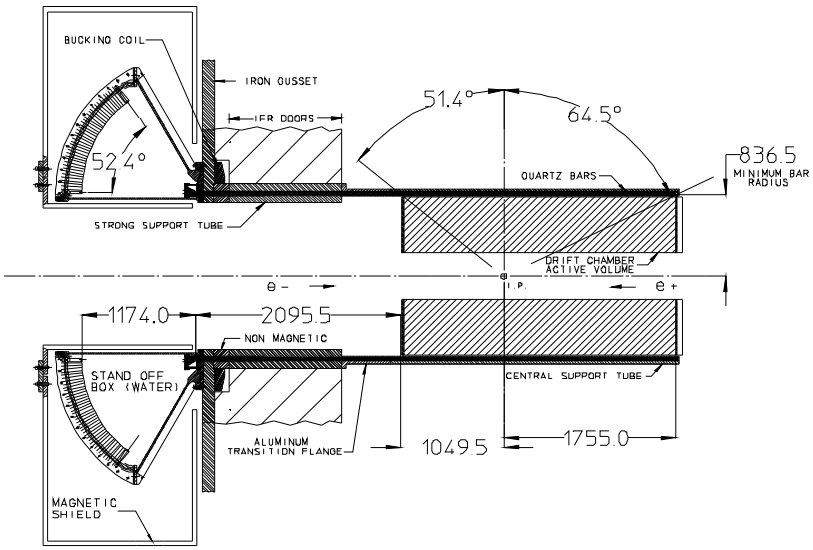


Figure 2.13. DIRC, longitudinal section.

that the particle is traversing, $\beta = v/c$, v is the speed of the particle, and c is the speed of light. The combined knowledge of the speed of the particle from the DIRC and the momentum information from the DCH and SVT is used to deduce the mass of the particle.

The DIRC is a novel ring-imaging Cherenkov detector. Its design is based on the principle that charged particles will emit Cherenkov light above a certain threshold and that the magnitudes of the Cherenkov angles will be preserved after reflection from a flat surface. Figure 2.14 is a schematic of the DIRC that depicts the process that occurs to transport and measure the Cherenkov light produced in the DIRC. Cherenkov light is produced in thin, 4.9 m long bars made of synthetic fused silica bars of rectangular cross section and dimensions 1.7 cm $\times$ 3.5 cm. These bars not only act as the radiators for the DIRC but also transport the light by total internal reflection to the instrumented end of the detector.

The DIRC bars are individually contained in 12 hermetically sealed boxes, which are housed in a 12-sided polygonal barrel. The DIRC was designed to be thin and uniform to prevent degradation of EMC performance. In addition, the small radius of the DIRC keeps its volume minimal and thus the size and cost of the EMC at a minimum. A mirror is

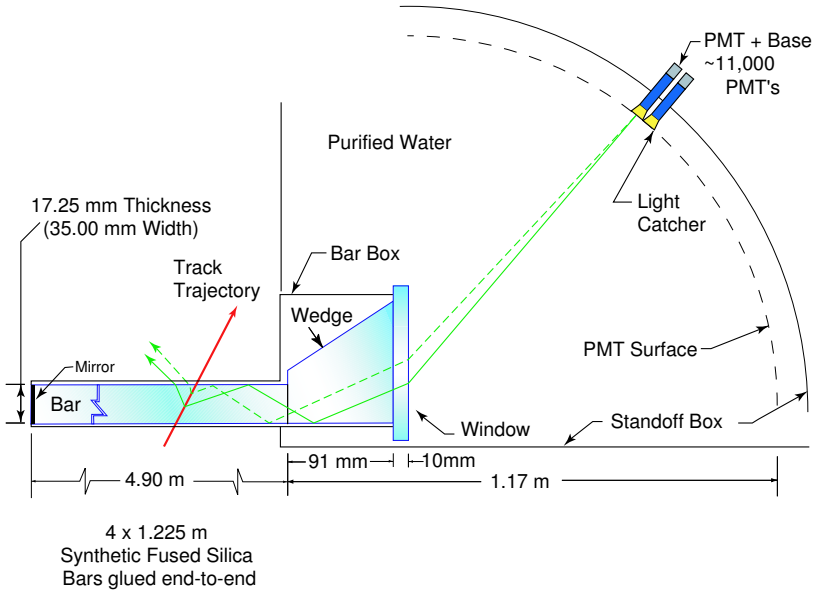


Figure 2.14. Schematic of the DIRC, depicting how Cherenkov light is produced, transported in the silica bars to the standoff box, and then detected by PMTs.

placed at the forward end of the DIRC to reflect incident photons to the backward end, so that only one end of the detector has to be instrumented with photon detectors.

At the instrumented end, the photons emerge into the standoff box, an expansion region filled with $6\,\mathrm{m}^3$ of water. In order to minimize the size of the detection surface and also recover photons that would be lost to internal reflection at the silica-water junction, a fused silica wedge is placed at the exit of the bar. The photons are then detected by a set of densely arranged photo-multiplier tubes (PMTs) about 1.2 m from the end of the bars. Each of the PMTs is also encased by reflecting light catcher cones to catch the light that would otherwise miss the active area of the PMT.

For each charged track, the DIRC reconstruction provides an estimate of the Cherenkov angle and the error, as well as a confidence level for the different mass hypotheses (e, μ, π, K, and p). The DIRC produces π-K separation of 4σ between the pion Cherenkov threshold up to $4.2\,\mathrm{GeV}/c$ for all tracks from the B decays. As mentioned earlier, PID information for particles with momentum lower than $700\,\mathrm{MeV}/c$ comes primarily from dE/dx measurements

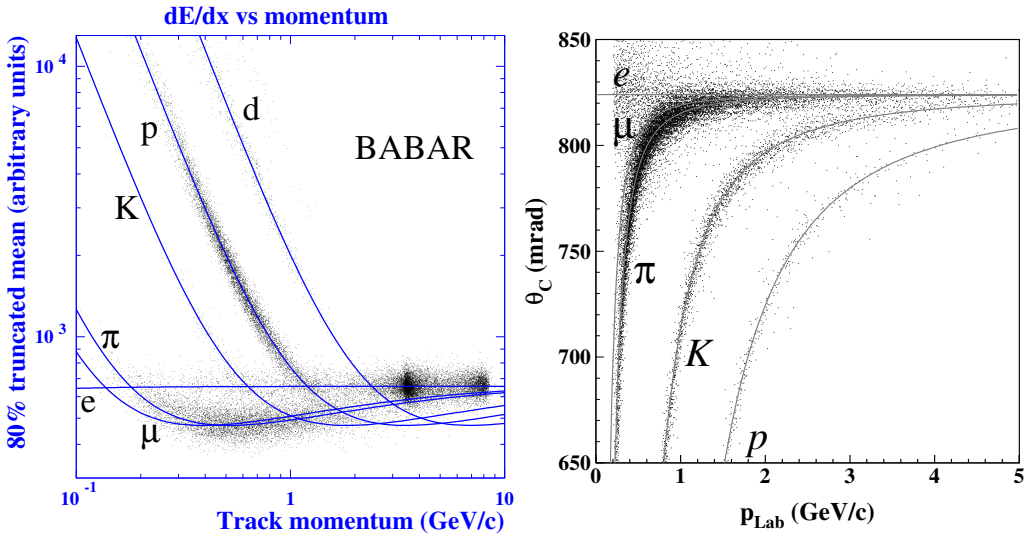


Figure 2.15. Performance of the *BABAR* PID system. dE/dx as a function of track momentum from the DCH using beam scan data, with overlaid parameterized Bethe-Block curves (left). Cherenkov angle as a function of momentum from the DIRC for an inclusive sample of multi-hadron events (right).

from the tracking system. Figure 2.15 shows particle identification performance plots for the DCH and DIRC.

2.5.1 Pion and Kaon Identification

The identification of pions and kaons is very important for *CP* studies that use pions and kaons for tagging, as well as for analyses that fully reconstruct B decays. The measurements of the branching ratios of $B^0 \to D_s^{(*)+}\pi^-$ and $B^0 \to D_s^{(*)-}K^+$ require very good π-K identification to reject cross-feeds as well as backgrounds from other B decays.

The PID requirements in this analysis use the method of likelihood selectors. The likelihood $\mathcal{L}_i$ is calculated for each particle hypothesis i:

$$\mathcal{L}_i = \mathcal{L}_{i\mathrm{DIRC}} \cdot \mathcal{L}_{i\mathrm{DCH}} \cdot \mathcal{L}_{i\mathrm{SVT}} \tag{2.1}$$

where the likelihoods from the DCH and SVT arise from comparing the measured dE/dx against the expected dE/dx from the Bethe-Bloch parametrization [20]. The DIRC like-

lihood is constructed from the Cherenkov angle, the number of photons, and the track quality. The likelihood selectors consist of different cuts on the likelihood ratio, such as $r_{K-\pi} = \mathcal{L}_{kaon}/(\mathcal{L}_{kaon} + \mathcal{L}_{pion})$ and $r_{K-p} = \mathcal{L}_{kaon}/(\mathcal{L}_{kaon} + \mathcal{L}_{proton})$. The different selectors are designed for different efficiency and mis-identification requirements.

For the kaon selectors, the requirements are:

- `NotPion`: $r_{K-\pi} > 0.20$ or $r_{p-\pi} > 0.20$
- `VeryLoose`: $r_{K-\pi} > 0.50$, $r_{K-p} > 0.018$, and $p < 0.40$ or does not pass tight electron selector
- `Loose`: $r_{K-\pi} > 0.8176$, $r_{K-p} > 0.2$, and $p < 0.40$ or does not pass tight electron selector
- `Tight`: $r_{K-\pi} > 0.90$, $r_{K-p} > 0.2$, and $p < 0.40$ or does not pass tight electron selector
- `VeryTight`: $r_{K-\pi} > 0.90$, and $p < 0.40$ or does not pass tight electron selector, and does not pass `VeryTight` muon selector

For the pion selectors, the requirements are:

- `VeryLoose`: $r_{K-\pi} < 0.98$, $r_{K-p} < 0.98$
- `Loose`: $r_{K-\pi} < 0.82$, $r_{K-p} < 0.98$
- `Tight`: $r_{K-\pi} < 0.5$, $r_{K-p} < 0.98$
- `VeryTight`: $r_{K-\pi} < 0.2$, $r_{K-p} < 0.5$, and does not pass `VeryTight` muon selector

These selectors are used in this analysis to discriminate against backgrounds, as described in Section 4.6. Of course, tighter requirements on PID may produce a cleaner sample, but lower efficiencies, so the requirements have been optimized to discriminate against backgrounds without sacrificing signal efficiency. The efficiencies and mis-identification rates

depend on the momentum range of interest. As an example, a `Tight` positive kaon identification is required for the kaon daughters of the D_s^+ for the $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$ and $D_s^+ \to \bar{K}^0K^+$ modes. This requirement has an efficiency of 85% and a pion misidentification rate of 5%.

2.6 Electromagnetic Calorimeter

The purpose of the electromagnetic calorimeter (EMC) is to detect electromagnetic showers with high efficiency and excellent energy and angular resolution, for energies of 20 MeV to 4 GeV. This range covers the detection of photons from low energy π^0 and η, as well as higher energy photons and electrons from electromagnetic, weak, and radiative processes. The tightest requirements for EMC energy resolution come from the study of rare B decays containing π^0, such as $B^0 \to \pi^0\pi^0$. For energies below 2 GeV, the π^0 mass resolution is dominated by the EMC energy resolution, and at higher energies, it is dominated by the angular resolution.

The main component of the EMC is a finely segmented array of 6,580 thallium-doped cesium iodide (CsI(Tl)) crystals. The EMC is divided into a cylindrical barrel section and a conical forward endcap. The barrel is located asymmetrically about the IP and has an

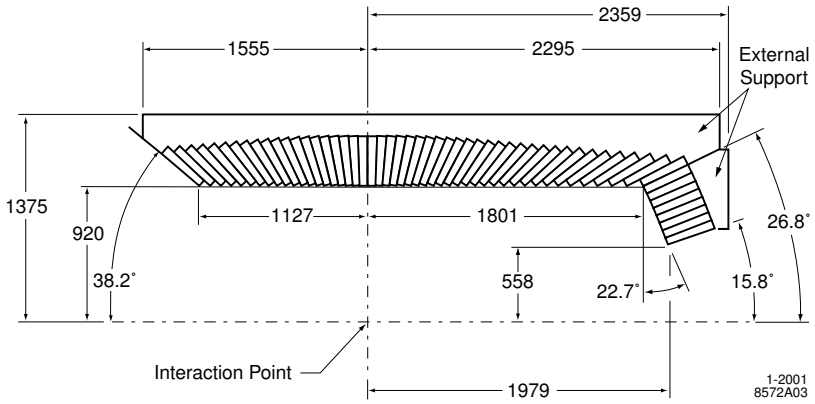


Figure 2.16. EMC, longitudinal section (top half only). Dimensions given in mm.

inner radius of 91 cm and an outer radius of 136 cm. The barrel contains 5,760 crystals arranged in 48 distinct rings with 120 identical crystals in each ring. The endcap has 820 crystals arranged in 8 rings. Figure 2.16 depicts a schematic cross-section of the EMC and shows the arrangement of the crystal rings. The barrel and endcap are each comprised of modules made of carbon-fiber epoxy composite. The modules are supported in the rear to minimize the material in front of the crystals. The barrel is divided into 280 modules, each with 21 crystals, and the endcap is made from 20 identical modules, each with 41 crystals.

Thallium-doped CsI crystals were chosen because their high light yield and small Molière radius provides excellent resolution, and the short radiation length allows us to contain showers while maintaining a compact design. The crystals, which have a tapered trapezoidal cross-section, increase in length toward the forward direction to minimize the effects of shower leakage. The crystals are supported at the outer radius, with a thin gas seal at the front, to limit pre-showering. The crystals serve as both the total absorption scintillating medium and as a light guide. Silicon photodiodes, matched to the spectrum of scintillation light, serve as the read-out for the crystals.

Precise and frequent calibrations of the electronics and energy response must be executed to guarantee accurate, optimal performance of the EMC. The electronics is calibrated using a charge injection system to linearize the response of the front-end electronics to better than 0.1%. Calibration of the energy response comes from a number of systems. A liquid radioactive source system uses 6.13 MeV photons from ^{16}N $\beta - \gamma$ cascades to set the initial energy scale per crystal to better than 0.5% and also to monitor long term, absolute changes in light collection. A light pulser system, which measures the response of individual crystals, is used to monitor short term changes to better than 0.5%. Lastly, a number of physics processes, such as Bhabha scattering events, are used to determine the energy scale for individual crystals to better than 0.25% and clusters to better than 0.5%.

2.7 Instrumented Flux Return

The Instrumented Flux Return (IFR) serves as the muon and neutral hadron detector for *BABAR*. It was designed to identify muons with large solid angle coverage, high efficiency, and good background rejection. Also, in conjunction with the calorimeter, the IFR is also used to detect neutral hadrons, such as K_L^0, over a large range of momenta and angles with high efficiency and angular resolution. The identification of muons is a crucial part of flavor-tagging neutral B mesons through semi-leptonic decays, reconstructing vector mesons, and also studying semi-leptonic and rare decays of the B, D, and τ. The detection of K_L^0 is important for the study of exclusive B decays.

The IFR system consists of the steel flux return of the magnet, which is segmented and instrumented with resistive plate chamber (RPC) detectors. The IFR is comprised of a barrel section and two end cap plugs. The steel is segmented into 18 plates, such that the thickness increases from 2 cm for the inner plates and 10 cm for the outer plates of the barrel. This graded segmentation was chosen to optimize K_L^0 detection and low momentum muon identification, without excessively increasing the number of layers. The RPCs are located in the gaps of the finely segmented steel of the barrel and end doors of the flux return. Figure 2.17 depicts a schematic of the IFR barrel and endcaps and shows the layout of the RPC modules. There are 19 RPC layers in the barrel and 18 layers in the endcaps, along with two layers of cylindrical RPCs surrounding the EMC to detect particles leaving the calorimeter.

RPCs detect streamers from ionizing particles as they pass through a gas filled chamber and produce a spark, and the signal is then read through capacitive readout strips. RPCs were chosen for their low cost and ability to cover many shapes, decreasing inactive space, in addition to large signals and fast response time. The planar RPCs consist of two 2 mm thick bakelite plates separated by 2 mm. The inner bakelite surfaces are treated with linseed oil to increase efficiency and decrease noise. The outer surfaces are covered with high surface resistivity graphite and then protected by an insulating film. One outer surface is connected

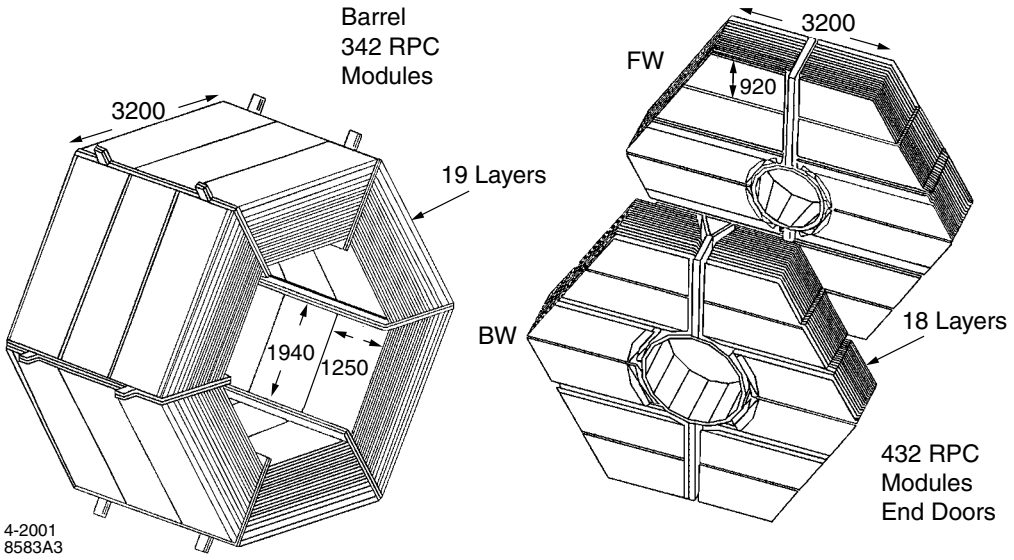


Figure 2.17. Schematic of the IFR: Barrel (left) and forward (FW) and backward (BW) end doors (right).

to high voltage while the other is kept at ground. The active volume between the plates is filled with a gas mixture of argon, freon, and a small amount of isobutane. The signals are read out from both surfaces with strip electrodes, which run lengthwise and crosswise along the RPC.

In total, there are 806 RPC modules. The barrel is divided into six sections of 57 RPC modules each. Each barrel module has 32 longitudinal strips running perpendicular to the beam axis for measuring z and 96 strips placed orthogonally extending over three modules for measuring ϕ. The endcaps are hexagonal and are divided vertically in half to act as doors allowing access to the inner detectors. Each of the four endcap half-sections half 108 RPC modules. The readout strips in the endcaps run horizontal and vertical readout strips for measuring x and y. Lastly there are 32 RPC modules in the two inner cylindrical layers. The inner layer has helical $u - v$ strips running parallel to the diagonals of the module and the outer layer has strips running parallel to z and ϕ.

Currently, the IFR barrel is being upgraded, replacing the barrel RPCs with limited-streamer tube (LST) technology. A decision was made to upgrade the IFR since the per-

formance of the RPCs was steadily declining with time, and the detection of muons and neutral hadrons is crucial to many *BABAR* analyses. The system was designed with a high degree of reliability and redundancy, and easy maintainability in mind. LSTs were chosen for their record of satisfactory performance with high efficiency and reliability, as evidenced by many other experiments. An LST cell consists of a silver-plated sense wire $100\mu m$ in diameter, located at the center of a cell of 9 mm square section. An extruded plastic structure, or "profile", contains 8 such cells and is open on one side. The profile is coated with a resistive layer of graphite and strung with wires, and then inserted in plastic tubes to contain the gas mixture. The *BABAR* detector continues to be upgraded with LSTs.

2.8 Trigger

At a luminosity of $10 \times 10^{33}\,\mathrm{cm}^{-2}s^{-1}$, PEP-II produces about 10 $e^+e^- \to \Upsilon(4S)$ events per second. A two-level trigger system is used to maximize the acceptance of physics data with respect to background, filtering out events that are no interesting for physics studies while maintaining a large acceptance of physics events.

The first-level, hardware-implemented trigger (L1) receives detector signals and removes beam-induced backgrounds. The L1 trigger reduces millions of events per second down to a design output rate of $\lesssim$ 2 kHz. This system uses information from the DCH, EMC, and the IFR, in addition to a Global Level 1 Trigger (GLT) which forms 24 trigger lines and decides whether to pass the events to the second trigger stage.

The second-level software trigger (L3) operates on a computing farm with an output rate of $\lesssim$ 200 Hz. The L3 trigger uses information from the DCH and EMC to form track and cluster objects, which are filtered to eliminate backgrounds such as tracks that don't arise from the interaction region. Events that pass the L3 trigger are stored as xtc files for later offline physics analysis.

The trigger is $\sim$ 99% efficient in selecting $e^+e^- \to \Upsilon(4S) \to B\bar{B}$ events. For $B^0\bar{B}^0$

events, each trigger is 99% efficient and > 99.9% efficient combined. The trigger also accepts other processes such as $e^+e^- \to q\bar{q}, \mu^+\mu^-, \tau^+\tau^-$, as well as $e^+e^- \to e^+e^-$ Bhabha scattering events and other physics processes.

2.9 Offline Processing

After passing through the trigger system, the data is sent through a set of offline filters before being fully reconstructed [21]. The first of these filters is the DigiFilter, which requires no reconstruction and uses information from the L1 and L3 triggers. The DigiFilter primarily filters out calibration events. The events that pass this filter are then sent through the BGFilter, which implements the first step of reconstruction. This involves track finding in the DCH and cluster finding in the EMC. The BGFilter classifies events based on these tracks and clusters and divides them into multi-hadron, $e^+e^- \to \mu^+\mu^-$, $e^+e^- \to \tau^+\tau^-$, two-prong events, two-photon events, radiative Bhabha events, etc. This information, in conjunction with information from L1 and L3 triggers, is then used to choose the events that will be fully reconstructed. Approximately 35% of the events written to xtc files are fully reconstructed.

Once the events are fully reconstructed, they are written to the database. The events that have been reconstructed are then "skimmed" into different types of physics events, to ease processing time for physics analyses. The skim used for this analysis is `BRecoToDsLight`, which uses a set of criteria, for instance on the mass of the D_s^+ and the momenta of the B^0 daughters, to filter $B^0 \to D_s^{(*)+}X,\ X = \pi^-, K^-, \rho^-, etc.$ events. Both data and Monte Carlo simulated events are passed through the skimming process.

The skimmed events are then processed with a combinatorics algorithm, which combines composite particles to form lists of particle "candidates". At this point, we apply a loose set of selection criteria, and vertex and kinematic constraints are also applied to improve position and four-momenta measurements. The data or Monte Carlo are then written to ROOT ntuples, which are used to refine and tighten the final selection criteria and then

extract the signal and background yields. These final analysis steps are described in the following chapters.

Chapter 3

Analysis Overview

This study measures the rates for rare decays with branching fractions on the order of $\mathcal{O}(10^{-5})$. The rareness of these decays poses a number of challenges for this analysis, the most difficult being the suppression of background events with respect to the signal. Background events can arise from random combinations of tracks from continuum events, as well as mis-reconstructed B decays that can be mistaken for signal events. The need to suppress such backgrounds requires that a relatively tight set of criteria is applied to the data sample in order to increase the statistical significance of the measurement. The significance corresponds to the probability that the signal measurement is not due to a statistical fluctuation of the background.

BABAR analyses are done in a "blind" manner, such that the signal region in the data sample is not revealed until the analysis is nearly complete and the selection criteria have been finalized. We conduct analyses in this manner to prevent bias arising from defining the signal region and selection criteria with prior knowledge of the experimental outcome. Because our analyses are blind, we rely on Monte Carlo simulation samples to study the selection efficiencies and background contamination before unveiling the signal region in data. We also use the data regions outside of the signal region to study background rates and detector resolutions before unblinding.

In this analysis, the initial data sample undergoes a preselection process using a set of loose selection criteria to skim the data set into events that may contain B decays of interest to our study. B mesons are then reconstructed in the modes $B^0 \to D_s^+\pi^-$, $B^0 \to D_s^{*+}\pi^-$, $B^0 \to D_s^- K^+$, and $B^0 \to D_s^{*-} K^+$ by combining the tracks and neutrals seen in the detector to build composite particles, such as B^0 and D_s^+. These decay channels are fully reconstructed, meaning that all of the final state particles are detected. After the particles are reconstructed, we apply a tighter set of selection criteria to decrease the number of background events and increase the expected signal to background ratio. Before unblinding the signal region in data, the number of signal and background events is estimated by using a large set of simulated events, as well as the data events outside of the signal box.

Once the selection is finalized, the data set is unblinded. The final data set, after applying the final selection criteria, is fitted to extract the signal yields and compute the branching fractions of the decay modes under study. We perform the yield extraction using a multi-dimensional unbinned maximum likelihood fit that simultaneously fits for all the D_s^+ modes. The method of maximum likelihoods is briefly reviewed in Appendix B. We perform all maximum likelihood fits using the RooFit package [22].

The steps described above are explained in detail in the following chapters. Chapter 4 describes the reconstruction, event pre-selection, selection optimization, and final candidate selection. Chapter 5 details the final yield and branching fraction measurement and the systematic errors.

Chapter 4

Event Selection

4.1 Data and Monte Carlo Samples

This study uses 208.7 fb^{-1} of *BABAR* data recorded at the $\Upsilon(4S)$ resonance. The data were recorded between 22 October 1999 and 31 July 2004, which consists of Runs 1-4 of the *BABAR* data set. This sample corresponds to 229,786,006 $B\bar{B}$ events.

The number of $B\bar{B}$ events is determined by counting the number of hadronic events in off-resonance and on-resonance data, assuming that the increase in the ratio between the number of hadronic events and the number of muon pairs between off-resonance to on-resonance data is due to $\Upsilon(4S)$ production [23]. The number of $\Upsilon(4S)$ events is then given by:

$$N_{\Upsilon(4S)} = N_{\text{hadronic}} - N_{\mu\mu} \cdot R_{\text{off}} \cdot \kappa, \tag{4.1}$$

where $N_{\Upsilon(4S)}$ is the number of hadronic events selected in the on-resonance sample, N_{hadronic} is the number of hadronic events selected in the on-resonance sample, $N_{\mu\mu}$ is the number of muon pairs selected in the off-resonance sample, R_{off} is the ratio between the number of hadronic events and the number of muon pairs in the off-resonance sample, and κ takes into account the dependence of the cross-section on the center-of-mass energy and any variation in selection efficiency. The true number of $\Upsilon(4S)$ mesons produced depends on the efficiency

of the hadronic selection of $B\bar{B}$ events, which is measured using Monte Carlo simulations. The number of $B\bar{B}$ events is then proportional to the number of $\Upsilon(4S)$ mesons produced.

For our Monte Carlo simulated samples, the detector response is simulated by the GEANT4 package [24]. GEANT4 models the interactions of particles traversing the detector, taking into account the varying detector conditions and beam backgrounds. Monte Carlo events are subject to the same reconstruction and subsequent event algorithm as the data.

The simulated samples used for this analysis include generic and exclusive MC samples. The generic MC samples contain a wide variety of decays and are mainly used to simulate combinatorial background events. The exclusive MC samples consist of events in which one of the B mesons is forced to decay into a particular decay channel, while the other B from the $\Upsilon(4S)$ decays generically. For instance, signal MC, in which the decays are generated in the signal modes, $B^0 \to D_s^+\pi^-$, $B^0 \to D_s^{*+}\pi^-$, $B^0 \to D_s^-K^+$ and $B^0 \to D_s^{*-}K^+$, are generated to emulate the signal events. There are also many exclusive MC samples that simulate backgrounds that arise from mis-reconstructed B decays. The generic and exclusive MC samples used in this study are listed in Appendix A.

4.2 Event Pre-selection

The event pre-selection first selects only multi-hadron events, since many of the events stored by the online data acquisition system come from beam-gas or other interactions that are not e^+e^- collisions. Multi-hadron events are required to have a minimum of three charged tracks in the fiducial region $0.41 < \theta_{lab} < 2.54$, where θ_{lab} is the polar angle of the track in the lab frame. These tracks are required to originate within 1.5 cm in the x-y plane (transverse to the beam axis) and 10 cm in z (along the beam axis) of the beam spot position. The tracks must also be reconstructed in the DCH. The primary vertex constructed from the tracks must be within 0.5 cm of the average IP position in the x-y plane and 6 cm in z. In addition, tracks with a large contribution to the χ^2 of the primary

vertex fit are removed until the χ^2 is greater than 1% or until only two tracks are left. Bumps, or local energy maxima, in the EMC not associated with charged tracks, with an energy greater than 30 MeV in the fiducial volume $0.41 < \theta_{lab} < 2.409$ and shower shape consistent with photon interactions, are assumed to be neutral candidates. Charged tracks and neutral candidates must have a total energy in the fiducial regions larger than 4.5 GeV.

One of the main backgrounds in our study are events in which random track combinations from *continuum* $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$) or generic $e^+e^- \to B\bar{B}$ are mistaken as signal events. Event topology can be utilized to discriminate against such backgrounds since the shapes of $q\bar{q}$ events differ from the the shapes of $B\bar{B}$ events. For signal events, an e^+e^- pair produces a $\Upsilon(4S)$, which decays into a $B\bar{B}$ pair. The B mesons have low momenta in the $\Upsilon(4S)$ frame, and the decay of each B is nearly isotropic. In addition, for a signal event, there is no correlation between the directions of the decay daughters coming from the B mesons. However, in a continuum $e^+e^- \to q\bar{q}$ event, the event is "jet-like", such that the event is characterized by a direction, called the jet axis. These types of events tend to be less isotropic in the $\Upsilon(4S)$ frame than real B events. Moreover, the directions of the decay products of the fake B mesons candidates from such continuum events tend to be correlated, lying within the two "jets".

R_2, the normalized second Fox-Wolfram moment of the event, is one such event shape variable used to separate background from signal events [25]. The l^{th} Fox-Wolfram moment, H_l, is the momentum-weighted sum of the l^{th} order Legendre polynomial computed from the cosine of the angle between all pairs of tracks:

$$H_l = \sum_{i,j} \frac{|p_i||p_j| P_l(\cos\theta_{ij})}{E_{vis}^2}, \tag{4.2}$$

where i and j are summed over all tracks, P_l is the l^{th} Legendre polynomial, $p_{i,j}$ is the momentum for track i, j, θ_{ij} is the opening angle between tracks i and j, and E_{vis} is the visible energy of the event. Energy-momentum conservation requires that $H_0 = 1$, neglecting the particle masses. For continuum $e^+e^- \to q\bar{q}$ events, $H_1 = 0$, $H_l \sim 1$ for even l and $H_l \sim 0$

for odd l. R_2 is defined as the ratio H_2/H_0. For the pre-selection an R_2 requirement of < 0.5 is applied, along with the multi-hadron selection.

The efficiency for selecting multi-hadron $B\bar{B}$ events with an $R_2 < 0.5$ is about 95%.

4.3 Intermediate Meson Reconstruction

Particles are reconstructed by combining decay daughters. Then we perform vertex and kinematic fits to improve position, energy, and momentum measurements. These methods help in dealing with complex decay chains in a straightforward way. Composite particles are constructed from their daughter particles, and then the composite particles replace the original particles in the subsequent fits and reconstruction procedures.

Kinematic fitting and vertexing are mathematical procedures in which physical constraints governing a particle interaction or decay are used to improve the measurements describing the process. Kinematic fitting uses kinematic constraints, such as invariant masses and energy-momentum conservation, to improve the measurements of the decay. Vertexing is the processing of accurately determining the three-dimensional point of intersection of a set of tracks. The best vertex hypothesis is determined by minimizing the sum of the least squares of the distance of closest approach of a set of tracks to a point. The problem is non-linear due to the curvature of the charged tracks in a magnetic field, but the problem is linearized to find a local solution. This process is iterated until it converges such that the χ^2 difference between two consecutive iterations is less than 0.01, with the maximum number of allowed iterations being six.

All final state particles are reconstructed for this study. The B mesons are reconstructed as $B^0 \to D_s^+\pi^-$, $B^0 \to D_s^{*+}\pi^-$, $B^0 \to D_s^-K^+$, and $B^0 \to D_s^{*-}K^+$. The D_s^{*+} candidates are reconstructed in the mode $D_s^{*+} \to D_s^+\gamma$. The D_s^+ candidates are reconstructed in the modes $D_s^+ \to \phi\pi^+$, $D_s^+ \to \bar{K}^{0*}K^+$, $D_s^+ \to \bar{K}^0K^+$. The ϕ, $\bar{K}^{0*}$, and $\overline{K}^0$ candidates are

Table 4.1. D_s^+ and D_s^{*+} decay modes and branching fractions used in this analysis. The middle column shows the branching fractions for the D_s^+ decay. The last column shows the branching fractions for the full sub-decay chain, including the branching fractions for the decays of the D_s^+ daughters.

Decay Chain	D_s^+ Decay $\mathcal{B}(\%)$	Full Decay $\mathcal{B}(\%)$
$D_s^{*+} \to D_s^+\gamma$	94.2 ± 2.5	-
$D_s^+ \to \phi\pi^+$, $\phi \to K^+K^-$	4.81 ± 0.64	2.36 ± 0.32
$D_s^+ \to \bar{K}^{0*}K^+$, $\bar{K}^{*0} \to K^-\pi^+$	-	2.67 ± 0.80
$D_s^+ \to \bar{K}^0K^+$, $\bar{K}^0 \to K_S^0$, $K_S^0 \to \pi^+\pi^-$	4.81 ± 1.47	1.66 ± 0.51

reconstructed as $\phi \to K^+K^-$, $\bar{K}^{0*} \to K^-\pi^+$, and $\overline{K}^0 \to K_S^0$, $K_S^0 \to \pi^+\pi^-$. The daughter decay chains and the associated branching fractions are shown in Table 4.1 [12].

The following requirements for the tracks and composite particles are used for reconstructing B mesons:

- Tracks: Every charged track is required to have momentum less than 10 GeV/c, a distance of closest approach to the beamspot less than 1.5 cm in the x-y plane, and a distance of closest approach to $z = 0$ of less than 10 cm.

- $\phi \to K^+K^-$: Candidates are made from pairs of oppositely charged tracks which pass the `NotPion` selection. The invariant mass of the two tracks is required to be within $\pm 30\,\mathrm{MeV}/c^2$ of the nominal ϕ mass of $1019\,\mathrm{MeV}/c^2$, which corresponds to about seven times the width of the ϕ resonance. The probability of the χ^2 of the vertex fit must be greater than 0.1%.

- $\bar{K}^{*0} \to K^-\pi^+$: Candidates are reconstructed from pairs of oppositely charged tracks with one of them passing the `NotPion` selection. The candidates must have an invariant mass within $\pm 75\,\mathrm{MeV}/c^2$ of the nominal $\bar{K}^{*0}$ mass of $896.10\,\mathrm{MeV}/c^2$, which corresponds to about 1.5 times the width of the $\bar{K}^{*0}$ resonance. The probability of the χ^2 of the vertex fit must be greater than 0.1%.

- $K_S^0 \to \pi^+\pi^-$: Candidates are reconstructed by combining pairs of oppositely charged tracks with an invariant mass window of $\pm 25\,\mathrm{MeV}/c^2$ of the nominal K_S^0 mass of

497.648 MeV/c^2. The probability of the χ^2 of the vertex fit must be greater than 0.1%.

- $D_s^+ \to \phi\pi^+$, $D_s^+ \to \bar{K}^{0*}K^+$, and $D_s^+ \to \bar{K}^0K^+$: Candidates are reconstructed from combinations of decay daughters that lie within an invariant mass window of $\pm 40\,\mathrm{MeV}/c^2$ of the nominal D_s^+ mass of $1968.3\,\mathrm{MeV}/c^2$. In addition, D_s^+ candidates are required to have a momentum of $p^* > 1.6\,\mathrm{GeV}/c^2$.

- $D_s^{*+} \to D_s^+\gamma$: Candidates are reconstructed from a D_s^+ candidate and a photon , with an invariant mass window of $\pm 500\,\mathrm{MeV}/c^2$ of the nominal D_s^{*+} value or $2112.1\,\mathrm{MeV}/c^2$. In addition, $\Delta M_{D_s^*}$, the difference between the masses of the D_s^{*+} and D_s^+ candidates, is required to be between 130 and $160\,\mathrm{MeV}/c^2$. The energy of the photon candidate is required to be greater than 100 MeV.

A plot of the invariant mass of $D_s^+ \to \phi\pi^+$ is shown in Figure 4.1. Plots of the other invariant masses can be found in Appendix C. It should be emphasized that these requirements are for candidate pre-selection and do not reflect the final set of selection criteria applied to choose the set of events used for the branching fraction measurement. We optimize the final criteria to produce the maximal possible significance, as described in Section 4.6.

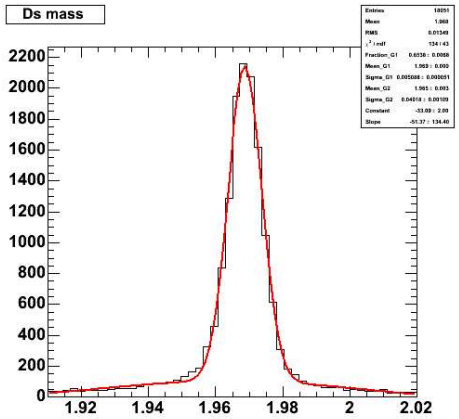


Figure 4.1. Mass of D_s^+ for $D_s^+ \to \phi\pi^+$mode from MC.

4.4 B^0 Reconstruction

B^0 mesons are reconstructed by combining a D_s^+ or D_s^{*+} candidates with a track of opposite charge. The D_s^+ and D_s^{*+} masses are constrained to the nominal values when reconstructing the B^0. The high momentum track used for the daughter of the B^0 is required to have a CM momentum of $p^* > 0.5\,\mathrm{GeV}/c$. The B^0 candidates are determined to be $B^0 \to D_s^{(*)+}\pi^-$ ($B^0 \to D_s^{(*)-}K^+$) if the charged track passes the `Tight` pion (kaon) selector criterion.

Two important kinematic variables, ΔE and m_{ES}, are used to select B candidates. These variables utilize the kinematic constraint from the initial $\Upsilon(4S)$ decay into a $B\bar{B}$ pair to discriminate against background and define the signal region. ΔE is defined in a Lorentz invariant way as:

$$\Delta E = (2q_B q_0 - s)/2\sqrt{s}, \tag{4.3}$$

where $\sqrt{s}$ is the total e^+e^- CM energy, q_B and q_0 are the Lorentz vectors representing the four-momenta of the B candidate and the e^+e^- system, respectively. In the CM frame, ΔE can be expressed as:

$$\Delta E = E_B^* - E_{beam}^*, \tag{4.4}$$

where $E_{beam}^* = \sqrt{s}/2$ and E_B^* is the B^0 candidate energy in the CM frame. In this form it is clear that ΔE corresponds to the difference between the reconstructed and expected energy of the B candidate in the CM frame, which peaks around zero for properly reconstructed signal events.

m_{ES} is the beam-energy-substituted mass:

$$m_{\mathrm{ES}} = \sqrt{(s/2 + \mathbf{p}_0 \cdot \mathbf{p}_B)^2/E_0^2 - \mathbf{p}_B^2}, \tag{4.5}$$

where $(E_0, \mathbf{p}_0)$ is the four-momentum of the initial e^+e^- system and $\mathbf{p}_B$ is the B^0 candidate momentum, both measured in the laboratory frame. In the CM frame, m_{ES} can be expressed as:

$$m_{\mathrm{ES}} = \sqrt{E_{beam}^* - p_B^{*2}}, \tag{4.6}$$

where p_B^* is the CM momentum of the B candidate, computed from the momenta of its decay products. $m_{\rm ES}$ for signal events peaks around the nominal B^0 mass of $5.279\,{\rm GeV}/c^2$.

ΔE and $m_{\rm ES}$ together form a nearly orthogonal set of variables that are used to define the signal region. By definition, they are dependent and not orthogonal. However, because the sources of experimental smearing that contribute to their resolutions are uncorrelated, ΔE and $m_{\rm ES}$ are practically uncorrelated. The ΔE resolution is dominated by detector resolution, while the resolution for $m_{\rm ES}$ is dominated by the spread of the beam energy.

Since the ΔE and $m_{\rm ES}$ distributions for background can differ significantly from signal, applying a cut on ΔE or $m_{\rm ES}$ can help in discriminating against background events. The ΔE distribution for signal events peaks around zero, while the ΔE distribution for combinatorial background events tend to be more flatly distributed. For mis-reconstructed B decays faking the signal, the ΔE distribution can have a peak that is shifted from zero. For instance, a $B^0 \to D_s^+\rho^-, \rho^- \to \pi^-\pi^0, \pi^0 \to \gamma\gamma$ decay can be mis-reconstructed as a $B^0 \to D_s^{*+}\pi^-$ event if one of the photons from the π^0 decay is lost. In this case, the ΔE distribution is shifted in the negative direction, the amount of the shift corresponding to the energy of the lost photon.

The $m_{\rm ES}$ distribution for signal events peaks around the nominal B^0 mass. Combinatorial backgrounds have an $m_{\rm ES}$ shape parametrized by the Argus threshold function c[30]:

$$f_{Argus}(x) = x\sqrt{1-(\frac{x}{E_b})^2}\; e^{\kappa(1-(\frac{x}{E_b})^2)}, \tag{4.7}$$

where E_b is the kinematic limit of $5.29\,{\rm GeV}$ and κ is the parameter that determines the overall shape of the Argus function. For background events coming from mis-reconstructed B decays, the $m_{\rm ES}$ distribution also peaks near the nominal B^0 mass value, though the peak for such backgrounds may be broader than that for signal.

4.4.1 Multiple Candidate Selection

Less than 20% of the selected events in $B^0 \to D_s^{*+}\pi^-$ and $B^0 \to D_s^{*-}K^+$ channels, and less than 4% for $B^0 \to D_s^{+}\pi^-$ and $B^0 \to D_s^{-}K^+$, contain two or more B candidates. In the case that there is an event with more than one B candidate for a particular mode passing the selection, the best candidate is selected based on the following criteria:

1. Lowest mass χ^2 value
2. $|\Delta E|$ closest to 0

where the mass χ^2 is defined as:

$$\chi^2 = \left(\frac{m_{D_s} - 1.9686}{\sigma(m_{D_s})}\right)^2 + \left(\frac{\Delta M - 0.1438}{\sigma(\Delta M)}\right)^2$$

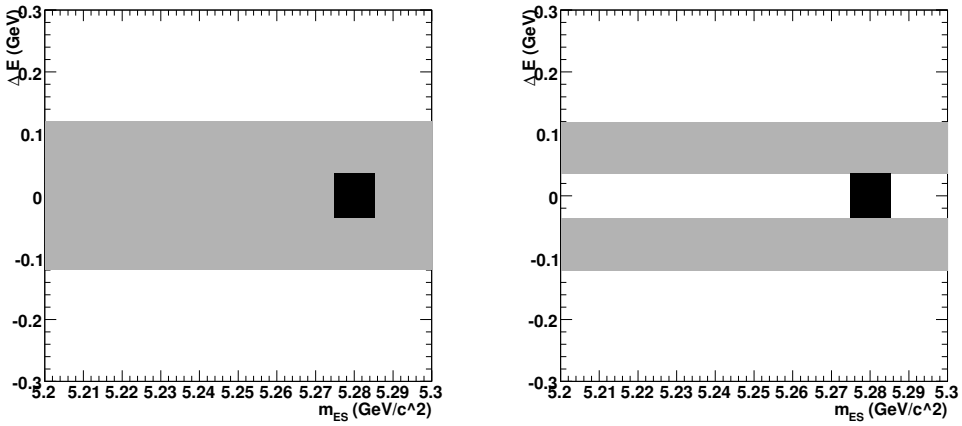
for the D_s^{*+} modes, and likewise without the second term for the D_s^{+} modes. $\sigma(m_{D_s})$ is computed for each candidate while $\sigma(\Delta M)$ is measured using signal Monte Carlo. If the mass χ^2 is identical for both candidates, then the candidate with $|\Delta E|$ closest to zero is chosen. The best candidate selection method was not found to bias the background distributions significantly.

4.5 Definitions of Signal and Sideband Regions

As mentioned in the previous section, the signal region is defined by the approximately orthogonal set of variables ΔE and $m_{\rm ES}$. In addition to the signal region, we also define *sideband* regions outside of the signal box used to estimate background contaminations.

In the $(\Delta E, m_{\rm ES})$ plane, four regions are defined as:

- Signal Box: $|\Delta E| < 0.036\,\text{GeV}$ and $|m_{\rm ES} - 5.28| < 0.0052\,\text{GeV}/c^2$. This corresponds to a 2σ window for ΔE and a 2σ window for $m_{\rm ES}$.
- Sideband: $0.036 < |\Delta E| < 0.120\,\text{GeV}$ in the full $m_{\rm ES}$ region excluding the signal box.



(a) Sideband (gray) and signal regions (black)

(b) ΔE sideband (gray) and signal regions (black)

Figure 4.2. Signal and sideband regions.

- $\underline{\Delta E \text{ Sideband}}$: $0.036\,\text{GeV} < |\Delta E| < 0.120\,\text{GeV}$.
- $\underline{\text{Grand Sideband}}$: $|\Delta E| < 0.3\,\text{GeV}$ and $5.2 < m_{\text{ES}} < 5.3\,\text{GeV}/c^2$, outside of the signal box region.

Figure 4.2 shows the signal and sideband regions in the $(\Delta E, m_{\text{ES}})$ plane. The Grand Sideband (GSB) region corresponds to the entire plane in these figures, except for the signal box. It should be emphasized that these regions are only used for estimating the significance before unblinding our analysis. The final method for computing yield involves a fit to the m_{ES} and $m_{D_s^+}$ distributions, after applying requirements on ΔE and other selection variables. The final yield extraction method is described in Section 5.1. However, defining these regions gives us an idea of the number of signal and background events, and hence the significance, expected in the data sample before unblinding. In addition, before unblinding, we use the data sideband regions to estimate the expected backgrounds in the unblinded data signal box.

Figures 4.3 to 4.8 show the m_{ES} and ΔE distributions from signal MC for $B^0 \to D_s^{(*)+}\pi^-$, after applying the final selection criteria and choosing the best B candidate

in the event, as described in Section 4.6.4. The ΔE distribution is fit to a double Gaussian, while the $m_{\rm ES}$ distribution is fit to a single Gaussian. No requirement was made on the MC truth of the decay for these distributions; in other words, these distributions may include signal events as well as mis-reconstructed background events. Tables 4.2–4.3 report ΔE and $m_{\rm ES}$ means and resolutions determined from the fits to MC.

Table 4.2. ΔE and $m_{\rm ES}$ resolutions in $B^0 \to D_s^+ \pi^-$ for the different D_s^+ modes

	$\langle \Delta E \rangle$ (MeV)	$\sigma(\Delta E)$ (MeV)	$\sigma(m_{\rm ES})$ (MeV)
$D_s^+ \to \phi\pi^+$	-0.7 ± 0.2	17.0 ± 0.2	2.50 ± 0.03
$D_s^+ \to \bar{K}^{0*} K^+$	-0.8 ± 0.3	16.7 ± 0.3	2.51 ± 0.03
$D_s^+ \to \bar{K}^0 K^+$	-0.4 ± 0.3	16.3 ± 0.2	2.51 ± 0.03

Table 4.3. ΔE and $m_{\rm ES}$ resolutions from $B^0 \to D_s^{*+} \pi^-$ MC for the different D_s^+ modes

	$\langle \Delta E \rangle$ (MeV)	$\sigma(\Delta E)$ (MeV)	$\sigma(m_{\rm ES})$ (MeV)
$D_s^+ \to \phi\pi^+$	-0.6 ± 0.3	15.6 ± 0.3	2.54 ± 0.04
$D_s^+ \to \bar{K}^{0*} K^+$	-0.03 ± 0.5	16.4 ± 0.5	2.61 ± 0.06
$D_s^+ \to \bar{K}^0 K^+$	-0.3 ± 0.4	15.1 ± 0.4	2.63 ± 0.06

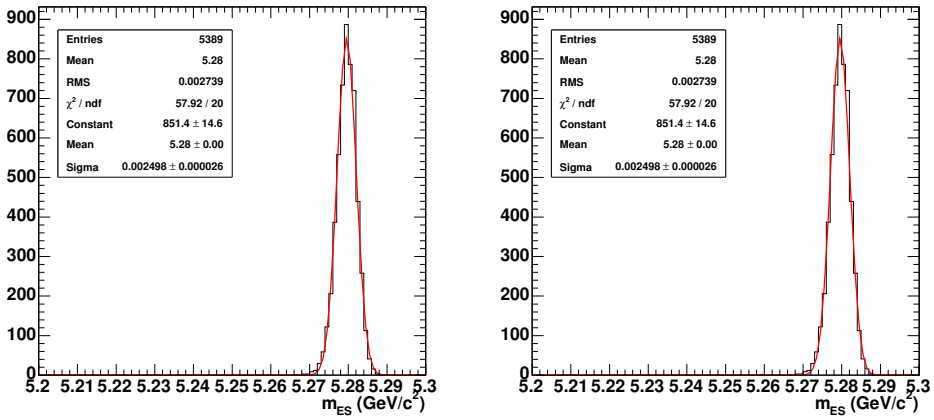


Figure 4.3. $m_{\rm ES}$ (left) and ΔE (right) distributions for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$ signal MC

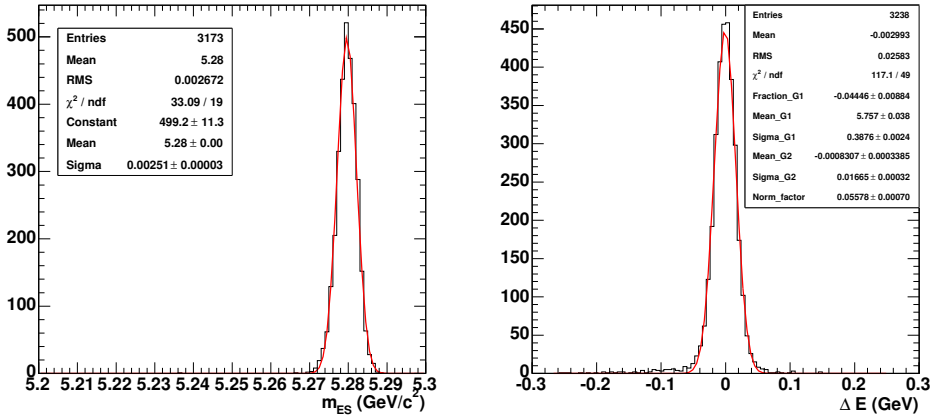


Figure 4.4. $m_{\rm ES}$ (left) and ΔE (right) distributions for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$ signal MC

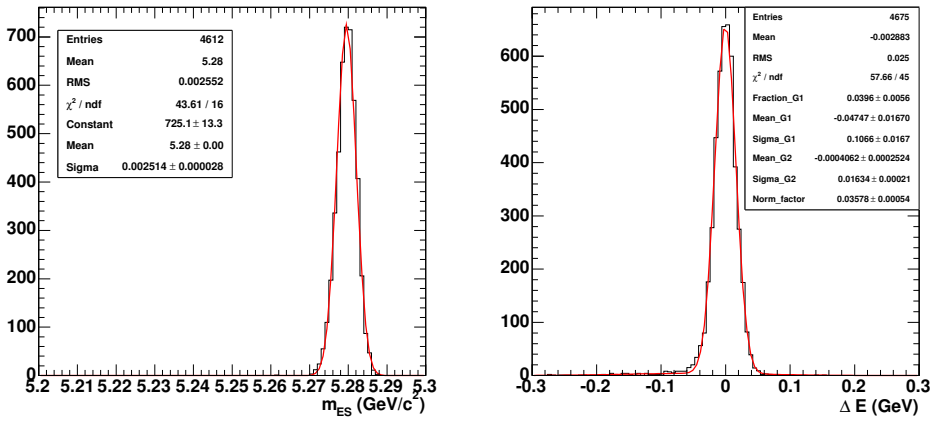


Figure 4.5. $m_{\rm ES}$ (left) and ΔE (right) distributions for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^0K^+$ signal MC

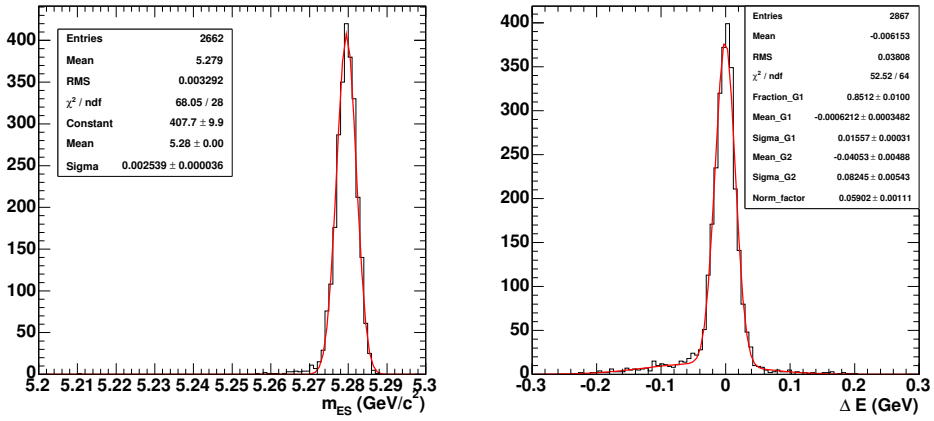


Figure 4.6. $m_{\rm ES}$ (left) and ΔE (right) distributions for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$ signal MC

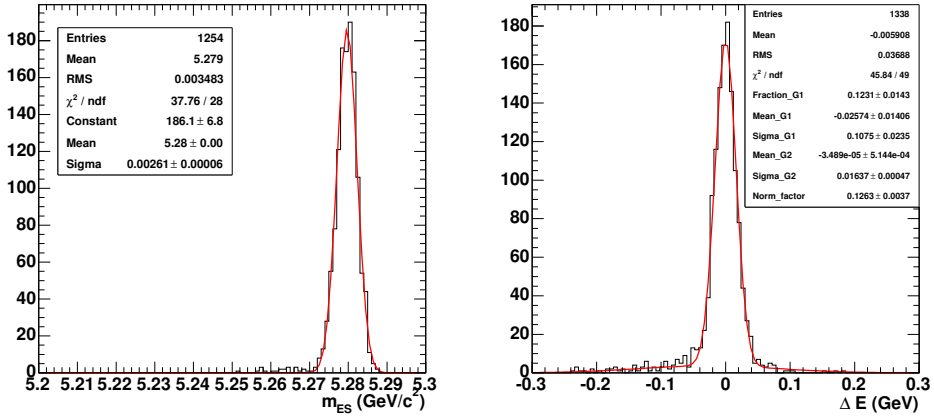


Figure 4.7. $m_{\rm ES}$ (left) and ΔE (right) distributions for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$ signal MC

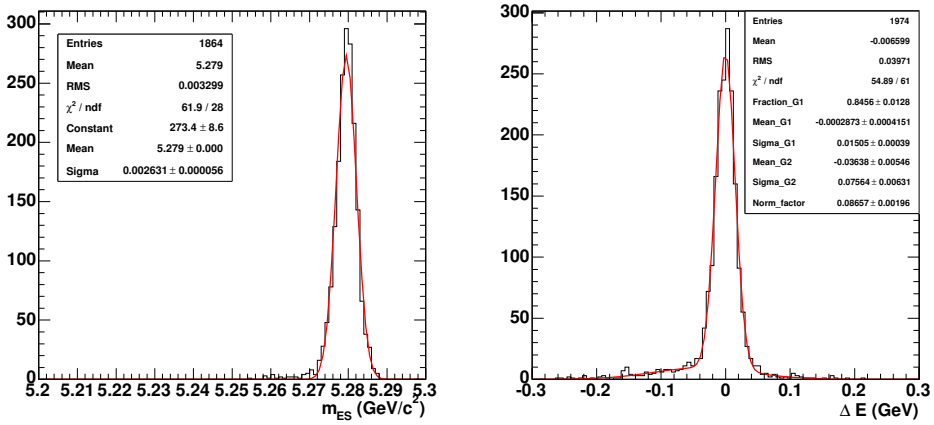


Figure 4.8. $m_{\rm ES}$ (left) and ΔE (right) distributions for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^0K^+$ signal MC

4.6 Final Selection Criteria

After the event reconstruction and pre-selection, a more stringent set of selection criteria is applied to increase the possible significance of the measurement. A combined likelihood selection method is used, in which a number of the selection criteria are merged into one variable. A few remaining cuts are applied individually. The following subsections describe these selection variables in detail, as well as the optimization process used to select the cuts that maximize the significance.

4.6.1 Discriminating Variables

The variables used for the final set of selection criteria are described below:

Legendre Fisher

A Fisher discriminant is a discriminant that finds the linear combination of a set of variables which optimally separates two or more events [27]. For this analysis, a Fisher discriminant is used to differentiate between $B\overline{B}$ and continuum background events by using the different shapes of energy flow for these two types of events. The Fisher discriminant, $\mathcal{F}$, is computed from a linear combination of discriminating variables x_i,

$$\mathcal{F} = \sum_{i=1}^{9} \alpha_i x_i, \tag{4.8}$$

where the coefficients α_i are trained using Monte Carlo to optimize the statistical separation between signal and background. The Fisher discriminant for this analysis is constructed from four quantities: the polar angles of the B momentum vector and the B-candidate thrust axis with respect to the beam axis in the CM frame, and the two Legendre moments L_0 and L_2 of the energy flow around the B-candidate thrust axis [29]. Because of the use of Legendre moments for this discriminant, it is referred to as the Legendre Fisher. Figure 4.6.1 shows the Fisher discriminant distributions for signal and generic Monte Carlo. It should

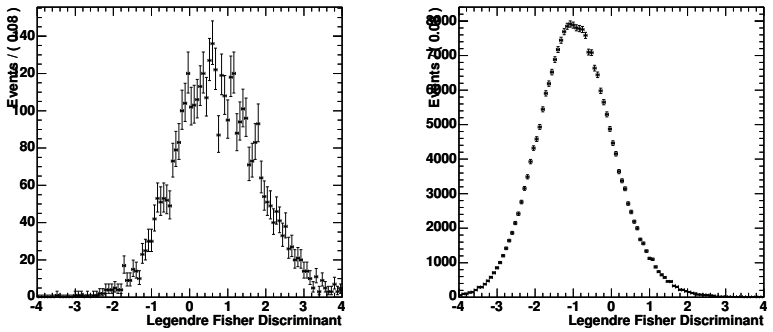


Figure 4.9. Legendre Fisher discriminant for signal Monte Carlo (left side) and continuum generic Monte Carlo (right).

be noted that the Fisher discriminant is not completely uncorrelated with other variables, such as the cosine of the thrust angle.

Thrust Angle, θ_{thrust}

The thrust axis for a set of particles is defined by the direction which maximizes the sum of the longitudinal momenta of the particles [28]. The thrust angle is defined as the angle between the thrust axis of the B candidate and the thrust axis of the rest of the event. The distribution of $\cos\theta_{\rm thrust}$ for $B\bar{B}$ events is uniform because there is no correlation between the thrust axes of the two B mesons. On the other hand, continuum events are jet-like and tend to have a distribution peaked at $|\cos\theta_{\rm thrust}| = 1$. The algorithm which computes the thrust angle arbitrarily assigns the sign of the cosine, and a cut is placed on the absolute value in order to reduce the systematics.

Flight Angle of the B^0, θ_B

The flight angle of the B, θ_B, is the angle of the B^0 in the CM system with respect to the direction of the electron. Since the e^+e^- collisions produce polarized virtual photons with spin ±1 and the B^0 has spin 0, the distribution of θ_B is $\sin^2\theta_{\rm B}$, while the combinatorial background has a uniform distribution for $\cos\theta_{\rm B}$.

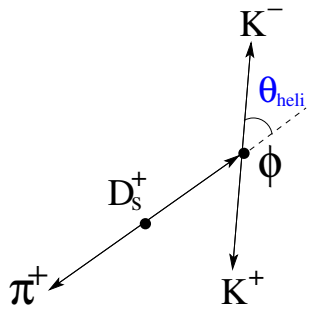


Figure 4.10. Helicity angle definition, for ϕ from $D_s^+ \to \phi\pi^+$

Probability of the B^0 and D_s^+ Vertices

A requirement on the probability of the vertex fit for B^0 and D_s^+ is applied. This helps in rejecting random combinations of a real D_s^+ and a real π or K from different B candidates, as well as random combinations from a single B meson, since such combinations do not originate from a single vertex. Similarly, the probability of the vertex fit for the D_s^+ can be used to discriminate against events that do not have a real D_s^+. A cut on the $-\log(\mathrm{Prob}(\chi^2_{vertex}))$ of the vertex fits is applied.

Helicity Angle of ϕ, $\bar{K}^{0*}$, and γ, θ_{hel}

The helicity angle of a particle is defined as the angle between the direction of an outgoing daughter of the particle and the flight direction of the particle in its rest frame. Fig. 4.10 depicts how the helicity angle is defined. Because of angular momentum conservation, the distribution of the helicity angle, θ_{hel}, depends on the spins of the particles. In the decay of a pseudo-scalar (PS) meson to a vector (V) and a pseudo-scalar, such as the decays $D_s^+ \to \phi\pi^+$ and $D_s^+ \to \bar{K}^{0*}K^+$, the helicity of the vector meson is zero and the distribution of the helicity angle is $\cos^2\theta_{\mathrm{hel}}$. The distribution of $\cos\theta_{\mathrm{hel}}$ for combinatorial background is nearly uniform, since these events arise from random combinations of tracks

for which the direction of the daughter decays are not correlated. Likewise, the helicity of the γ from $D_s^{*+} \to D_s^+\gamma$ which is a $V \to PSV$ decay, is distributed as $\sin^2\theta_{\rm hel}$.

K_S^0 Flight Angle

$\theta_{\rm flight}$ is the angle between the direction of the K_S^0 momentum vector and the flight direction, which is defined by the vector drawn from the interaction point to the K_S^0 vertex. For true K_S^0 candidates, $\theta_{\rm flight}$ peaks near zero and therefore $|\cos\theta_{\rm flight}|$ peaks near 1, while the distribution for combinatorial background is more uniformly distributed. A cut is applied on $|\cos\theta_{\rm flight}|$.

Mass Difference of the D_s^{*+}, $\Delta m_{D_s^{*+}}$

$\Delta m_{D_s^{*+}}$ is the difference between the invariant masses of the D_s^{*+} and D_s^+: $\Delta m(D_s^{*+}) = m(D_s^{*+}) - m(D_s^+)$. For events with a real D_s^{*+}, $\Delta m_{D_s^{*+}}$ has a peak near the nominal value of this difference, $143.8\,{\rm MeV}/c^2$ [12]. The distribution is uniform for events which do not contain a real D_s^{*+}, such as events mis-reconstructed from a random D_s^+ and a random γ not associated with a D_s^{*+}.

Invariant Masses

The invariant masses for D_s^+, ϕ, $\bar{K}^{0*}$, and K_S^0 also provide discriminating power against combinatorial backgrounds. The mass distributions for combinatorial events is mostly uniform, with the possibility of a peak from events with a real D_s^+, ϕ, $\bar{K}^{0*}$, or K_S^0. The mass of the D_s^+ is not included in the selection criteria for $B^0 \to D_s^+\pi^-$ and $B^0 \to D_s^-K^+$, since a fit to the D_s^+ distribution is used in the final yield extraction.

Particle Identification

Particle Identification (PID) is very powerful in discriminating against backgrounds where a pion is mistaken for a kaon, or vice versa. For instance, a $B^0 \to D_s^- K^+$ event can be mistaken for a $B^0 \to D_s^+ \pi^-$ event in this manner, as well as the opposite case where a $B^0 \to D_s^+ \pi^-$ event is mistaken as $B^0 \to D_s^- K^+$. Applying a relatively tight PID selection on the high momentum track from the B^0 decay can help distinguish such events. Likewise, a PID requirement on the kaon daughter in the decay $\bar{K}^{0*} \to K^- \pi^+$ can help distinguish the $\bar{K}^{0*}$ from other particles with a $\pi^+ \pi^-$ final state. A PID requirement on the kaon daughters of the $\phi \to K^+ K^-$ can also help reduce backgrounds, although the $D_s^+ \to \phi \pi^+$ mode is relatively clean with respect to the other D_s^+ modes and does not require a very tight PID selection. Similarly, a requirement on the PID of the kaon daughter of the D_s^+ for $D_s^+ \to \bar{K}^{0*} K^+$ and $D_s^+ \to \bar{K}^0 K^+$ events reduces backgrounds from D^+ decays. The different PID selection criteria are described in Section 2.5.1.

4.6.2 Likelihood Selection

Several selection criteria are merged into one combined likelihood variable. The method of maximum likelihoods is briefly discussed in Appendix B. The goal of using this likelihood selection, rather than applying all the cuts separately as is typically done are three-fold. The first goal is to simplify the selection optimization, since in this case, only a few variables have to be optimized. The likelihood selection also makes our selection optimization less sensitive to fluctuations from low statistics in our MC samples. Lastly, this method can provide a better signal-to-background ratio when optimizing the selection.

In this likelihood selection method, we parametrize each selection variable to be included in the likelihood by fitting the distribution of the variable in Monte Carlo to a probability density function (PDF). After we fit all the selection variables, the likelihood variable R_L

is constructed as:

$$\mathcal{L}_{sig} = \prod_{i=0}^{N} \mathcal{P}_{sig,i}, \tag{4.9}$$

$$\mathcal{L}_{bkg} = \prod_{i=0}^{N} \mathcal{P}_{bkg,i}, \tag{4.10}$$

$$R_L = \mathcal{L}_{sig}/(\mathcal{L}_{sig} + \mathcal{L}_{bkg}), \tag{4.11}$$

where i is indexed over all selection variables used, and N is the total number of selection variables included in the likelihood. $\mathcal{P}_{sig,i}$ corresponds to the signal PDF for variable i determined from signal Monte Carlo, and $\mathcal{P}_{bkg,i}$ corresponds to the background PDF for the variable i determined from generic Monte Carlo. $\mathcal{L}_{sig}$ and $\mathcal{L}_{bkg}$ are normalized to 1 over the range of interest, and this range is the same for both signal and background components. The distribution of R_L in Monte Carlo is strongly peaked at 1 for signal MC and 0 for generic MC, which represents our combinatorial background, as can be seen in Figures 4.12-C.49.

The selection variables included in the likelihood are:

- Legendre Fisher discriminant
- -log (B^0 vertex probability)
- $|\cos\theta_{\rm thrust}|$
- $\cos\theta_{\rm B}$
- -log (D_s^+ vertex probability)
- $\cos\theta_{\rm hel}$ of ϕ and $\bar{K}^{0*}$
- Invariant mass of ϕ, $\bar{K}^{0*}$, K_s

Additional variables for the D_s^{*+} modes are:

- $\Delta M_{D_s^*}$
- $\cos\theta_{\rm hel}$ of the photon

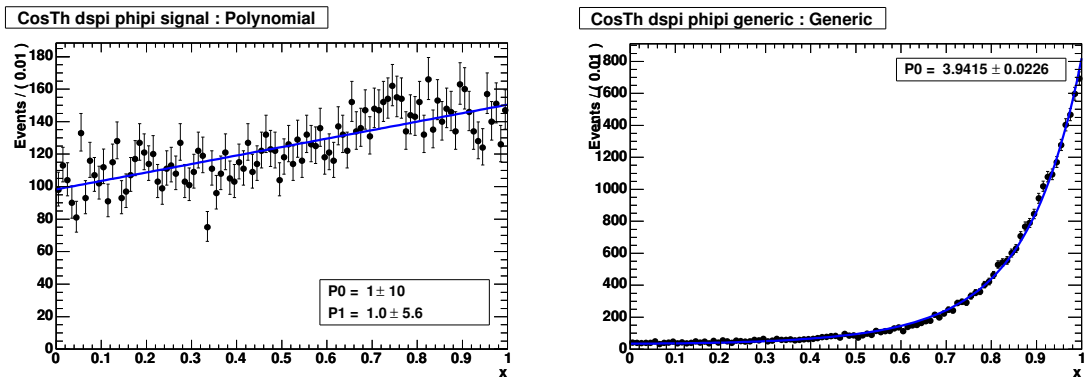


Figure 4.11. $\cos\theta_{\text{thrust}}$ for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$, signal (left) and generic (right) MC

Table 4.4 describes the PDFs used for fitting the distributions of the discriminating variables included in R_L. An example of the PDF fits for signal and background can be seen in Figure 4.11. Plots of all the fitted PDFs can be seen in Figures C.1 - C.46 of Appendix C. In cases such as invariant mass and $\Delta M_{D_s^*}$, where we expect the generic MC distribution to include a small signal peak, the values of the signal components are fixed to the values determined by the signal MC fit. In all cases, we use the simplest and most intuitive PDF that produced the best fit. In addition, we made an effort to maintain similar PDFs across the three D_s modes for consistency. To create the PDFs, a loose set of criteria—the preselection plus the other selection variables not included in the likelihood—are applied to the Monte Carlo samples.

A few of the selection variables described earlier are not included in the combined likelihood, and for these variables we apply a straight cut. $\cos\theta_{\text{flight}}$ is not included in the likelihood because it is highly peaked for signal events, and it was clear where to apply a cut. The PID is applied separately because there is no natural way to fit these to PDFs. The invariant mass of the D_s^+ is not included in R_L because for some of the modes, a fit to the D_s^+ distribution is used to extract the yield and branching fraction. For the purposes of estimating the expected number of signal and background events in our signal box, we apply a m_{Ds} window of 10 MeV around the nominal D_s^+ mass.

Once the PDFs are fitted, the combined likelihood variable R_L is constructed, as de-

Table 4.4. Probability density functions used to describe signal and backgrounds for the combined likelihood variable. Gauss is an abbreviation for Gaussian, Poly for Polynomial, and BW for Breit-Wigner function.

Variable	Signal PDF	Background PDF
Legendre Fisher	Gauss + Bifurcated Gauss	Gauss + Bifurcated Gauss
-log (B^0 vtx prob)	$e^{(-P1*x^{P2})}$	$e^{(-P1*x^{P2})}$
$\cos\theta_{\rm thrust}$	1st order Poly	$e^{(P1*x^2)}$
$\cos\theta_{\rm B}$	2nd order Poly	1st order Poly
-log (D_s vtx prob)	$e^{(-P1*x^{P2})}$	$e^{(-P1*x^{P2})}$
$\cos\theta_{\rm hel}$ of ϕ, $\bar{K}^{0*}$	2nd order Poly	2nd order Poly
Mass of ϕ	Gauss + BW	Gauss + BW + 2nd order Poly
Mass of $\bar{K}^{0*}$	Double Gauss	Double Gauss + 2nd order Poly
Mass of K_s	Double Gauss	Double Gauss + 2nd order Poly
$\Delta M_{D_s^*}$	Double Gauss + 1st order Poly	Double Gauss + 1st order Poly
$\cos\theta_{\rm hel}$ of photon	2nd order Poly	2nd order Poly

scribed in Equation 4.11. Figures 4.12-4.14 show the distributions of R_L in signal and generic Monte Carlo for $B^0 \to D_s^+\pi^-$. Similar figures for the other modes can be seen in Appendix C. We can see in these figures that the likelihood is highly peaked at 1 for signal events, and peaked at 0 for background events. The distribution is much more peaked for $c\bar{c}$ and $u\bar{u}$,$d\bar{d}$,$s\bar{s}$ events than for $B\bar{B}$ events, since there are many $B\bar{B}$ background events with real B^0, $D_s^{(*)+}$, ϕ, $\overline{K}^{*0}$ and K_S^0 candidates. The signal events are filtered from the $B\bar{B}$ generic MC in these plots.

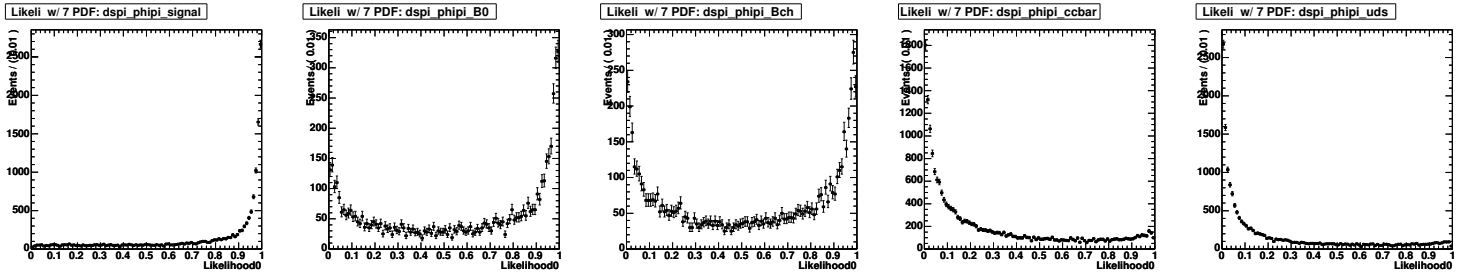


Figure 4.12. Likelihood distribution for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$: signal, B^0, B^+, $c\bar{c}$, uds MC

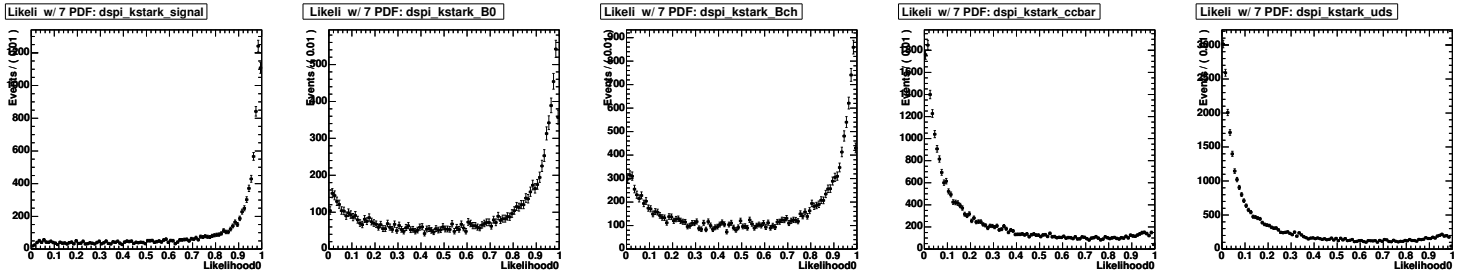


Figure 4.13. Likelihood distribution for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$: signal, B^0, B^+, $c\bar{c}$, uds MC

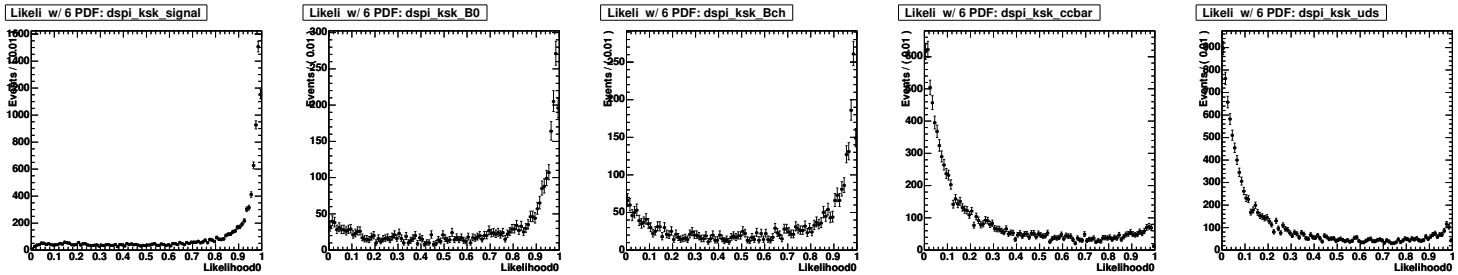


Figure 4.14. Likelihood distribution for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^0K^+$: signal, B^0, B^+, $c\bar{c}$, uds MC

4.6.3 Control Sample

A "control sample" is a set of events that is kinematically similar to the signal mode under study. We utilize one such set of events to check the validity of the PDFs for $B^0 \to D_s^{(*)+}\pi^-$ and determine how much the PDFs from Monte Carlo can be expected to differ from data. A control sample is used because the branching fractions for the signal decay modes are exceedingly small, rendering a large statistics study using real data unfeasible. Additionally, this analysis is completed in a blind manner, which prevented any study using the signal region in data before unblinding.

The control sample for this study consists of data and exclusive Monte Carlo in the modes:

- $B^0 \to D^-\pi^+$, $D^- \to K^+\pi^-\pi^-, \bar{K}^0\pi^-$ (for D_s^+ modes)
- $B^+ \to D^{*0}\pi^+$, $D^{*0} \to D^0\gamma, D^0 \to K^-\pi^+$ (for D_s^{*+} modes).

These modes are chosen because the B and D meson decays are kinematically similar to the signal modes.

First, the PDFs determined from the $D_s^{(*)+}\pi^-$ MC are compared to the distributions seen in the control sample MC, for those variables common to both modes, to see if they are at all similar. To this end, we overlay the $D_s^{(*)+}\pi^-$ PDFs on top of the control sample signal MC distributions. The overlaid plots can be seen in Figures C.56-C.69 of Appendix C. Since the PDFs look acceptable without refitting to the control sample distributions, the $D_s^{(*)+}\pi^-$ PDFs, using only the variables common to the control sample, are used for this study.

Then the R_L distributions in the control sample MC are cross-checked, to make sure they are sensible (Figures C.70-C.72 of Appendix C). We then construct R_L for the control samples using the PDFs that are common between the $D^{(*)+}\pi^-$ and $D_s^{(*)+}\pi^-$ modes, without refitting.

Next, the signal yields as a function of R_L in both the MC and data for the control sample are studied. Since the sample in data is not as pure as in the MC, we fit for the yields, rather than simply cutting and counting the number of events that fall into the signal region. To extract the yield, we fit the $m_{\rm ES}$ distribution to a Gaussian (signal) plus an Argus function (background), and these yields are normalized by the yield obtained when there is no requirement on R_L. These distributions can be seen in Figures 4.15-4.17 below. These figures also show the percent difference between the data and Monte Carlo yields.

These plots show that there is good agreement between the normalized signal yields for data and Monte Carlo for the control sample. For the $D_s^{*+}\pi^-$ and corresponding $D^{*0}\pi$ samples, photon smearing, the process of deliberately smearing the photon resolution of the EMC in Monte Carlo, is applied to disentangle the effect of discrepancies in photon resolution in data and Monte Carlo from the effect of the differences in PDFs for data and Monte Carlo. As can be seen in the figures, the agreement is on the order of a few percent, with slightly larger discrepancies at tighter cuts on R_L. Although this difference is small, these effects are included as part of the systematic errors.

The percent difference between the MC and data efficiencies from the control samples for the particular likelihood cut applied can be read from Figures 4.15-4.17 to determine the systematic error to assign. The systematics assigned depend on the value of the R_L cut, which is 0.75 for $D_s^+\pi^-$ and 0.8 for $D_s^{*+}\pi^-$, as shown in Tables 4.6-4.7. However, because the control samples use a reduced likelihood (that is, the likelihood is constructed out of PDFs common between the signal modes and control sample modes), the error may be smaller than if the difference is read from the plot at the value of the full likelihood cut applied.

To quantify this effect, we study the correlation between the full likelihood (which includes all the $D_s^{(*)}\pi$ PDFs) and reduced likelihood (which uses only the PDFs common between $D_s^{(*)}\pi$ and $D^{(*)}\pi$ modes) in the $B^0 \to D_s^{(*)+}\pi^-, D_s^+ \to \phi\pi^+$ signal Monte Carlo.

These correlation plots can be seen in Figures 4.18-4.19. The left hand plots are similar to the previous plots for the control sample. The yield as a function of R_L is extracted from fits to the $m_{\rm ES}$ distribution, normalized to the yield with no requirement on R_L. The blue asterisks correspond to the reduced likelihood distribution and red triangles correspond to the full likelihood distribution. Looking at these plots, it can be seen that for $D_s^+\pi^-$, the efficiency from a full likelihood cut of 0.75 corresponds to a reduced likelihood value of 0.7 and for $D_s^{*+}\pi^-$, a full likelihood cut of 0.8 corresponds to the same efficiency as a reduced likelihood cut of 0.7.

We then assign the systematic error by looking at the control sample plots of Figures 4.15-4.17, using the cuts shifted by the full-reduced efficiency correlation plots. For $D_s^+\pi^-$, a reduced likelihood cut of 0.7 corresponds to a systematic error of about 3% from the $D^-\pi^+$ distributions. For $D_s^{*+}\pi^-$, a reduced likelihood cut of 0.7 corresponds to a systematic error of about 7% from the $D^{*0}\pi^+$ plots. The same systematic errors are assigned to $B^0 \to D_s^{(*)-}K^+$ since the distributions are very similar to $B^0 \to D_s^{(*)+}\pi^-$.

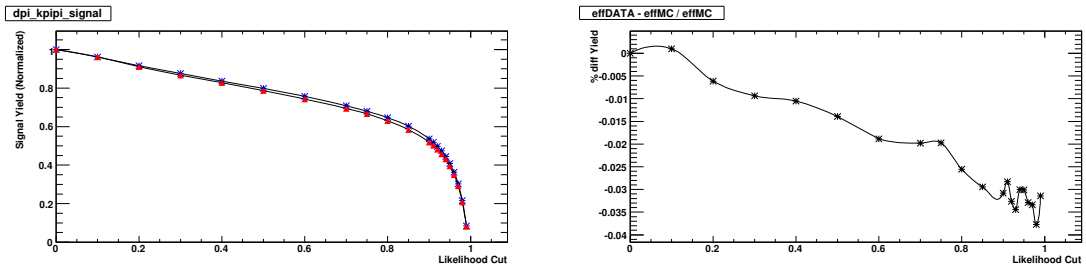


Figure 4.15. Signal yield (normalized to yield with no cut) for $B^0 \to D^-\pi^+$, $D^- \to K^+\pi^-\pi^-$ MC (blue asterisks) and data (red triangles) (Left). Percent Difference between data and MC (Right).

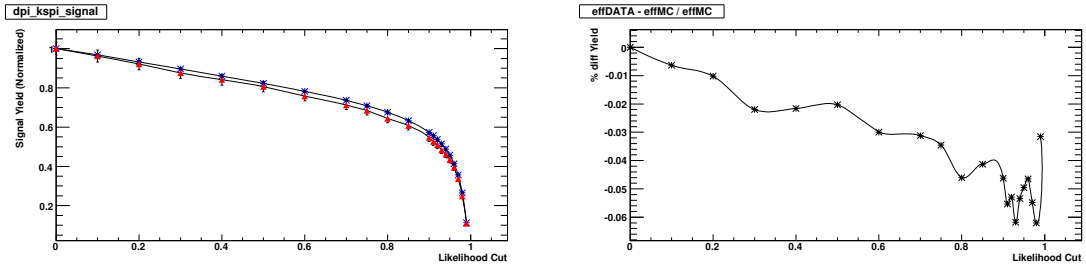


Figure 4.16. Signal yield (normalized to yield with no cut) for $B^0 \to D^-\pi^+$, $D^- \to \bar{K}^0\pi^-$ MC (blue asterisks) and data (red triangles) (Left). Percent Difference between data and MC (Right).

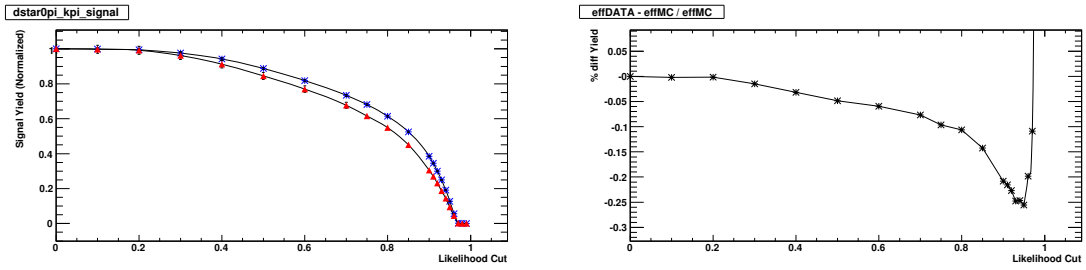


Figure 4.17. Signal yield (normalized to yield with no cut) for $B^+ \to D^{*0}\pi^+$, $D^{*0} \to D^0\gamma, D^0 \to K^-\pi^+$ MC (blue asterisks) and data (red triangles) (Left). Percent Difference between data and MC (Right).

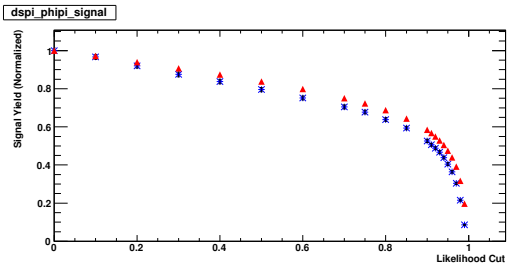


Figure 4.18. Signal yield (normalized to yield with no cut) for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$ MC for reduced likelihood (blue asterisks) and full likelihood (red triangles).

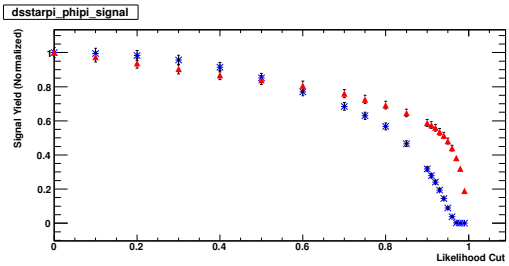


Figure 4.19. Signal yield (normalized to yield with no cut) for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$ MC for reduced likelihood (blue asterisks) and full likelihood (red triangles).

4.6.4 Cut Optimization

Once the PDF parameters are determined and the combined likelihood variable constructed, we optimize the likelihood variable R_L using a cut-and-count method to maximize the ratio $S/\sqrt{S+B}$, the inverse of the statistical error on the measured branching fraction, or $S/\sqrt{B}$, the statistical significance on observation of the decay, depending on the mode. B is the number of background events estimated from the generic Monte Carlo sidebands, scaled to the area in the signal box and the data luminosity. S is the number of signal events estimated from Monte Carlo as:

$$S = \epsilon_{MC} \cdot N_{\text{expect}}, \tag{4.12}$$

where ϵ_{MC} is the selection efficiency computed from signal Monte Carlo, N_{expect} is the number of events expected in the data sample before any selection is applied. N_{expect} can be computed as:

$$N_{\text{expect}} = N_{B\bar{B}} \cdot \mathcal{B}, \tag{4.13}$$

where $N_{B\bar{B}}$ is the total number of $B\bar{B}$ events in the data sample, and $\mathcal{B}$ is the branching fraction of the full decay chain being studied. For estimating N_{expect}, we use the branching fractions measured by *BABAR*, as shown in Table 1.2. N_{expect} for $D_s^+\pi^-$ and $D_s^{*+}\pi^-$ are shown in Table 4.5. The numbers for $B^0 \to D_s^{(*)-}K^+$ are the same as those shown for $B^0 \to D_s^{(*)+}\pi^-$, since the branching fractions are assumed to have the same values in the optimization.

The results of the optimization are shown in Figures 4.6.4-4.6.4. The plots in the top row correspond to the expected significance as a function of the likelihood cut. The middle row of plots correspond to number of signal events, as estimated from the number of signal MC events falling into the signal box after the selection is applied, and the number of background events, as estimated from the generic MC sidebands, ΔE sidebands, and signal box.

From these plots we see that there is good agreement between the estimates from the

sidebands and ΔE sidebands. We also see that the signal box estimates suffer from statistical fluctuations. These various methods of estimating the background are shown only as a cross check; to compute significance shown in Figures 4.6.4, 4.6.4, 4.6.4, and 4.6.4, the estimates from the sidebands are used for B. The last row of plots is the signal efficiency as a function of the likelihood cut, as estimated from the signal Monte Carlo. The efficiency is simply the number of signal events that pass the selection, scaled by the total number of generated signal events in the MC sample.

For each B^0 mode, the same cut for all three D_s modes is chosen for sake of consistency. A conservative cut of 0.75 is chosen to maintain a high $S/\sqrt{S+B}$ without sacrificing signal efficiency for $B^0 \to D_s^+\pi^-$, for which evidence has already been seen. For $B^0 \to D_s^{*+}\pi^-$, which has not yet been discovered, a tighter cut of 0.8 is selected as a compromise between maximizing $S/\sqrt{B}$ and avoiding the rapidly falling edge for $S/\sqrt{S+B}$. After examining the optimization plots, similar cuts are chosen for $B^0 \to D_s^- K^+$ and $B^0 \to D_s^{*-} K^+$ for sake of consistency.

Table 4.5. Expected number of events in 208.7 fb^{-1} for $B^0 \to D_s^{(*)+}\pi^-$ in the different D_s^+ decay modes (before any selection is applied).

Mode	$N_{\rm exp}(D_s^+\pi^-)$	$N_{\rm exp}(D_s^{*+}\pi^-)$
$D_s^+ \to \phi\pi^+$	174	97
$D_s^+ \to \bar{K}^{0*}K^+$	162	90
$D_s^+ \to \bar{K}^0 K^+$	91	51

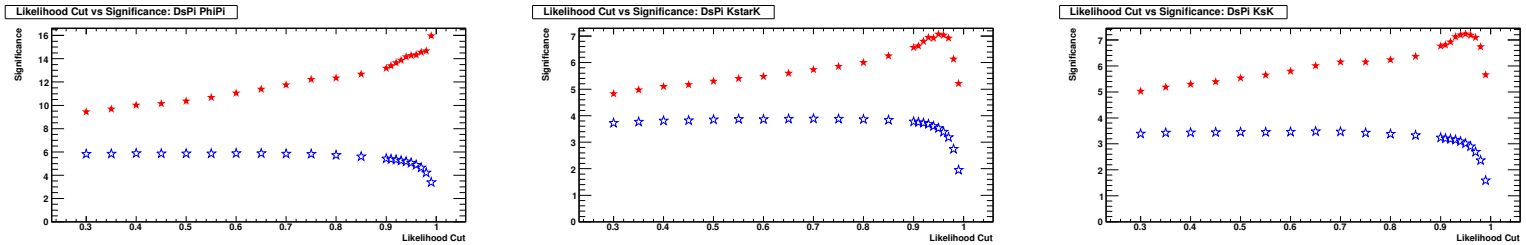


Figure 4.20. Significance vs likelihood cut for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+, \bar{K}^{0*}K^+, K_sK^+$. Red filled stars are $S/\sqrt{(B)}$ and open blue stars are $S/\sqrt{(S+B)}$, where S is estimated from signal MC signal box and B is from generic MC sidebands

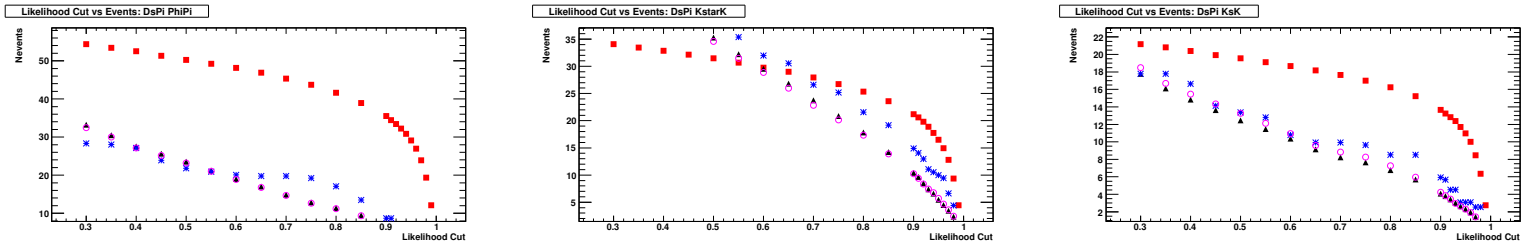


Figure 4.21. Number of events vs likelihood cut for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+, \bar{K}^{0*}K^+, K_sK^+$. Red squares are signal from signal MC, black triangles are background from generic MC sidebands, pink circles are background from generic MC ΔE sidebands, and blue asterisks are background from generic MC signal box.

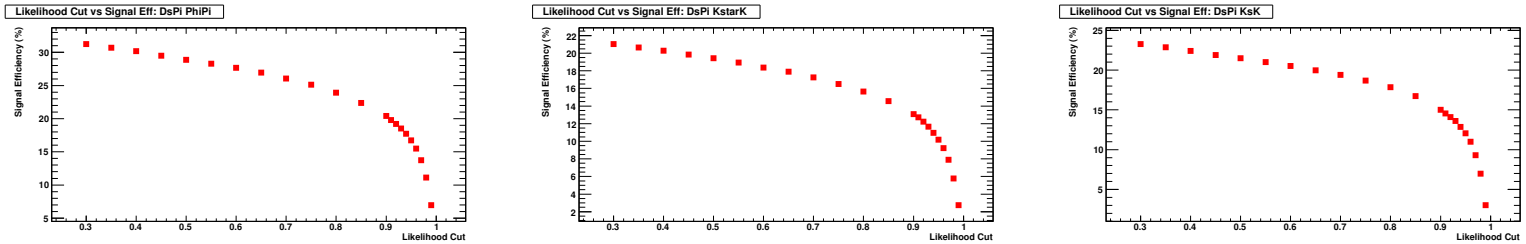


Figure 4.22. Signal efficiency vs likelihood cut for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+, \bar{K}^{0*}K^+, K_sK^+$.

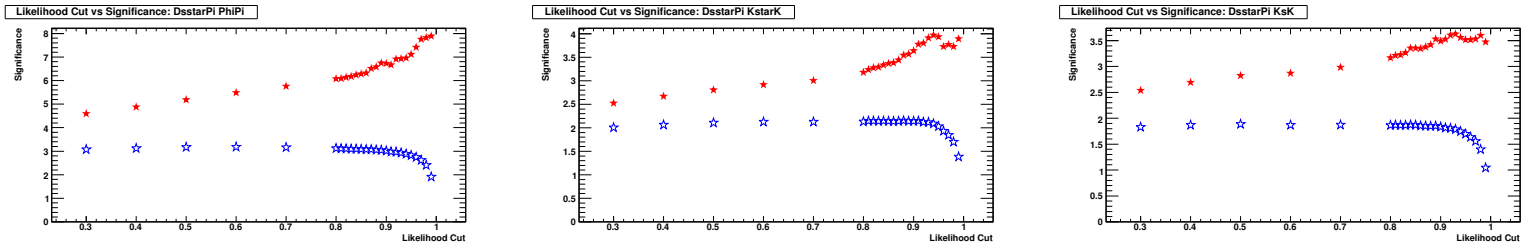


Figure 4.23. Significance vs likelihood cut for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+, \bar{K}^{0*}K^+, K_sK^+$. Red filled stars are $S/\sqrt{(B)}$ and open blue stars are $S/\sqrt{(S+B)}$, where S is estimated from signal MC signal box and B is estimated from generic MC sidebands

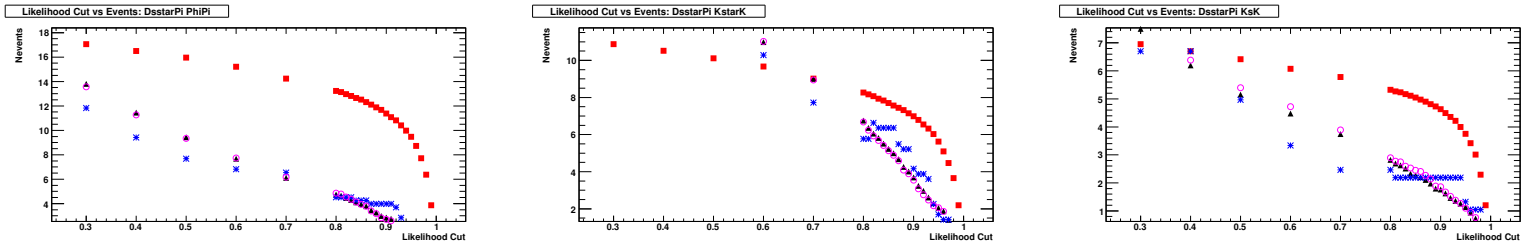


Figure 4.24. Number of events vs likelihood cut for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+, \bar{K}^{0*}K^+, K_sK^+$. Red squares are signal from signal MC, black triangles are background from generic MC sidebands, pink circles are background from generic MC ΔE sidebands, and blue asterisks are background from generic MC signal box.

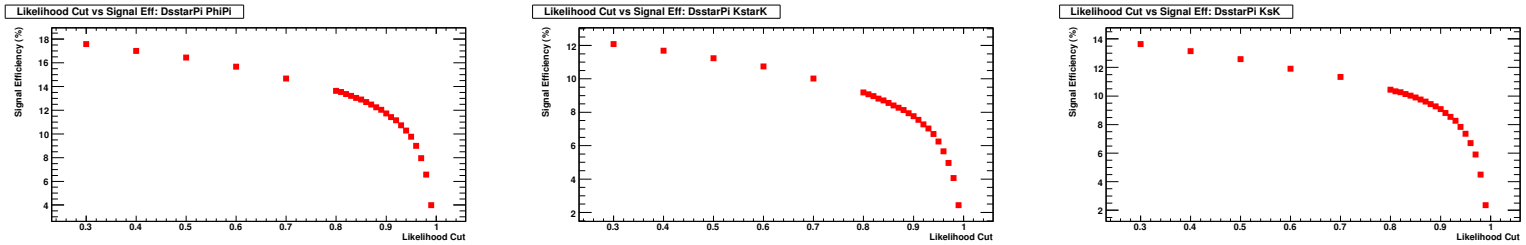


Figure 4.25. Signal efficiency vs likelihood cut for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+, \bar{K}^{0*}K^+, K_sK^+$.

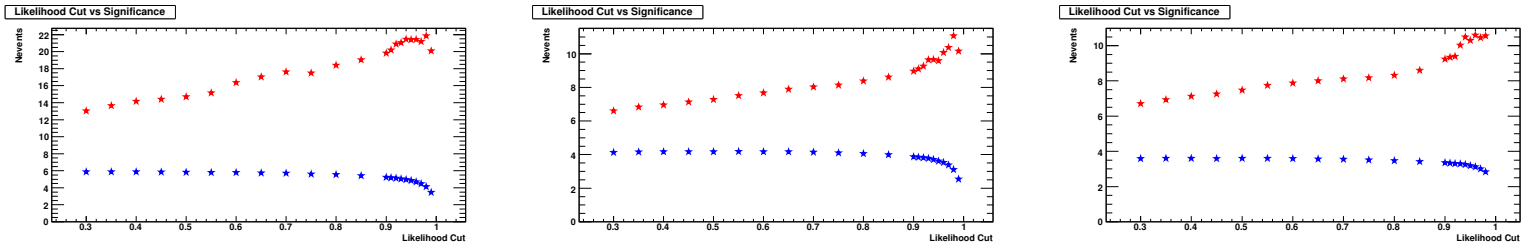


Figure 4.26. Significance vs likelihood cut for $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+, \bar{K}^{0*}K^+, K_sK^+$. Red filled stars are $S/\sqrt{(B)}$ and open blue stars are $S/\sqrt{(S+B)}$, where S is estimated from signal MC signal box and B is from generic MC sidebands

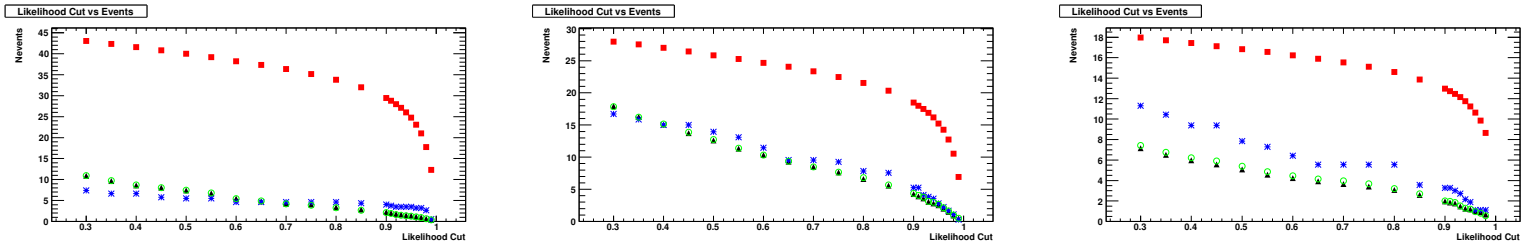


Figure 4.27. Number of events vs likelihood cut for $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+, \bar{K}^{0*}K^+, K_sK^+$. Red squares are signal from signal MC, black triangles are background from generic MC sidebands, pink circles are background from generic MC ΔE sidebands, and blue asterisks are background from generic MC signal box.

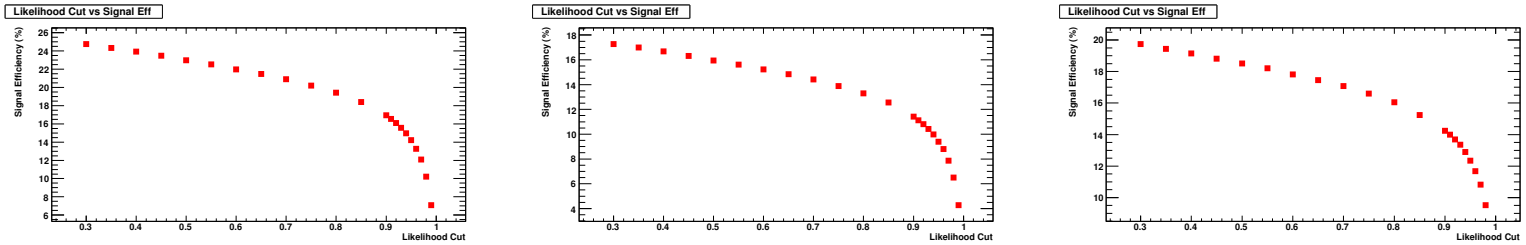


Figure 4.28. Signal efficiency vs likelihood cut for $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+, \bar{K}^{0*}K^+, K_sK^+$.

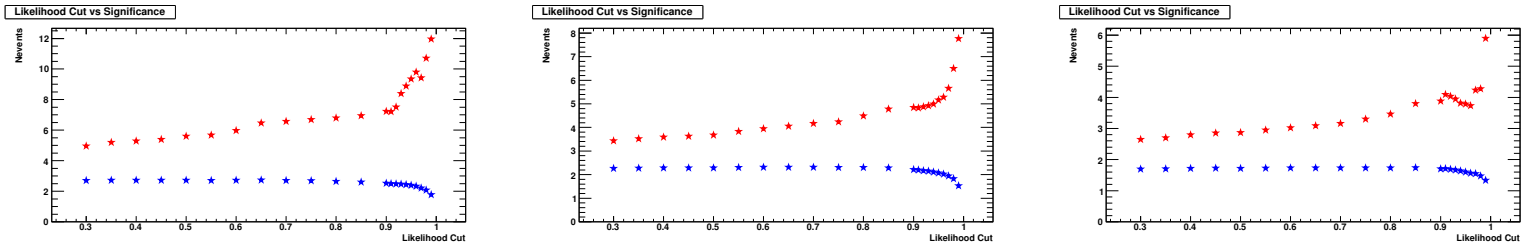


Figure 4.29. Significance vs likelihood cut for $B^0 \to D_s^{*-}K^+$, $D_s^+ \to \phi\pi^+, \bar{K}^{0*}K^+, K_sK^+$. Red filled stars are $S/\sqrt{(B)}$ and open blue stars are $S/\sqrt{(S+B)}$, where S is estimated from signal MC signal box and B is estimated from generic MC sidebands

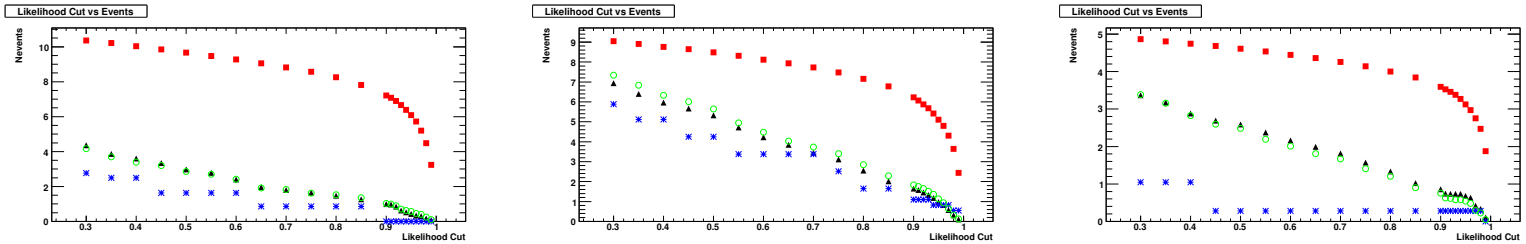


Figure 4.30. Number of events vs likelihood cut for $B^0 \to D_s^{*-}K^+$, $D_s^+ \to \phi\pi^+, \bar{K}^{0*}K^+, K_sK^+$. Red squares are signal from signal MC, black triangles are background from generic MC sidebands, pink circles are background from generic MC ΔE sidebands, and blue asterisks are background from generic MC signal box.

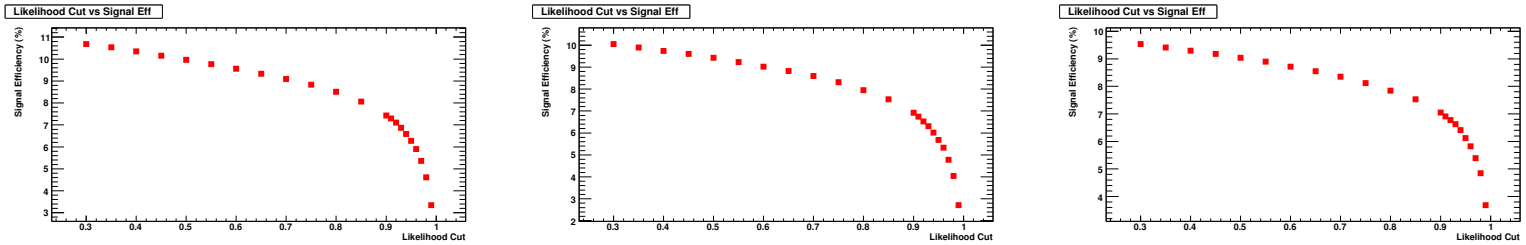


Figure 4.31. Signal efficiency vs likelihood cut for $B^0 \to D_s^{*-}K^+$, $D_s^+ \to \phi\pi^+, \bar{K}^{0*}K^+, K_sK^+$.

4.6.5 $B^0 \to D_s^{(*)+}\pi^-$ and $B^0 \to D_s^{(*)-}K^+$ Final Selection

The final selection criteria and the resulting performances are shown in Table 4.6 for $B^0 \to D_s^+\pi^-$ and $B^0 \to D_s^- K^+$ and Table 4.7 for $B^0 \to D_s^{*+}\pi^-$and $B^0 \to D_s^{*-}K^+$. The number of signal events is estimated from signal Monte Carlo, scaled to the data luminosity, as described in the previous section. The number of background events is estimated using a large sample of generic Monte Carlo, scaling the number of generic background events that fall into the sidebands to the area in the signal box and the data luminosity. For $B^0 \to D_s^{(*)-}K^+$ we apply the same selection as $B^0 \to D_s^{(*)+}\pi^-$, except that the requirement on the sister of the D_s^+ is switched from a `Tight` requirement on the pion to a `Tight` requirement on the kaon.

It should be emphasized that, although the selection presented here is the final one applied to measure the branching fractions, this is not the final method used to compute the number of signal and background events. We use this approach to estimate the significance of our measurement before the analysis is unblinded. For the final result, we perform a fit to the data distribution to extract the number of signal and background events. In this sense, the cuts for $m_{\rm ES}$ and D_s^+ mass are here only to estimate S and B, and these cuts are loosened for the final fit.

Table 4.6. Optimized selection criteria for $B^0 \to D_s^+\pi^-$ (and $B^0 \to D_s^-K^+$).

Cut	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
D_s^+ mass (MeV/c^2)	± 10	± 10	± 10
$\lvert\cos\theta_{\mathrm{flight}}\rvert$	-	-	< 0.98
Hard π^- PID (K^- PID)	`Tight`	`Tight`	`Tight`
Bachelor Kaon PID	`NotPion`	`Tight` (fast)	`Tight`
Daughter Kaon PID	`NotPion`	`NotPion` (slow)	-
Likelihood	0.75	0.75	0.75
$\lvert m_{\mathrm{ES}} - 5.28\ \mathrm{GeV}/c^2 \rvert$ (MeV/c^2)	± 5.2	± 5.2	± 5.2
ΔE (MeV)	± 36	± 36	± 36
Efficiency (%)	25.1 ± 0.3	16.5 ± 0.3	18.7 ± 0.3
Signal	43.7	26.7	17.0
Generic Background Sidebands	12.0	20.5	7.3
$S/\sqrt{B}$	12.6	5.9	6.3
$S/\sqrt{S+B}$	5.9	3.9	3.4

Table 4.7. Optimized selection criteria for $B^0 \to D_s^{*+}\pi^-$ (and $B^0 \to D_s^{*-}K^+$).

Cut	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
D_s^+ mass (MeV/c^2)	± 10	± 10	± 10
$\lvert\cos\theta_{\mathrm{flight}}\rvert$	-	-	< 0.98
Hard π^- PID (K^- PID)	`Tight`	`Tight`	`Tight`
Bachelor Kaon PID	`NotPion`	`Tight` (fast)	`Tight`
Daughter Kaon PID	`NotPion`	`Loose` (slow)	-
Likelihood	0.8	0.8	0.8
$\lvert m_{\mathrm{ES}} - 5.28\ \mathrm{GeV}/c^2 \rvert$ (MeV/c^2)	± 5.2	± 5.2	± 5.2
ΔE (MeV/c^2)	± 36	± 36	± 36
Efficiency (%)	13.2 ± 0.2	9.2 ± 0.2	10.4 ± 0.2
Signal	11.4	8.3	5.3
Generic Background Sidebands	4.5	6.4	2.7
$S/\sqrt{B}$	5.4	3.3	3.2
$S/\sqrt{S+B}$	2.9	2.2	1.9

4.7 Background Studies

As mentioned earlier, one of the main challenges of measuring rare branching fractions, such as those for $B^0 \to D_s^{(*)+}\pi^-$ and $B^0 \to D_s^{(*)-}K^+$, is the high levels of background that can overwhelm the signals from these decays.

The backgrounds for this analysis are divided into two categories:

- Combinatorial Background, which come from random combinations of tracks and neutrals that can emulate the signal events. These backgrounds arise from continuum $c\bar{c}$ and $u\bar{u}$,$d\bar{d}$,$s\bar{s}$, or $B\bar{B}$ events. Combinatorial backgrounds can be parameterized by an Argus function in $m_{\rm ES}$ and are typically uniformly distributed in ΔE. However, combinatorial events that happen to have a real D_s^+ are peaking in m_{Ds}.

- Peaking Background, which peaks in $m_{\rm ES}$, mainly come from mis-reconstructed B candidates from $B\bar{B}$ events. Because they are from mis-reconstructed events, the ΔE distribution is often shifted with respect to the signal distribution, and applying a cut or fitting for ΔE helps to fight against such backgrounds. Since these types of events often have a real D_s^+, they may have a peak in m_{Ds}.

The following sections describe the background studies done before unblinding the data.

4.7.1 Background Estimates on Simulated Continuum

The background from continuum events has been studied using generic $c\bar{c}$, $u\bar{u}$, $d\bar{d}$ and $s\bar{s}$ Monte Carlo events. The equivalent luminosities of these samples are 271 fb^{-1} for $c\bar{c}$ and 241 fb^{-1} for $u\bar{u}/d\bar{d}/s\bar{s}$.

Tables 4.8-4.9 shows the number of continuum background events for 208.7 fb^{-1} of data. Figures D.1-D.8 in Appendix D show the simulated continuum background distributions in the $(\Delta E, m_{\rm ES})$ plane after applying the final selection. The numbers for the signal box are computed by counting the number of the generic MC events in the signal box after the

final selection and scaling the results to the data luminosity. The number of background events is also computed by counting the events in the sideband and ΔE sideband regions and scaling by the ratio of the areas in the $(\Delta E, m_{\rm ES})$ plane and to the data luminosity. The sideband and ΔE sideband regions are defined in Section 4.5. The estimates from the sidebands are scaled assuming a uniform distribution, not taking into account the shapes of the ΔE and $m_{\rm ES}$ distributions in the sidebands, which can result in an *overestimate* of the background extrapolated into the signal box, in particular for $c\bar{c}$ background events.

4.7.2 Background Estimates on Simulated $B\bar{B}$

The background from generic $B\bar{B}$ events is estimated using $744\,\mathrm{fb}^{-1}$ $B^0\overline{B}^0$ simulated events and $763\,\mathrm{fb}^{-1}$ for B^+B^- simulated events. Tables 4.8-4.9 show signal box, sideband, and ΔE sideband estimates for the expected number of $B^0\overline{B}^0$ and B^+B^- background events, scaled to the data luminosity. Figures D.9-D.16 in Appendix D show the simulated $B\bar{B}$ background distributions in the $(\Delta E, m_{\rm ES})$ plane after applying the final selection. Again, the estimates from sidebands are scaled assuming a uniform distribution.

For $B\bar{B}$ events, the majority of events are peaking in $m_{\rm ES}$ but uniform in ΔE. A naive counting and scaling from sidebands to signal box area in this case can *underestimate* the background. On the other hand, there are also $B\bar{B}$ backgrounds that peak in the ΔE sidebands but not in the signal box, which leads to an *overestimate*. In summary, for $B\bar{B}$ backgrounds, the scale in $m_{\rm ES}$ may be underestimated and the scale in ΔE overestimated.

Table 4.8. Number of generic background events scaled to 208.7 fb^{-1} for $B^0 \to D_s^+\pi^-$

$B^0\overline{B}^0$	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Signal Box	3.1 ± 0.9	7.3 ± 1.4	3.4 ± 1.0
Sidebands	1.1 ± 0.1	3.2 ± 0.1	0.9 ± 0.0
ΔE Sidebands	1.1 ± 0.1	3.5 ± 0.1	1.1 ± 0.1
B^+B^-	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Signal Box	1.9 ± 0.7	7.4 ± 1.4	1.1 ± 0.5
Sidebands	1.3 ± 0.1	4.0 ± 0.1	0.9 ± 0.0
ΔE Sidebands	1.4 ± 0.1	3.9 ± 0.1	0.9 ± 0.1
uds	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Signal Box	4.3 ± 1.9	6.9 ± 2.4	0.0 ± 0.0
Sidebands	1.7 ± 0.2	5.7 ± 0.4	1.9 ± 0.2
ΔE Sidebands	1.6 ± 0.2	5.4 ± 0.4	2.1 ± 0.3
$c\bar{c}$	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Signal Box	9.2 ± 2.7	3.8 ± 1.7	4.6 ± 1.9
Sidebands	7.9 ± 0.4	7.6 ± 0.4	3.6 ± 0.3
ΔE Sidebands	7.7 ± 0.5	7.0 ± 0.4	3.8 ± 0.3

Table 4.9. Number of generic background events scaled to 208.7 fb^{-1} for $B^0 \to D_s^{*+}\pi^-$

$B^0\overline{B}^0$	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Signal Box	0.6 ± 0.4	1.4 ± 0.6	0.6 ± 0.4
Sidebands	0.5 ± 0.0	1.0 ± 0.0	0.3 ± 0.0
ΔE Sidebands	0.5 ± 0.0	1.0 ± 0.1	0.4 ± 0.0
B^+B^-	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Signal Box	1.4 ± 0.6	1.1 ± 0.5	0.3 ± 0.3
Sidebands	0.6 ± 0.0	1.9 ± 0.1	0.5 ± 0.0
ΔE Sidebands	0.6 ± 0.0	1.9 ± 0.1	0.5 ± 0.0
uds	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Signal Box	0.0 ± 0.0	1.7 ± 1.2	0.9 ± 0.9
Sidebands	0.5 ± 0.1	1.0 ± 0.2	0.5 ± 0.1
ΔE Sidebands	0.5 ± 0.1	0.9 ± 0.2	0.4 ± 0.1
$c\bar{c}$	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Signal Box	2.3 ± 1.3	1.5 ± 1.1	0.8 ± 0.8
Sidebands	2.9 ± 0.2	2.5 ± 0.2	1.4 ± 0.2
ΔE Sidebands	2.8 ± 0.3	2.6 ± 0.3	1.5 ± 0.2

Table 4.10. Number of generic background events scaled to 208.7 fb^{-1} for $B^0 \to D_s^- K^+$

$B^0\overline{B}^0$	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Signal Box	2.5 ± 0.8	4.2 ± 1.1	2.5 ± 0.8
Sidebands	0.1 ± 0.0	0.4 ± 0.0	0.1 ± 0.0
ΔE Sidebands	0.0 ± 0.0	0.4 ± 0.0	0.0 ± 0.0
B^+B^-	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Signal Box	0.3 ± 0.3	0.5 ± 0.4	0.3 ± 0.3
Sidebands	0.2 ± 0.0	0.9 ± 0.0	0.2 ± 0.0
ΔE Sidebands	0.2 ± 0.0	0.9 ± 0.1	0.2 ± 0.0
uds	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Signal Box	1.7 ± 1.2	0.9 ± 0.9	0.9 ± 0.9
Sidebands	0.9 ± 0.2	3.1 ± 0.3	1.5 ± 0.2
ΔE Sidebands	0.8 ± 0.2	3.3 ± 0.3	1.6 ± 0.2
$c\bar{c}$	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Signal Box	0.0 ± 0.0	2.3 ± 1.3	0.0 ± 0.0
Sidebands	1.7 ± 0.2	3.2 ± 0.2	0.8 ± 0.1
ΔE Sidebands	1.7 ± 0.2	2.9 ± 0.3	0.8 ± 0.1

Table 4.11. Number of generic background events scaled to 208.7 fb^{-1} for $B^0 \to D_s^{*-} K^+$

$B^0\overline{B}^0$	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Signal Box	0.0 ± 0.0	0.3 ± 0.3	0.3 ± 0.3
Sidebands	0.0 ± 0.0	0.1 ± 0.0	0.0 ± 0.0
ΔE Sidebands	0.0 ± 0.0	0.1 ± 0.0	0.0 ± 0.0
B^+B^-	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Signal Box	0.0 ± 0.0	1.4 ± 0.6	0.0 ± 0.0
Sidebands	0.1 ± 0.0	0.3 ± 0.0	0.0 ± 0.0
ΔE Sidebands	0.0 ± 0.0	0.3 ± 0.0	0.0 ± 0.0
uds	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Signal Box	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Sidebands	0.1 ± 0.0	0.6 ± 0.1	0.3 ± 0.1
ΔE Sidebands	0.0 ± 0.0	0.6 ± 0.1	0.4 ± 0.1
$c\bar{c}$	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Signal Box	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Sidebands	0.5 ± 0.1	0.7 ± 0.1	0.3 ± 0.1
ΔE Sidebands	0.6 ± 0.1	0.8 ± 0.1	0.2 ± 0.1

4.7.3 Peaking Background Studies

We also perform additional studies on various exclusive samples of two-body B^0 and B^+ decays with a potential for peaking in the signal region. The samples we consider are listed in Appendix A. Figures D.17-D.3.4 in Appendix D show the distributions in the $(\Delta E, m_{\rm ES})$ plane for the different modes after applying the final selection.

For many peaking backgrounds, the background peaks in the $m_{\rm ES}$ signal region but is shifted in ΔE. An example is $B^0 \to D_s^{*+}\pi^-$ as a background for $B^0 \to D_s^+\pi^-$, as seen in the Figure D.17 of Appendix D. In this case, a photon is accidentally lost in the reconstruction of $B^0 \to D_s^{*+}\pi^-$, and the event is mistaken as $B^0 \to D_s^+\pi^-$. The ΔE distribution, then, is shifted toward negative values of ΔE and falls rapidly approaching the signal box. Another example of a peaking background for $B^0 \to D_s^+\pi^-$ is $B^0 \to D_s^-K^+$, when the fast kaon is misidentified as a pion.

Tables 4.12-4.23 summarize the expected number of peaking background events for the samples that have nonzero entries in the signal box, scaled to the data luminosity. The first error quoted is the statistical error from Monte Carlo. The second error is from the error on the branching ratios.

In order to scale these numbers to the data luminosity, we must make an assumption for the branching fractions. The branching ratios listed in the tables are for the B^0 decay and do not include the sub-decay branching ratios, although they are included in the calculation of the expected number of background events. The *BABAR* measurements for branching ratios are used where possible; otherwise values from the PDG are used [12]. For the $B \to D_s^{+(*)}X$ modes (X=light meson), the expected branching ratios are similar to the $B^0 \to D_s^+\pi^-$ branching ratio, and therefore the equivalent luminosities of the Monte Carlo samples are thousands times that for data. For those modes with unknown branching ratios, we scale known branching ratios to estimate the expected background. We assume a 100% branching ratio error for those modes for with unknown branching ratios.

In addition to the peaking backgrounds listed here, "charmless" modes that do not have a D or D_s^+ meson, such as $B \to K^{(*)}K\pi$ or $B \to K^{(*)}KK$, can be a dangerous background since they can peak in the signal box for both $m_{\rm ES}$ and ΔE. Their branching fractions are largely unmeasured, so it is difficult to make accurate estimates for their contributions using Monte Carlo. In order to constrain the contribution of charmless backgrounds to the measured yields for $B^0 \to D_s^+\pi^-$ and $B^0 \to D_s^- K^+$, we perform a two-dimensional fit over $m_{\rm ES}$ and m_{Ds} variables, as described in Section 5.2. For $B^0 \to D_s^{*+}\pi^-$ and $B^0 \to D_s^{*-}K^+$, the additional photon from the D_s^{*+} decay gives us a greater ability to constrain the background and a multi-dimensional fit is not required.

Another type of possible peaking background are "reflection" backgrounds from $B^0 \to D^{(*)-}\pi^+$ and $B^0 \to D^{*-}K^+$ decays. These backgrounds may peak in $m_{\rm ES}$ and have a broad D_s^+ mass peak. However, these reflections have the branching fractions that are relatively well known ($\sim \mathcal{O}(10^{-3})$) and can therefore be constrained in the fit. Because these types of decays are already included in the generic $B\bar{B}$ Monte Carlo samples, their contributions are not included here to avoid double counting. However, the reflection backgrounds are taken into account in the final fit for yield for $B^0 \to D_s^{(*)+}\pi^-$. In the case of $B^0 \to D_s^{(*)-}K^+$, reflection backgrounds are not deemed a large contribution since the branching fractions for modes such as $B^0 \to D^{*-}K^+$ are on the order of $\mathcal{O}(10^{-4})$, an order of magnitude smaller than for the $B^0 \to D^{(*)-}\pi^+$ modes.

Any significant peaking background contributions are taken into account in the yield extraction, as described in Section 5.7. Peaking backgrounds that posed only a small but nonzero background contamination are not included in the fit but included in the systematic errors. The most significant peaking backgrounds for $B^0 \to D_s^+\pi^-$ are from $B^0 \to D_s^- K^+$, charmless backgrounds, and $B^0 \to D^-\pi^+$ and $B^0 \to D^{*-}\pi^+$ reflection backgrounds. For $B^0 \to D_s^{*+}\pi^-$, the only significant peaking backgrounds are from $B^0 \to D^{*-}\pi^+$ and $B^0 \to D^-\rho^+$ reflections. For $B^0 \to D_s^- K^+$, the main peaking backgrounds are from $B^0 \to D_s^+\pi^-$ and charmless backgrounds. No significant peaking backgrounds are seen for $B^0 \to D_s^{*-}K^+$.

Table 4.12. Number of peaking background events in 208.7 fb^{-1}, estimated from exclusive MC, for $B^0 \to D_s^+\pi^-, D_s^+ \to \phi\pi^+$

	$B.F.(10^{-5})$	N_{raw}	N_{scaled}
$B^0 \to D_s^{*+}\pi^-$	1.9 ± 1.3	6	0.02 ± 0.01 ± 0.01
$B^0 \to D_s^+\rho^-$	0.2 ± 0.7	1	0.00 ± 0.00 ± 0.00
$B^0 \to D_s^- K^+$	3.2 ± 1.0	179	0.52 ± 0.04 ± 0.16
$B^0 \to D_s^{*-} K^+$	3.2 ± 1.0	1	0.00 ± 0.00 ± 0.00
$B^+ \to D_s^+\pi^0$	1.6 ± 1.0	1	0.00 ± 0.00 ± 0.00
$Total$	-	-	0.54 ± 0.04 ± 0.16

Table 4.13. Number of peaking background events in 208.7 fb^{-1}, estimated from exclusive MC, for $B^0 \to D_s^+\pi^-, D_s^+ \to \bar{K}^{*0}K^+$

	$B.F.(10^{-5})$	N_{raw}	N_{scaled}
$B^0 \to D_s^{*+}\pi^-$	1.9 ± 1.3	4	0.02 ± 0.01 ± 0.01
$B^0 \to D_s^+\rho^-$	0.2 ± 0.7	2	0.00 ± 0.00 ± 0.00
$B^0 \to D_s^- K^+$	3.2 ± 1.0	137	0.45 ± 0.04 ± 0.14
$B^0 \to D_s^{*-} K^+$	0.6 ± 0.9	1	0.00 ± 0.00 ± 0.00
$B^0 \to D_s^- K^{*+}$	5.9 ± 5.9	1	0.01 ± 0.01 ± 0.01
$B^0 \to D_s^{*+} K^{*-}$	5.9 ± 5.9	1	0.01 ± 0.01 ± 0.01
$B^+ \to D_s^{*+}\rho^0$	0.1 ± 0.1	1	0.00 ± 0.00 ± 0.00
$Total$	-	-	0.49 ± 0.04 ± 0.14

Table 4.14. Number of peaking background events in 208.7 fb^{-1}, estimated from exclusive MC, for $B^0 \to D_s^+\pi^-, D_s^+ \to \bar{K}^0K^+$

	$B.F.(10^{-5})$	N_{raw}	N_{scaled}
$B^0 \to D_s^{*+}\pi^-$	1.9 ± 1.3	6	0.00 ± 0.00 ± 0.01
$B^0 \to D_s^- K^+$	3.2 ± 1.0	147	0.30 ± 0.02 ± 0.09
$Total$	-	-	0.30 ± 0.02 ± 0.09

Table 4.15. Number of peaking background events in 208.7 fb^{-1}, estimated from exclusive MC, for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$

	$B.F.(10^{-5})$	N_{raw}	N_{scaled}
$B^0 \to D_s^+\pi^-$	3.2 ± 0.9	12	0.07 ± 0.02 ± 0.02
$B^0 \to D_s^+\rho^-$	0.2 ± 0.7	19	0.01 ± 0.00 ± 0.02
$B^0 \to D_s^{*+}\rho^-$	0.2 ± 0.7	2	0.00 ± 0.00 ± 0.00
$B^0 \to D_s^- K^+$	3.2 ± 1.0	16	0.05 ± 0.01 ± 0.01
$B^0 \to D_s^{*-} K^+$	0.6 ± 0.9	67	0.03 ± 0.00 ± 0.05
$B^+ \to D_s^+\rho^0$	0.1 ± 0.1	14	0.00 ± 0.00 ± 0.00
$B^+ \to D_s^{*+}\rho^0$	0.1 ± 0.1	2	0.00 ± 0.00 ± 0.00
Total	-	-	0.16 ± 0.02 ± 0.06

Table 4.16. Number of peaking background events in 208.7 fb^{-1}, estimated from exclusive MC, for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$

	$B.F.(10^{-5})$	N_{raw}	N_{scaled}
$B^0 \to D_s^+\pi^-$	3.2 ± 0.9	7	0.05 ± 0.02 ± 0.01
$B^0 \to D_s^+\rho^-$	0.2 ± 0.7	18	0.01 ± 0.00 ± 0.03
$B^0 \to D_s^{*+}\rho^-$	0.2 ± 0.7	1	0.00 ± 0.00 ± 0.00
$B^0 \to D_s^- K^+$	3.2 ± 1.0	11	0.04 ± 0.01 ± 0.01
$B^0 \to D_s^{*-} K^+$	0.6 ± 0.9	50	0.03 ± 0.00 ± 0.04
$B^+ \to D_s^+\rho^0$	0.1 ± 0.1	16	0.01 ± 0.00 ± 0.01
$B^+ \to D_s^{*+}\rho^0$	0.1 ± 0.1	2	0.00 ± 0.00 ± 0.00
Total	-	-	0.12 ± 0.02 ± 0.06

Table 4.17. Number of peaking background events in 208.7 fb^{-1}, estimated from exclusive MC, for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^0K^+$

	$B.F.(10^{-5})$	N_{raw}	N_{scaled}
$B^0 \to D_s^+\pi^-$	3.2 ± 0.9	4	0.01 ± 0.01 ± 0.00
$B^0 \to D_s^+\rho^-$	0.2 ± 0.7	28	0.01 ± 0.00 ± 0.03
$B^0 \to D_s^{*+}\rho^-$	0.2 ± 0.7	1	0.00 ± 0.00 ± 0.00
$B^0 \to D_s^- K^+$	3.2 ± 1.0	18	0.04 ± 0.01 ± 0.01
$B^0 \to D_s^{*-} K^+$	0.6 ± 0.9	49	0.02 ± 0.00 ± 0.03
$B^+ \to D_s^+\rho^0$	0.1 ± 0.1	17	0.00 ± 0.00 ± 0.00
Total	-	-	0.08 ± 0.01 ± 0.04

Table 4.18. Number of peaking background events in 208.7 fb^{-1}, estimated from exclusive MC, for $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$

	$B.F.(10^{-5})$	N_{raw}	N_{scaled}
$B^0 \to D_s^+\pi^-$	3.2 ± 0.9	35	0.20 ± 0.03 ± 0.06
$B^0 \to D_s^{*+}\pi^-$	1.9 ± 1.3	4	0.01 ± 0.01 ± 0.01
$B^0 \to D_s^{*-}K^+$	3.2 ± 1.0	6	0.00 ± 0.00 ± 0.00
$B^0 \to D_s^- K^{*+}$	5.9 ± 5.9	1	0.01 ± 0.01 ± 0.01
$B^0 \to D^- K^+$, $D^- \to K^+\pi^-\pi^-$	20.0 ± 6.0	2	0.07 ± 0.05 ± 0.02
$B^0 \to D^- K^+$, $D^- \to K^{*0}\pi^-$	20.0 ± 6.0	3	0.01 ± 0.01 ± 0.00
Total	-	-	0.31 ± 0.06 ± 0.06

Table 4.19. Number of peaking background events in 208.7 fb^{-1}, estimated from exclusive MC, for $B^0 \to D_s^- K^+$, $D_s^+ \to \bar{K}^{0*}K^+$

	$B.F.(10^{-5})$	N_{raw}	N_{scaled}
$B^0 \to D_s^+\pi^-$	3.2 ± 0.9	21	0.14 ± 0.03 ± 0.04
$B^0 \to D_s^{*+}\pi^-$	1.9 ± 1.3	4	0.02 ± 0.01 ± 0.01
$B^0 \to D_s^{*-}K^+$	3.2 ± 1.0	5	0.00 ± 0.00 ± 0.00
$B^0 \to D^- K^+$, $D^- \to K^+\pi^-\pi^-$	20.0 ± 6.0	6	0.22 ± 0.09 ± 0.07
$B^0 \to D^- K^+$, $D^- \to K^{*0}\pi^-$	20.0 ± 6.0	17	0.06 ± 0.01 ± 0.02
Total	-	-	0.45 ± 0.10 ± 0.08

Table 4.20. Number of peaking background events in 208.7 fb^{-1}, estimated from exclusive MC, for $B^0 \to D_s^- K^+$, $D_s^+ \to \bar{K}^0 K^+$

	$B.F.(10^{-5})$	N_{raw}	N_{scaled}
$B^0 \to D_s^+\pi^-$	3.2 ± 0.9	28	0.11 ± 0.02 ± 0.03
$B^0 \to D_s^{*+}\pi^-$	1.9 ± 1.3	3	0.01 ± 0.00 ± 0.00
$B^0 \to D_s^{*-}K^+$	3.2 ± 1.0	1	0.00 ± 0.00 ± 0.00
$B^0 \to D^- K^+$, $D^- \to K_s\pi^-$	20.0 ± 6.0	30	0.06 ± 0.01 ± 0.02
Total	-	-	0.17 ± 0.01 ± 0.04

Table 4.21. Number of peaking background events in 208.7 fb^{-1}, estimated from exclusive MC, for $B^0 \to D_s^{*-} K^+$, $D_s^+ \to \phi\pi^+$

	$B.F.(10^{-5})$	N_{raw}	N_{scaled}
$B^0 \to D_s^{*+}\pi^-$	1.9 ± 1.3	12	0.04 ± 0.01 ± 0.03
$B^0 \to D_s^- K^+$	3.2 ± 1.0	2	0.01 ± 0.00 ± 0.00
$B^0 \to D_s^- K^{*+}$	5.9 ± 5.9	10	0.10 ± 0.03 ± 0.10
$B^0 \to D_s^{*+} K^{*-}$	5.9 ± 5.9	1	0.01 ± 0.01 ± 0.01
$Total$	-	-	0.15 ± 0.04 ± 0.11

Table 4.22. Number of peaking background events in 208.7 fb^{-1}, estimated from exclusive MC, for $B^0 \to D_s^{*-} K^+$, $D_s^+ \to \bar{K}^{0*} K^+$

	$B.F.(10^{-5})$	N_{raw}	N_{scaled}
$B^0 \to D_s^{*+}\pi^-$	1.9 ± 1.3	15	0.06 ± 0.01 ± 0.04
$B^0 \to D_s^- K^+$	3.2 ± 1.0	2	0.01 ± 0.00 ± 0.00
$B^0 \to D_s^- K^{*+}$	5.9 ± 5.9	8	0.10 ± 0.04 ± 0.10
$B^0 \to D_s^{*+} K^{*-}$	5.9 ± 5.9	1	0.01 ± 0.01 ± 0.01
$Total$	-	-	0.18 ± 0.04 ± 0.11

Table 4.23. Number of peaking background events in 208.7 fb^{-1}, estimated from exclusive MC, for $B^0 \to D_s^{*-} K^+$, $D_s^+ \to \bar{K}^0 K^+$

	$B.F.(10^{-5})$	N_{raw}	N_{scaled}
$B^0 \to D_s^{*+}\pi^-$	1.9 ± 1.3	16	0.03 ± 0.01 ± 0.02
$B^0 \to D_s^+ \rho^-$	0.2 ± 0.7	1	0.00 ± 0.00 ± 0.03
$B^0 \to D_s^- K^+$	3.2 ± 1.0	2	0.00 ± 0.00 ± 0.00
$B^0 \to D_s^- K^{*+}$	5.9 ± 5.9	5	0.04 ± 0.02 ± 0.04
$Total$	-	-	0.08 ± 0.02 ± 0.04

4.7.4 Background Estimates on Data Sidebands

In addition to the studies using generic Monte Carlo, We cross-check the background estimates using the data sidebands before unblinding. The numbers are scaled by the ratio of the area of the signal box to the area of the sidebands. The results are shown in Tables 4.24 and 4.25. The agreement between data and generic MC sidebands is reasonable, though, as mentioned earlier the numbers from generic MC can be overestimates in some cases.

Table 4.24. Number of background events in data sidebands for $B^0 \to D_s^+ \pi^-$

	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0 K^+$
Data Sidebands	10.6 ± 0.6	15.8 ± 0.7	5.3 ± 0.4
Data ΔE Sidebands	10.7 ± 0.7	15.6 ± 0.8	5.3 ± 0.5

Table 4.25. Number of background events in data sidebands for $B^0 \to D_s^{*+} \pi^-$

	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0 K^+$
Data Sidebands	4.0 ± 0.4	5.8 ± 0.4	1.6 ± 0.2
Data ΔE Sidebands	4.2 ± 0.4	5.9 ± 0.5	1.6 ± 0.3

Table 4.26. Number of background events in data sidebands for $B^0 \to D_s^- K^+$

	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0 K^+$
Data Sidebands	3.5 ± 0.3	7.6 ± 0.5	1.7 ± 0.2
Data ΔE Sidebands	3.2 ± 0.4	7.9 ± 0.6	2.0 ± 0.3

Table 4.27. Number of background events in data sidebands for $B^0 \to D_s^{*-} K^+$

	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0 K^+$
Data Sidebands	0.6 ± 0.1	1.6 ± 0.2	0.4 ± 0.1
Data ΔE Sidebands	0.5 ± 0.1	1.8 ± 0.3	0.3 ± 0.1

4.7.5 Summary of Estimated Backgrounds

A summary of the background estimates is given in Tables 4.28-4.31. The numbers for generic MC are computed from the sum of the contributions for $c\overline{c}$, $u\bar{u}$,$d\bar{d}$,$s\bar{s}$, and $B\bar{B}$ MC. The peaking background numbers are computed as the sum of all peaking background contributions listed in Tables 4.12-4.23.

Table 4.28. Summary of background estimates for $B^0 \to D_s^+\pi^-$ in 208.7 fb^{-1} from generic Monte Carlo, peaking background Monte Carlo, and data.

	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Generic Signal Box	18.5 ± 12.2	25.4 ± 12.6	9.1 ± 4.9
Generic Sidebands	12.0 ± 0.2	20.5 ± 0.3	7.3 ± 0.1
Generic ΔE Sidebands	11.8 ± 0.3	19.8 ± 0.3	7.9 ± 0.2
Peaking Signal Box	$0.54 \pm 0.04 \pm 0.16$	$0.49 \pm 0.04 \pm 0.14$	$0.30 \pm 0.02 \pm 0.09$
Data Sidebands	10.6 ± 0.6	15.8 ± 0.7	5.3 ± 0.4
Data ΔE Sidebands	10.7 ± 0.7	15.6 ± 0.8	5.3 ± 0.5

Table 4.29. Summary of background estimates for $B^0 \to D_s^{*+}\pi^-$ in 208.7 fb^{-1} from generic Monte Carlo, peaking background Monte Carlo, and data.

	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Generic Signal Box	4.3 ± 2.2	5.7 ± 3.3	2.6 ± 1.7
Generic Sidebands	4.5 ± 0.1	6.4 ± 0.1	2.7 ± 0.1
Generic ΔE Sidebands	4.4 ± 0.1	6.4 ± 0.2	2.8 ± 0.1
Peaking Signal Box	$0.16 \pm 0.02 \pm 0.06$	$0.12 \pm 0.02 \pm 0.06$	$0.08 \pm 0.01 \pm 0.04$
Data Sidebands	4.0 ± 0.4	5.8 ± 0.4	1.6 ± 0.2
Data ΔE Sidebands	4.2 ± 0.4	5.9 ± 0.5	1.6 ± 0.3

Table 4.30. Summary of background estimates for $B^0 \to D_s^- K^+$ in 208.7 fb^{-1} from generic Monte Carlo, peaking background Monte Carlo, and data.

	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0 K^+$
Generic Signal Box	4.5 ± 2.2	7.9 ± 3.9	3.7 ± 1.5
Generic Sidebands	2.9 ± 0.1	7.6 ± 0.1	2.6 ± 0.1
Generic ΔE Sidebands	2.7 ± 0.1	7.5 ± 0.2	2.6 ± 0.1
Peaking Signal Box	$0.31 \pm 0.06 \pm 0.06$	$0.45 \pm 0.10 \pm 0.08$	$0.17 \pm 0.01 \pm 0.04$
Data Sidebands	3.5 ± 0.3	7.6 ± 0.5	1.7 ± 0.2
Data ΔE Sidebands	3.2 ± 0.4	7.9 ± 0.6	2.0 ± 0.3

Table 4.31. Summary of background estimates for $B^0 \to D_s^{*-} K^+$ in 208.7 fb^{-1} from generic Monte Carlo, peaking background Monte Carlo, and data.

	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0 K^+$
Generic Signal Box	0.0 ± 0.0	1.7 ± 0.5	0.3 ± 0.1
Generic Sidebands	0.7 ± 0.0	1.7 ± 0.0	0.6 ± 0.0
Generic ΔE Sidebands	0.6 ± 0.0	1.8 ± 0.0	0.6 ± 0.0
Peaking Signal Box	$0.15 \pm 0.04 \pm 0.11$	$0.18 \pm 0.04 \pm 0.11$	$0.08 \pm 0.02 \pm 0.04$
Data Sidebands	0.6 ± 0.1	1.6 ± 0.2	0.4 ± 0.1
Data ΔE Sidebands	0.5 ± 0.1	1.8 ± 0.3	0.3 ± 0.1

Chapter 5

Yields and Branching Fraction Measurement

5.1 Yield Extraction Overview

The method for extracting the signal yields and branching fractions differ between the modes, depending on the level of peaking backgrounds. In all cases, we perform an unbinned extended maximum likelihood fit. In addition, we fit to all three D_s decay modes simultaneously, constraining $\mathcal{B}(B^0 \to D_s^{*+}\pi^-)$ to be the same for all D_s three modes. An extended fit directly determines the yields, in addition to the PDF parameters. A description of extended likelihood PDFs can be found in Appendix B.

For the modes with significant peaking backgrounds, $B^0 \to D_s^+\pi^-$ and $B^0 \to D_s^-K^+$, a two-dimensional unbinned extended maximum likelihood fit is performed over $m_{\rm ES}$ and m_{Ds} variables in the region $5.2 < m_{\rm ES} < 5.3$ GeV, $|m_{Ds} - 1.9683| < 50$ MeV after applying a cut of $|\Delta E| < 36$ MeV. The two-dimensional fit constrains the signal yield as well as peaking background contributions from charmless B decays.

Since the backgrounds for $B^0 \to D_s^{*+}\pi^-$ and $B^0 \to D_s^{*-}K^+$ are much less significant (as

shown in Section 4.7), we perform a 1-dimensional extended unbinned maximum likelihood fit to the $m_{\rm ES}$ distribution, fitting for all three D_s^+ modes simultaneously.

The fits for the $B^0 \to D_s^+\pi^-$ and $B^0 \to D_s^{*+}\pi^-$ yields are described in Sections 5.2 and 5.3, respectively. The fits for the $B^0 \to D_s^-K^+$ and $B^0 \to D_s^{*-}K^+$ yields are described in Sections 5.4 and 5.5, respectively. All of the studies in these sections are performed using Monte Carlo, before unblinding.

5.2 Yield Extraction Method for $B^0 \to D_s^+\pi^-$

In order to constrain peaking backgrounds, a two-dimensional fit over $m_{\rm ES}$ and m_{Ds} variables is performed for $B^0 \to D_s^+\pi^-$. As mentioned earlier, charmless backgrounds can peak in both $m_{\rm ES}$ and ΔE. However, we can distinguish the signal mode from the charmless background by the peak in m_{Ds}.

The PDF for the $B^0 \to D_s^+\pi^-$ fit consists of the following additive components:

- Signal: single Gaussian in $m_{\rm ES}$ and double Gaussian in m_{Ds}.
- Combinatorial Background: Argus function in $m_{\rm ES}$ and a first-order polynomial plus a double Gaussian peak (constrained to the same shape as signal PDF) in m_{Ds}. The Gaussian is included to account for the fraction of combinatorial events with a real D_s and a random fast pion.
- Peaking backgrounds:
  - $B^0 \to D^{(*)-}X$ reflection background from $B^0 \to D^-\pi^+$ and $B^0 \to D^{*-}\pi^+$, $D^{*-} \to D^0\pi^-$: Crystal Ball in $m_{\rm ES}$ and a single Gaussian in m_{Ds}.
  - Charmless background: single Gaussian in $m_{\rm ES}$, constrained to the same shape as the signal PDF. First-order polynomial in m_{Ds}.
  - $B^0 \to D_s^-K^+$ background: same PDF as signal.

We determine the shapes for these components from exclusive and generic MC.

The event likelihood is constructed for each decay mode as

$$\begin{aligned}
\mathcal{L} &= n_{\rm sig}\,\mathcal{S}(m_{\rm ES})\mathcal{S}(m_{Ds}) \\
&+ n_{\rm comb}\,\mathcal{C}(m_{\rm ES})\,(\eta_p\mathcal{S}(m_{Ds}) + (1-\eta_p)\mathcal{C}(m_{Ds})) \\
&+ n_{\rm refl}\,\mathcal{R}(m_{\rm ES})\mathcal{R}(m_{Ds}) \\
&+ n_{\rm charmless}\,\mathcal{S}(m_{\rm ES})\mathcal{C}(m_{Ds}) \\
&+ n_{DsK}\,\mathcal{S}(m_{\rm ES})\mathcal{S}(m_{Ds})
\end{aligned} \tag{5.1}$$

where $\mathcal{S}, \mathcal{C}, \mathcal{R}$ correspond to the signal, combinatorial background, and reflection background PDFs, and:

$$\begin{aligned}
\mathcal{S}(m_{\rm ES}) &= \mathcal{G}(m_{\rm ES}; \mu_S^{m_{\rm ES}}, \sigma_S^{m_{\rm ES}}) \\
\mathcal{S}(m_{Ds}) &= \left(\eta_S\mathcal{G}(m_{Ds}-1.97; \mu_{1S}^{m_{Ds}}, \sigma_{1S}^{m_{\rm ES}}) + (1-\eta_S)\mathcal{G}(m_{Ds}-1.97; \mu_{2S}^{m_{Ds}}, \sigma_{2S}^{m_{\rm ES}}\right) \\
\mathcal{C}(m_{\rm ES}) &= \mathcal{A}(m_{\rm ES}; \kappa) \\
\mathcal{C}(m_{Ds}) &= (1 + p_1(m_{\rm ES}-1.97) + p_2(m_{\rm ES}-1.97)^2) \\
\mathcal{R}(m_{\rm ES}) &= \mathcal{CB}(m_{\rm ES}; \mu_r^{m_{\rm ES}}, \sigma_r^{m_{\rm ES}}, \alpha_r^{m_{\rm ES}}, n_r^{m_{\rm ES}}) \\
\mathcal{R}(m_{Ds}) &= \mathcal{G}(m_{Ds}-1.97; \mu_r^{m_{Ds}}, \sigma_r^{m_{Ds}})
\end{aligned} \tag{5.2}$$

where $\mathcal{G}$ is a Gaussian, $\mathcal{A}$ is an Argus function, and $\mathcal{CB}$ is a Crystal Ball function. Event yields $n_{\rm sig}$, $n_{\rm refl}$, and n_{DsK} are expressed in terms of the corresponding branching fractions:

$$\begin{aligned}
n_{\rm sig} &= n_{BB}\cdot\epsilon_i(B^0\to D_s^+\pi^-)\cdot\mathcal{B}_i\cdot\mathcal{B}(B^0\to D_s^+\pi^-) \\
n_{\rm refl} &= n_{BB}\cdot\epsilon_i(B^0\to D^{(*)-}X)\cdot\mathcal{B}(B^0\to D^{(*)-}X) \\
n_{\rm DsK} &= n_{BB}\cdot\epsilon_i(B^0\to D_s^-K^+)\cdot\mathcal{B}_i\cdot\mathcal{B}(B^0\to D_s^-K^+)
\end{aligned} \tag{5.3}$$

where ϵ_i(decay) is the efficiency to reconstruct a particular decay as $B^0\to D_s^+\pi^-$ (or $B^0\to D_s^-K^+$) in a particular D_s decay mode i, n_{BB} is the number of $B\overline{B}$ events in the data sample, $\mathcal{B}_i$ is the branching fraction for the $D_s^+\to\phi\pi^+$, $D_s^+\to\bar{K}^{0*}K^+$, or $D_s^+\to\bar{K}^0K^+$

modes, and $\mathcal{B}$ is the branching fraction for the particular B mode. A simultaneous fit to all three D_s decay modes is performed, constraining $\mathcal{B}(B^0 \to D_s^+\pi^-)$ to be the same for all D_s three modes. Branching fractions of all peaking backgrounds, except for the charmless modes, are fixed to measured values [11, 12].

Where appropriate, the parameters of the PDFs are fixed to the values obtained from the Monte Carlo simulations. The list of parameters is summarized in Table 5.1. Since the fit operates over a large $m_{\rm ES}$ and m_{Ds} range, the signal and peaking background efficiencies are somewhat larger than quoted in Table 4.6. The efficiencies are summarized in Table 5.2.

The results of the signal Monte Carlo fits are shown in Fig. 5.1-5.3. For the reflection backgrounds, the PDF shape parameters are determined from a fit to the $B^0 \to D^{(*)}X$ "cocktail" Monte Carlo (Fig. 5.4- 5.6). Then the efficiencies for the reflections are determined from fits to the $B^0 \to D^-\pi^+$ (Fig. 5.7- 5.9) and $B^0 \to D^{*-}\pi^+, D^{*-} \to D^0\pi^-$ (Fig. 5.10-5.11). The peaking background fits for $B^0 \to D_s^- K^+$ are shown in Fig. 5.13-5.15.

To check the performance of the fit, we test the fit on a sample of generic Monte Carlo events equivalent to approximately 209 fb^{-1}, constructed from appropriately scaled $B^0\bar{B}^0$, B^+B^-, $c\bar{c}$, and $u\bar{u}$,$d\bar{d}$,$s\bar{s}$ samples. Since the generic Monte Carlo also includes the signal mode, as well as the reflection backgrounds, this sample is representative of what is expected in the data. The generic $B^0\overline{B}^0$ samples were generated with the branching fraction for $B^0 \to D_s^+\pi^-$ mode set to $\mathcal{B}(B^0 \to D_s^+\pi^-) = 3.2{\cdot}10^{-5}$. The plots of the fit to this sample are shown in Fig. 5.16-5.18. The result of the fit is $\mathcal{B}(B^0 \to D_s^+\pi^-)^{\rm fit} = (2.8 \pm 0.5) \cdot 10^{-5}$, with the $B^0 \to D_s^+\pi^-$ signal established at 7.4σ significance.

For comparison with a larger statistics sample, a fit on the full generic Monte Carlo sample is also performed, with approximately 744 fb^{-1} sample of generic $B^0\bar{B}^0$ decays. This fit returns $\mathcal{B}(B^0 \to D_s^+\pi^-)^{\rm fit} = (3.0 \pm 0.2) \cdot 10^{-5}$, and the plots can be seen in Fig. 5.19-5.21.

Table 5.1. Parameters of the $B^0 \to D_s^+\pi^-$ PDFs, determined from Monte Carlo simulations.

$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0K^+$
Signal		
$\mathcal{S}(m_{\rm ES})$ = Single Gaussian		
$\mu_S^{m_{\rm ES}} = 5.2796 \pm 0.00002$ GeV		
$\sigma_S^{m_{\rm ES}} = 2.553 \pm 0.014$ MeV		
$\mathcal{S}(m_{Ds})$ = Double Gaussian		
$\eta_S = 0.842 \pm 0.009$	$\eta_S = 0.816 \pm 0.012$	$\eta_S = 0.876 \pm 0.012$
$\mu_{1S}^{m_{Ds}} = 0.6 \pm 0.1$ MeV	$\mu_{1S}^{m_{Ds}} = 0.5 \pm 0.1$ MeV	$\mu_{1S}^{m_{Ds}} = 1.3 \pm 0.1$ MeV
$\sigma_{1S}^{m_{\rm ES}} = 4.8 \pm 0.1$ MeV	$\sigma_{1S}^{m_{\rm ES}} = 5.0 \pm 0.1$ MeV	$\sigma_{1S}^{m_{\rm ES}} = 5.9 \pm 0.1$ MeV
$\mu_{2S}^{m_{Ds}} = -2.3 \pm 0.6$ MeV	$\mu_{2S}^{m_{Ds}} = -2.4 \pm 0.7$ MeV	$\mu_{2S}^{m_{Ds}} = -1.0 \pm 0.8$ MeV
$\sigma_{2S}^{m_{\rm ES}} = 18.2 \pm 0.6$ MeV	$\sigma_{2S}^{m_{\rm ES}} = 20.0 \pm 0.8$ MeV	$\sigma_{2S}^{m_{\rm ES}} = 19.7 \pm 1.0$ MeV
Reflection		
$\mathcal{R}(m_{\rm ES})$ = Crystal Ball		
$\mu_r^{m_{\rm ES}} = 5.2801 \pm 0.0001$ GeV		
$\sigma_r^{m_{\rm ES}} = 2.57 \pm 0.12$ MeV		
$\alpha_r^{m_{\rm ES}} = 1.4 \pm 0.3$		
$n_r^{m_{\rm ES}} = 3.1 \pm 2.8$		
$\mathcal{R}(m_{Ds})$ = Single Gaussian		
$\mu_r^{m_{Ds}} = 0.070 \pm 0.015$ GeV	$\mu_r^{m_{Ds}} = 0.024 \pm 0.003$ GeV	$\mu_r^{m_{Ds}} = 0.038 \pm 0.019$ GeV
$\sigma_r^{m_{Ds}} = 0.039 \pm 0.033$ GeV	$\sigma_r^{m_{Ds}} = 0.026 \pm 0.003$ GeV	$\sigma_r^{m_{Ds}} = 0.048 \pm 0.015$ GeV

Table 5.2. Efficiencies for $B^0 \to D_s^+\pi^-$ in a fit region $5.2 < m_{\rm ES} < 5.3$ GeV, $|m_{Ds} - 1.9683| < 50$ MeV, $|\Delta E| < 36$ MeV

	$\epsilon(\phi\pi^+)$ (%)	$\epsilon(\bar{K}^{0*}K^+)$ (%)	$\epsilon(K_S^0K^+)$ (%)
Signal	30.2 ± 0.3	20.6 ± 0.3	23.3 ± 0.3
$B^0 \to D^-\pi^+$	$(7.7 \pm 0.7) \cdot 10^{-5}$	$(18.8 \pm 1.1) \cdot 10^{-5}$	$(8.3 \pm 0.7) \cdot 10^{-5}$
$B^0 \to D^{*-}\pi^+, D^{*-} \to D^0\pi^-$	$(0.04 \pm 0.07) \cdot 10^{-5}$	$(0.4 \pm 0.2) \cdot 10^{-5}$	$(1.2 \pm 0.3) \cdot 10^{-5}$
$B^0 \to D_s^-K^+$	$(4.1 \pm 0.3) \cdot 10^{-3}$	$(3.2 \pm 0.2) \cdot 10^{-3}$	$(3.4 \pm 0.2) \cdot 10^{-3}$

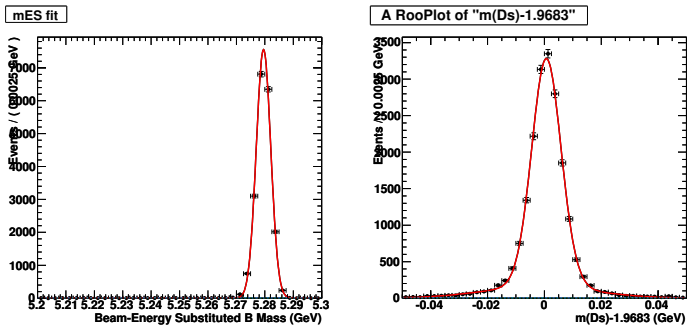


Figure 5.1. Projections of the two-dimensional fit to $B^0 \to D_s^+\pi^-$ signal Monte Carlo events, all D_s^+ modes combined. $m_{\rm ES}$ projection on the left and m_{Ds} projection on the right.

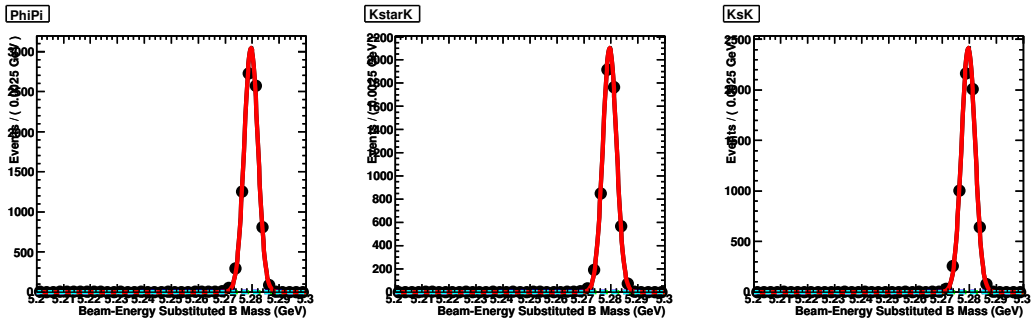


Figure 5.2. $m_{\rm ES}$ projections of the two-dimensional fit to $B^0 \to D_s^+\pi^-$ signal Monte Carlo events. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right).

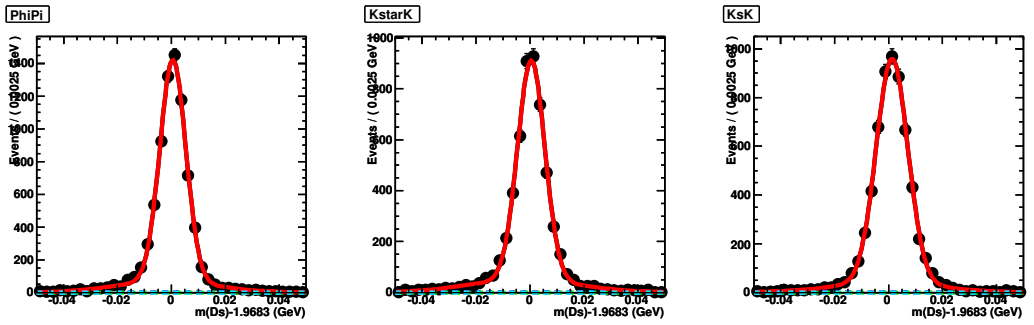


Figure 5.3. m_{Ds} projections of the two-dimensional fit to $B^0 \to D_s^+\pi^-$ signal Monte Carlo events. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right).

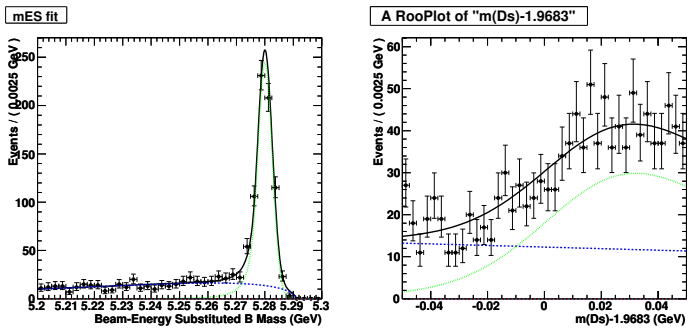


Figure 5.4. Projections of the two-dimensional fit to $B^0 \to D^{(*)}X$ reflection background events, all D_s^+ modes combined. $m_{\rm ES}$ projection on the left and m_{Ds} projection on the right. Dashed line shows combinatorial (non-peaking) background contribution, and the dotted shows the contribution from $B^0 \to D^{(*)-}\pi^+$ reflections.

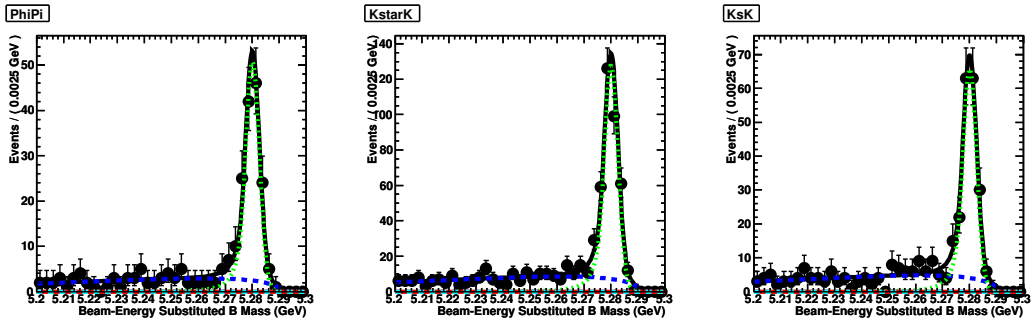


Figure 5.5. $m_{\rm ES}$ projections of the two-dimensional fit to $B^0 \to D^{(*)}X$ reflection background events. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right). Dashed line shows combinatorial (non-peaking) background contribution, and the dotted shows the contribution from $B^0 \to D^{(*)-}\pi^+$ reflections.

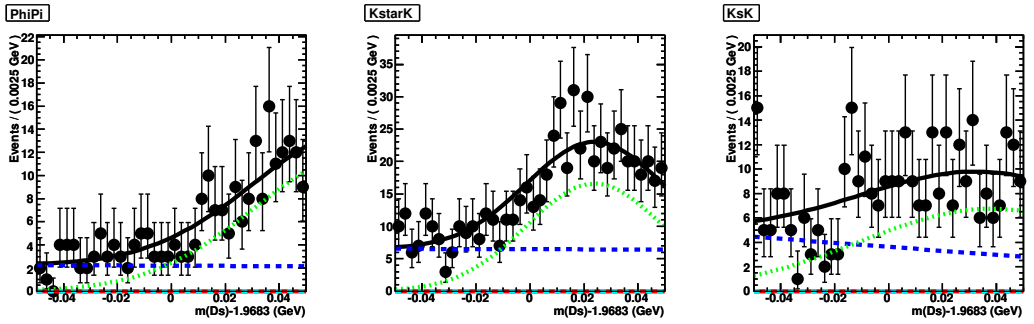


Figure 5.6. m_{Ds} projections of the two-dimensional fit to $B^0 \to D^{(*)}X$ reflection background events. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right). Dashed line shows combinatorial (non-peaking) background contribution, and the dotted shows the contribution from $B^0 \to D^{(*)-}\pi^+$ reflections.

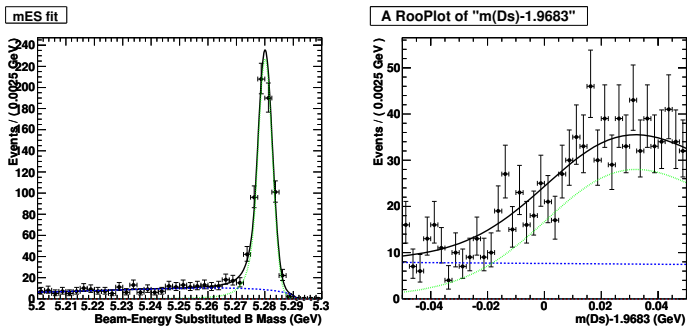


Figure 5.7. Projections of the two-dimensional fit to $B^0 \to D^- \pi^+$ reflection background events, all D_s^+ modes combined. $m_{\rm ES}$ projection on the left and m_{Ds} projection on the right. Dashed line shows combinatorial (non-peaking) background contribution, and the dotted shows the contribution from $B^0 \to D^{(*)-}\pi^+$ reflections.

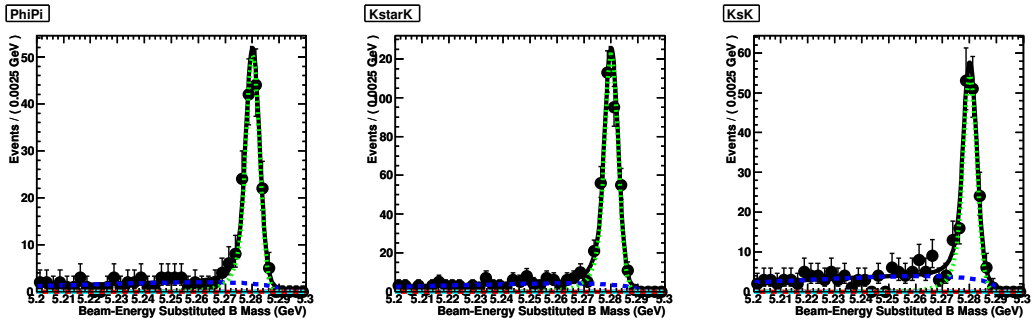


Figure 5.8. $m_{\rm ES}$ projections of the two-dimensional fit to $B^0 \to D^- \pi^+$ reflection background events. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0 K^+$ (right). Dashed line shows combinatorial (non-peaking) background contribution, and the dotted shows the contribution from $B^0 \to D^{(*)-}\pi^+$ reflections.

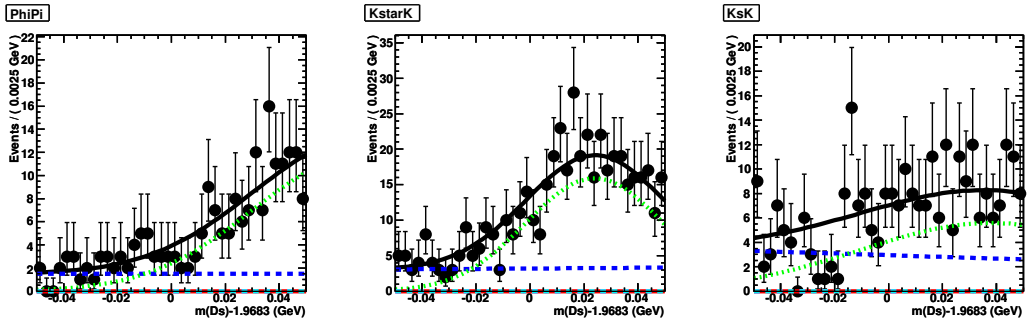


Figure 5.9. m_{Ds} projections of the two-dimensional fit to $B^0 \to D^- \pi^+$ reflection background events. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0 K^+$ (right). Dashed line shows combinatorial (non-peaking) background contribution, and the dotted shows the contribution from $B^0 \to D^{(*)-}\pi^+$ reflections.

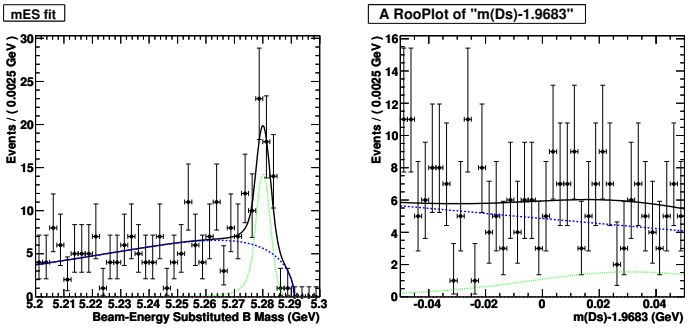


Figure 5.10. Projections of the two-dimensional fit to $B^0 \to D^{*-}\pi^+, D^{*-} \to D^0\pi^-$ reflection background events, all D_s^+ modes combined. $m_{\rm ES}$ projection on the left and m_{Ds} projection on the right. Dashed line shows combinatorial (non-peaking) background contribution, and the dotted shows the contribution from $B^0 \to D^{(*)-}\pi^+$ reflections.

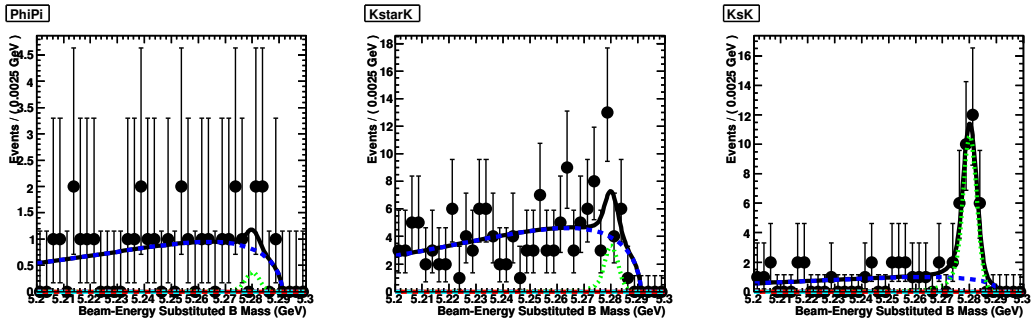


Figure 5.11. $m_{\rm ES}$ projections of the two-dimensional fit to $B^0 \to D^{*-}\pi^+, D^{*-} \to D^0\pi^-$ reflection background events. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right). Dashed line shows combinatorial (non-peaking) background contribution, and the dotted shows the contribution from $B^0 \to D^{(*)-}\pi^+$ reflections.

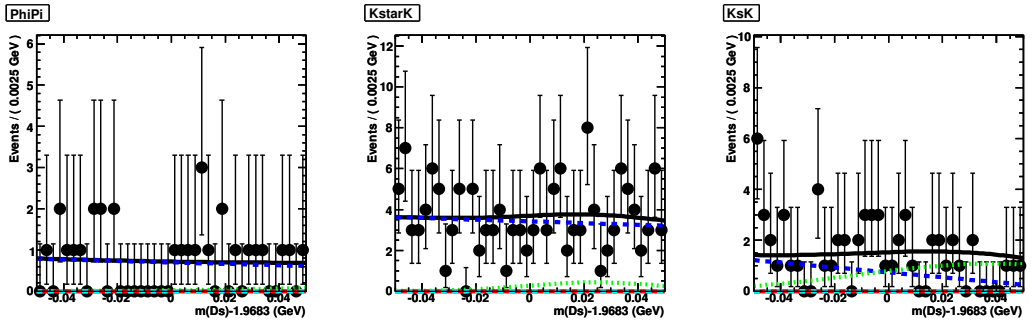


Figure 5.12. m_{Ds} projections of the two-dimensional fit to $B^0 \to D^{*-}\pi^+, D^{*-} \to D^0\pi^-$ reflection background events. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right). Dashed line shows combinatorial (non-peaking) background contribution, and the dotted shows the contribution from $B^0 \to D^{(*)-}\pi^+$ reflections.

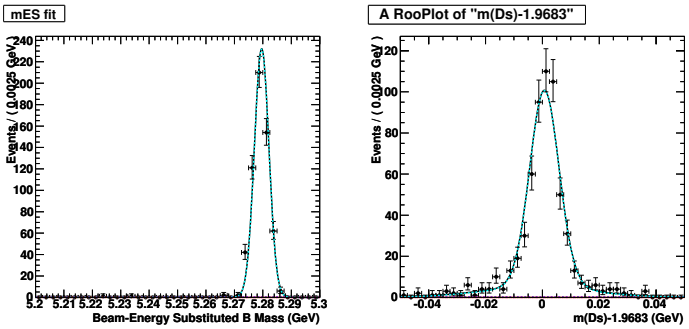


Figure 5.13. Projections of the two-dimensional fit to $B^0 \to D_s^- K^+$ reflection background events, all D_s^+ modes combined. $m_{\rm ES}$ projection on the left and m_{Ds} projection on the right.

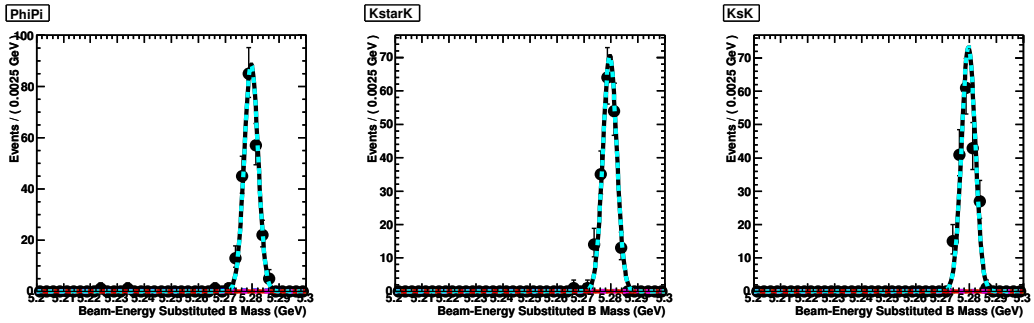


Figure 5.14. $m_{\rm ES}$ projections of the two-dimensional fit to $B^0 \to D_s^- K^+$ reflection background events. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*} K^+$ (middle), $D_s^+ \to \bar{K}^0 K^+$ (right).

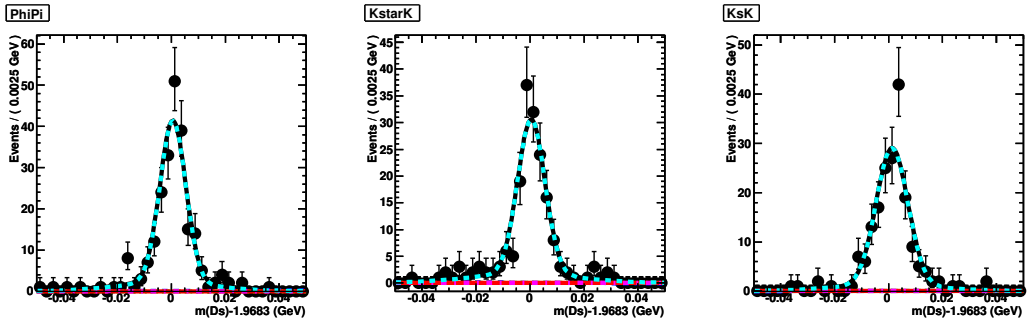


Figure 5.15. m_{Ds} projections of the two-dimensional fit to $B^0 \to D_s^- K^+$ reflection background events. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*} K^+$ (middle), $D_s^+ \to \bar{K}^0 K^+$ (right).

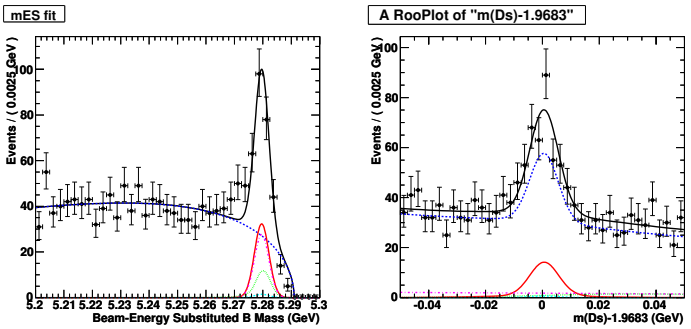


Figure 5.16. Projections of the two-dimensional fit to the generic Monte Carlo sample approximately equivalent to 209 fb^{-1} (properly scaled $B^0\bar{B}^0$, B^+B^-, $c\bar{c}$ and $u\bar{u}/d\bar{d}/s\bar{s}$ samples), all D_s^+ modes combined. $m_{\rm ES}$ projection on the left and m_{Ds} projection on the right. Dashed line shows combinatorial (non-peaking) background contribution, and the dotted shows the contribution from $B^0 \to D^{(*)-}\pi^+$ reflections.

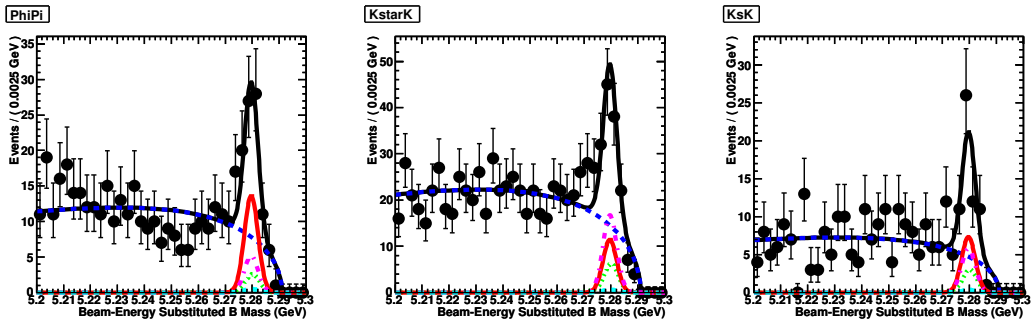


Figure 5.17. $m_{\rm ES}$ projections of the two-dimensional fit to the generic Monte Carlo sample approximately equivalent to 209 fb^{-1} (properly scaled $B^0\bar{B}^0$, B^+B^-, $c\bar{c}$ and $u\bar{u}/d\bar{d}/s\bar{s}$ samples). $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right). Dashed line shows combinatorial (non-peaking) background contribution, and the dotted shows the contribution from $B^0 \to D^{(*)-}\pi^+$ reflections.

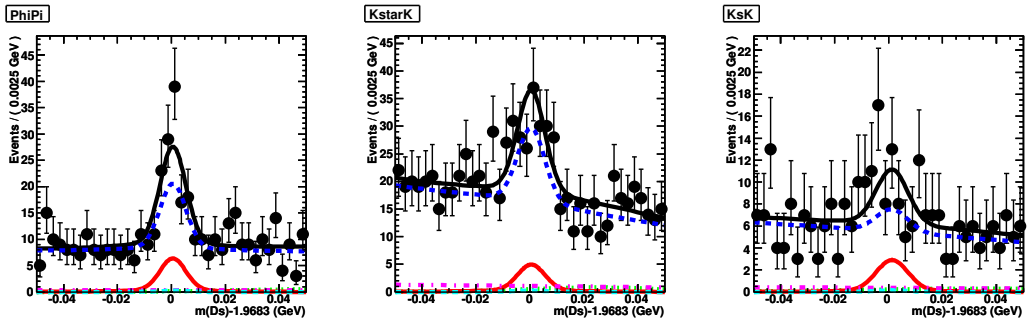


Figure 5.18. m_{Ds} projections of the two-dimensional fit to the generic Monte Carlo sample approximately equivalent to 209 fb^{-1} (properly scaled $B^0\bar{B}^0$, B^+B^-, $c\bar{c}$ and $u\bar{u}/d\bar{d}/s\bar{s}$ samples). $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right). Dashed line shows combinatorial (non-peaking) background contribution, and the dotted shows the contribution from $B^0 \to D^{(*)-}\pi^+$ reflections.

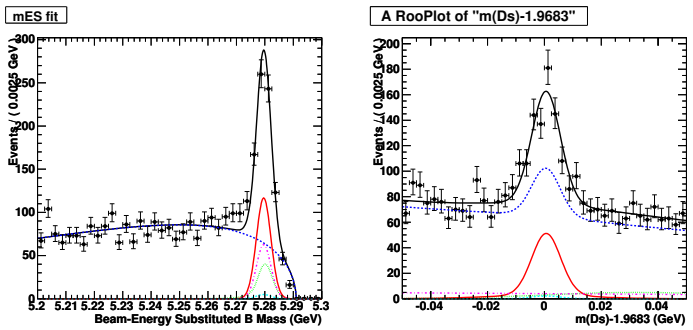


Figure 5.19. Projections of the two-dimensional fit to the generic Monte Carlo sample approximately equivalent to 744 fb^{-1} , all D_s^+ modes combined. $m_{\rm ES}$ projection on the left and m_{Ds} projection on the right. Dashed line shows combinatorial background contribution, and the dotted shows the contribution from $B^0 \to D^{(*)-}\pi^+$ reflections.

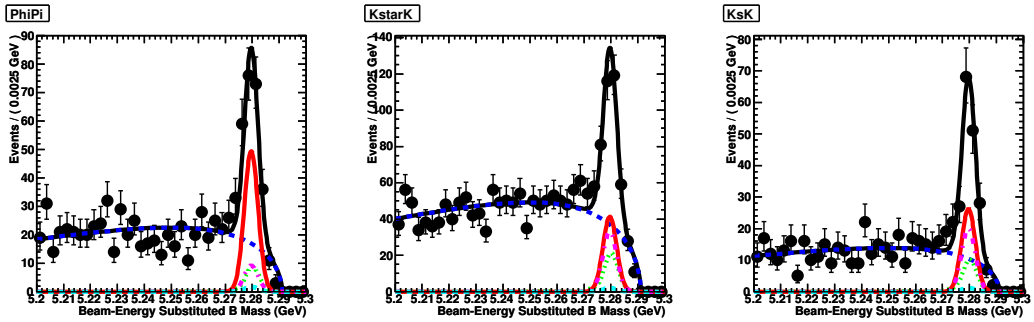


Figure 5.20. $m_{\rm ES}$ projections of the two-dimensional fit to the generic Monte Carlo sample approximately equivalent to 744 fb^{-1}. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right). Dashed line shows combinatorial background contribution, and the dotted shows the contribution from $B^0 \to D^{(*)-}\pi^+$ reflections.

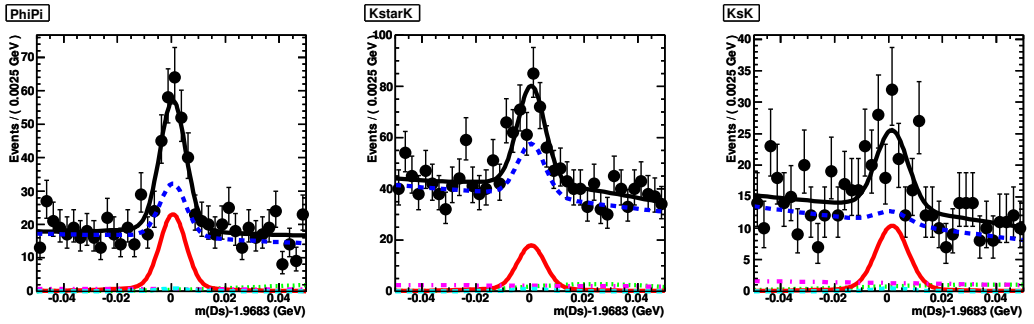


Figure 5.21. m_{Ds} projections of the two-dimensional fit to the generic Monte Carlo sample approximately equivalent to 744 fb^{-1}. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right). Dashed line shows combinatorial background contribution, and the dotted shows the contribution from $B^0 \to D^{(*)-}\pi^+$ reflections.

5.3 Yield Extraction Method for $B^0 \to D_s^{*+}\pi^-$

The yield for $B^0 \to D_s^{*+}\pi^-$ is obtained using a 1-dimensional extended maximum likelihood fit to the $m_{\rm ES}$ distribution, fitting directly for the branching fraction using an extended fit and fitting for all three D_s^+ modes simultaneously. A much simpler fit is chosen for $B^0 \to D_s^{*+}\pi^-$ since the peaking background contributions are much smaller than in the $B^0 \to D_s^+\pi^-$ case.

The PDF for $B^0 \to D_s^{*+}\pi^-$ consists of three additive components:

- $B^0 \to D_s^{*+}\pi^-$ Signal: Crystal Ball function
- Combinatorial Background: Argus function
- $B^0 \to D^{*-}\pi^+$ and $B^0 \to D^-\rho^+$ Reflection Backgrounds: Gaussian

Though the contribution from the reflection backgrounds are small for $B^0 \to D_s^{*+}\pi^-$, they are included for completeness. The branching fractions of the peaking backgrounds are fixed to measured values [11, 12].

The event likelihood is constructed for each decay mode as

$$\mathcal{L} = n_{\rm sig}\,\mathcal{S}(m_{\rm ES}) + n_{\rm comb}\,\mathcal{C}(m_{\rm ES}) + n_{\rm refl}\,\mathcal{R}(m_{\rm ES}) \tag{5.4}$$

where $\mathcal{S}, \mathcal{C}, \mathcal{R}$ correspond to the signal, combinatorial background, and reflection background PDFs, and:

$$\begin{aligned} \mathcal{S}(m_{\rm ES}) &= \mathcal{CB}(m_{\rm ES}; \mu_S^{m_{\rm ES}}, \sigma_S^{m_{\rm ES}}, \alpha_S^{m_{\rm ES}}, n_S^{m_{\rm ES}}) \\ \mathcal{C}(m_{\rm ES}) &= \mathcal{A}(m_{\rm ES}; \kappa) \\ \mathcal{R}(m_{\rm ES}) &= \mathcal{G}(m_{\rm ES}; \mu_r^{m_{\rm ES}}, \sigma_r^{m_{\rm ES}}) \end{aligned} \tag{5.5}$$

where $\mathcal{G}$ is a Gaussian, $\mathcal{A}$ is an Argus function, and $\mathcal{CB}$ is a Crystal Ball function. Event yields $n_{\rm sig}$ and $n_{\rm refl}$ are expressed in terms of the corresponding branching fractions:

$$\begin{aligned} n_{\rm sig} &= n_{BB} \cdot \epsilon_i(B^0 \to D_s^{*+}\pi^-) \cdot \mathcal{B}_i \cdot \mathcal{B}(B^0 \to D_s^+\pi^-) \\ n_{\rm refl} &= n_{BB} \cdot \epsilon_i(B^0 \to D^{(*)-}X) \cdot \mathcal{B}(B^0 \to D^{(*)-}X) \end{aligned}$$

where ϵ_i(decay) is the efficiency to reconstruct a particular decay as $B^0 \to D_s^{*+}\pi^-$ in a particular D_s decay mode i, n_{BB} is the number of $B\overline{B}$ events in the data sample, $\mathcal{B}_i$ is the branching fraction for the $D_s^+ \to \phi\pi^+$, $D_s^+ \to \bar{K}^{0*}K^+$, or $D_s^+ \to \bar{K}^0K^+$ modes, and $\mathcal{B}$ is the branching fraction for the particular B mode.

Again, many of the PDF parameters are determined from Monte Carlo. The efficiencies and $m_{\rm ES}$ shape parameters for signal are fixed to values obtained from signal Monte Carlo. The Argus shape parameter for the combinatorial background component is left floating in the final fit. The fits used to obtain the shape parameters and efficiencies are shown in Figures 5.22-5.25. The shape parameters and efficiencies are reported in Tables 5.3 and 5.4. Photon smearing has been applied to the $B^0 \to D_s^{*+}\pi^-$ Monte Carlo to better emulate the shapes and efficiencies that are seen data. As in the case of $B^0 \to D_s^+\pi^-$, the reflection PDF parameters are determined from a fit to the $B^0 \to D^{(*)}X$ Cocktail MC (Fig. 5.24-5.25). Then the $B^0 \to D^{*-}\pi^+$ and $B^0 \to D^-\rho^+$ samples are fitted separately to get the efficiencies for the different reflection modes (Fig. 5.26-5.29).

After fixing the shape parameters and efficiencies for signal and reflection backgrounds, we perform the full fit on $209\,\mathrm{fb}^{-1}$ of generic Monte Carlo. The fit is shown in Figures 5.30-5.31. The resulting branching ratio from the fit to generic Monte Carlo is $\mathcal{B}(B^0 \to D_s^{*+}\pi^-)^{\rm fit} = 3.1 \pm 0.6 \cdot 10^{-5}$, compared to the value in the generic decay file of $3.2 \cdot 10^{-5}$, with a significance of 6.5σ. For comparison, a fit on the full $B^0\bar{B}^0$ generic Monte Carlo sample available, with approximately 744 fb^{-1}, is performed. The fit returns $\mathcal{B}(B^0 \to D_s^{*+}\pi^-)^{\rm fit} = (3.1 \pm 0.3) \cdot 10^{-5}$, and the corresponding plots can be seen in Fig. 5.32-5.33.

Table 5.3. $B^0 \to D_s^{*+}\pi^-$ $m_{\rm ES}$ fit parameters, obtained from Monte Carlo.

	Parameters
Signal	Crystal Ball $\mu = 5.2796 \pm 0.00005$ GeV $\sigma = 2.56 \pm 0.04$ MeV $\alpha = 1.72 \pm 0.10$ $n = 3.1 \pm 0.4$
Combinatorial	Argus $\kappa =$ floating
Reflection	Gaussian $\mu = 5.2762 \pm 0.0012$ GeV $\sigma = 5.9 \pm 1.0$ MeV

Table 5.4. Efficiencies for $B^0 \to D_s^{*+}\pi^-$ (for $5.2 < m_{\rm ES} < 5.3$ GeV) obtained from fits to Monte Carlo. The efficiencies for signal have been determined after photon smearing has been applied to the Monte Carlo.

	$\epsilon(\phi\pi^+)$ (%)	$\epsilon(\bar{K}^{0*}K^+)$ (%)	$\epsilon(K_S^0K^+)$ (%)
Signal	13.9 ± 0.4 (%)	9.5 ± 0.3 (%)	10.4 ± 0.3 (%)
$B^0 \to D^{*-}\pi^+$Reflection	$0.1 \pm 0.1 \cdot 10^{-5}$	$1.5 \pm 0.4 \cdot 10^{-5}$	$0.9 \pm 0.4 \cdot 10^{-5}$
$B^0 \to D^-\rho^+$Reflection	$0.00 \pm 0.01 \cdot 10^{-5}$	$0.22 \pm 0.10 \cdot 10^{-5}$	$0.17 \pm 0.07 \cdot 10^{-5}$

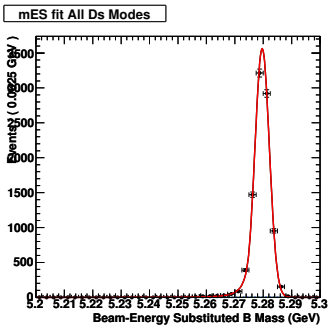


Figure 5.22. $m_{\rm ES}$ fit to signal Monte Carlo for $B^0 \to D_s^{*+}\pi^-$, all D_s^+ modes combined

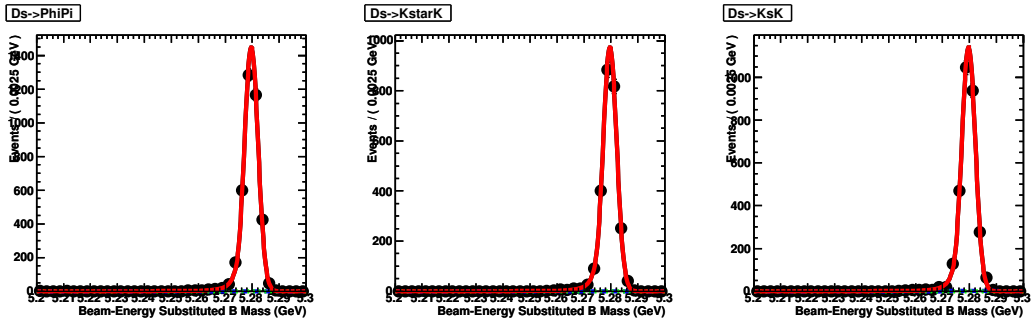


Figure 5.23. $m_{\rm ES}$ fit to signal Monte Carlo for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0 K^+$ (right)

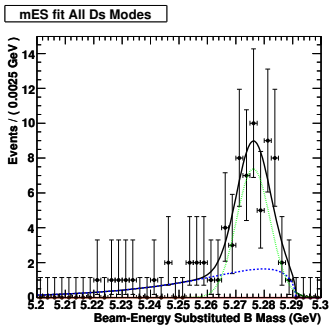


Figure 5.24. $m_{\rm ES}$ fit to $B^0 \to D^{(*)}X$ cocktail Monte Carlo for $B^0 \to D_s^{*+}\pi^-$, all D_s^+ modes combined. Dashed line shows combinatorial (non-peaking) background contribution. Dotted line shows reflection background contribution.

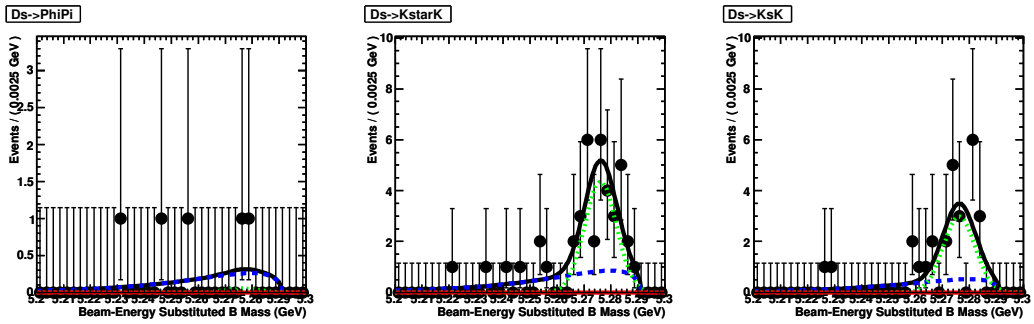


Figure 5.25. $m_{\rm ES}$ fit to $B^0 \to D^{(*)}X$ cocktail Monte Carlo for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right). Dashed line shows combinatorial (non-peaking) background contribution. Dotted line shows reflection background contribution.

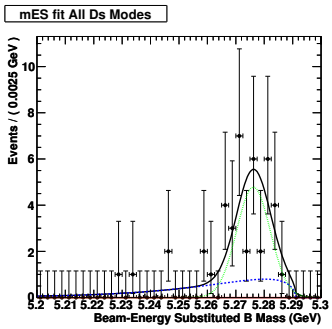


Figure 5.26. $m_{\rm ES}$ fit to $B^0 \to D^{*-}\pi^+$ from cocktail Monte Carlo for $B^0 \to D_s^{*+}\pi^-$, all D_s^+ modes combined. Dashed line shows combinatorial (non-peaking) background contribution. Dotted line shows reflection background contribution.

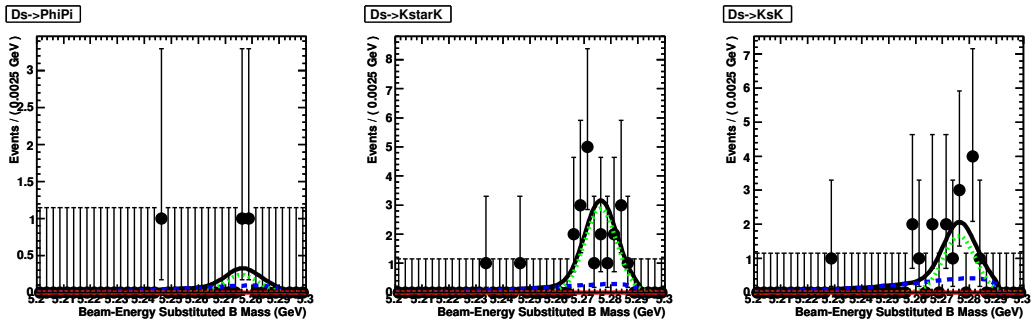


Figure 5.27. $m_{\rm ES}$ fit to $B^0 \to D^{*-}\pi^+$ from cocktail Monte Carlo for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right). Dashed line shows combinatorial (non-peaking) background contribution. Dotted line shows reflection background contribution.

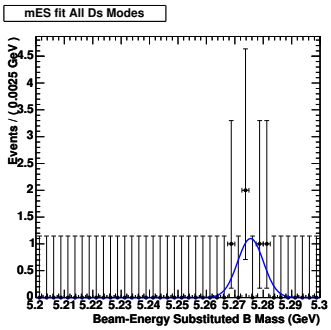


Figure 5.28. $m_{\rm ES}$ fit to $B^0 \to D^- \rho^+$ from cocktail Monte Carlo for $B^0 \to D_s^{*+} \pi^-$, all D_s^+ modes combined. Dashed line shows combinatorial (non-peaking) background contribution. Dotted line shows reflection background contribution.

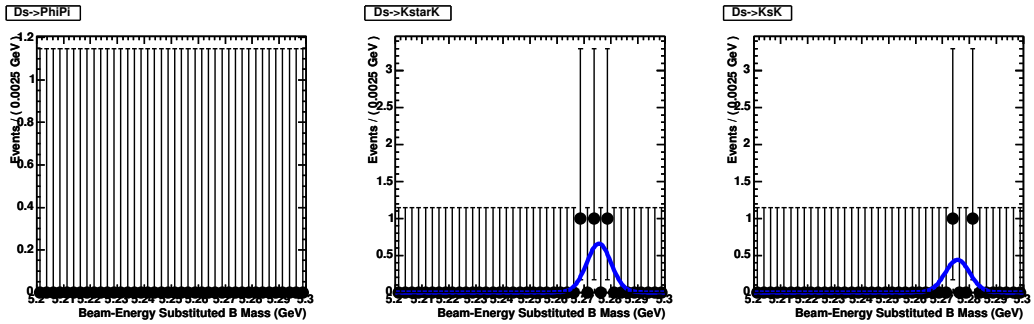


Figure 5.29. $m_{\rm ES}$ fit to $B^0 \to D^- \rho^+$ from cocktail Monte Carlo for $B^0 \to D_s^{*+} \pi^-$, $D_s^+ \to \phi \pi^+$ (left), $D_s^+ \to \bar{K}^{0*} K^+$ (middle), $D_s^+ \to \bar{K}^0 K^+$ (right). Dashed line shows combinatorial (non-peaking) background contribution. Dotted line shows reflection background contribution.

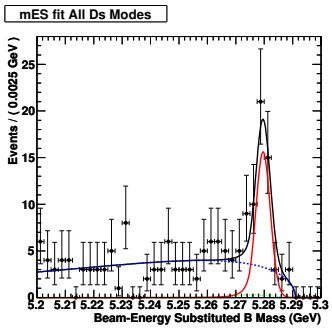


Figure 5.30. $m_{\rm ES}$ fit to 209 fb^{-1} generic Monte Carlo for $B^0 \to D_s^{*+}\pi^-$, all D_s^+ modes combined. Dashed line shows combinatorial (non-peaking) background contribution. Dotted line shows reflection background contribution.

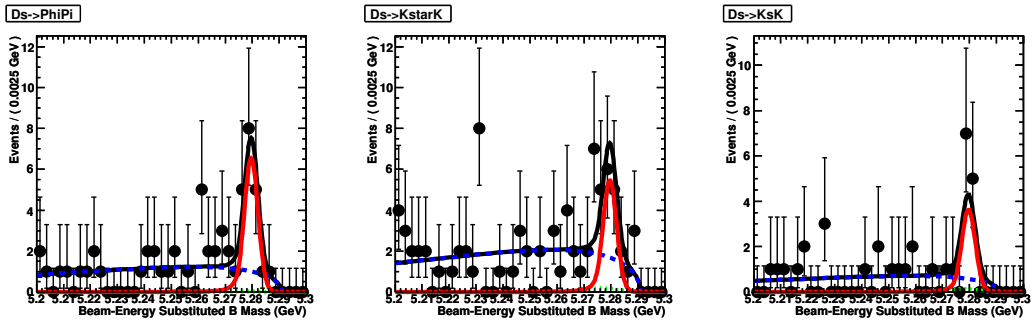


Figure 5.31. $m_{\rm ES}$ fit to 209 fb^{-1} generic Monte Carlo for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right). Dashed line shows combinatorial (non-peaking) background contribution. Dotted line shows reflection background contribution.

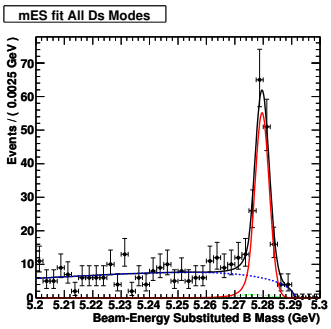


Figure 5.32. $m_{\rm ES}$ fit to 744 fb^{-1} generic $B^0\bar{B}^0$ Monte Carlo for $B^0 \to D_s^{*+}\pi^-$, all D_s^+ modes combined ($\mathcal{B}(B^0 \to D_s^{*+}\pi^-)$ fixed to zero). Dashed line shows combinatorial (non-peaking) background contribution. Dotted line shows reflection background contribution.

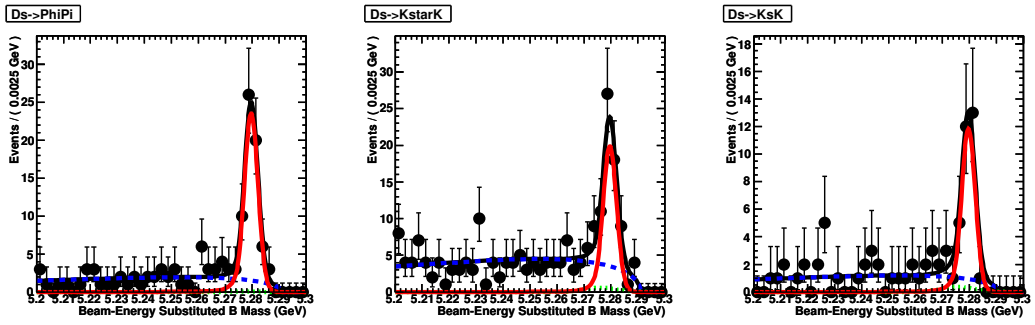


Figure 5.33. $m_{\rm ES}$ fit to 744 fb^{-1} generic $B^0\bar{B}^0$ Monte Carlo for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right). ($\mathcal{B}(B^0 \to D_s^{*+}\pi^-)$ fixed to zero). Dashed line shows combinatorial (non-peaking) background contribution. Dotted line shows reflection background contribution.

5.4 Yield Extraction Method for $B^0 \to D_s^- K^+$

The yield extraction for $B^0 \to D_s^- K^+$ is very similar to that for $B^0 \to D_s^+ \pi^-$, except that there is no PDF component for reflection backgrounds. The $B^0 \to D_s^- K^+$ fit is a 2-dimensional unbinned extended maximum likelihood fit to $m_{\rm ES}$ and m_{Ds}, and the PDF consists of the following additive components:

- Signal: single Gaussian in $m_{\rm ES}$ and double Gaussian in m_{Ds}.
- Combinatorial Background: Argus function in $m_{\rm ES}$ and a first-order polynomial plus a double Gaussian peak (constrained to the same shape as signal PDF) in m_{Ds}.
- Peaking backgrounds:
  - Charmless background: single Gaussian in $m_{\rm ES}$, constrained to the same shape as the signal PDF. First-order polynomial in m_{Ds}.
  - $B^0 \to D_s^+ \pi^-$ background: same PDF as signal.

The parameters for the PDF, determined from Monte Carlo, are shown in Table 5.5. The efficiencies for $B^0 \to D_s^- K^+$ in the fit region are shown in Table 5.6. The plots of the fits used to determine these parameters and efficiencies are shown in Figures 5.34-5.36.

Table 5.5. Parameters of the $B^0 \to D_s^- K^+$ PDFs, determined from Monte Carlo simulations.

$D_s^+ \to \phi\pi^+$	$D_s^+ \to \bar{K}^{0*}K^+$	$D_s^+ \to \bar{K}^0 K^+$
Signal		
$\mathcal{S}(m_{\rm ES})$ = Single Gaussian		
$\mu_S^{m_{\rm ES}} = 5.2795 \pm 0.00001$ GeV		
$\sigma_S^{m_{\rm ES}} = 2.560 \pm 0.010$ MeV		
$\mathcal{S}(m_{Ds})$ = Double Gaussian		
$\eta_S = 0.845 \pm 0.007$	$\eta_S = 0.814 \pm 0.009$	$\eta_S = 0.846 \pm 0.009$
$\mu_{1S}^{m_{Ds}} = 0.2 \pm 0.1$ MeV	$\mu_{1S}^{m_{Ds}} = 0.2 \pm 0.1$ MeV	$\mu_{1S}^{m_{Ds}} = 1.2 \pm 0.1$ MeV
$\sigma_{1S}^{m_{\rm ES}} = 4.8 \pm 0.1$ MeV	$\sigma_{1S}^{m_{\rm ES}} = 5.1 \pm 0.1$ MeV	$\sigma_{1S}^{m_{\rm ES}} = 5.7 \pm 0.1$ MeV
$\mu_{2S}^{m_{Ds}} = -2.4 \pm 0.5$ MeV	$\mu_{2S}^{m_{Ds}} = -2.4 \pm 0.5$ MeV	$\mu_{2S}^{m_{Ds}} = -1.8 \pm 0.5$ MeV
$\sigma_{2S}^{m_{\rm ES}} = 18.7 \pm 0.5$ MeV	$\sigma_{2S}^{m_{\rm ES}} = 19.7 \pm 0.6$ MeV	$\sigma_{2S}^{m_{\rm ES}} = 19.1 \pm 0.6$ MeV

Table 5.6. Efficiencies for $B^0 \to D_s^- K^+$ in a fit region $5.2 < m_{\rm ES} < 5.3$ GeV, $|m_{Ds} - 1.9683| < 50$ MeV, $|\Delta E| < 36$ MeV

	$\epsilon(\phi\pi^+)$ (%)	$\epsilon(\bar{K}^{0*}K^+)$ (%)	$\epsilon(K_S^0 K^+)$ (%)
Signal	23.4 ± 0.2	17.6 ± 0.2	19.0 ± 0.3
$B^0 \to D_s^+\pi^-$	$(1.5 \pm 0.2) \cdot 10^{-3}$	$(1.0 \pm 0.2) \cdot 10^{-3}$	$(1.1 \pm 0.2) \cdot 10^{-3}$

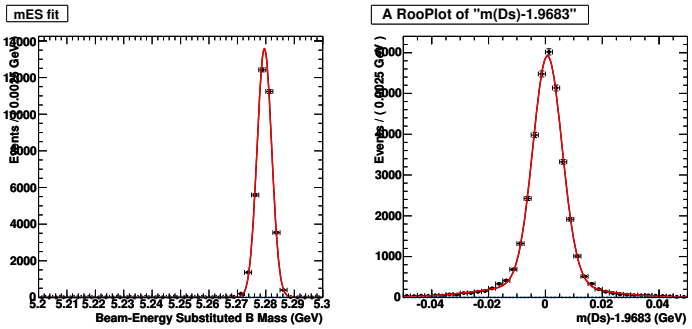


Figure 5.34. Projections of the two-dimensional fit to $B^0 \to D_s^+\pi^-$ signal Monte Carlo events, all D_s^+ modes combined. $m_{\rm ES}$ projection on the left and m_{Ds} projection on the right.

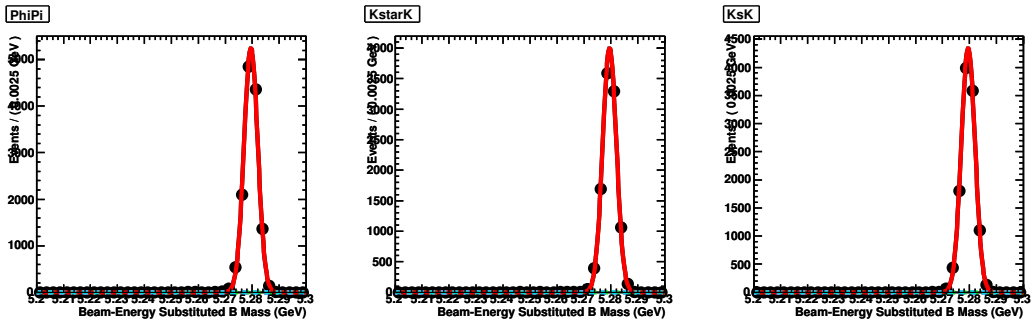


Figure 5.35. $m_{\rm ES}$ projections of the two-dimensional fit to $B^0 \to D_s^-K^+$ signal Monte Carlo events. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right).

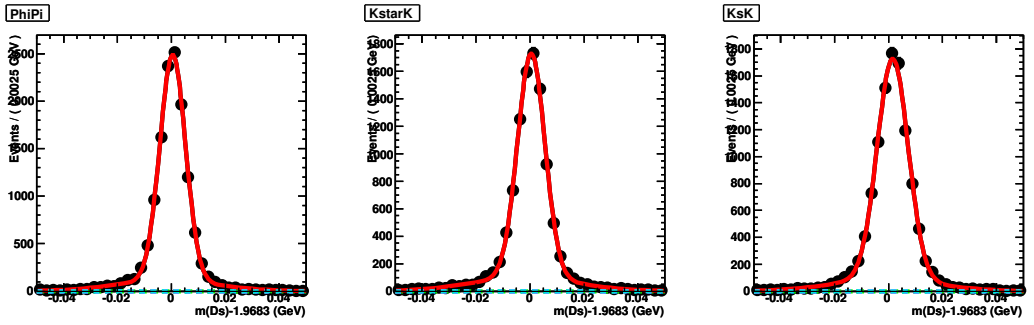


Figure 5.36. m_{Ds} projections of the two-dimensional fit to $B^0 \to D_s^-K^+$ signal Monte Carlo events. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right).

5.5 Yield Extraction Method for $B^0 \to D_s^{*-} K^+$

The yield extraction for $B^0 \to D_s^{*-} K^+$ is very similar to that for $B^0 \to D_s^{*+} \pi^-$, except that no reflection component is included. The $B^0 \to D_s^{*-} K^+$ fit is a 1-dimensional unbinned extended maximum likelihood fit. It is the simplest of the four fits. The PDF for $B^0 \to D_s^{*-} K^+$ consists of two additive components:

- $B^0 \to D_s^{*+} \pi^-$ Signal: Crystal Ball function
- Combinatorial Background: Argus function

There are no peaking background components, since none are deemed significant for $B^0 \to D_s^{*-} K^+$. The fit parameters determined from Monte Carlo are shown in Table 5.7, and the efficiencies in the fit region are shown in Table 5.8. The plots of the fits used to determine these parameters and efficiencies are shown in Figures 5.37-5.38.

Table 5.7. $B^0 \to D_s^{*-} K^+$ $m_{\rm ES}$ fit parameters, obtained from Monte Carlo.

	Parameters
Signal	Crystal Ball $\mu = 5.27967 \pm 0.00007$ GeV $\sigma = 3.06 \pm 0.06$ MeV $\alpha = 1.61 \pm 0.11$ $n = 1.5 \pm 0.3$
Combinatorial	Argus $\kappa =$ floating

Table 5.8. Efficiencies for $B^0 \to D_s^{*-} K^+$ (for $5.2 < m_{\rm ES} < 5.3$ GeV) obtained from fits to Monte Carlo. The efficiencies for signal have been determined after photon smearing has been applied to the Monte Carlo.

	$\epsilon(\phi\pi^+)$ (%)	$\epsilon(\bar{K}^{0*} K^+)$ (%)	$\epsilon(K_S^0 K^+)$ (%)
Signal	8.9 ± 0.1 (%)	6.6 ± 0.1 (%)	9.5 ± 0.2 (%)

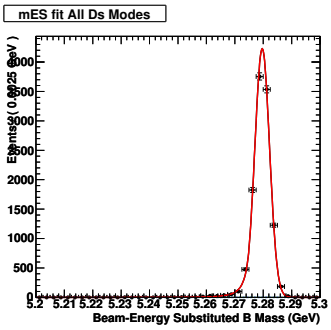


Figure 5.37. $m_{\rm ES}$ fit to signal Monte Carlo for $B^0 \to D_s^{*-} K^+$, all D_s^+ modes combined

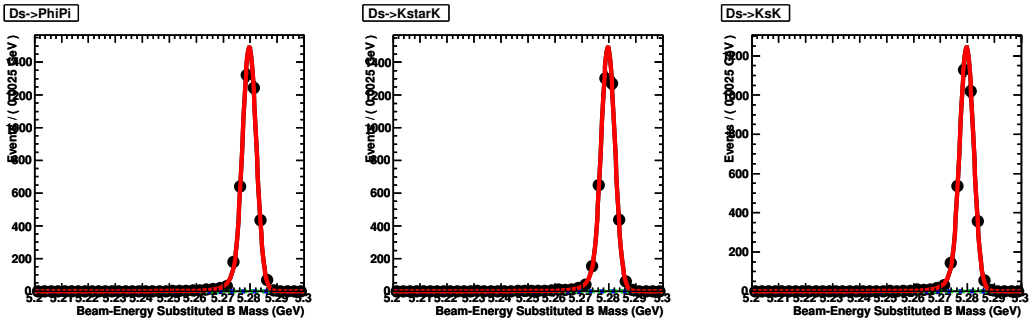


Figure 5.38. $m_{\rm ES}$ fit to signal Monte Carlo for $B^0 \to D_s^{*-} K^+$, $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*} K^+$ (middle), $D_s^+ \to \bar{K}^0 K^+$ (right)

5.6 $B \to D^{(*)}\pi$ Control Samples

We use a control sample of on-resonance data reconstructed as $B^0 \to D^-\pi^+$ (to compare with $B^0 \to D_s^+\pi^-$) and $B^+ \to \bar{D}^{0*}\pi^+$ (to compare with $B^0 \to D_s^{*+}\pi^-$) to study the means and resolutions of the ΔE distribution in data, as well as the differences in PDF shapes between data and Monte Carlo. The sample consists of:

- $207\,\text{fb}^{-1}$ data, reconstructed in the modes $B^0 \to D^-\pi^+$, with $D^+ \to K^-\pi^+\pi^+$, $\bar{K}^0\pi^+$ and $B^+ \to \bar{D}^{0*}\pi^+$, with $\bar{D}^{*0} \to \bar{D}^0\gamma$, $\bar{D}^0 \to K^-\pi^+$
- 466.0k events $B^0 \to D^-\pi^+$, $D^+ \to K^-\pi^+\pi^+$ Monte Carlo
- 169.0k events $B^0 \to D^-\pi^+$, $D^+ \to \bar{K}^0\pi^+$ Monte Carlo
- 192.5k events $B^+ \to \bar{D}^{0*}\pi^+$, $\bar{D}^{*0} \to \bar{D}^0\gamma$, $\bar{D}^0 \to K^-\pi^+$ Monte Carlo

A summary of the shifts seen in the control samples and applied to our data sample can be found in Table 5.14.

5.6.1 Control Sample ΔE Distributions

The ΔE resolutions and shifts are studied using $B^0 \to D^-\pi^+$ and $B^+ \to \bar{D}^{0*}\pi^+$ control samples, after applying a selection similar to our final selection for $B^0 \to D_s^{(*)+}\pi^-$. The results are listed in Tables 5.9-5.10, and the corresponding plots can be found in Figures 5.39-5.40. A 2-3 MeV shift is observed in the data ΔE distribution, and is taken into account in the final fit to the data.

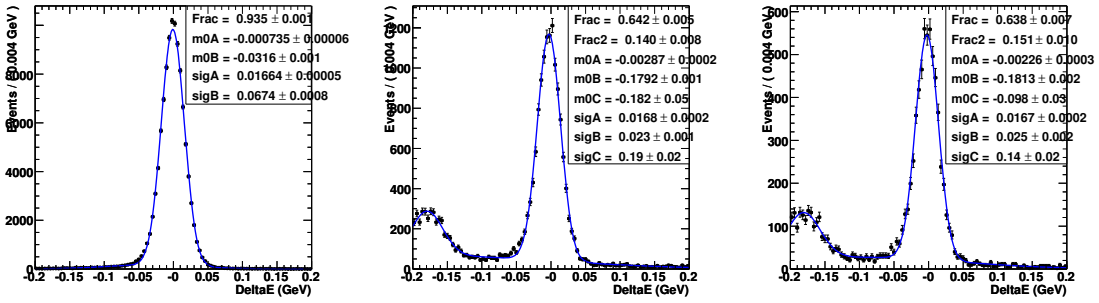
5.6.2 Control Sample PDF Shapes

In addition to checking the ΔE distribution in data, we also study the differences between PDF shapes in data and Monte Carlo. The same fitting code written to fit to

$B^0 \to D_s^{(*)+}\pi^-$ data (as described in Sections 5.2 and 5.3) is used, and the PDFs are fitted to the data control sample and corresponding Monte Carlo samples.

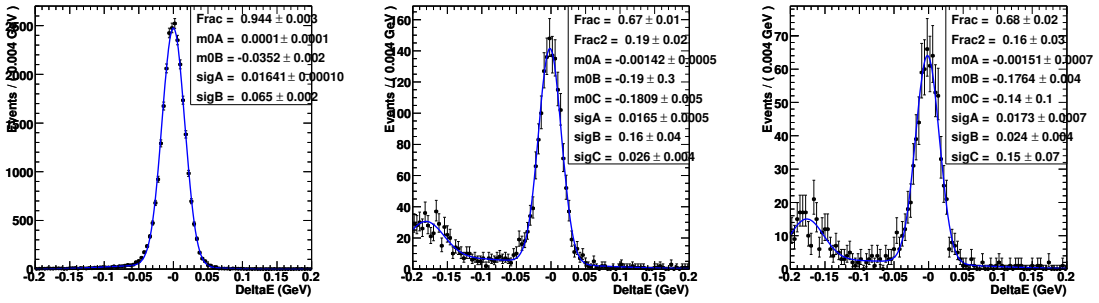
The fits to the data control samples are shown in Figures 5.42-5.44. A comparison of PDF parameters from the fits is shown in Tables 5.11-5.13. In some cases, the parameter in the data fit is fixed so that the fit would converge. There is generally good agreement between the parameters determined from the data and Monte Carlo control sample fits, although the distributions are wider in data than Monte Carlo. Since the $m_{\rm ES}$ PDF shape determined from the $D^{*0}\pi$ control sample is significantly wider than for the PDF determined from MC, a wider $m_{\rm ES}$ shape is used in the final $B^0 \to D_s^{*+}\pi^-$ and $B^0 \to D_s^{*-}K^+$ fits.

In addition to fitting for the PDF parameters, the branching ratios in data are fitted for as a check on the validity of our fitting method. The results are $\mathcal{B}(B^0 \to D^-\pi^+) = (2.68 \pm 0.02) \cdot 10^{-3}$ (for $D^+ \to K^-\pi^+\pi^+$) and $\mathcal{B}(B^0 \to D^-\pi^+) = (2.95 \pm 0.08) \cdot 10^{-3}$ (for $D^+ \to \bar{K}^0\pi^+$), compared to the PDG value of $(2.76 \pm 0.25) \cdot 10^{-3}$, and $\mathcal{B}(B^+ \to \bar{D}^{0*}\pi^+) = (6.5 \pm 0.2) \cdot 10^{-3}$, compared to the PDG value of $(4.6 \pm 0.4) \cdot 10^{-3}$ [12].



(a) MC (b) Run1-4 Data (c) Run4 Data

Figure 5.39. $B^0 \to D^-\pi^+$, $D^+ \to K^-\pi^+\pi^+$ Control sample ΔE distributions

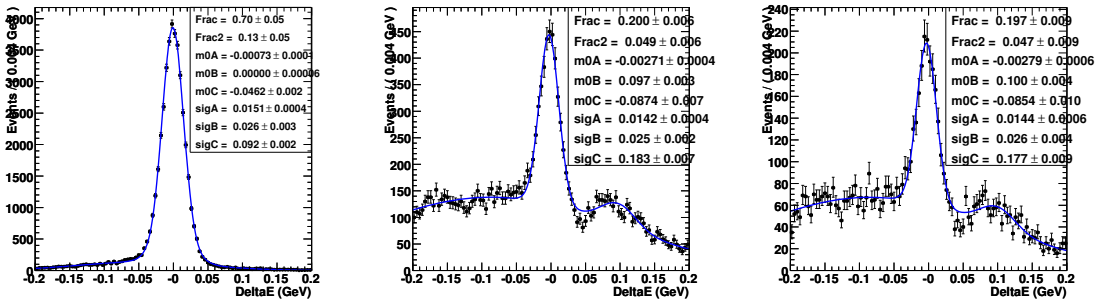


(a) MC (b) Run1-4 Data (c) Run4 Data

Figure 5.40. $B^0 \to D^-\pi^+$, $D^+ \to \bar{K}^0\pi^+$ Control sample ΔE distributions

Table 5.9. ΔE resolution from data and MC control samples.

$D^+ \to K^-\pi^+\pi^+$	Monte Carlo	Run 1-4 Data	Run 4 Data
$\langle\Delta E\rangle$ (MeV)	-0.74 ± 0.06	-2.9 ± 0.2	-2.3 ± 0.3
$\sigma(\Delta E)$ (MeV)	16.64 ± 0.05	16.8 ± 0.2	16.7 ± 0.2
$D^+ \to \bar{K}^0\pi^+$	Monte Carlo	Run 1-4 Data	Run 4 Data
$\langle\Delta E\rangle$ (MeV)	0.1 ± 0.1	-1.4 ± 0.5	-1.5 ± 0.7
$\sigma(\Delta E)$ (MeV)	16.4 ± 0.1	16.5 ± 0.5	17.3 ± 0.7



(a) MC (b) Run1-4 Data (c) Run4 Data

Figure 5.41. Control sample ΔE distributions, $B^+ \to \bar{D}^{0*}\pi^+$, $\bar{D}^{*0} \to \bar{D}^0\gamma$, $\bar{D}^0 \to K^-\pi^+$.

Table 5.10. ΔE resolution from data and MC control samples for $\bar{D}^{*0} \to \bar{D}^0\gamma$, $\bar{D}^0 \to K^-\pi^+$

$\bar{D}^{*0} \to \bar{D}^0\gamma$, $\bar{D}^0 \to K^-\pi^+$	Monte Carlo	Run 1-4 Data	Run 4 Data
$\langle\Delta E\rangle$ (MeV)	-0.5 ± 0.1	-1.4 ± 0.5	-1.4 ± 0.8
$\sigma(\Delta E)$ (MeV)	15.1 ± 0.4	14.2 ± 0.4	14.4 ± 0.6

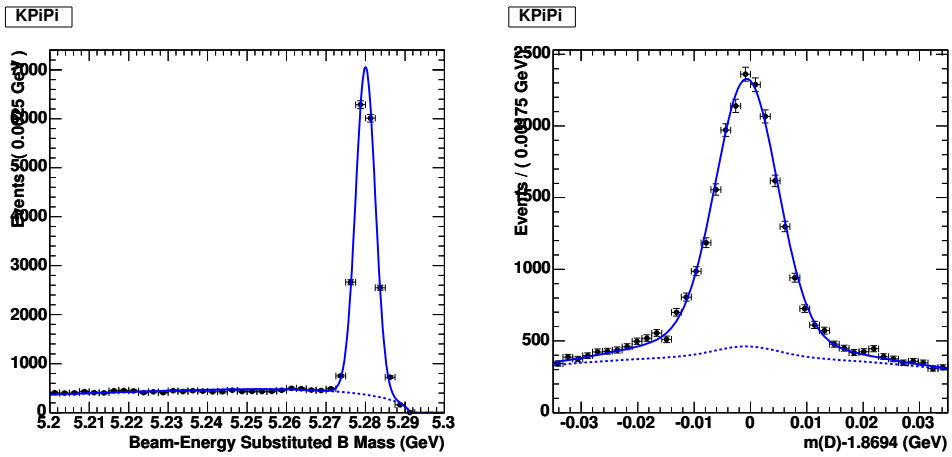


Figure 5.42. $B^0 \to D_s^+\pi^-$ fit to $B^0 \to D^-\pi^+$, $D^+ \to K^-\pi^+\pi^+$ data control sample

Table 5.11. Parameters of the $B^0 \to D_s^+\pi^-$ PDFs, determined from the $B^0 \to D^-\pi^+$, $D^+ \to K^-\pi^+\pi^+$ control sample

Monte Carlo	Data
$\mathcal{S}(m_{\rm ES})$ = Single Gaussian	
$\mu_S^{m_{\rm ES}} = 5.2795 \pm 0.000007$ GeV	$\mu_S^{m_{\rm ES}} = 5.2800 \pm 0.000021$ GeV
$\sigma_S^{m_{\rm ES}} = 2.533 \pm 0.005$ MeV	$\sigma_S^{m_{\rm ES}} = 2.486 \pm 0.017$ MeV
$\mathcal{S}(m_{Ds})$ = Double Gaussian	
$\eta_S = 0.805 \pm 0.004$	$\eta_S = 0.805$ (fixed)
$\mu_{1S}^{m_{Ds}} = -0.02 \pm 0.02$ MeV	$\mu_{1S}^{m_{Ds}} = 0.05 \pm 0.05$ MeV
$\sigma_{1S}^{m_{\rm ES}} = 5.12 \pm 0.02$ MeV	$\sigma_{1S}^{m_{\rm ES}} = 5.44 \pm 0.05$ MeV
$\mu_{2S}^{m_{Ds}} = -3.3 \pm 0.1$ MeV	$\mu_{2S}^{m_{Ds}} = -3.3$ MeV (fixed)
$\sigma_{2S}^{m_{\rm ES}} = 15.6 \pm 0.2$ MeV	$\sigma_{2S}^{m_{\rm ES}} = 15.6$ MeV (fixed)

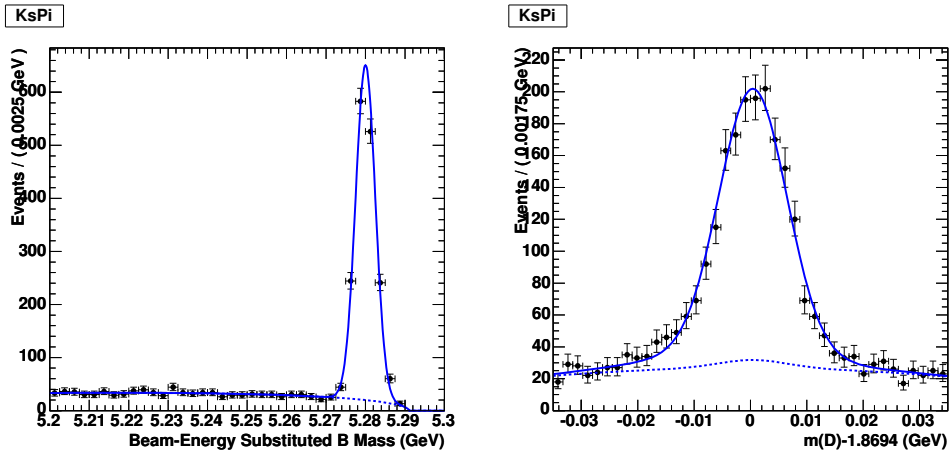


Figure 5.43. $B^0 \to D_s^+\pi^-$ fit to $B^0 \to D^-\pi^+$, $D^+ \to \bar{K}^0\pi^+$ data control sample

Table 5.12. Parameters of the $B^0 \to D_s^+\pi^-$ PDFs, determined from the $B^0 \to D^-\pi^+$, $D^+ \to \bar{K}^0\pi^+$ control sample

Monte Carlo	Data
$\mathcal{S}(m_{\rm ES})$ = Single Gaussian	
$\mu_S^{m_{\rm ES}} = 5.2795 \pm 0.00002$ GeV	$\mu_S^{m_{\rm ES}} = 5.2800 \pm 0.00007$ GeV
$\sigma_S^{m_{\rm ES}} = 2.52 \pm 0.01$ MeV	$\sigma_S^{m_{\rm ES}} = 2.50 \pm 0.05$ MeV
$\mathcal{S}(m_{Ds})$ = Double Gaussian	
$\eta_S = 0.878 \pm 0.009$	$\eta_S = 0.878$ (fixed)
$\mu_{1S}^{m_{Ds}} = 1.00 \pm 0.05$ MeV	$\mu_{1S}^{m_{Ds}} = 0.46 \pm 0.19$ MeV
$\sigma_{1S}^{m_{\rm ES}} = 5.84 \pm 0.05$ MeV	$\sigma_{1S}^{m_{\rm ES}} = 6.00 \pm 0.16$ MeV
$\mu_{2S}^{m_{Ds}} = -3.5 \pm 0.4$ MeV	$\mu_{2S}^{m_{Ds}} = -3.5$ MeV (fixed)
$\sigma_{2S}^{m_{\rm ES}} = 15.1 \pm 0.5$ MeV	$\sigma_{2S}^{m_{\rm ES}} = 15.1$ MeV (fixed)

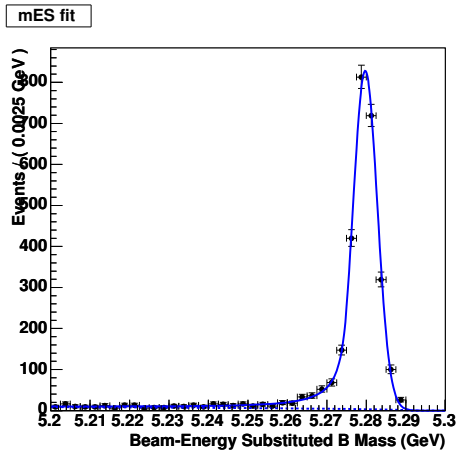


Figure 5.44. $B^0 \to D_s^{*+}\pi^-$ fit to $B^+ \to \bar{D}^{0*}\pi^+$ data control sample

Table 5.13. Parameters of the $B^0 \to D_s^{*+}\pi^-$ PDFs, determined from the $B^+ \to \bar{D}^{0*}\pi^+$, $\bar{D}^{*0} \to \bar{D}^0\gamma$, $\bar{D}^0 \to K^-\pi^+$ control sample

Monte Carlo	Data
$\mathcal{S}(m_{\rm ES})$ = Crystal Ball	
$\mu = 5.2793 \pm 0.00002$ GeV	$\mu = 5.2797 \pm 0.00007$ GeV
$\sigma = 2.45 \pm 0.03$ MeV	$\sigma = 3.06 \pm 0.06$ MeV
$\alpha = 1.91 \pm 0.05$	$\alpha = 1.61 \pm 0.11$
$n = 3.1 \pm 0.2$	$n = 1.5 \pm 0.3$

Table 5.14. Summary of shifts in data wrt MC, from control sample comparison.

	$\langle \Delta E \rangle$ (MeV)	$\mu_{1S}^{m_{Ds}}$ (MeV)
$D^+ \to K^-\pi^+\pi^+$	-2.2	0
$D^+ \to \bar{K}^0\pi^+$	-1.5	-0.54
$\bar{D}^{*0} \to \bar{D}^0\gamma,\ \bar{D}^0 \to K^-\pi^+$	-0.9	–

5.7 Unblinded Results on 208.7 fb^{-1} of Data

The unblinded $m_{\rm ES}$ and m_{Ds} distributions in 208.7 fb^{-1} of data are shown in Figures 5.45 and 5.46 for $B^0 \to D_s^+\pi^-$ and $B^0 \to D_s^- K^+$, respectively. The unblinded $m_{\rm ES}$ distributions in 208.7 fb^{-1} of data are shown in Figure 5.47 for $B^0 \to D_s^{*+}\pi^-$ and $B^0 \to D_s^{*-} K^+$.

As described in the Sections 5.2, we extract the branching fraction for $B^0 \to D_s^+\pi^-$ from the data sample using a 2-dimensional unbinned extended maximum likelihood fit to $m_{\rm ES}$ and m_{Ds}, fitting for all three D_s^+ modes simultaneously. This fit accounts for signal and combinatorial background, as well as peaking background contributions. In the case of $B^0 \to D_s^{*+}\pi^-$, we utilize a 1-dimensional unbinned extended maximum likelihood fit to $m_{\rm ES}$ to extract the yields and branching fraction, accounting for the signal, combinatorial background, and $D^{(*)}\pi/\rho$ reflection background. For $B^0 \to D_s^- K^+$ the selection is identical to the selection for $B^0 \to D_s^+\pi^-$, except that the PID requirement on the daughter of the B^0 is for a kaon rather than a pion. The fitting technique for $B^0 \to D_s^- K^+$ is also similar to that for $B^0 \to D_s^+\pi^-$, as described in Section 5.4. Similarly, the selection and fitting method for $B^0 \to D_s^{*-} K^+$ are nearly identical to those for $B^0 \to D_s^{*+}\pi^-$, except for the PID requirement on the daughter of the B^0.

The fits, as well as the extracted yields and branching fractions, are shown in the following subsections. Since the largest systematic error arises from $\mathcal{B}(D_s^+ \to \phi\pi^+)$, we quote both $\mathcal{B}$ and $\mathcal{B} \cdot \mathcal{B}(D_s^+ \to \phi\pi^+)$. Yields in the full fit region as well as scaled to the signal region are listed in the tables shown in the following sections. The plots shown are projections made with a cut applied to the orthogonal variable. They show the different PDF components overlaid, with the black solid line corresponding to the total PDF, the red solid line the signal component, the blue dashed line the combinatorial background, the dash-dotted pink line the charmless background, and the dotted green line the reflection background.

The results shown include corrections on the efficiency due to discrepancies between data and Monte Carlo efficiencies for the likelihood selection (3% for $D_s^+\pi^-$ and $D_s^- K^+$ and 7%

for $D_s^{*+}\pi^-$ and $D_s^{*-}K^+$) and K_s efficiency (2.5% for $B^0 \to D_s^+\pi^-, D_s^+ \to \bar{K}^0K^+$). Also, for $D_s^{*+}\pi^-$ and $D_s^{*-}K^+$modes, the $m_{\rm ES}$ PDF shape determined from the $D^{*0}\pi$ control sample is used since the width of this PDF is significantly larger than for the PDF determined from MC. Using the broader $m_{\rm ES}$ PDF resulted in larger central values for the branching fraction and also slightly larger systematic errors due to the uncertainty from the PDF shapes. In addition, we iterate the fits for $D_s^+\pi^-$ and $D_s^-K^+$ since these modes are peaking backgrounds with respect to one another. The fits converged for $\mathcal{B}(B^0 \to D_s^+\pi^-)$ and $\mathcal{B}(B^0 \to D_s^-K^+)$ within a couple iterations. Lastly, the systematic errors as described in Section 5.8 have been folded into the calculation of the significance quoted in the tables below.

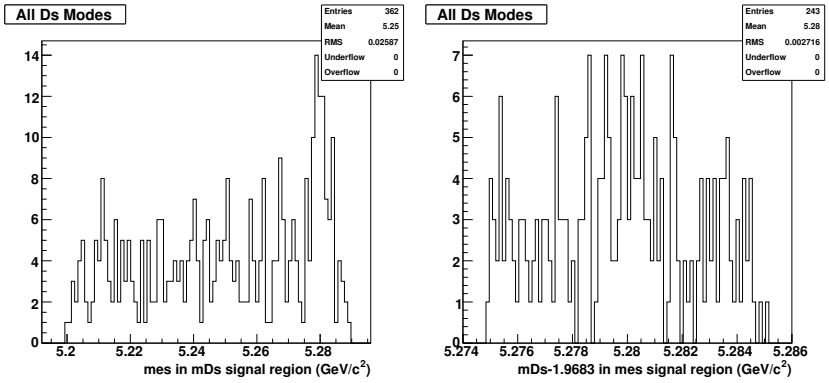


Figure 5.45. Unblinded $m_{\rm ES}$ (in m_{Ds} signal region $|m_{Ds} - 1.9683| < 0.010$ GeV) (left) and $m_{Ds} - 1.9683$ (in $m_{\rm ES}$ signal region $5.2748 < m_{\rm ES} < 5.2852$ GeV) (right) for $B^0 \to D_s^+\pi^-$ with all D_s modes combined.

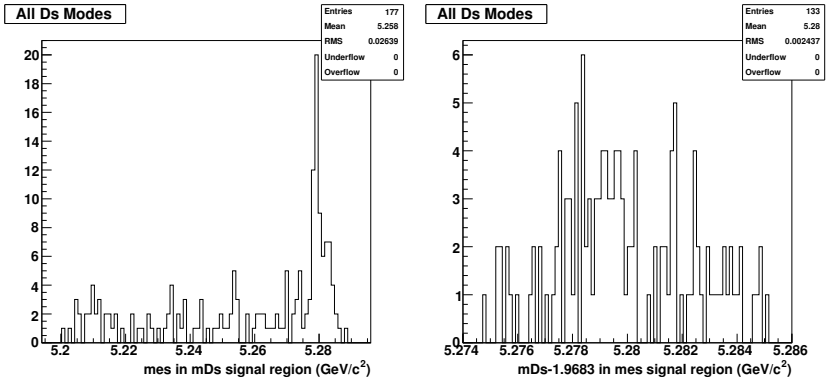


Figure 5.46. Unblinded $m_{\rm ES}$ (in m_{Ds} signal region $|m_{Ds} - 1.9683| < 0.010$ GeV) (left) and $m_{Ds} - 1.9683$ (in $m_{\rm ES}$ signal region $5.2748 < m_{\rm ES} < 5.2852$ GeV) (right) for $B^0 \to D_s^-K^+$ with all D_s modes combined.

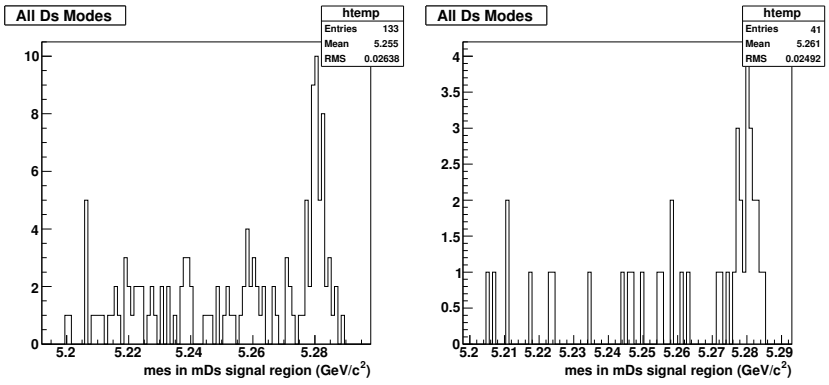


Figure 5.47. Unblinded $m_{\rm ES}$ distribution with all D_s modes combined, for $B^0 \to D_s^{*+}\pi^-$ (left) and $B^0 \to D_s^{*-}K^+$ (right)

5.7.1 Final Results for $B^0 \to D_s^+\pi^-$

Table 5.15. Branching Fraction for $B^0 \to D_s^+\pi^-$ from 208.7 fb^{-1} of data.

	(10^{-5})
$\mathcal{B}(B^0 \to D_s^+\pi^-)$	1.3 ± 0.3 (5.1σ)
$\mathcal{B}(B^0 \to D_s^+\pi^-) \cdot \mathcal{B}(D_s^+ \to \phi\pi^+)$	$(0.6 \pm 0.2) \cdot 10^{-1}$

Table 5.16. Yields for $B^0 \to D_s^+\pi^-$ from 2-dim $m_{\rm ES}$-m_{Ds} fit in 208.7 fb^{-1} of data. Yields are for the fit region, $5.2 < m_{\rm ES} < 5.3$ GeV and $|m_{Ds} - 1.9683| < 0.050$ GeV.

	N_{raw}	N_{signal}	$N_{combinatorial}$	$N_{charmless}$	N_{refl}	N_{DsK}	N_{total}^{fit}
$\phi\pi^+$	405	20.8 ± 4.9	364.1 ± 20.2	12.9 ± 8.4	7.1	0.5	405.5 ± 22.4
$\bar{K}^{0*}K^+$	677	16.1 ± 3.8	603.5 ± 26.0	39.6 ± 11.7	17.5	0.5	677.2 ± 28.8
$\bar{K}^0K^+$	223	11.0 ± 2.6	197.2 ± 14.7	7.3 ± 6.1	8.2	0.3	224.0 ± 16.2

Table 5.17. Yields for $B^0 \to D_s^+\pi^-$ from 2-dim $m_{\rm ES}$-m_{Ds} fit in 208.7 fb^{-1} of data, scaled to the signal box region, $5.2748 < m_{\rm ES} < 5.2852$ GeV and $|m_{Ds} - 1.9683| < 0.010$ GeV.

	N_{raw}	N_{signal}	$N_{combinatorial}$	$N_{charmless}$	N_{refl}	N_{DsK}	N_{total}^{fit}
$\phi\pi^+$	29	15.3 ± 3.6	9.6 ± 0.5	2.6 ± 1.7	0.9	0.4	28.8 ± 4.0
$\bar{K}^{0*}K^+$	40	11.0 ± 2.6	16.0 ± 0.7	7.9 ± 2.3	3.6	0.3	38.8 ± 3.5
$\bar{K}^0K^+$	14	8.2 ± 1.9	5.2 ± 0.4	1.5 ± 1.2	1.6	0.2	16.6 ± 2.3

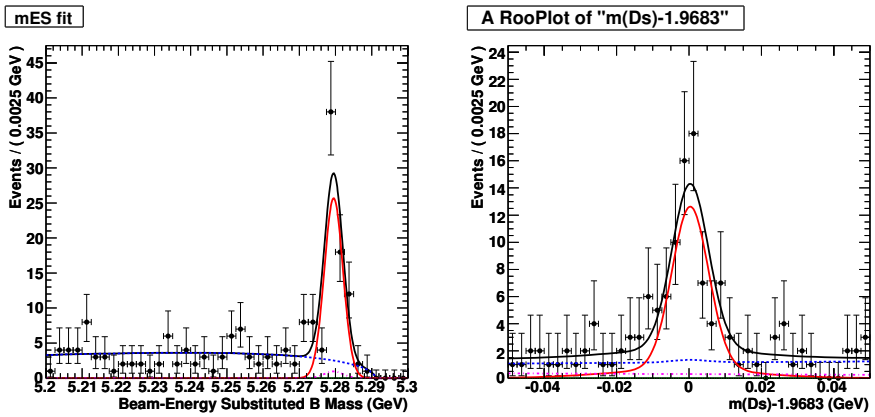


Figure 5.48. Projections of the two-dim fit 208.7 fb^{-1} of data for $B^0 \to D_s^+\pi^-$, all D_s^+ modes combined. $m_{\rm ES}$ projection (with $|m_{Ds} - 1.9683| < 0.010$ GeV) (left) and m_{Ds} projection (with $5.2748 < m_{\rm ES} < 5.2852$ GeV) (right). The black solid line corresponds to the total PDF, red solid to signal, blue dashed to combinatorial background, green dotted to reflection background, magenta dash-dotted to charmless background.

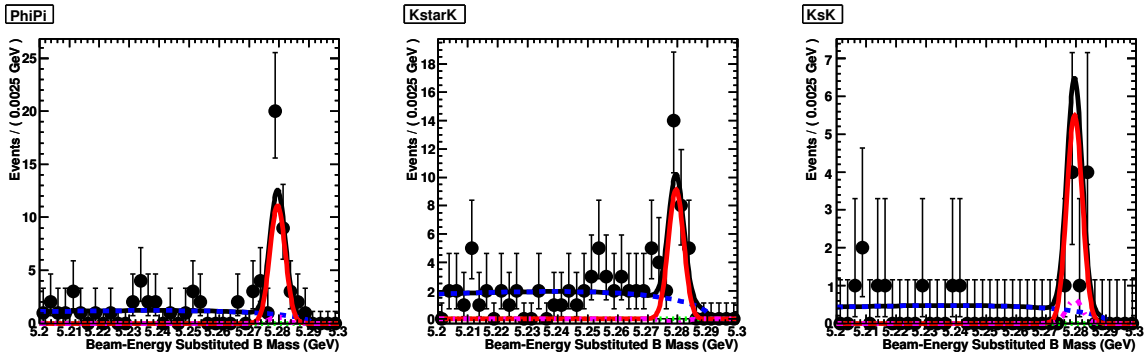


Figure 5.49. $m_{\rm ES}$ projections (with $|m_{Ds} - 1.9683| < 0.010$ GeV) of the two-dim fit to 208.7 fb^{-1} of data for $B^0 \to D_s^+\pi^-$. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right).

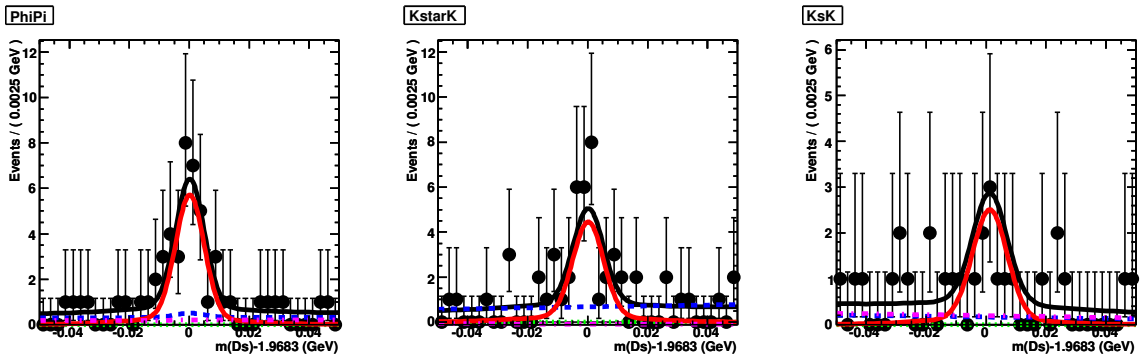


Figure 5.50. m_{Ds} projections (with $5.2748 < m_{\rm ES} < 5.2852$ GeV) of the two-dim fit to 208.7 fb^{-1} of data for $B^0 \to D_s^+\pi^-$. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0K^+$ (right).

5.7.2 Final Results for $B^0 \to D_s^{*+}\pi^-$

Table 5.18. Branching Fraction for $B^0 \to D_s^{*+}\pi^-$ from 208.7 fb^{-1} of data.

	(10^{-5})
$\mathcal{B}(B^0 \to D_s^{*+}\pi^-)$	2.8 ± 0.6 (5.9σ)
$\mathcal{B}(B^0 \to D_s^{*+}\pi^-) \cdot \mathcal{B}(D_s^+ \to \phi\pi^+)$	$(1.3 \pm 0.3) \cdot 10^{-1}$

Table 5.19. Yields for $B^0 \to D_s^{*+}\pi^-$ in 208.7 fb^{-1} of data. Yields are for the fit region, $5.2 < m_{\rm ES} < 5.3$ GeV and $|m_{Ds} - 1.9683| < 0.010$ GeV.

	N_{raw}	N_{signal}	$N_{combinatorial}$	N_{refl}	N_{total}^{fit}
$\phi\pi^+$	46	18.3 ± 3.7	28.5 ± 5.9	0.0	46.9 ± 6.9
$\bar{K}^{0*}K^+$	67	14.1 ± 2.9	48.2 ± 7.8	0.9	63.2 ± 8.3
$\bar{K}^0K^+$	19	9.5 ± 1.9	12.0 ± 3.8	0.6	22.1 ± 4.2

Table 5.20. Yields for $B^0 \to D_s^{*+}\pi^-$ in 208.7 fb^{-1} of data. Yields are for the signal box region.

	N_{raw}	N_{signal}	$N_{combinatorial}$	N_{refl}	N_{total}^{fit}
$\phi\pi^+$	18	15.0 ± 3.0	2.7 ± 0.6	0.0	17.7 ± 3.1
$\bar{K}^{0*}K^+$	22	11.6 ± 2.3	4.6 ± 0.7	0.5	16.6 ± 2.5
$\bar{K}^0K^+$	8	7.7 ± 1.6	1.1 ± 0.4	0.3	9.2 ± 1.6

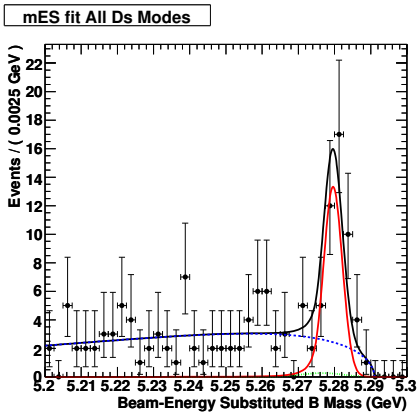


Figure 5.51. $m_{\rm ES}$ fit to $208.7\,{\rm fb}^{-1}$ of data for $B^0 \to D_s^{*+}\pi^-$, all D_s^+ modes combined. The black solid line corresponds to the total PDF, red solid to signal, blue dashed to combinatorial background, green dotted to reflection background.

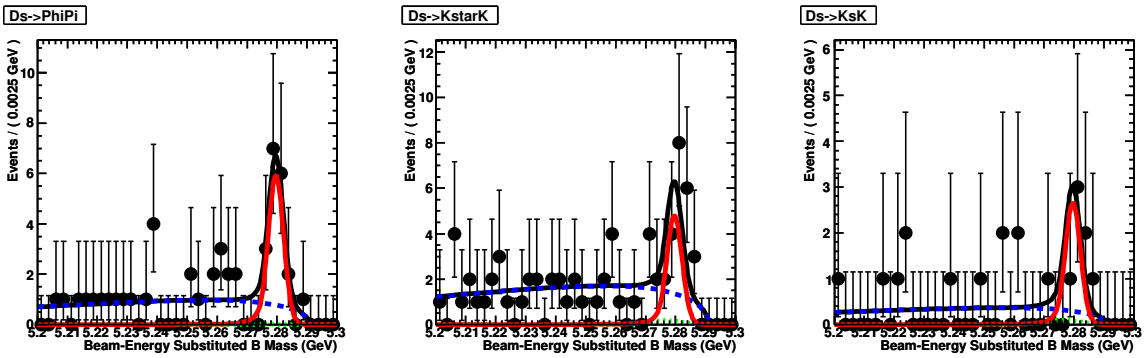


Figure 5.52. $m_{\rm ES}$ fit to $208.7\,{\rm fb}^{-1}$ of data for $B^0 \to D_s^{*+}\pi^-$. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*}K^+$ (middle), $D_s^+ \to \bar{K}^0 K^+$ (right).

5.7.3 Final Results for $B^0 \to D_s^- K^+$

Table 5.21. Branching Fraction for $B^0 \to D_s^- K^+$ from 208.7 fb^{-1} of data.

	(10^{-5})
$\mathcal{B}(B^0 \to D_s^- K^+)$	2.5 ± 0.4 (9.3σ)
$\mathcal{B}(B^0 \to D_s^- K^+) \cdot \mathcal{B}(D_s^+ \to \phi\pi^+)$	$(1.2 \pm 0.2) \cdot 10^{-1}$

Table 5.22. Yields for $B^0 \to D_s^- K^+$ from 2-dim $m_{\rm ES}$-m_{Ds} fit in 208.7 fb^{-1} of data. Yields are for the fit region, $5.2 < m_{\rm ES} < 5.3$ GeV and $|m_{Ds} - 1.9683| < 0.050$ GeV.

	N_{raw}	N_{signal}	$N_{combinatorial}$	$N_{charmless}$	N_{refl}	$N_{Ds\pi}$	N_{total}^{fit}
$\phi\pi^+$	197	32.0 ± 4.5	151.4 ± 13.0	7.9 ± 5.8	-	0.1	191.4 ± 14.9
$\bar{K}^{0*}K^+$	331	27.2 ± 3.9	306.4 ± 18.3	-3.7 ± 6.1	-	0.1	330.0 ± 19.7
$\bar{K}^0K^+$	101	18.2 ± 2.6	81.6 ± 9.5	8.8 ± 4.7	-	0.1	108.6 ± 11.0

Table 5.23. Yields for $B^0 \to D_s^- K^+$ from 2-dim $m_{\rm ES}$-m_{Ds} fit, in 208.7 fb^{-1} of data. Yields have been scaled to the signal box region, $5.2748 < m_{\rm ES} < 5.2852$ GeV and $|m_{Ds} - 1.9683| < 0.010$ GeV.

	N_{raw}	N_{signal}	$N_{combinatorial}$	$N_{charmless}$	N_{refl}	$N_{Ds\pi}$	N_{total}^{fit}
$\phi\pi^+$	34	23.3 ± 3.3	3.9 ± 0.3	1.6 ± 1.2	-	0.1	28.9 ± 3.5
$\bar{K}^{0*}K^+$	29	18.7 ± 2.7	7.8 ± 0.5	-0.7 ± 1.2	-	0.1	25.8 ± 3.0
$\bar{K}^0K^+$	10	13.1 ± 1.9	2.1 ± 0.2	1.8 ± 0.9	-	0.0	16.9 ± 2.1

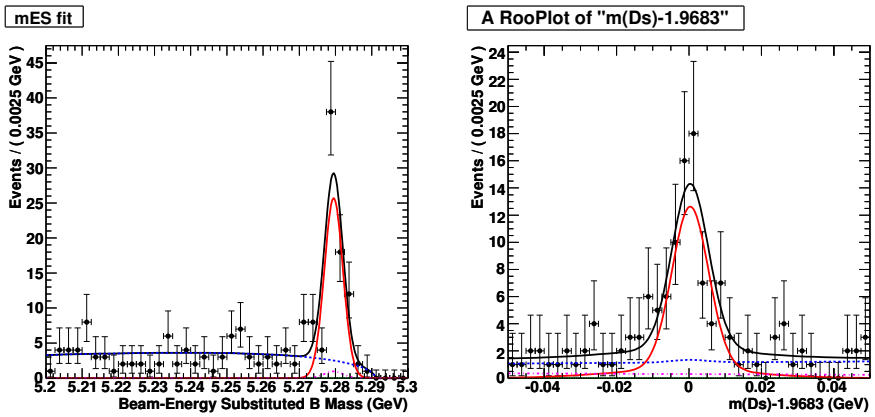


Figure 5.53. Projections of the two-dim fit 208.7 fb^{-1} of data for $B^0 \to D_s^- K^+$, all D_s^+ modes combined. $m_{\rm ES}$ projection (with $|m_{Ds} - 1.9683| < 0.010$ GeV) (left) and m_{Ds} projection (with $5.2748 < m_{\rm ES} < 5.2852$ GeV) (right). The black solid line corresponds to the total PDF, red solid to signal, blue dashed to combinatorial background, green dotted to reflection background, magenta dash-dotted to charmless background.

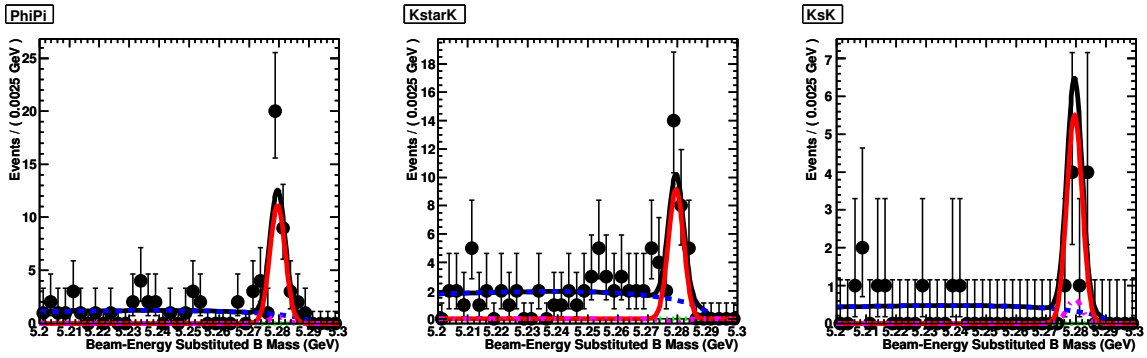


Figure 5.54. $m_{\rm ES}$ projections (with $|m_{Ds} - 1.9683| < 0.010$ GeV) of the two-dim fit to 208.7 fb^{-1} of data for $B^0 \to D_s^- K^+$. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*} K^+$ (middle), $D_s^+ \to \bar{K}^0 K^+$ (right).

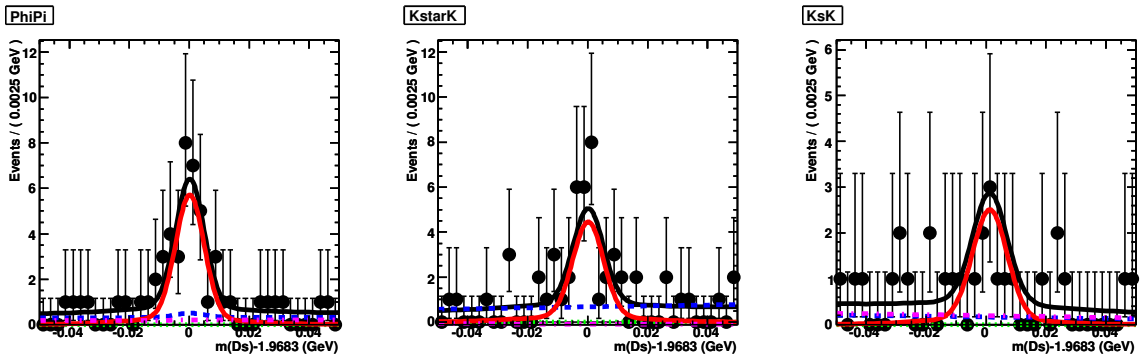


Figure 5.55. m_{Ds} projections (with $5.2748 < m_{\rm ES} < 5.2852$ GeV) of the two-dim fit to 208.7 fb^{-1} of data for $B^0 \to D_s^- K^+$. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*} K^+$ (middle), $D_s^+ \to \bar{K}^0 K^+$ (right).

5.7.4 Final Results for $B^0 \to D_s^{*-} K^+$

Table 5.24. Branching Fraction for $B^0 \to D_s^{*-} K^+$ from 208.7 fb^{-1} of data.

	(10^{-5})
$\mathcal{B}(B^0 \to D_s^{*-} K^+)$	2.0 ± 0.5 (4.7σ)
$\mathcal{B}(B^0 \to D_s^{*-} K^+) \cdot \mathcal{B}(D_s^+ \to \phi\pi^+)$	$(1.0 \pm 0.3) \cdot 10^{-1}$

Table 5.25. Yields for $B^0 \to D_s^{*-} K^+$ from 1-dim $m_{\rm ES}$ fit in 208.7 fb^{-1} of data. Yields are for the fit region, $5.2 < m_{\rm ES} < 5.3$ GeV and $|m_{Ds} - 1.9683| < 0.010$ GeV.

	N_{raw}	N_{signal}	$N_{combinatorial}$	N_{refl}	N_{total}^{fit}
$\phi\pi^+$	15	9.1 ± 2.3	7.5 ± 3.0	-	16.6 ± 3.8
$\bar{K}^{0*} K^+$	16	7.7 ± 1.9	7.4 ± 3.0	-	15.1 ± 3.6
$\bar{K}^0 K^+$	10	4.8 ± 1.2	4.5 ± 2.6	-	9.3 ± 2.9

Table 5.26. Yields for $B^0 \to D_s^{*-} K^+$ from 1-dim $m_{\rm ES}$-m_{Ds} fit in 208.7 fb^{-1} of data. Yields have been scaled to the signal box region, $5.2748 < m_{\rm ES} < 5.2852$ GeV and $|m_{Ds} - 1.9683| < 0.010$ GeV.

	N_{raw}	N_{signal}	$N_{combinatorial}$	N_{refl}	N_{total}^{fit}
$\phi\pi^+$	7	7.5 ± 1.9	0.7 ± 0.3	-	8.1 ± 1.9
$\bar{K}^{0*} K^+$	8	6.3 ± 1.6	0.7 ± 0.3	-	6.9 ± 1.6
$\bar{K}^0 K^+$	4	3.9 ± 1.0	0.4 ± 0.2	-	4.3 ± 1.0

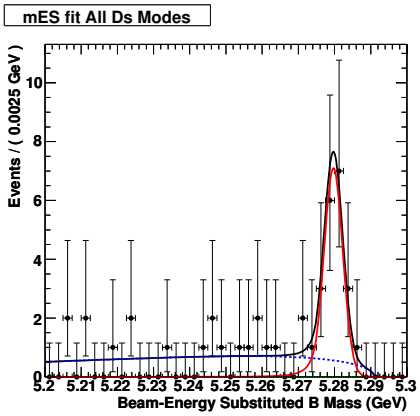


Figure 5.56. $m_{\rm ES}$ fit to $208.7\,{\rm fb}^{-1}$ of data for $B^0 \to D_s^{*-} K^+$, all D_s^+ modes combined. The black solid line corresponds to the total PDF, red solid to signal, blue dashed to combinatorial background, green dotted to reflection background.

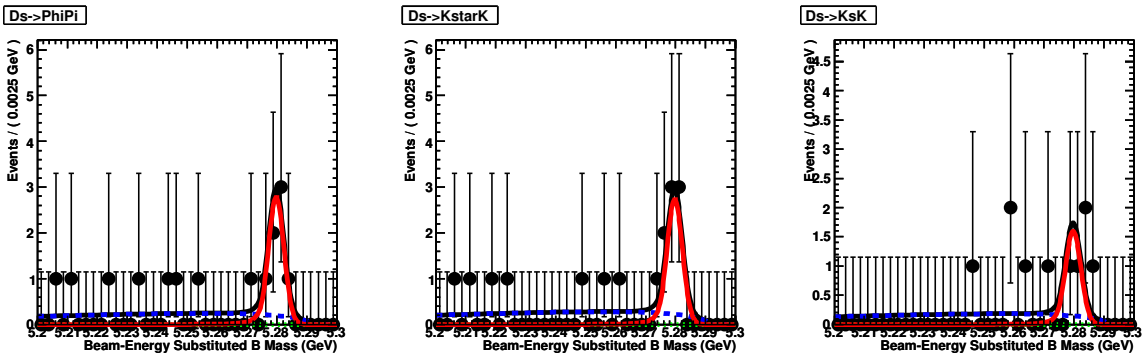


Figure 5.57. $m_{\rm ES}$ fit to $208.7\,{\rm fb}^{-1}$ of data for $B^0 \to D_s^{*-} K^+$. $D_s^+ \to \phi\pi^+$ (left), $D_s^+ \to \bar{K}^{0*} K^+$ (middle), $D_s^+ \to \bar{K}^0 K^+$ (right).

5.7.5 Fit Bias Studies from Monte Carlo

"Toy" Monte Carlo tests are used to check the validity of the fitting method for $B^0 \to D_s^+\pi^-$ and $B^0 \to D_s^{*+}\pi^-$. Toy Monte Carlo samples are simulated events that do not simulate the full detector response and therefore are much simpler to generate. To run these tests, a large number of toy MC experiments are generated. For each toy experiment, signal, combinatorial, and peaking background distributions are generated using the PDFs from the data fit. The number of events for each distribution is taken from the data fit, and smeared according to a Poisson distribution.

Figure 5.58 shows the distribution of the difference between the fitted values of the branching fraction $\mathcal{B}_{fit}$ and the branching fraction used to generate the $B^0 \to D_s^+\pi^-$ events $\mathcal{B}_{generated} = 2.1 \cdot 10^{-5}$, from 400 toy experiments. The RMS of the distribution, $3 \cdot 10^{-6}$, is the average statistical error expected from the fit, and this agrees well with the value from the actual fit to data. Figure 5.62 shows a similar distribution for $B^0 \to D_s^{*+}\pi^-$, with $\mathcal{B}_{generated} = 2.1 \cdot 10^{-5}$. The RMS of the distribution for $B^0 \to D_s^{*+}\pi^-$ is $4 \cdot 10^{-6}$, and this, too, agrees well with the value from the actual fit to data. The distribution here shows that there is a slight bias of $-0.56 \cdot 10^{-6}$ for $B^0 \to D_s^{*+}\pi^-$. This may be due to the low statistics of the toy MC sample.

Figure 5.59 shows the fitted value of the branching function as a fraction of the generated value from 5k toy experiments for $B^0 \to D_s^+\pi^-$. The signal branching fraction is varied randomly from 0 to $1 \cdot 10^{-4}$. The linear fit shows an intercept below $1 \cdot 10^{-7}$ and a slope of 0.997 ± 0.002 for $B^0 \to D_s^+\pi^-$, consistent with the fit not biasing the branching fraction result. Figure 5.63 shows the distribution for $B^0 \to D_s^{*+}\pi^-$, the linear fit for which shows an intercept below $-2.9 \cdot 10^{-7}$ and a slope of 0.991 ± 0.003, also consistent with the fit not biasing the branching fraction result.

Figures 5.60 and 5.64 show the distributions of pulls, $(\mathcal{B}_{fit} - \mathcal{B}_{generated})/\text{error}_{fit}$ for $B^0 \to D_s^+\pi^-$ and $B^0 \to D_s^{*+}\pi^-$, respectively. The means of the distributions are zero with an RMS of one, again showing no bias in the fits. Figures 5.61 and 5.65 show the distribution

of $-log(Likelihood)$, for $B^0 \to D_s^+\pi^-$ and $B^0 \to D_s^{*+}\pi^-$, respectively, relative to the value obtained in the data fit.

For $B^0 \to D_s^+\pi^-$, similar toy Monte Carlo tests are run with the charmless background fixed to zero, to see whether an overestimation of the charmless background in the fits may be taking events away from the signal, thereby artificially decreasing the branching fraction. From these tests, no bias in the fitted signal branching fraction is seen. On average, the fit returns the same value as the branching ratio used to generate the events.

In conclusion, the fitting method works as expected, and does not bias the extracted value of the branching fraction in the case of $B^0 \to D_s^+\pi^-$. For $B^0 \to D_s^{*+}\pi^-$, there is a slight bias of $-0.56 \cdot 10^{-6}$ according to these toy MC tests.

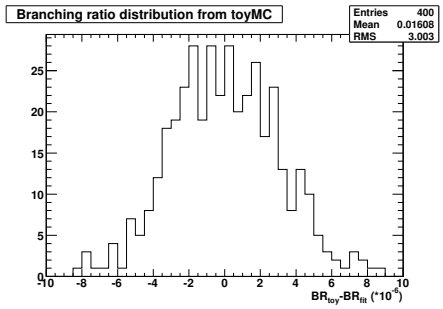


Figure 5.58. Difference between the fitted values of the branching fraction and branching fraction used for generating the events, from 400 toy MC experiments. The branching fraction used for generating is $\mathcal{B}(B^0 \to D_s^+\pi^-) = 2.1 \cdot 10^{-5}$.

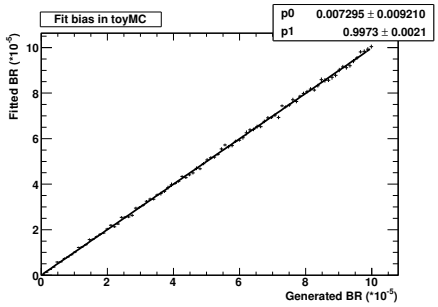


Figure 5.59. Fitted value of the $B^0 \to D_s^+\pi^-$ branching fraction as a function of the generated value, from 5k toy MC experiments.

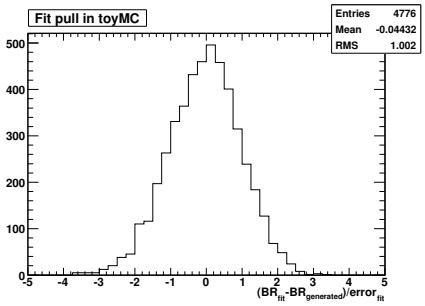


Figure 5.60. Distribution of pulls for the $\mathcal{B}(B^0 \to D_s^+\pi^-)$ fit, from 5k toy MC experiments.

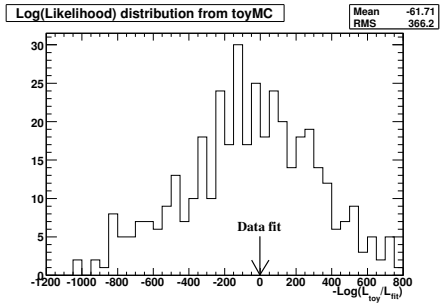


Figure 5.61. Distribution of $-log(Likelihood)$ from the $\mathcal{B}(B^0 \to D_s^{*+}\pi^-)$ fit, from 5k toy MC experiments (relative to the value obtained in the data fit).

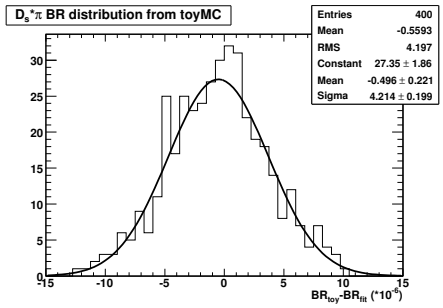


Figure 5.62. Difference between the fitted values of the branching fraction and branching fraction used for generating the events, from 400 toy MC experiments. The branching fraction used for generating is $\mathcal{B}(B^0 \to D_s^{*+}\pi^-) = 2.1 \cdot 10^{-5}$.

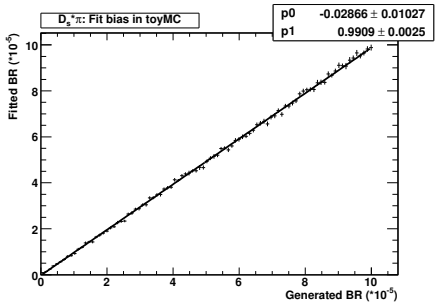


Figure 5.63. Fitted value of the $B^0 \to D_s^{*+}\pi^-$ branching fraction as a function of the generated value, from 5k toy MC experiments.

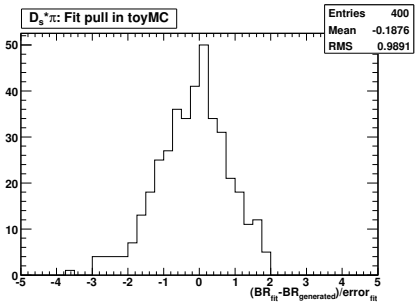


Figure 5.64. Distribution of pulls for the $\mathcal{B}(B^0 \to D_s^{*+}\pi^-)$ fit, from 5k toy MC experiments.

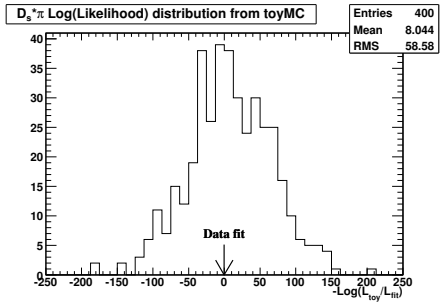


Figure 5.65. Distribution of $-log(Likelihood)$ from the $\mathcal{B}(B^0 \to D_s^{*+}\pi^-)$ fit, from 5k toy MC experiments (relative to the value obtained in the data fit).

5.8 Systematic Errors

The systematic uncertainties considered are listed below. They have been divided into errors that are correlated and uncorrelated between the three D_s^+ modes. The uncorrelated errors are treated by refitting for the branching ratio after smearing the parameters (for instance, the PDF shape parameters) by their errors, assuming they are Gaussian distributed. We repeat this smearing and refitting 100 times and take the resulting width of the fitted branching fraction as the systematic error. This method automatically weighs the errors between the three D_s^+ modes. The correlated and uncorrelated errors are then added in quadrature. The final errors are summarized as relative errors in Table 5.27 and as absolute errors on the branching ratios in Table 5.28.

Correlated Errors:

- B Counting. 1.1% uncertainty on the number of $B^0\overline{B}^0$ events produced.
- Tracking efficiency. 1.3% per track for tracks, as recommended by the *BABAR* tracking group.Since we have 4 such tracks for all D_s^+ modes, the total error is 5.2%.
- PID efficiency. Since the efficiencies are measured using Monte Carlo, the discrepancy between particle identification efficiency in data and in Monte Carlo must be accounted for. An overall efficiency correction scale factor is not applied because these are all found to be 1.0 within 1σ. However, a 1% error on the PID efficiency corrections is included as a systematic error, based on recommendations from the *BABAR* PID group.
- γ efficiency. 1.8% uncertainty for single γ efficiency, for D_s^{*+} modes, based on a study of based π^0 efficiency in τ decays [32].
- PDF shapes - likelihood variable. Uncertainty due to the differences between Monte Carlo and data for PDF shapes used in constructing the likelihood selection variable. We compute this in Section 4.6.3 to be 3% for $D_s^+\pi^-$ and 7% for $D_s^{*+}\pi^-$, using a

control sample of $B \to D^{(*)}\pi$ events. Similarly, we include 3% error for $D_s^- K^+$ and 7% for $D_s^{*-} K^+$.

- $\mathcal{B}(D_s^+ \to \phi\pi^+)$. 13% uncertainty on $\mathcal{B}(D_s^+ \to \phi\pi^+)$ [33]. Since the $\mathcal{B}(D_s^+ \to \bar{K}^{0*} K^+)$ and $\mathcal{B}(D_s^+ \to \bar{K}^0 K^+)$ are measured relative to $\mathcal{B}(D_s^+ \to \phi\pi^+)$, this is a correlated error for all three modes.

- $\mathcal{B}(D_s^{*+} \to D_s^+\gamma)$. 2.7% uncertainty on $\mathcal{B}(D_s^{*+} \to D_s^+\gamma)$ for D_s^{*+} modes.

- ΔE cut. Uncertainty on ΔE resolution, as determined from the $D^{(*)}\pi$ control samples. This systematic is small ($< 1\%$), but included for completeness. The uncertainty due to the uncertainty in the ΔE shift is not included because it is even smaller ($< 0.01\%$) and deemed negligible.

- Mass D_s cut. Uncertainty on D_s mass mean and resolution, as determined from the $D\pi$ control sample. This systematic is small $< 1\%$. For the D_s^* modes, the numbers are taken from the $D^+\pi$ control sample rather than the $D^{*0}\pi$ control sample, since the kinematics of the D^+ decay better mimic the D_s decays.

Uncorrelated Errors:

- MC Statistics. Uncertainties in signal and peaking background efficiencies due to limited MC statistics.

- K_s efficiency. 2.5% uncertainty for the K_s efficiency, following the recommendations of the *BABAR* Tracking Efficiency Task Force.

- PDF shapes - yield extraction. Uncertainty due to the errors on signal and peaking background PDF shape parameters.

- Peaking background. Uncertainty due to peaking backgrounds not included in our final PDF. The contribution to the uncertainty from each of these background modes is estimated by including the numbers quoted in Tables 4.12-4.14 into the peaking

background PDF and refitting for the yield. We quote the systematic error as the difference in the branching fraction from this fit and the original fit.

- $\mathcal{B}$ of sub-decays. Although the $\mathcal{B}(D_s^+ \to \phi\pi^+)$ is considered as a correlated error between the three D_s^+modes, we also consider the error due to the uncertainties on the of $\mathcal{B}(D_s^+ \to \bar{K}^{0*}K^+)$ and $\mathcal{B}(D_s^+ \to \bar{K}^0K^+)$ relative to $\mathcal{B}(D_s^+ \to \phi\pi^+)$. The relative errors, as quoted from the PDG, are $\Gamma(\bar{K}^{0*}K)/\Gamma(\phi\pi) = 17\%$ and $\Gamma(K_sK)/\Gamma(\phi\pi) = 16\%$.

- $\mathcal{B}$ of reflection background. Uncertainty on the branching ratio of the reflection backgrounds. These are on the order of 10%.

- $\mathcal{B}(D_s^-K^+)$ and $\mathcal{B}(D_s^+\pi^-)$. Uncertainty on the branching ratio of the $D_s^-K^+$ peaking background for $D_s^+\pi^-$, and vice versa for $D_s^+\pi^-$.

- $\bar{K}^{*0}$ mass lineshape. Uncertainty due to the difference in $\bar{K}^{*0}$ mass shapes in MC and data. We utilize a study of D_sD control samples to estimate this uncertainty [34]. This study exhibits a 4 MeV shift in the K^{*0} lineshape in MC with respect to on-resonance data. We estimate the uncertainty on our measurement by recomputing the signal efficiency before and after shifting the $\bar{K}^{*0}$ PDF by ± 4 MeV and taking the relative difference in efficiencies as the relative error on the branching fraction. We also performed a similar check on the ϕ lineshape, but the resulting systematic error is small and deemed negligible.

Table 5.27. Systematic uncertainties as relative errors. (%)

	$B^0 \to D_s^+\pi^-$	$B^0 \to D_s^{*+}\pi^-$	$B^0 \to D_s^-K^+$	$B^0 \to D_s^{*-}K^+$
B counting	1.1			
Tracking efficiency	5.2			
PID efficiency	1.0			
PDF shape, likelihood	3.0	7.0	3.0	7.0
$\mathcal{B}(D_s^+ \to \phi\pi^+)$	13.4			
$\mathcal{B}(D_s^{*+} \to D_s^+\gamma)$	-	2.7	-	2.7
γ efficiency	-	1.8	-	1.8
ΔE resolution	0.5	0.4	0.5	0.4
m_{D_s} resolution	0.9	0.9	0.9	0.9
m_{D_s} shift	-	0.1	-	0.1
Total Correlated	14.8	16.4	14.8	16.4
MC statistics	0.8	1.5	0.6	0.9
Ks efficiency	0.5	0.4	0.4	0.2
PDF shape, yield fit	1.6	4.0	0.5	5.2
Peaking background	0.1	0.5	0.8	1.1
$\mathcal{B}$(sub-decays)	5.1	5.6	6.4	7.3
$\mathcal{B}$(refl. background)	0.3	0.1	-	-
$\mathcal{B}(D_s^-K^+)$	0.4	-	0.1	-
K^{*0} lineshape	0.1	0.1	0.1	0.2
Total Uncorrelated	5.4	6.9	6.5	9.1
TOTAL	15.8	17.4	16.1	18.3

Table 5.28. Systematic uncertainties as absolute errors on the branching ratio. (10^{-5})

	$B^0 \to D_s^+ \pi^-$	$B^0 \to D_s^{*+} \pi^-$	$B^0 \to D_s^- K^+$	$B^0 \to D_s^{*-} K^+$
$\mathcal{B}$ Result	1.3	2.8	2.5	2.0
B counting	0.01	0.03	0.03	0.02
Tracking efficiency	0.07	0.14	0.13	0.10
PID efficiency	0.01	0.03	0.03	0.02
PDF shape, likelihood	0.04	0.19	0.08	0.14
$\mathcal{B}(D_s^+ \to \phi\pi^+)$	0.18	0.37	0.34	0.27
$\mathcal{B}(D_s^{*+} \to D_s^+\gamma)$	-	0.07	-	0.05
γ efficiency	-	0.05	-	0.04
ΔE resolution	0.01	0.01	0.01	0.01
m_{D_s} resolution	0.01	0.02	0.02	0.02
m_{D_s} shift	-	0.00	-	0.00
Total Correlated	0.19	0.45	0.37	0.33
MC statistics	0.01	0.04	0.01	0.02
Ks efficiency	0.01	0.01	0.01	0.00
PDF shape, yield fit	0.02	0.10	0.01	0.11
Peaking background	0.00	0.01	0.02	0.02
$\mathcal{B}$(sub-decays)	0.07	0.15	0.16	0.15
$\mathcal{B}$(refl. background)	0.00	0.00	-	-
$\mathcal{B}(D_s^- K^+)$	0.01	-	0.00	-
K^{*0} lineshape	0.00	0.00	0.00	0.00
Total Uncorrelated	0.07	0.19	0.16	0.18
TOTAL	0.21	0.48	0.41	0.37

5.9 Calculation of $r_{D^{(*)}\pi}$

From the branching fractions of $B^0 \to D_s^+\pi^-$ and $B^0 \to D_s^{*+}\pi^-$, we can compute the value of $r_{D^{(*)}\pi}$ using the equation first described in Section 1.5:

$$\mathcal{B}(B^0 \to D_s^{(*)+}\pi^-) = \frac{\mathcal{B}(B^0 \to D^{(*)-}\pi^+)}{\tan^2\theta_{\rm C}}\ r^2_{D^{(*)}\pi}(\frac{f_{D_s^{(*)}}}{f_{D^{(*)}}})^2, \tag{5.6}$$

where $\theta_{\rm C}$ is the Cabibbo angle and

$$r_{D^{(*)}\pi} = \frac{A(B^0 \to D^{(*)+}\pi^-)}{A(B^0 \to D^{(*)-}\pi^+)}. \tag{5.7}$$

$\frac{f_{D_s^{(*)}}}{f_{D^{(*)}}}$ is the ratio of the decay constants that takes into account the core of the $SU(3)$ symmetry breaking. In computing $r_{D^{(*)}\pi}$, we use the value:

$$\frac{f_{D_s^{(*)}}}{f_{D^{(*)}}} = 1.22 \pm 0.04\ [13]. \tag{5.8}$$

Using Equations 5.6 and 5.8 and the branching fractions of the CKM-favored decays $\mathcal{B}(B^0 \to D^-\pi^+) = (2.76 \pm 0.25) \cdot 10^{-3}$ and $\mathcal{B}(B^0 \to D^{*-}\pi^+) = (2.76 \pm 0.21) \cdot 10^{-3}$, in addition to $\tan\theta_{\rm C} = 0.2286 \pm 0.0024$, we arrive at [12]:

$$r_{D\pi} = 0.013 \pm 0.002\ (stat) \pm 0.001\ (syst) \tag{5.9}$$

and

$$r_{D^*\pi} = 0.017 \pm 0.002\ (stat) \pm 0.002\ (syst). \tag{5.10}$$

These values of $r_{D^{(*)}\pi}$ are slightly smaller than the original estimate of ~ 0.02, computed from the CKM matrix elements. The fact that $r_{D^{(*)}\pi}$ is small indicates limited sensitivity to *CP* asymmetries in $B^0 \to D^{(*)\mp}\pi^\pm$ decays.

5.10 Summary of Results

The table below summarizes the measured yields and branching fractions. Also shown are the significances, as well as the corresponding probabilities of the background fluctuating upward to produce the signal, with the systematic errors folded into the calculation.

Table 5.29. The number of reconstructed candidates ($N_{\rm raw}$), the signal yield ($N_{\rm sig}$), combinatorial background ($N_{\rm comb}$), and the sum of charmless, reflection, and cross-feed contributions ($N_{\rm peak}$), extracted from the likelihood fit. The contributions correspond to the signal region $5.2 < m_{\rm ES} < 5.3$ GeV/c^2, $|\Delta E| < 36$ MeV, and $|m_{Ds} - m_{Ds}^{\rm PDG}| < 50$ (10) MeV/c^2 for $B^0 \to D_s^+\pi^-$ and $B^0 \to D_s^- K^+$ ($B^0 \to D_s^{*+}\pi^-$and $B^0 \to D_s^{*-}K^+$). Also given are the reconstruction efficiency (ε), the significance, and the probability ($P_{\rm bckg}$) of the data being consistent with the background in the absence of signal, and the measured branching fraction $\mathcal{B}$. The first uncertainty is statistical, and the second is systematic.

B mode	D_s mode	$N_{\rm raw}$	$N_{\rm sig}$	$N_{\rm comb}$	$N_{\rm peak}$	$\varepsilon(\%)$	Signif.	$P_{\rm bckg}$	$\mathcal{B}(10^{-5})$	$\mathcal{B} \times \mathcal{B}(D_s^+ \to \phi\pi^+)$ (10^{-6})
$B^0 \to D_s^+\pi^-$	$\phi\pi^+$	405	21 ± 5	364 ± 20	21 ± 8	29.3	5.1σ	$3 \cdot 10^{-6}$	$1.3 \pm 0.3 \pm 0.2$	$0.63 \pm 0.15 \pm 0.05$
	$\overline{K}^{*0}K^+$	677	16 ± 4	604 ± 26	58 ± 12	20.0				
	$K_S^0 K^+$	223	11 ± 3	197 ± 15	16 ± 6	22.1				
$B^0 \to D_s^{*+}\pi^-$	$\phi\pi^+$	46	18 ± 4	29 ± 6	0	13.0	5.9σ	$3 \cdot 10^{-8}$	$2.8 \pm 0.6 \pm 0.5$	$1.32 \pm 0.27 \pm 0.15$
	$\overline{K}^{*0}K^+$	67	14 ± 3	48 ± 8	1	8.9				
	$K_S^0 K^+$	19	10 ± 2	12 ± 4	1	9.6				
$B^0 \to D_s^- K^+$	$\phi\pi^+$	197	32 ± 5	151 ± 13	8 ± 6	23.4	9.3σ	$3 \cdot 10^{-19}$	$2.5 \pm 0.4 \pm 0.4$	$1.21 \pm 0.17 \pm 0.11$
	$\overline{K}^{*0}K^+$	331	27 ± 4	306 ± 18	-4 ± 6	17.6				
	$K_S^0 K^+$	101	18 ± 3	82 ± 10	9 ± 5	19.0				
$B^0 \to D_s^{*-} K^+$	$\phi\pi^+$	15	9 ± 2	8 ± 3	-	8.9	4.7σ	$2 \cdot 10^{-5}$	$2.0 \pm 0.5 \pm 0.4$	$0.97 \pm 0.24 \pm 0.12$
	$\overline{K}^{*0}K^+$	16	8 ± 2	7 ± 3	-	6.6				
	$K_S^0 K^+$	10	5 ± 1	5 ± 3	-	6.7				

Chapter 6

Conclusions

This thesis presents measurements of the branching fractions of the decays $B^0 \to D_s^+\pi^-$, $B^0 \to D_s^{*+}\pi^-$, $B^0 \to D_s^- K^+$, and $B^0 \to D_s^{*-} K^+$. The results for the branching fractions, as well as the results for $r_{D^{(*)}\pi}$, are summarized in the table below. This is the first observation of the decays $B^0 \to D_s^{*+}\pi^-$ and $B^0 \to D_s^{*-} K^+$. As expected, the branching fractions for $B^0 \to D_s^{(*)-} K^+$ are small compared to the dominant decays $B^0 \to D^{(*)-}\pi^+$, implying relatively insignificant contributions from the color-suppressed W-exchange diagrams and therefore supporting our initial assumption in relating $B^0 \to D_s^+\pi^-$ to $B^0 \to D^+\pi^-$. In addition, the smallness of $r_{D^{(*)}\pi}$, originally estimated to be ~ 0.02, implies small CP asymmetries in $B^0 \to D^{(*)\mp}\pi^\pm$ decays.

In addition to providing a necessary input for the $\sin(2\beta+\gamma)$ analyses, the measurement

Table 6.1. Branching fractions and $r_{D^{(*)}\pi}$ results from 208.7 fb^{-1} of data.

Mode	$\mathcal{B}$ (10^{-5})	$\mathcal{B} \cdot \mathcal{B}(D_s^+ \to \phi\pi^+)$ (10^{-6})	Significance	$r_{D^{(*)}\pi}$ (10^{-2})
$B^0 \to D_s^+\pi^-$	$1.3 \pm 0.3 \pm 0.2$	$0.63 \pm 0.15 \pm 0.05$	5.1σ	$1.3 \pm 0.2 \pm 0.1$
$B^0 \to D_s^{*+}\pi^-$	$2.8 \pm 0.6 \pm 0.5$	$1.32 \pm 0.27 \pm 0.15$	5.9σ	$1.9 \pm 0.2 \pm 0.2$
$B^0 \to D_s^- K^+$	$2.5 \pm 0.4 \pm 0.4$	$1.21 \pm 0.17 \pm 0.11$	9.3σ	-
$B^0 \to D_s^{*-} K^+$	$2.0 \pm 0.5 \pm 0.4$	$0.97 \pm 0.24 \pm 0.12$	4.7σ	-

of $B^0 \to D_s^+\pi^-$ can also provide constraints on $|V_{ub}|$. The measurement of $|V_{ub}|$ through exclusive non-leptonic B decays into two meson final states has been explored by a number of theory papers [36, 37]. Final states in which no quark-antiquark pair has the same flavor have the least theoretical difficulty in terms of non-spectator diagrams. $B^0 \to D_s^+\pi^-$ is the cleanest of the processes explored without any penguin corrections.

The decay rate of $B^0 \to D_s^+\pi^-$ can be related to $|V_{ub}|$ by the equation:

$$\Gamma(B^0 \to D_s^+\pi^-) = \frac{G_F^2}{2}\,|V_{ub}V_{cs}^*|^2\,\frac{p_c^\pi}{8\pi m_B^2}\,(m_B^2 - m_\pi^2)^2\,[a_1 f_{D_s} F_0^{B\pi}(m_{D_s}^2)]^2, \tag{6.1}$$

assuming factorization. G_F is the Fermi constant, p_c^π is the momentum of the pion in the CM frame, a_1 is the effective parameter derived from the Wilson coefficients, f_{D_s} is the D_s decay constant, and $F_0^{B\pi}(m_{D_s}^2)$ is the form factor evaluated at $q^2 = m_{D_s}^2$ [37]. The largest uncertainties arise from the measurement of the $B^0 \to D_s^+\pi^-$ branching fraction, as well as the hadronic form factors and a_1. After inputting the values of the Fermi constant, V_{cs}, the momentum of the pion, and the meson masses, this equation can be written as:

$$\Gamma(B^0 \to D_s^+\pi^-) = (1.065 \cdot 10^{-12})|V_{ub}|^2 \left[\frac{a_1}{1.059}\right]^2 \left[\frac{f_{D_s}}{0.240\,\text{GeV}}\right]^2 \left[\frac{F_0^{B\pi}(m_{D_s}^2)}{0.319}\right]^2 \text{GeV}. \tag{6.2}$$

Using the values of a_1 and $F_0^{B\pi}(m_{D_s}^2)$ quoted in [37] and the updated value of f_{D_s} from [13], in addition to total width of the B^0 [12], we arrive at:

$$|V_{ub}| = (2.1 \pm 0.4) \cdot 10^{-3}. \tag{6.3}$$

The value from other measurements quoted in the PDG is $|V_{ub}| = (3.67 \pm 0.47) \cdot 10^{-3}$ [12]. The deviation of our measurement from this central value suggests a breakdown of the factorization assumption for $B^0 \to D_s^+\pi^-$, which must be further understood.

Recent studies have analyzed the $SU(3)$ violating effects in using the branching fractions of $B^0 \to D_s^{(*)+}\pi^-/\rho^-$ to compute the branching fractions of the CKM-suppressed decays $B^0 \to D^{(*)+}\pi^-/\rho^-$ and have calculated a 10.5% error for residual $SU(3)$ breaking in non-factorizable corrections [35]. These results present limits on γ from $B^0 \to D^{(*)\mp}\pi^\pm/\rho^\pm$, using the latest branching fractions shown in Table 6.1 averaged with the results from other

experiments. The results for the 90% confidence regions for γ are [-0.19,1.84] rad, including the $SU(3)$ breaking error, and [-0.18,1.83] rad without the error. These calculations include a flat 5% error from neglecting the W-exchange diagram contributions as well as a flat theoretical error of 30% from the original $SU(3)$ assumption.

In the near future, we hope to update these branching fraction measurements to the full dataset available, which is twice as large as the one used in this study. The updated measurement will include the modes studied here, as well as $B^0 \to D_s^{(*)+}\rho^-$ and $B^0 \to D_s^{(*)-}K^{*+}$, both of which have yet to be observed. Increased statistics will help to increase the significance of the measurements. In addition, an improved measurement of $\mathcal{B}(D_s^+ \to \phi\pi^+)$, which contributes the single largest systematic, will help to decrease our errors.

Bibliography

[1] T.D. Lee and C.N. Yang, Phys. Rev. **104**, 254 (1956).

[2] C.S. Wu *et al.*, Phys. Rev. Lett. **105**, 1413 (1957).

[3] J.H. Christenson, J.W. Cronin, V.L. Fitch, and R. Turlay, Phys. Rev. Lett. **13**, 138 (1964).

[4] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **86**, 2515 (2001).

[5] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963);

[6] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[7] C. Jarlskog, in *CP Violation*, C. Jarlskog ed., World Scientific, Singapore (1988).

[8] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).

[9] M. Bona *et al.*, UT fit group, http://www.utfit.org.

[10] I. Dunietz, Phys. Lett. **B427**, 179 (1998).

[11] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett **90**, 181803 (2003).

[12] S. Eidelman *et al.*(Particle Data Group), Phys. Lett. **B592**, 1 (2004) and 2005 partial update for edition 2006.

[13] D. Becirevic, preprint hep-ph/0310072, eConf **C0304052**, WG202 (2003).

[14] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **92**, 251801 (2004); Belle Collaboration, T. Sarangi *et al.*, Phys. Rev. Lett. **93**, 031802 (2004); *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. **D71**, 112003 (2005).

[15] B. Block, M. Gronau, and J.L. Rosner, Phys. Rev. Lett. **78**, 3999 (1997).

[16] B. Blok, M. Gronau, and J. Rosner, hep-ph/9701396 (2001).

[17] C. Lu and K. Ukai, hep-ph/0210206 (2002).

[18] PEP-II Conceptual Design Report, SLAC-0418 (1993).

[19] *BABAR* Collaboration, B. Aubert *et al.*, Nucl. Instrum. Methods **A479**, 1 (2002).

[20] http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/Pid/Hadrons/pid_summarytable.html

[21] R. Bartoldus, A. Eisner, A. Ryd, and D. Su, *Trigger and Filter Documentation for Run1*, *BABAR* Analysis Document #194, 2002.

[22] The RooFit Tookit for Data Modeling, http://roofit.sourceforge.net.

[23] C. Hearty, *Measurement of the number of $\Upsilon(4S)$ mesons produced in Run1 (B counting)*, *BABAR* Analysis Document #134, 2001.

[24] GEANT4 Collaboration, S. Agostinelli *et al.*, Nucl. Instr. Methods Phys. Res., Sect. A**479**, 1 (2002).

[25] G.C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).

[26] ARGUS Collaboration, H. Albrecht *et al.*, Z. Phys. **C48**, 543 (1990).

[27] R. A. Fisher, Annals Eugen. **7**, 179 (1936).

[28] S. Brandt *et al.*, Phys. Lett. **12**, 57 (1964); E. Fahiri, Phys. Rev. Lett. **39**, 1587 (1977).

[29] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **89**, 281802 (2002); Phys. Rev. **D70**, 032006 (2004).

[30] ARGUS Collaboration, H. Albrecht *et al.*, Z. Phys. **C48**, 543 (1990).

[31] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. **D71**, 091104 (2005).

[32] M.T. Allen, M.T. Naisbit, A. Roodman, and S. Banerjee, *A Study of Pi-zero Efficiency*, *BABAR* Analysis Document #870, 2004.

[33] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. D **71**, 091104 (2005).

[34] S. Kaiser, H.M. Lacker, and A. Volk, *Measurement of the Decay $B^+ \to D_s\pi^0$*, *BABAR* Analysis Document #1329, 2006.

[35] M. Baak and D. Pirjol, *$SU(3)$ Breaking and rescattering in $B^0 \to D_s^{(*)+}\pi^-/\rho^-$ and $B^0 \to D^{(*)+}\pi^-/\rho^-$*, *BABAR* Analysis Document #1517, 2006.

[36] Y. Koide, Phys. Rev. D **39**, 3500 (1989); D. Choudhury, D. Indumati, A. Soni, and S.U. Sankar, Phys. Rev. D **45**, 217 (1992); I. Dunietz and J.L. Rosner, CERN-TH-5899-90.

[37] C.S. Kim, Y. Kwon, J. Lee, W. Namgung, Phys. Rev. D **63**, 094506, (2001).

Appendix A

Monte Carlo Samples

A.1 Generic Monte Carlo Samples

1. 390,668,000 events: $B^0\overline{B}^0$ generic (equivalent luminosity ~ 744 fb^{-1})
2. 400,622,000 events: B^+B^- generic (equivalent luminosity ~ 763 fb^{-1})
3. 352,834,000 events: $c\bar{c}$ generic (equivalent luminosity ~ 271 fb^{-1})
4. 503,482,000 events: $u\bar{u}/s\bar{s}/d\bar{d}$ generic events (equivalent luminosity ~ 241 fb^{-1})

A.2 Exclusive Monte Carlo Samples

1. 171000 events: $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+, \bar{K}^{0*}K^+, \bar{K}^0K^+, \phi\rho^+, \phi\pi^+\pi^0$
2. 173000 events: $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+, \bar{K}^{0*}K^+, \bar{K}^0K^+, \phi\rho^+, \phi\pi^+\pi^0$
3. 159000 events: $B^0 \to D_s^+\rho^-$, $D_s^+ \to \phi\pi^+, \bar{K}^{0*}K^+, \bar{K}^0K^+, \phi\rho^+, \phi\pi^+\pi^0$
4. 176000 events: $B^0 \to D_s^{*+}\rho^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+, \bar{K}^{0*}K^+, \bar{K}^0K^+, \phi\rho^+, \phi\pi^+\pi^0$, longitudinal polarization
5. 176000 events: $B^0 \to D_s^{*+}\rho^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+, \bar{K}^{0*}K^+, \bar{K}^0K^+, \phi\rho^+, \phi\pi^+\pi^0$, transverse polarization
6. 60000 events: $B^0 \to D_s^-K^+$, $D_s^+ \to \phi\pi^+$, $\phi \to K^+K^-$.

7. 60000 events: $B^0 \to D_s^- K^+$, $D_s^+ \to \bar{K}^{0*} K^+$ $K^{*0} \to K^+ \pi^-$.

8. 60000 events: $B^0 \to D_s^- K^+$, $D_s^+ \to \bar{K}^0 K^+$, $K_s \to \pi^+ \pi^-$.

9. 60000 events: $B^0 \to D_s^{*-} K^+$, $D_s^{*+} \to D_s^+ \gamma$, $D_s^+ \to \phi \pi^+$, $\phi \to K^+ K^-$.

10. 60000 events: $B^0 \to D_s^{*-} K^+$, $D_s^{*+} \to D_s^+ \gamma$, $D_s^+ \to \bar{K}^{0*} K^+$ $K^{*0} \to K^+ \pi^-$.

11. 60000 events: $B^0 \to D_s^{*-} K^+$, $D_s^{*+} \to D_s^+ \gamma$, $D_s^+ \to \bar{K}^0 K^+$, $K_s \to \pi^+ \pi^-$.

12. 174000 events: $B^0 \to D_s^- K^{*+}$, $D_s^+ \to \phi \pi^+, \bar{K}^{0*} K^+, \bar{K}^0 K^+, \phi \rho^+, \phi \pi^+ \pi^0$

13. 175000 events: $B^0 \to D_s^{*+} K^{*-}$, $D_s^{*+} \to D_s^+ \gamma$, $D_s^+ \to \phi \pi^+, \bar{K}^{0*} K^+, \bar{K}^0 K^+, \phi \rho^+, \phi \pi^+ \pi^0$, longitudinal polarization

14. 173000 events: $B^0 \to D_s^{*+} K^{*-}$, $D_s^{*+} \to D_s^+ \gamma$, $D_s^+ \to \phi \pi^+, \bar{K}^{0*} K^+, \bar{K}^0 K^+, \phi \rho^+, \phi \pi^+ \pi^0$, transverse polarization

15. 348000 events: $B^+ \to D_s^+ \pi^0$, $D_s^+ \to \phi \pi^+, \bar{K}^{0*} K^+, \bar{K}^0 K^+, \phi \rho^+, \phi \pi^+ \pi^0$

16. 347000 events: $B^+ \to \bar{D}^{0*} \pi^+$, $D_s^{*+} \to D_s^+ \gamma$, $D_s^+ \to \phi \pi^+, \bar{K}^{0*} K^+, \bar{K}^0 K^+, \phi \rho^+, \phi \pi^+ \pi^0$

17. 115000 events: $B^+ \to D_s^+ \rho^0$, $D_s^+ \to \phi \pi^+, \bar{K}^{0*} K^+, \bar{K}^0 K^+, \phi \rho^+, \phi \pi^+ \pi^0$

18. 115000 events: $B^+ \to D_s^{*+} \rho^0$, $D_s^{*+} \to D_s^+ \gamma$, $D_s^+ \to \phi \pi^+, \bar{K}^{0*} K^+, \bar{K}^0 K^+, \phi \rho^+, \phi \pi^+ \pi^0$, longitudinal polarization

19. 116000 events: $B^+ \to D_s^{*+} \rho^0$, $D_s^{*+} \to D_s^+ \gamma$, $D_s^+ \to \phi \pi^+, \bar{K}^{0*} K^+, \bar{K}^0 K^+, \phi \rho^+, \phi \pi^+ \pi^0$, transverse polarization

20. 28,954,000 events: "cocktail" $B^0 \to D^- X$ and $B^0 \to D^{*-} X$, where $X = \pi$, a_1, ρ. The fraction of events with $B^0 \to D^- \pi^+$ is 6.3% and the fraction of $B^0 \to D^- \rho^+$ is 16.4%.

21. 6,282,000 events: "cocktail" $B^+ \to \bar{D}^0 \pi^+$ and $B^+ \to \bar{D}^{0*} \pi^+$. The fraction of events with $B^+ \to \bar{D}^0 \pi^+$ is 53.4%.

22. 7,322,000 events: "cocktail" $B^+ \to D^{(*)0} X$, where $X = \rho$, a_1, $\rho(2S)$, 4π .

23. 114,000 events: $B^0 \to D^- K^+, D^- \to K^+ \pi^- \pi^-$

24. 232,000 events: $B^0 \to D^- K^+, D^- \to K^0 \pi^-$

25. 114,000 events: $B^0 \to D^- K^+, D^- \to K^{*0} \pi^-$

A.3 Control Samples

1. $207\,\mathrm{fb}^{-1}$ of skimmed data, reconstructed in the modes $B^0 \to D^-\pi^+$, with $D^+ \to K^-\pi^+\pi^+$, $\bar{K}^0\pi^+$ and $B^+ \to \bar{D}^{0*}\pi^+$, with $\bar{D}^{*0} \to \bar{D}^0\gamma$, $\bar{D}^0 \to K^-\pi^+$
2. 466,000 events: $B^0 \to D^-\pi^+$, $D^+ \to K^-\pi^+\pi^+$
3. 169,000 events: $B^0 \to D^-\pi^+$, $D^+ \to \bar{K}^0\pi^+$
4. 192,500 events: $B^+ \to \bar{D}^{0*}\pi^+$, $\bar{D}^{*0} \to \bar{D}^0\gamma$, $\bar{D}^0 \to K^-\pi^+$

Appendix B

Maximum Likelihood Fits

B.1 Likelihood Functions

If x is the possible outcome of an observation and the range of x is continuous, then the probability that the outcome of a measurement will lie between x and $x + dx$ can be written as $\mathcal{P}(x;\theta)dx$ [12]. $\mathcal{P}(x;\theta)$ is called the probability density function (PDF), which can depend on one or more parameters θ. PDFs are always normalized to unit area.

Next we assume a set of N independently measured quantities x_i come from a PDF, $\mathcal{P}(x_i;\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a set of n parameters $\boldsymbol{\theta} = (\theta_1, ..., \theta_n)$ with undetermined values. In the method of maximum likelihood, the estimators $\hat{\boldsymbol{\theta}}$ are taken to be the values of $\boldsymbol{\theta}$ which maximize the likelihood function, defined as:

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{i=1}^{N} \mathcal{P}(x_i;\boldsymbol{\theta}). \tag{B.1}$$

Typically what is maximized is the logarithm of the likelihood in determining the maximum likelihood estimators:

$$\frac{\partial \mathcal{L}}{\partial \theta_i} = 0, \quad i = 1, \ldots, n. \tag{B.2}$$

B.2 Extended Maximum Likelihood Fits

The fits used to extract the yields and branching fractions for this analysis are unbinned extended maximum likelihood fits. The likelihood function that is maximized in the fits is:

$$\mathcal{L}(\boldsymbol{\theta}) = \exp\left(-\sum_k n_k\right)\prod_{i=1}^{N}\left(\sum_j n_j \mathcal{P}_j(\vec{x}_i, \boldsymbol{\theta})\right), \tag{B.3}$$

where n_j is the yield for event type j =signal, combinatorial background, charmless background, reflection background, etc. $\mathcal{P}_j$ is the PDF for j and is a function of the variables $\vec{x}_i$. $\vec{x}_i = \{m_{\rm ES}, m_{Ds}\}$ for the $B^0 \to D_s^+\pi^-$ and $B^0 \to D_s^- K^+$ fits, and $\vec{x}_i = \{m_{\rm ES}\}$ for the $B^0 \to D_s^{*+}\pi^-$ and $B^0 \to D_s^{*-} K^+$ fits.

B.3 Significance of the Measurement

To compute the significance, the likelihood fit is compared to the result of the fit assuming a null hypothesis for the signal yield. In other words, the significance is computed by comparing the likelihood of the result to the likelihood of the background fluctuating upward to emulate the signal events.

Assuming the likelihood distribution is Gaussian distributed:

$$\mathcal{L} = C\,\exp\left((n_{sig} - \hat{n}_{sig})^2/2\sigma^2\right), \tag{B.4}$$

where C is a normalization constant, n_{sig} is the observable number of signal events, $\hat{n}_{sig}$ is the estimator for the number of signal events, and σ is the error on the number of signal events. The ratio of the likelihood of the fit, $\mathcal{L}_0$, to the likelihood of the fit assuming $n_{sig} = 0$, $\mathcal{L}$, is equivalent to the inverse of the probability that the background fluctuated upward to produce the signal:

$$1/\text{Probability of fluctuation} = \mathcal{L}_0/\mathcal{L} = \exp\left(-n_{sig}^2/2\sigma^2\right). \tag{B.5}$$

The significance is then computed as:

$$\text{Significance} = n_{sig}/\sigma = \sqrt{-2\,\ln\left(\mathcal{L}_0/\mathcal{L}\right)}. \tag{B.6}$$

Appendix C

Likelihood Plots

C.1 PDF Fits for Signal and Generic MC

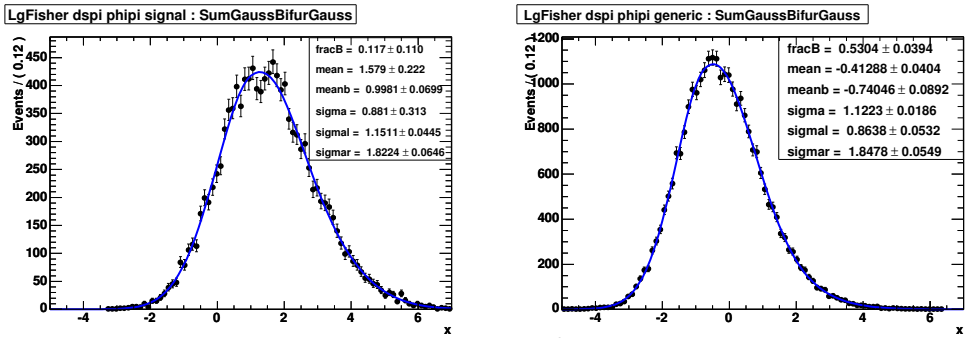


Figure C.1. Legendre Fisher discriminant for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$, signal (left) and generic (right) MC

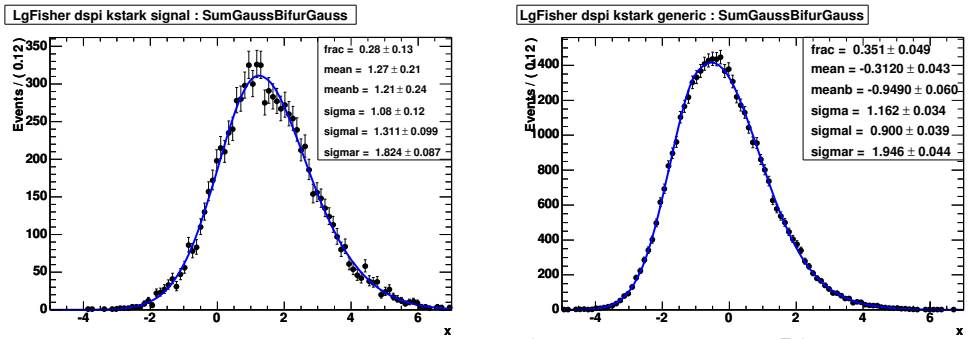


Figure C.2. Legendre Fisher discriminant for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$, signal (left) and generic (right) MC

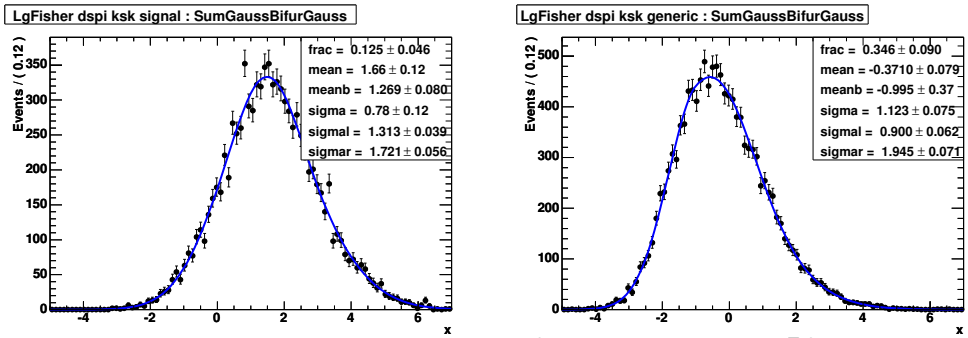


Figure C.3. Legendre Fisher discriminant for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^0K^+$, signal (left) and generic (right) MC

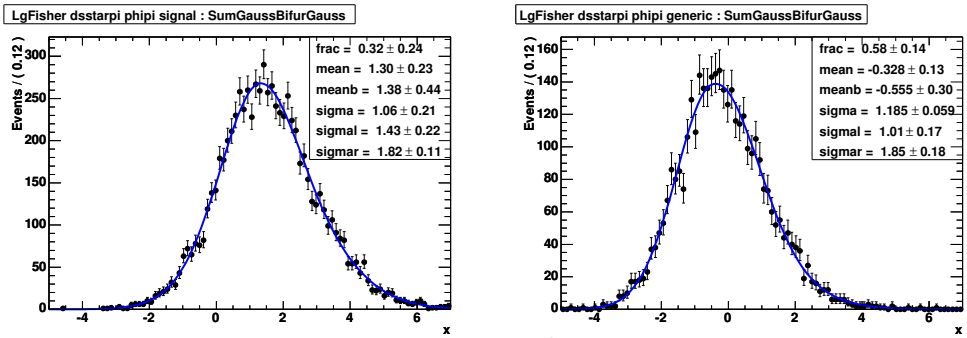


Figure C.4. Legendre Fisher discriminant for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$, signal (left) and generic (right) MC

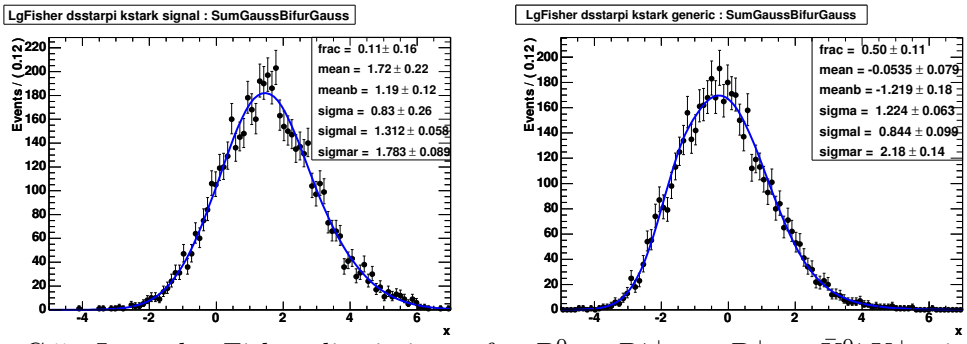


Figure C.5. Legendre Fisher discriminant for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$, signal (left) and generic (right) MC

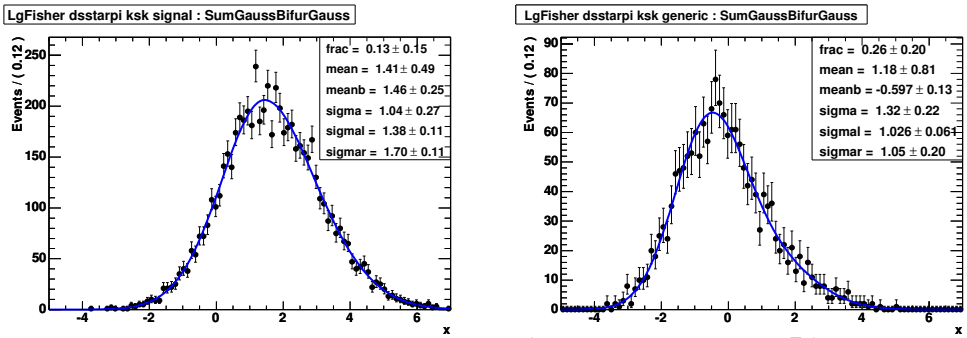


Figure C.6. Legendre Fisher discriminant for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^0K^+$, signal (left) and generic (right) MC

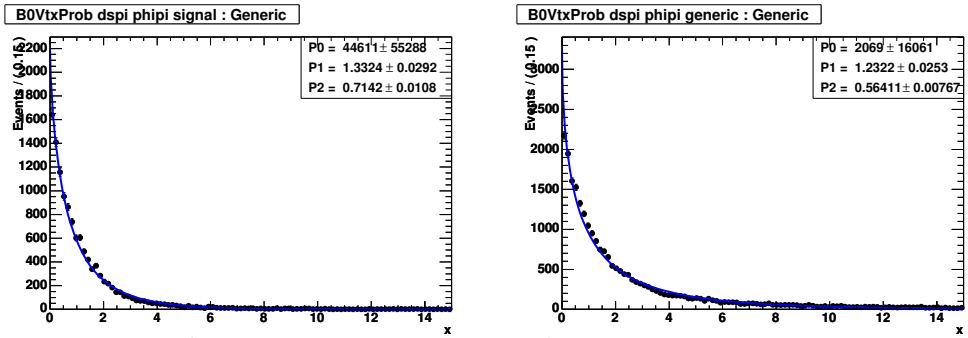


Figure C.7. -log(B^0 vertex probability) for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$, signal (left) and generic (right) MC

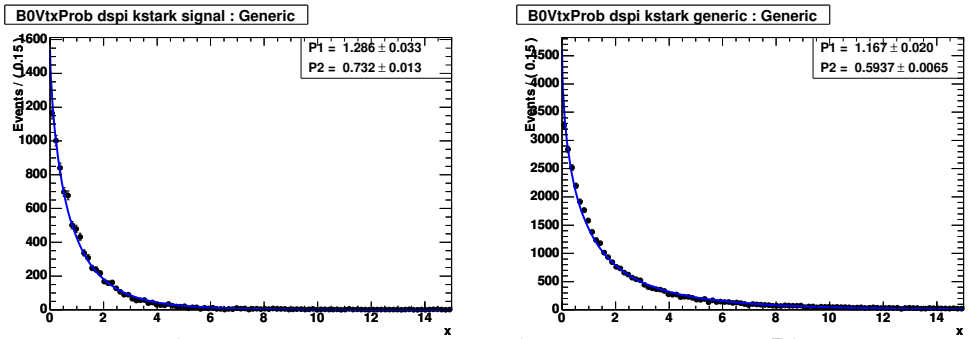


Figure C.8. -log(B^0 vertex probability) for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$, signal (left) and generic (right) MC

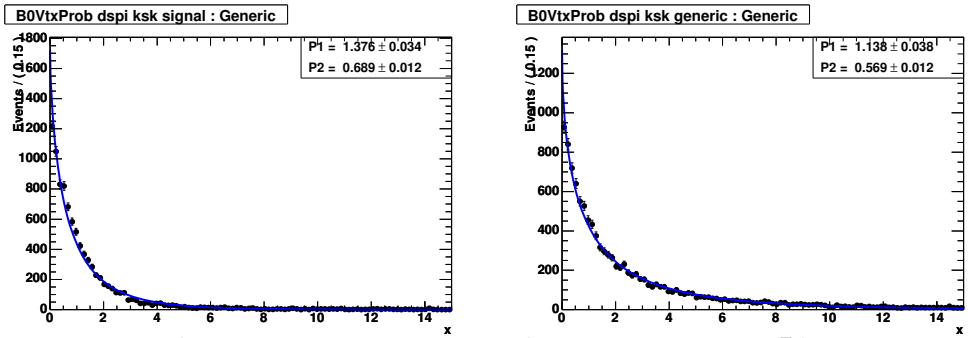


Figure C.9. -log(B^0 vertex probability) for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^0K^+$, signal (left) and generic (right) MC

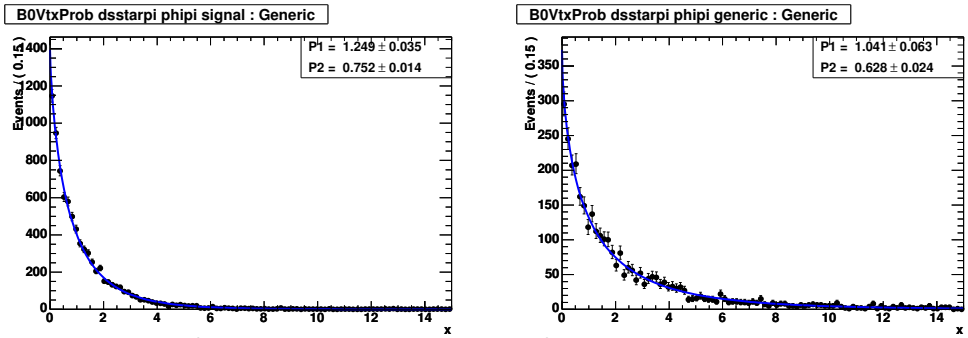


Figure C.10. -log(B^0 vertex probability) for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$, signal (left) and generic (right) MC

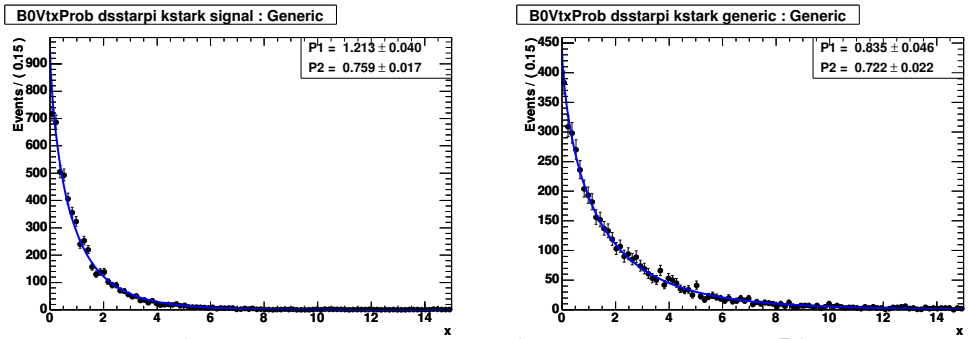


Figure C.11. -log(B^0 vertex probability) for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$, signal (left) and generic (right) MC

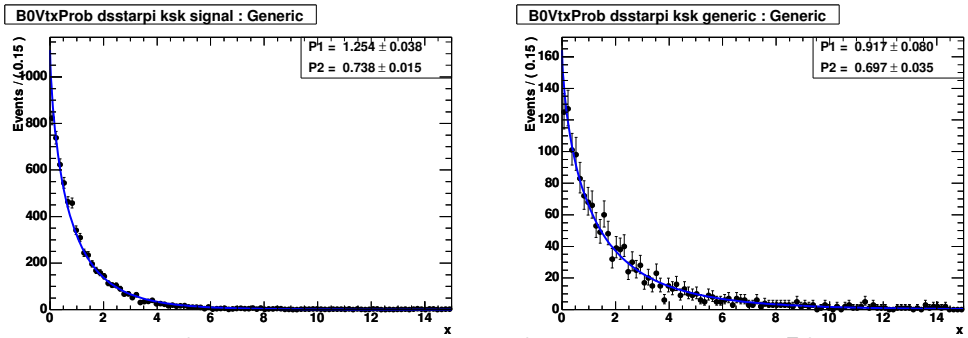


Figure C.12. -log(B^0 vertex probability) for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^0K^+$, signal (left) and generic (right) MC

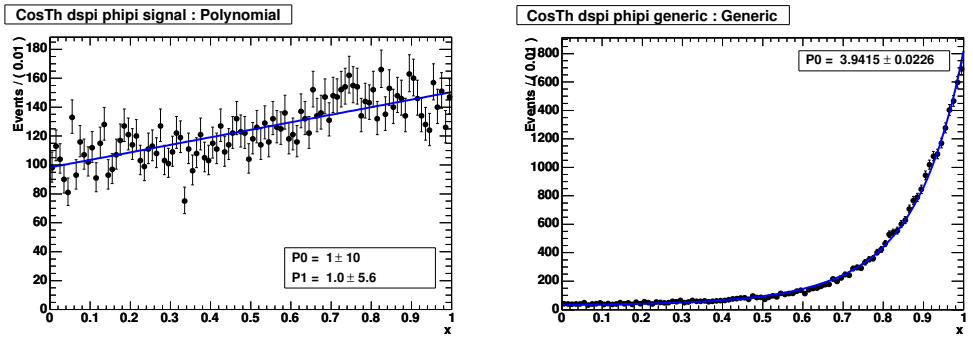


Figure C.13. $\cos\theta_{\rm thrust}$ for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$, signal (left) and generic (right) MC

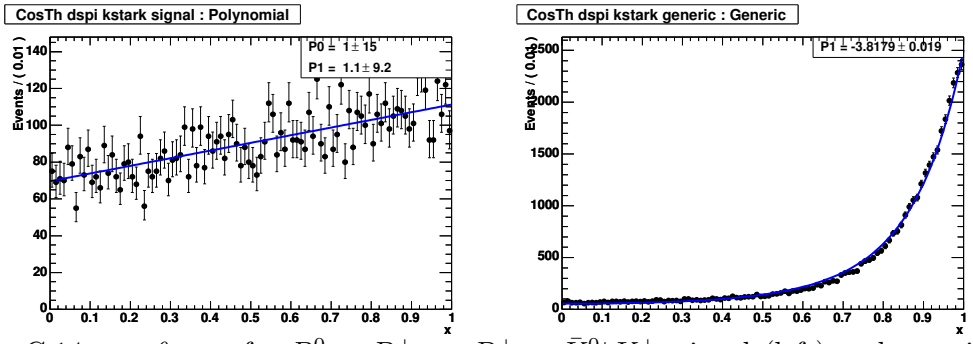


Figure C.14. $\cos\theta_{\rm thrust}$ for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$, signal (left) and generic (right) MC

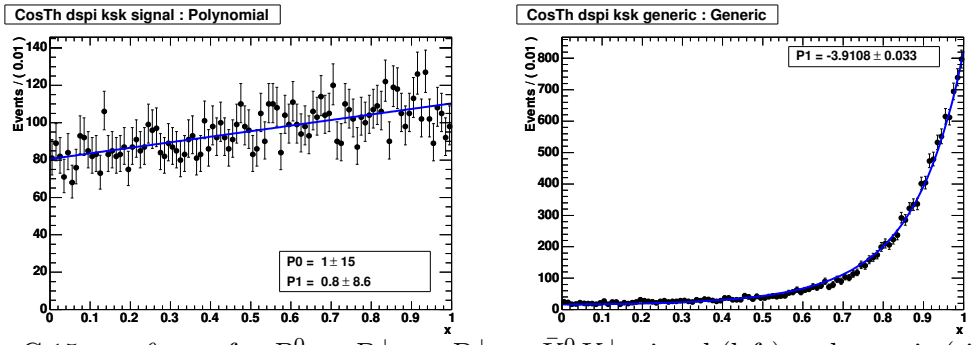


Figure C.15. $\cos\theta_{\rm thrust}$ for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^0K^+$, signal (left) and generic (right) MC

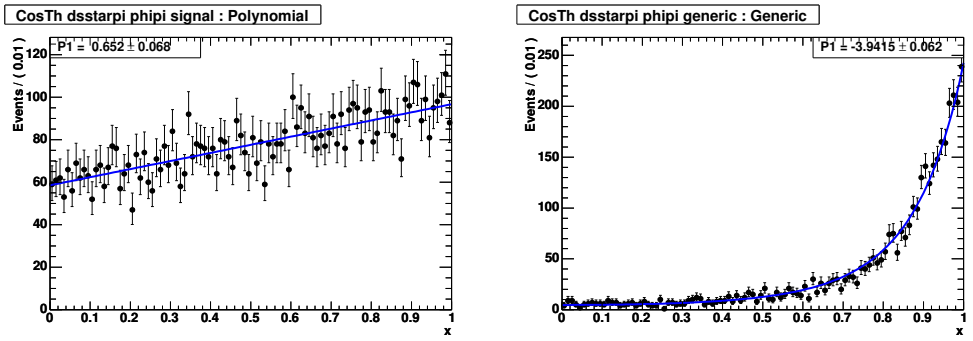


Figure C.16. $\cos\theta_{\rm thrust}$ for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$, signal (left) and generic (right) MC

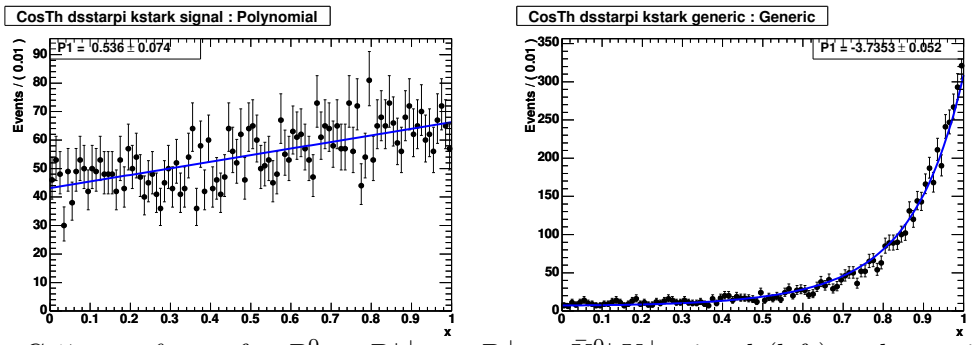


Figure C.17. $\cos\theta_{\rm thrust}$ for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$, signal (left) and generic (right) MC

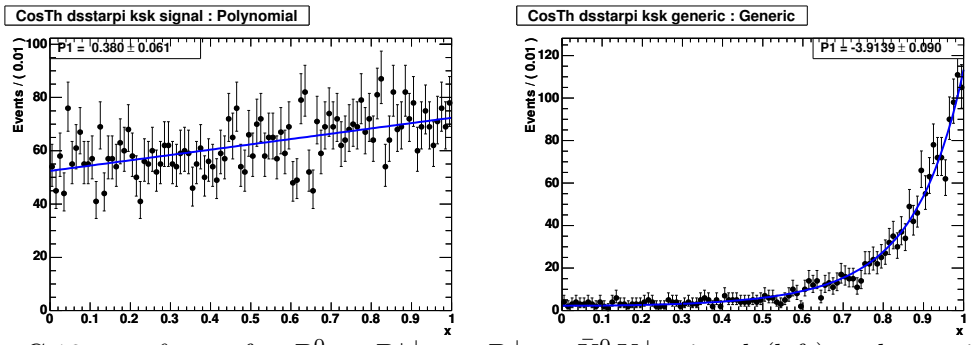


Figure C.18. $\cos\theta_{\rm thrust}$ for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^0K^+$, signal (left) and generic (right) MC

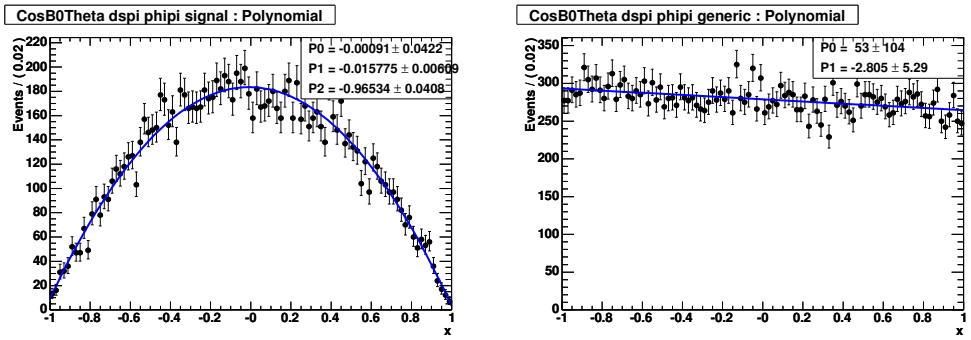


Figure C.19. $\cos\theta_{\rm B}$ for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$, signal (left) and generic (right) MC

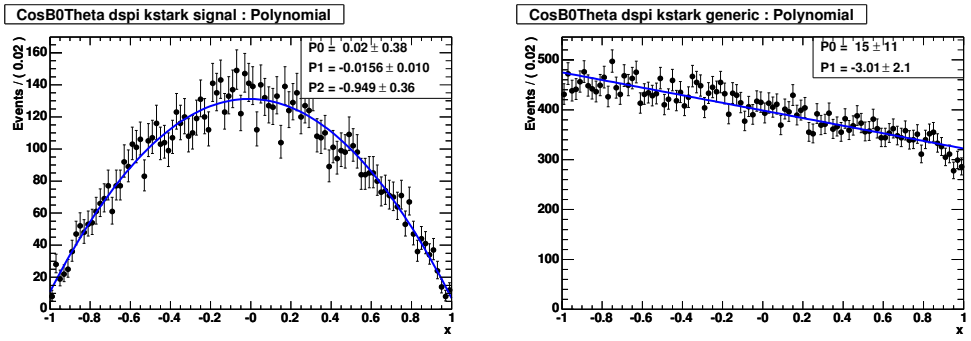


Figure C.20. $\cos\theta_{\rm B}$ for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$, signal (left) and generic (right) MC

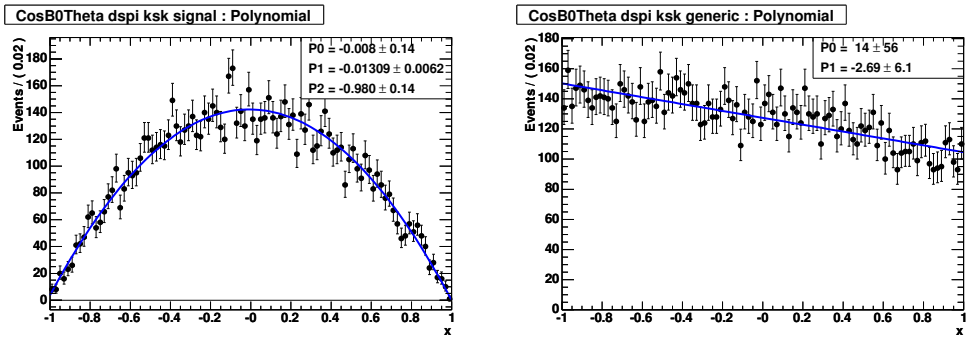


Figure C.21. $\cos\theta_{\rm B}$ for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^0K^+$, signal (left) and generic (right) MC

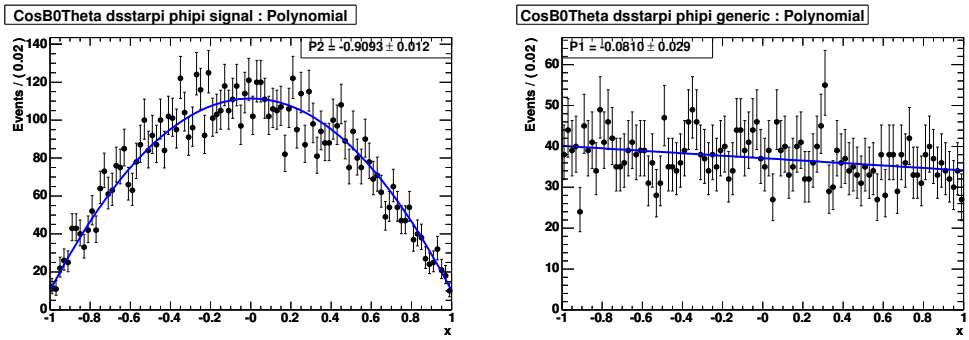


Figure C.22. $\cos\theta_{\rm B}$ for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$, signal (left) and generic (right) MC

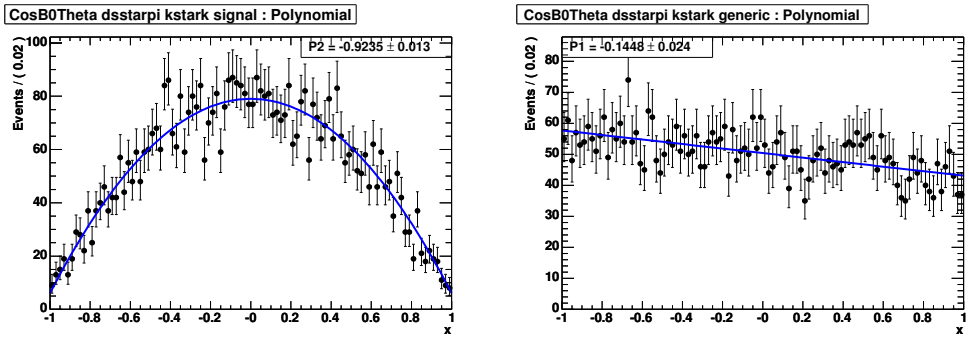


Figure C.23. $\cos\theta_{\rm B}$ for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$, signal (left) and generic (right) MC

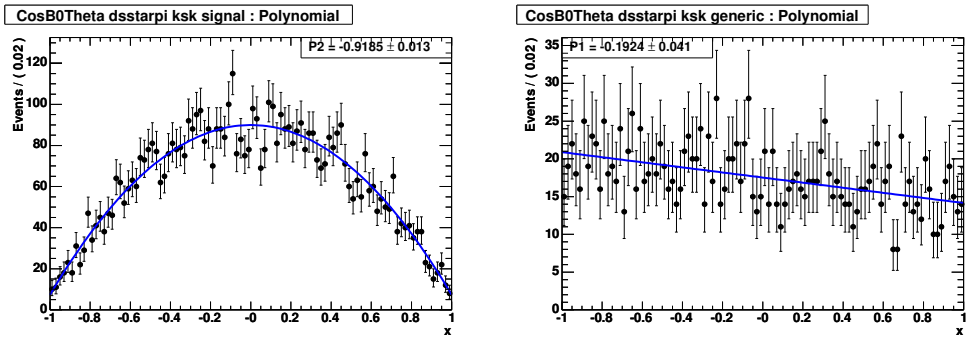


Figure C.24. $\cos\theta_{\rm B}$ for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^0K^+$, signal (left) and generic (right) MC

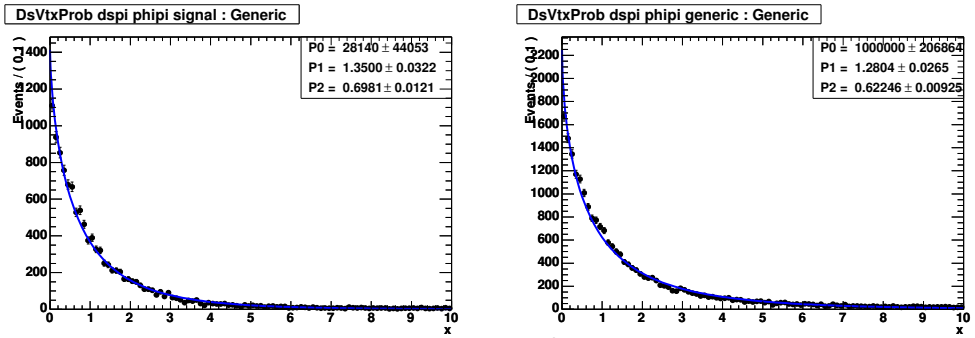


Figure C.25. -log(D_s vertex probability) for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$, signal (left) and generic (right) MC

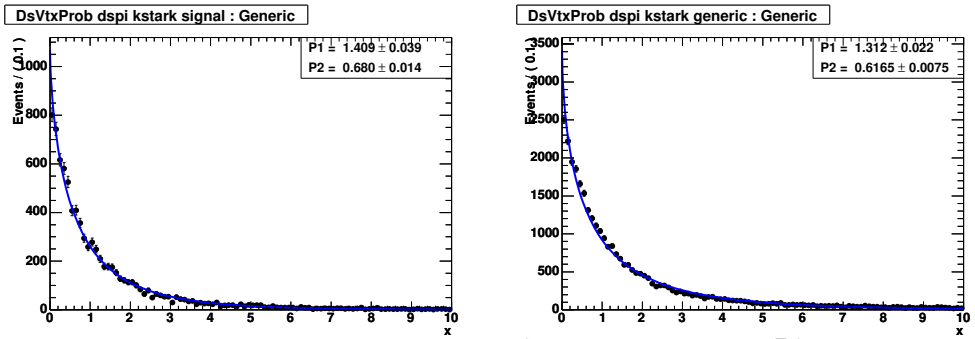


Figure C.26. -log(D_s vertex probability) for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$, signal (left) and generic (right) MC

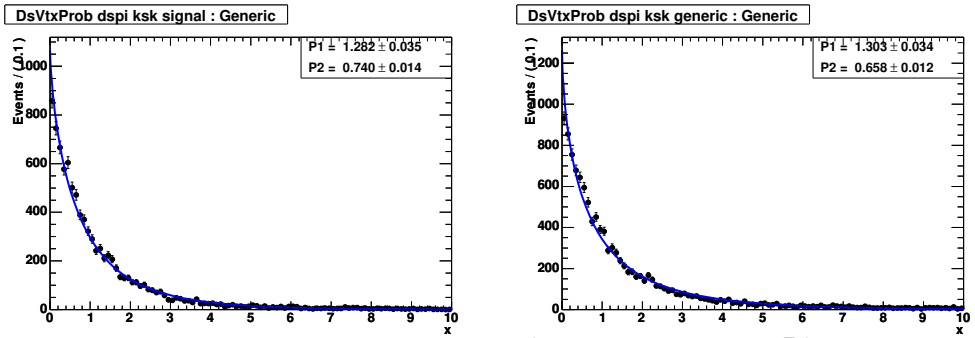


Figure C.27. -log(D_s vertex probability) for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^0K^+$, signal (left) and generic (right) MC

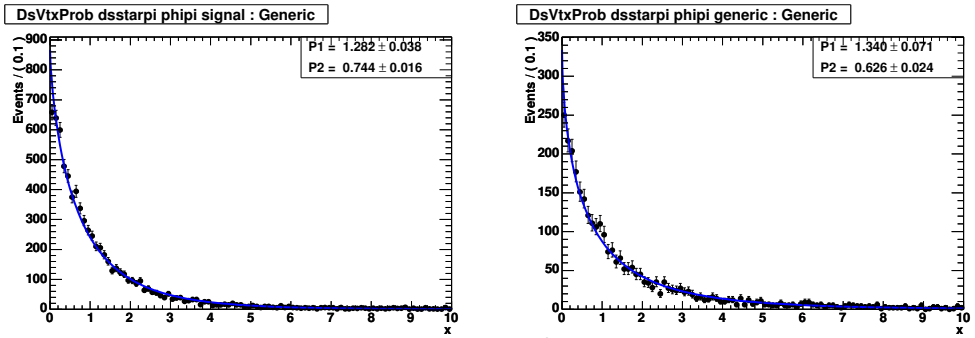


Figure C.28. -log(D_s vertex probability) for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$, signal (left) and generic (right) MC

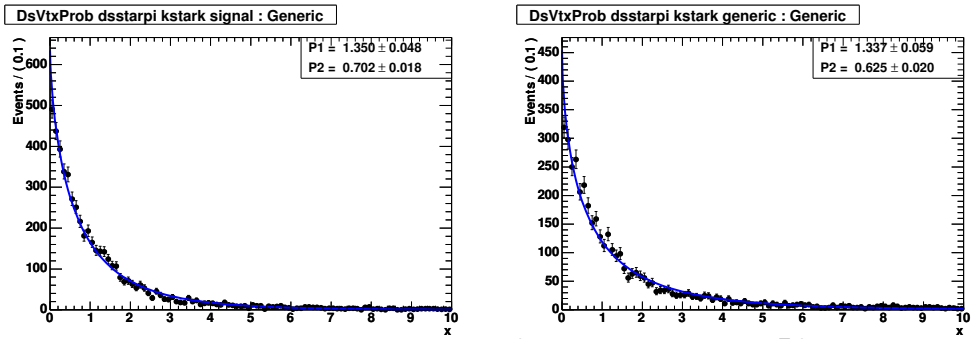


Figure C.29. -log(D_s vertex probability) for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$, signal (left) and generic (right) MC

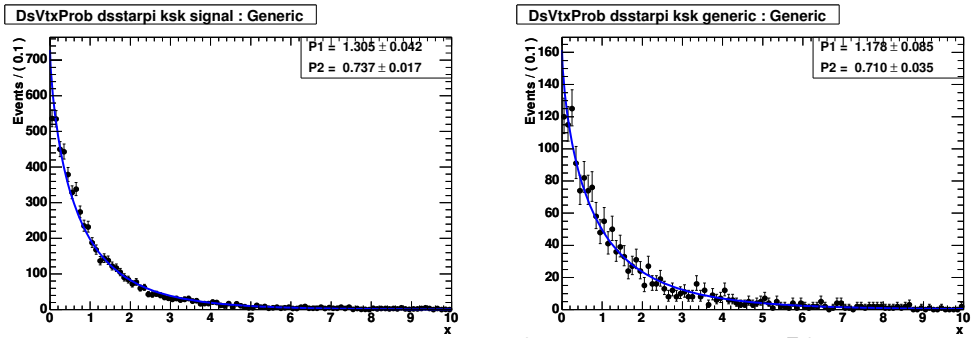


Figure C.30. -log(D_s vertex probability) for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^0K^+$, signal (left) and generic (right) MC

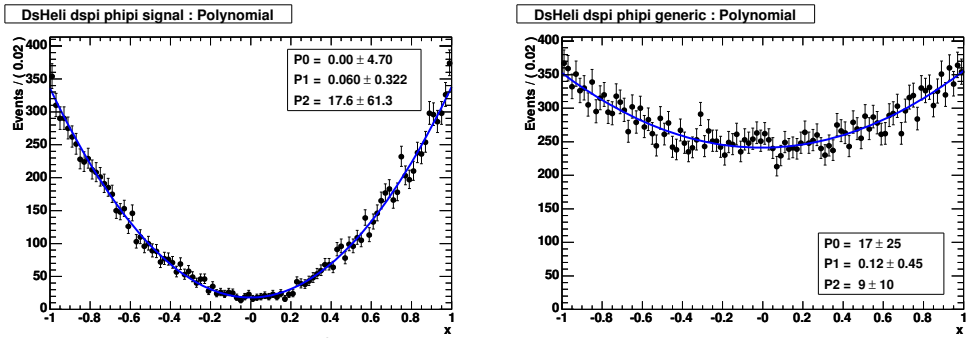


Figure C.31. $\cos\theta_{\rm hel}$ of ϕ for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$, signal (left) and generic (right) MC

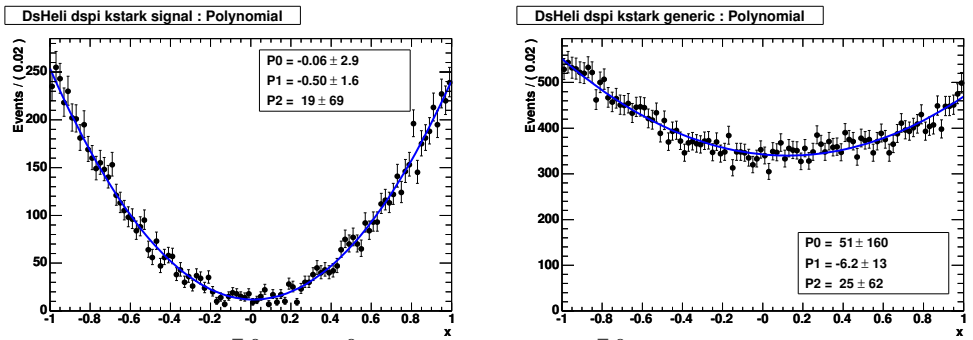


Figure C.32. $\cos\theta_{\rm hel}$ of $\bar{K}^{0*}$ for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$, signal (left) and generic (right) MC

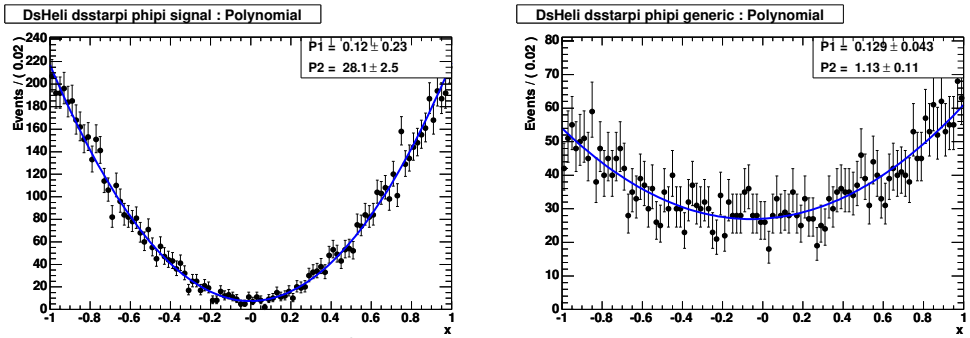


Figure C.33. $\cos\theta_{\rm hel}$ of ϕ for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$, signal (left) and generic (right) MC

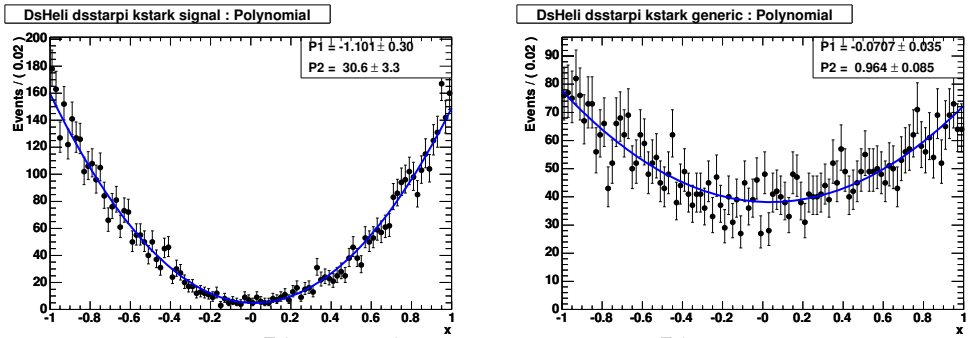


Figure C.34. $\cos\theta_{\rm hel}$ of $\bar{K}^{0*}$ for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$, signal (left) and generic (right) MC

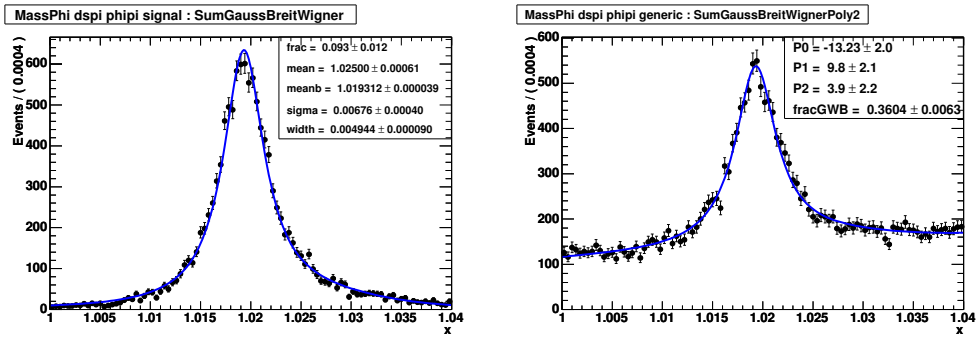


Figure C.35. Mass ϕ for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$, signal (left) and generic (right) MC

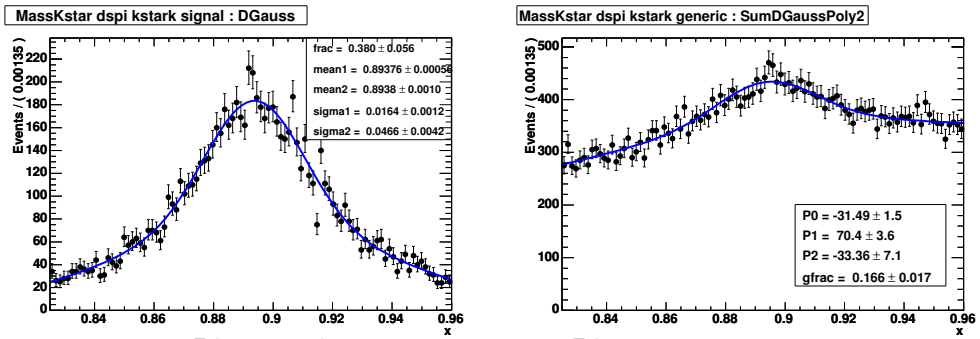


Figure C.36. Mass $\bar{K}^{0*}$ for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$, signal (left) and generic (right) MC

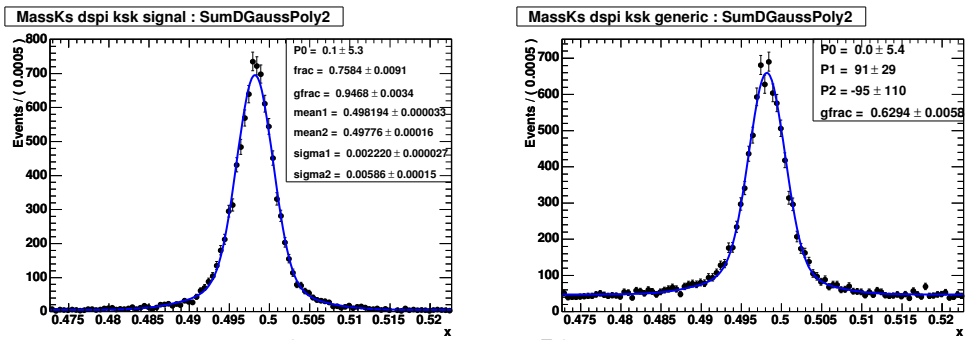


Figure C.37. Mass K_s for $B^0 \to D_s^+\pi^-$, $D_s^+ \to \bar{K}^0K^+$, signal (left) and generic (right) MC

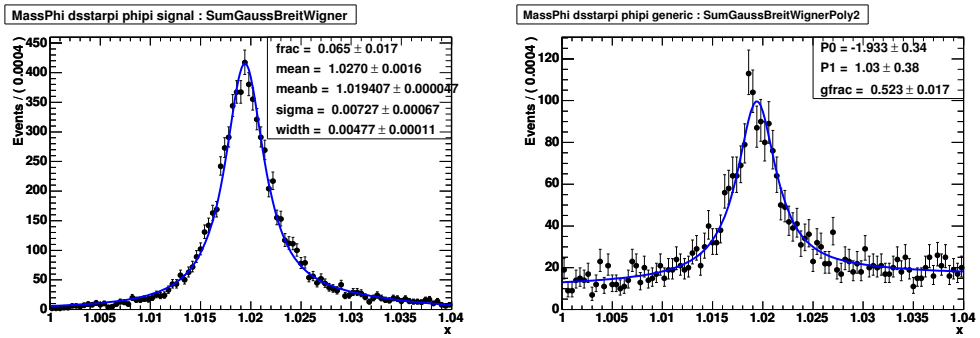


Figure C.38. Mass ϕ for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$, signal (left) and generic (right) MC

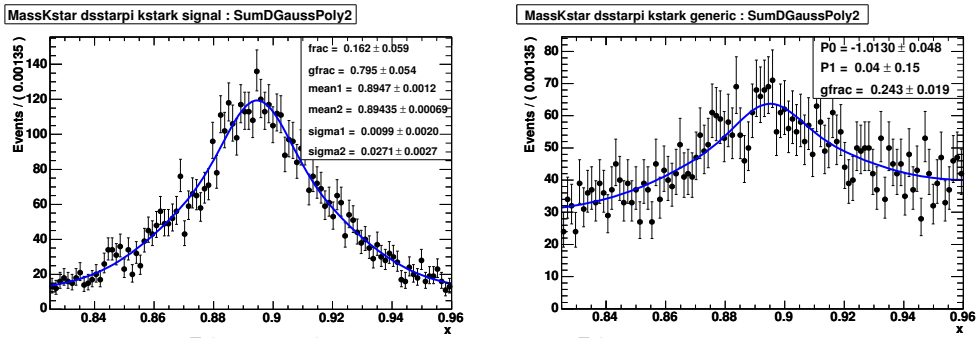


Figure C.39. Mass $\bar{K}^{0*}$ for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$, signal (left) and generic (right) MC

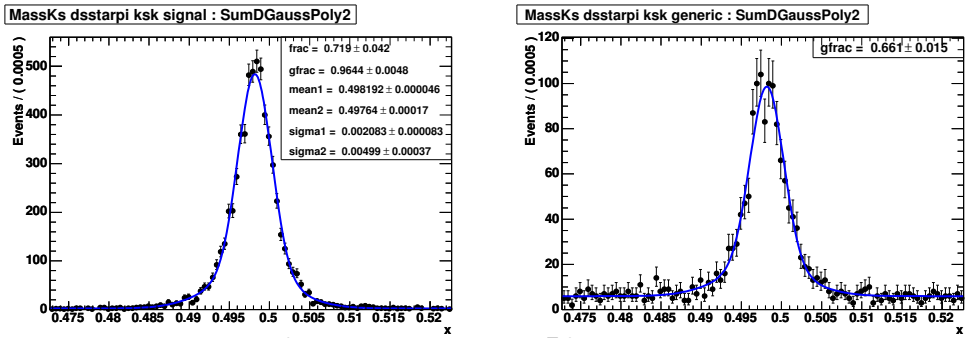


Figure C.40. Mass K_s for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^0K^+$, signal (left) and generic (right) MC

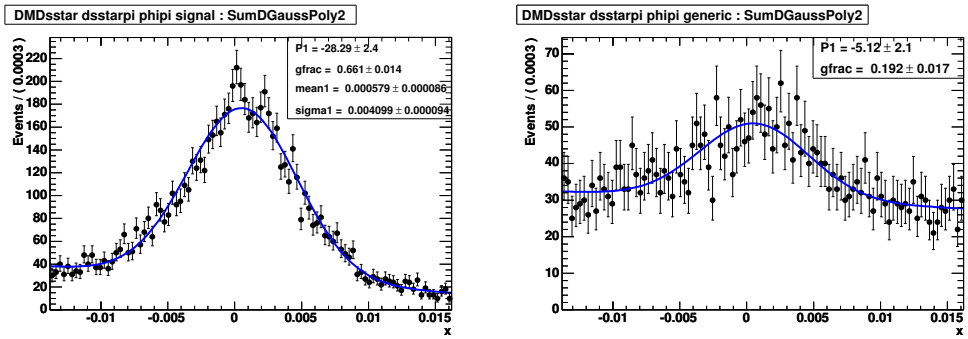


Figure C.41. $\Delta m_{D_s^*}$ for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$, signal (left) and generic (right) MC

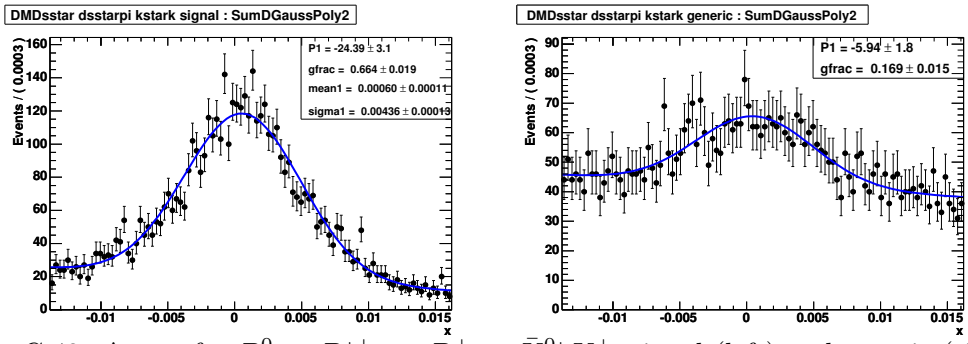


Figure C.42. $\Delta m_{D_s^*}$ for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$, signal (left) and generic (right) MC

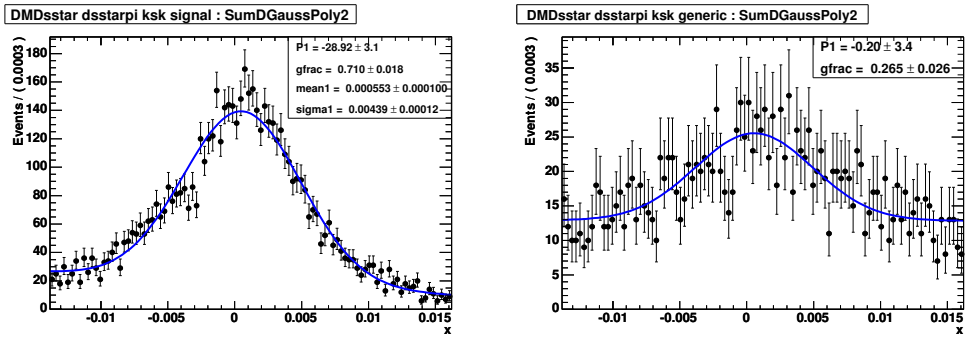


Figure C.43. $\Delta m_{D_s^*}$ for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^0K^+$, signal (left) and generic (right) MC

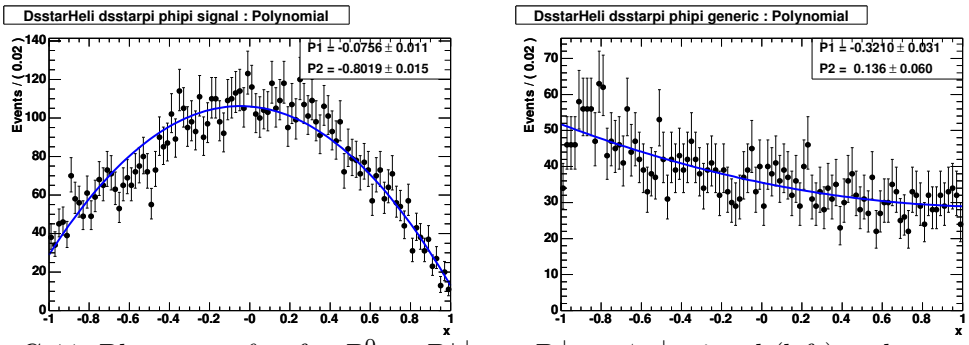


Figure C.44. Photon $\cos\theta_{\rm hel}$ for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$, signal (left) and generic (right) MC

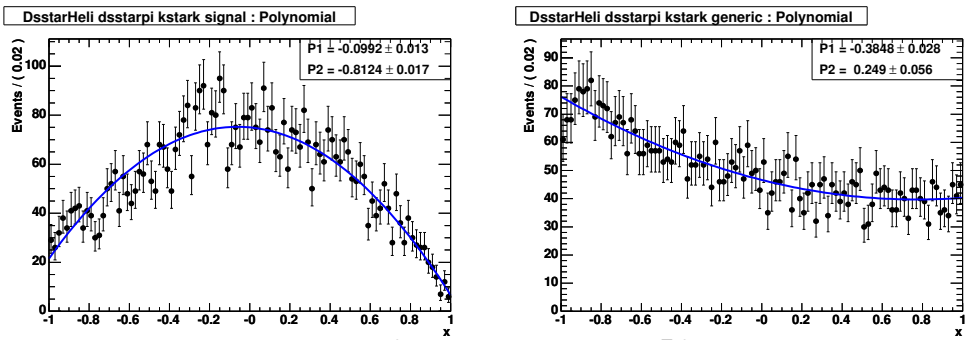


Figure C.45. Photon $\cos\theta_{\rm hel}$ for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$, signal (left) and generic (right) MC

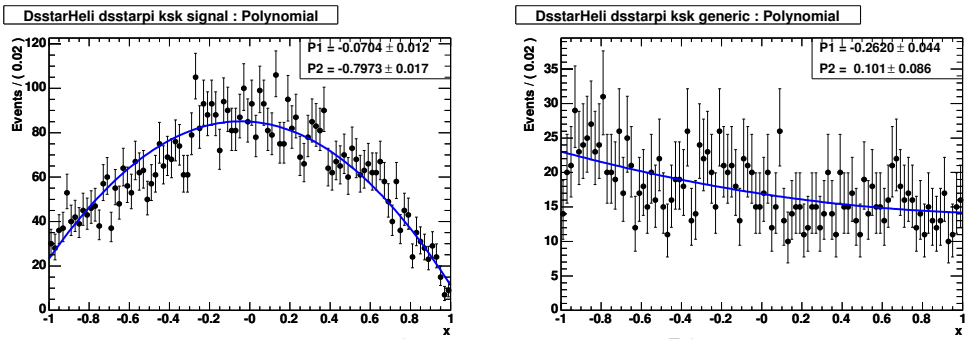


Figure C.46. Photon $\cos\theta_{\rm hel}$ for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^0K^+$, signal (left) and generic (right) MC

C.2 Likelihood Distributions

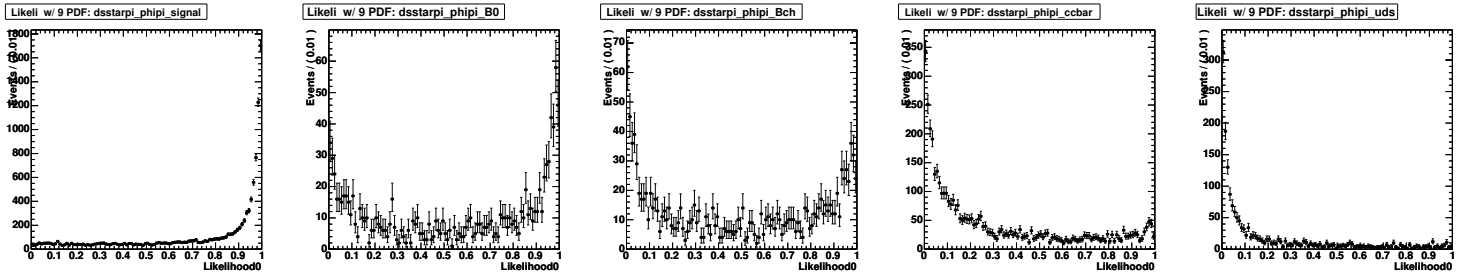


Figure C.47. Likelihood distribution for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \phi\pi^+$: signal, B^0, B^+, $c\bar{c}$, uds MC

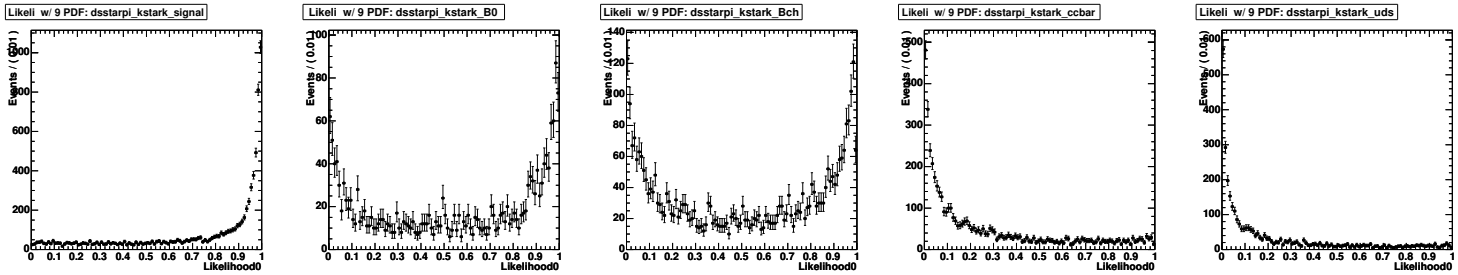


Figure C.48. Likelihood distribution for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^{0*}K^+$: signal, B^0, B^+, $c\bar{c}$, uds MC

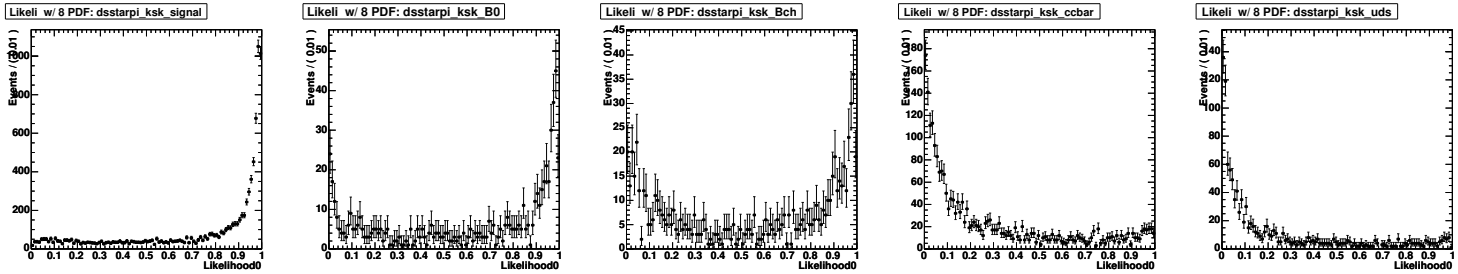


Figure C.49. Likelihood distribution for $B^0 \to D_s^{*+}\pi^-$, $D_s^+ \to \bar{K}^0K^+$: signal, B^0, B^+, $c\bar{c}$, uds MC

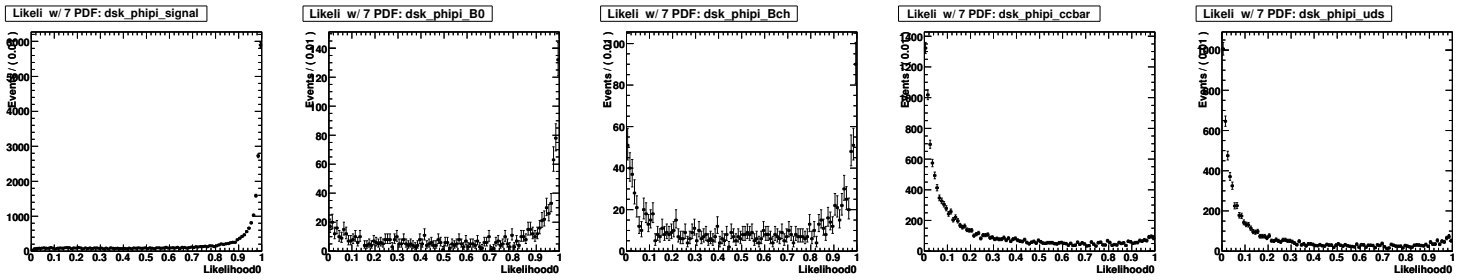


Figure C.50. Likelihood distribution for $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$: signal, B^0, B^+, $c\bar{c}$, uds MC

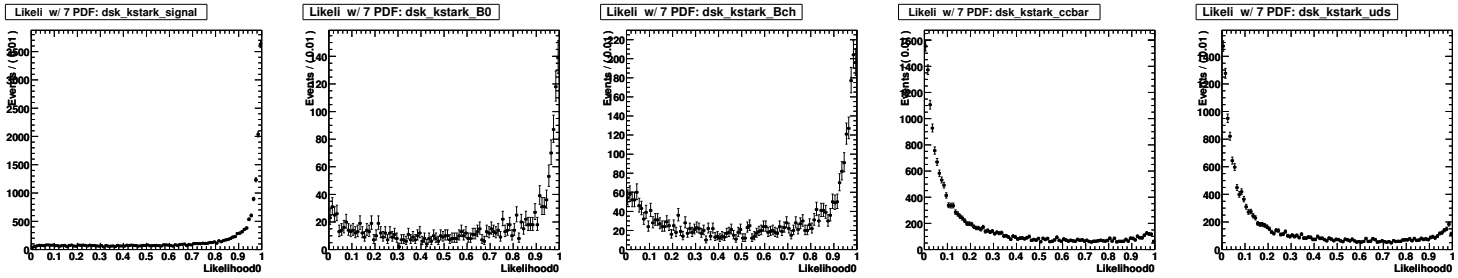


Figure C.51. Likelihood distribution for $B^0 \to D_s^- K^+$, $D_s^+ \to \bar{K}^{0*} K^+$: signal, B^0, B^+, $c\bar{c}$, uds MC

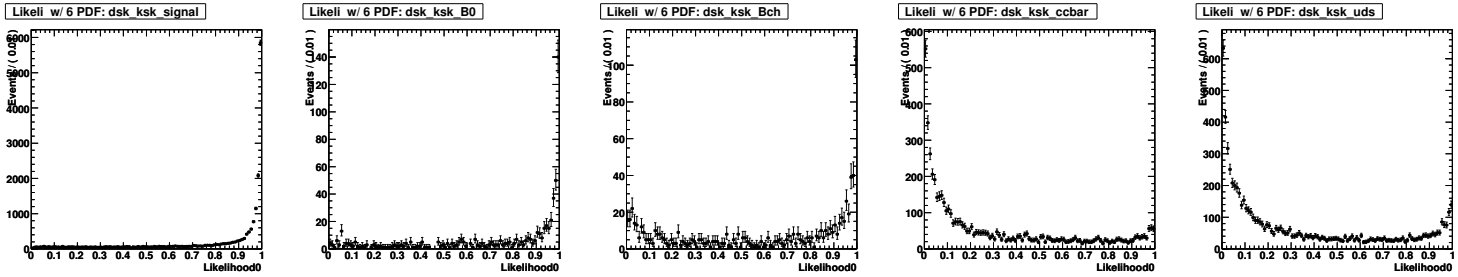


Figure C.52. Likelihood distribution for $B^0 \to D_s^- K^+$, $D_s^+ \to \bar{K}^0 K^+$: signal, B^0, B^+, $c\bar{c}$, uds MC

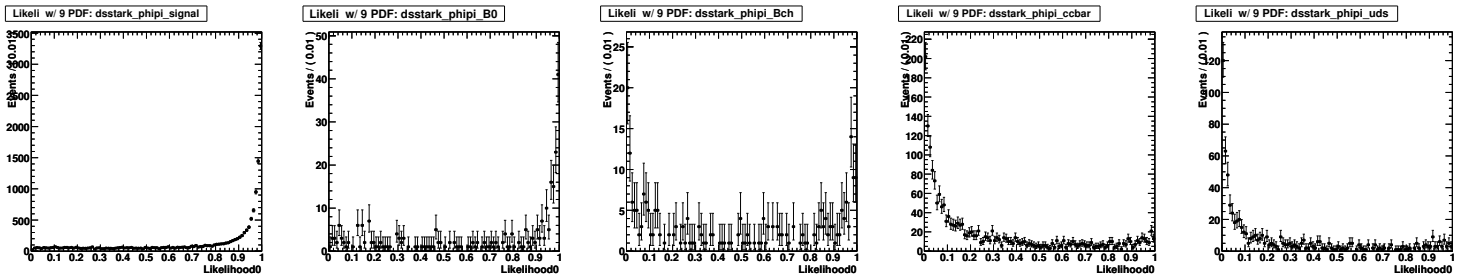


Figure C.53. Likelihood distribution for $B^0 \to D_s^{*-} K^+$, $D_s^+ \to \phi\pi^+$: signal, B^0, B^+, $c\bar{c}$, uds MC

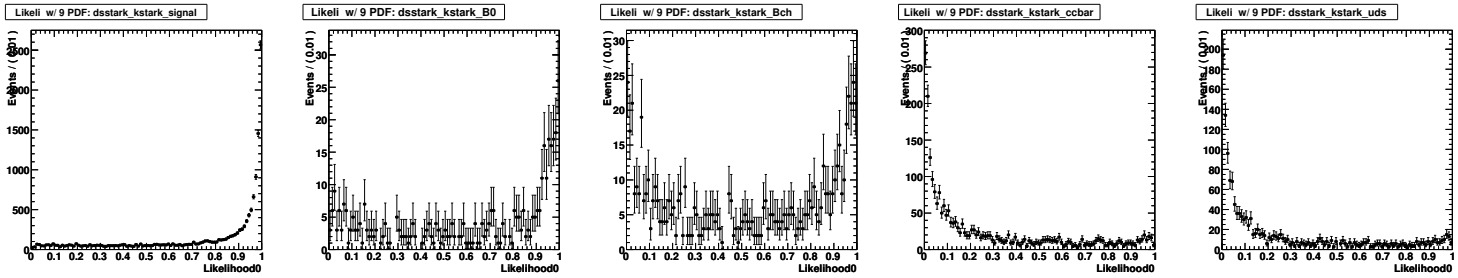


Figure C.54. Likelihood distribution for $B^0 \to D_s^{*-} K^+$, $D_s^+ \to \bar{K}^{0*} K^+$: signal, B^0, B^+, $c\bar{c}$, uds MC

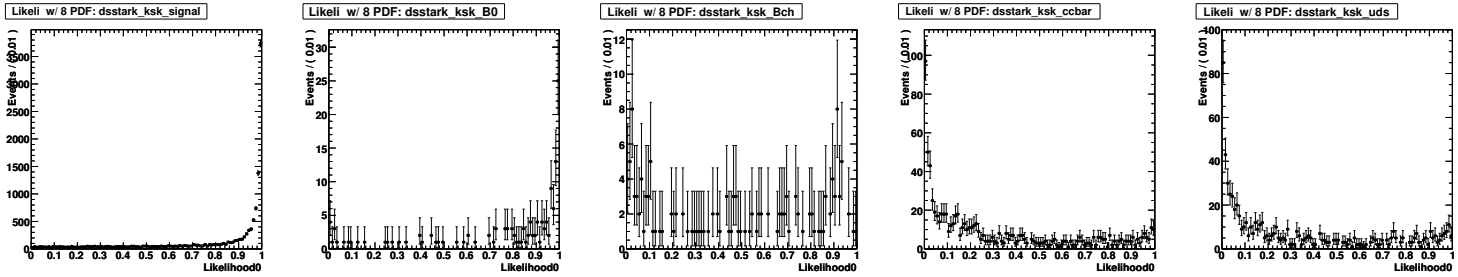


Figure C.55. Likelihood distribution for $B^0 \to D_s^{*-} K^+$, $D_s^+ \to \bar{K}^0 K^+$: signal, B^0, B^+, $c\bar{c}$, uds MC

C.3 Control Sample Checks for Likelihood Selection

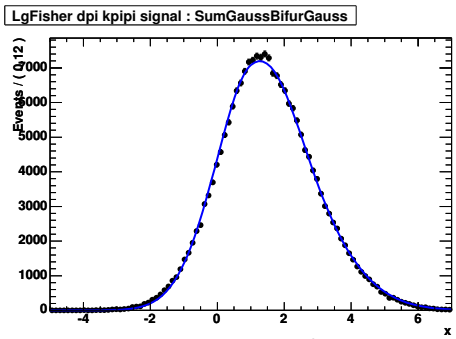


Figure C.56. Legendre Fisher discriminant for $B^0 \to D^-\pi^+$, $D^- \to K^+\pi^-\pi^-$ signal MC with $D_s\pi$ PDF overlaid.

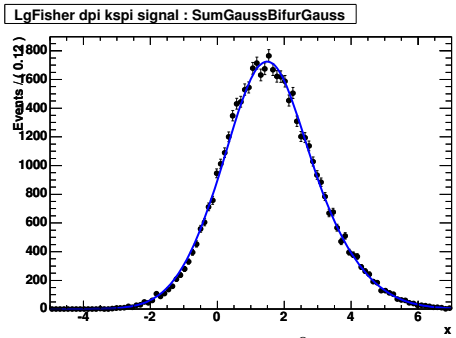


Figure C.57. Legendre Fisher discriminant for $B^0 \to D^-\pi^+$, $D^- \to \bar{K}^0\pi^-$ signal MC with $D_s\pi$ PDF overlaid.

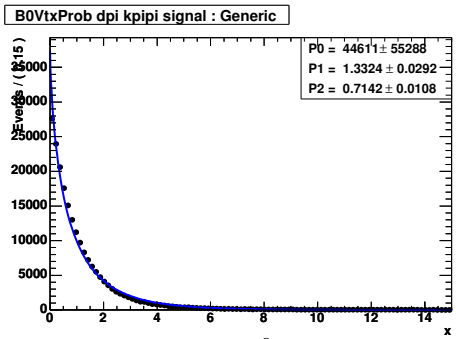


Figure C.58. -log(B^0 vertex probability) for $B^0 \to D^-\pi^+$, $D^- \to K^+\pi^-\pi^-$ signal MC with $D_s\pi$ PDF overlaid.

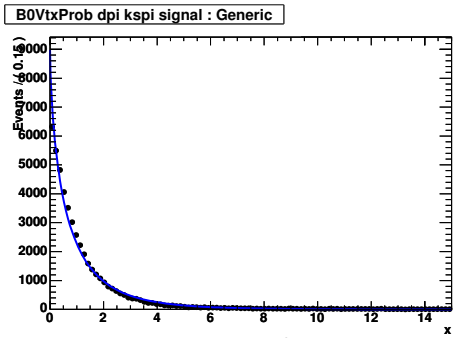


Figure C.59. -log(B^0 vertex probability) for $B^0 \to D^-\pi^+$, $D^- \to \bar{K}^0\pi^-$ signal MC with $D_s\pi$ PDF overlaid.

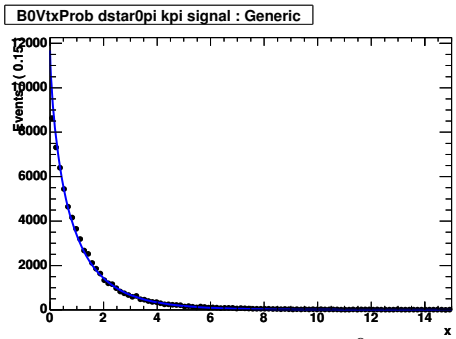


Figure C.60. -log(B^0 vertex probability) for $B^+ \to D^{*0}\pi^+$, $D^{*0} \to D^0\gamma, D^0 \to K^-\pi^+$ signal MC with $D_s\pi$ PDF overlaid.

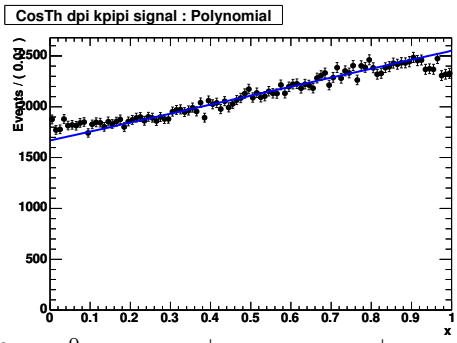


Figure C.61. $\cos\theta_{\rm thrust}$ for $B^0 \to D^-\pi^+$, $D^- \to K^+\pi^-\pi^-$ signal MC with $D_s\pi$ PDF overlaid.

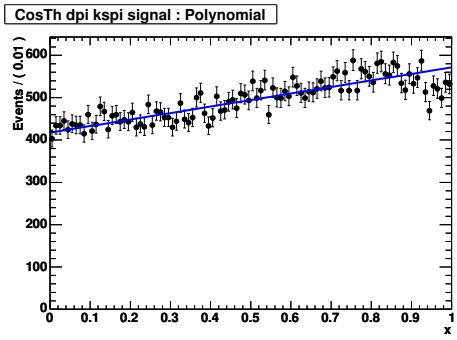


Figure C.62. $\cos\theta_{\rm thrust}$ for $B^0 \to D^-\pi^+$, $D^- \to \bar{K}^0\pi^-$ signal MC with $D_s\pi$ PDF overlaid.

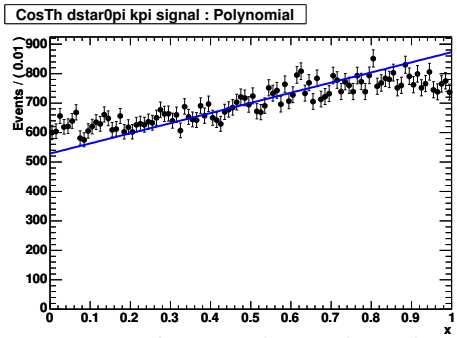


Figure C.63. $\cos\theta_{\rm thrust}$ for $B^+ \to D^{*0}\pi^+$, $D^{*0} \to D^0\gamma, D^0 \to K^-\pi^+$ signal MC with $D_s\pi$ PDF overlaid.

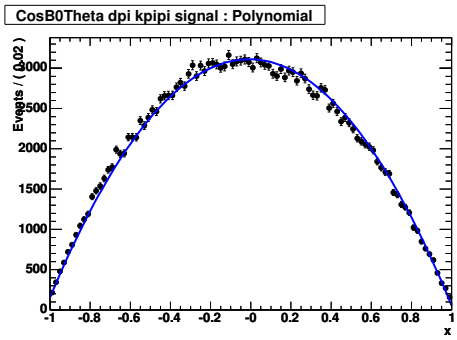


Figure C.64. $\cos\theta_{\rm B}$ for $B^0 \to D^-\pi^+$, $D^- \to K^+\pi^-\pi^-$ signal MC with $D_s\pi$ PDF overlaid.

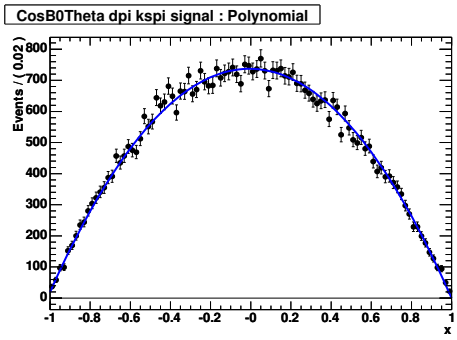


Figure C.65. $\cos\theta_{\rm B}$ for $B^0 \to D^-\pi^+$, $D^- \to \bar{K}^0\pi^-$ signal MC with $D_s\pi$ PDF overlaid.

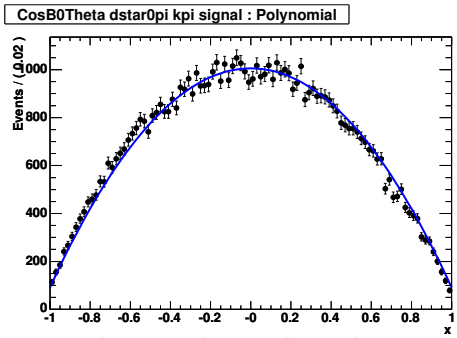


Figure C.66. $\cos\theta_{\rm B}$ for $B^+ \to D^{*0}\pi^+$, $D^{*0} \to D^0\gamma, D^0 \to K^-\pi^+$ signal MC with $D_s\pi$ PDF overlaid.

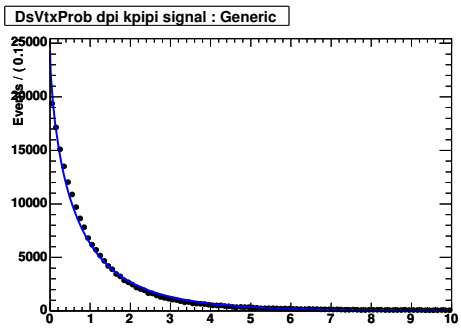


Figure C.67. -log(D_s vertex probability) for $B^0 \to D^-\pi^+$, $D^- \to K^+\pi^-\pi^-$ signal MC with $D_s\pi$ PDF overlaid.

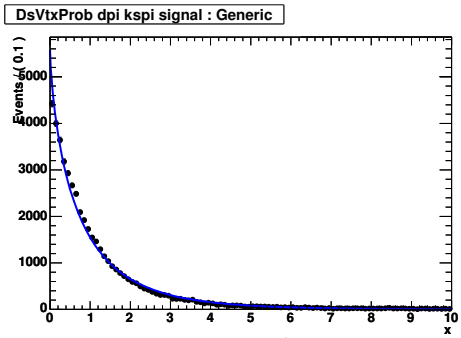


Figure C.68. -log(D_s vertex probability) for $B^0 \to D^-\pi^+$, $D^- \to \bar{K}^0\pi^-$ signal MC with $D_s\pi$ PDF overlaid.

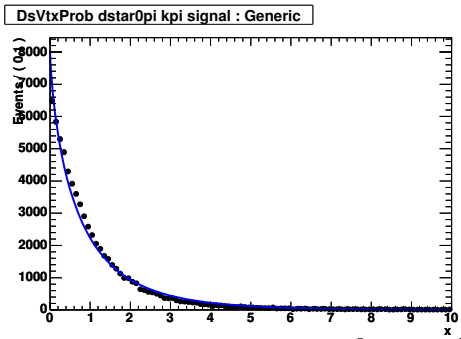


Figure C.69. -log(D_s vertex probability) for $B^+ \to D^{*0}\pi^+$, $D^{*0} \to D^0\gamma, D^0 \to K^-\pi^+$ signal MC with $D_s\pi$ PDF overlaid.

C.4 Control Sample Likelihood Distribution

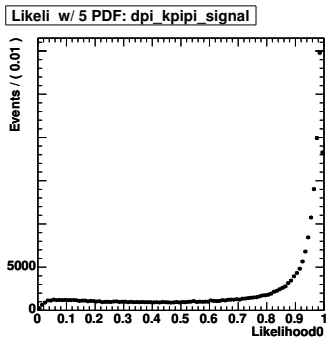


Figure C.70. Likelihood distribution for $B^0 \to D^-\pi^+$, $D^- \to K^+\pi^-\pi^-$ signal MC.

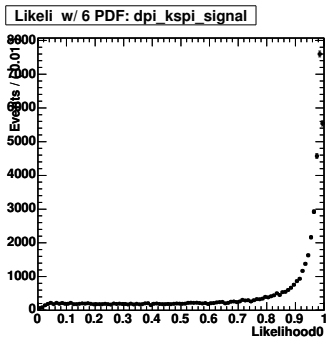


Figure C.71. Likelihood distribution for $B^0 \to D^-\pi^+$, $D^- \to \bar{K}^0\pi^-$ signal MC.

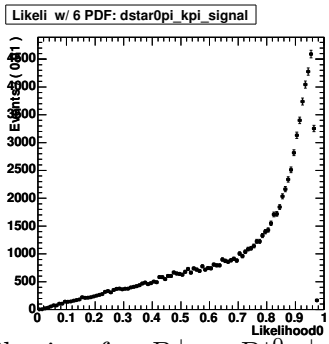


Figure C.72. Likelihood distribution for $B^+ \to D^{*0}\pi^+$, $D^{*0} \to D^0\gamma, D^0 \to K^-\pi^+$ signal MC.

Appendix D

Background Studies

This Appendix summarizes background studies performed on the various generic and exclusive Monte Carlo samples listed in Appendix A. The plots show ΔE versus m_{ES} for candidates reconstructed as signal events. In all plots the large blue box delimits the sideband region and the small red box delimits the signal region.

D.1 Generic Continuum Background Studies

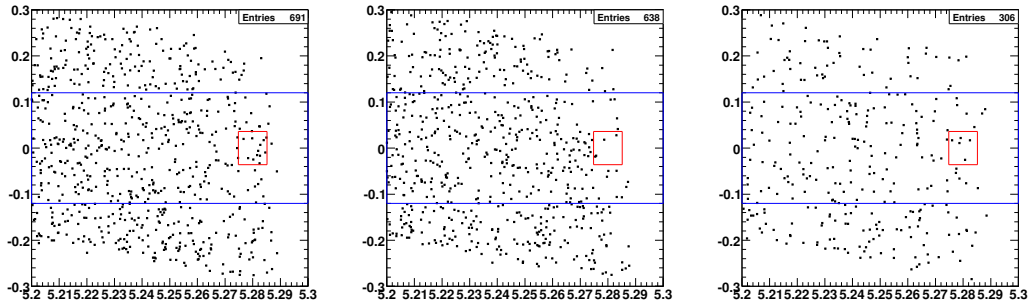


Figure D.1. Simulated $c\overline{c}$ events (L_{eq}=271 fb^{-1}), $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

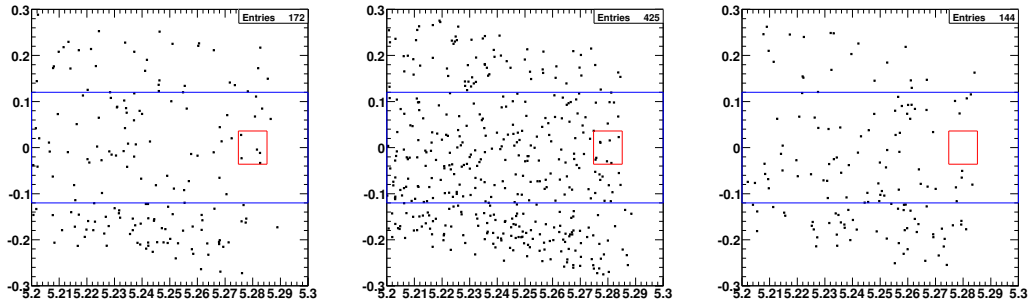


Figure D.2. Simulated $u\bar{u}$,$d\bar{d}$,$s\bar{s}$ events (L_{eq}=241 fb^{-1}), $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

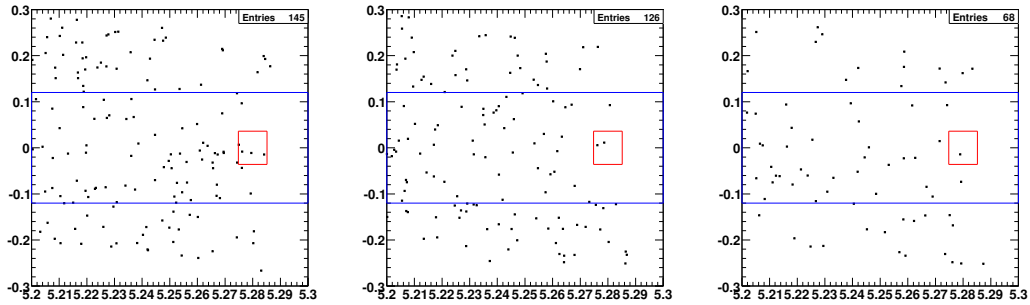


Figure D.3. Simulated $c\bar{c}$ events (L_{eq}=271 fb^{-1}), $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

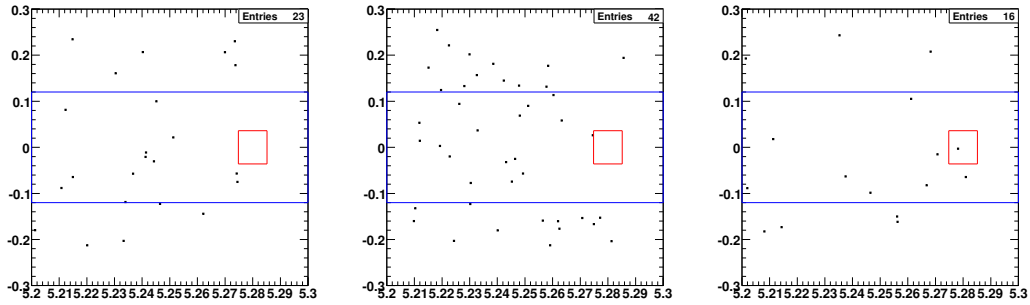


Figure D.4. Simulated $u\bar{u}$,$d\bar{d}$,$s\bar{s}$ events (L_{eq}=241 fb^{-1}), $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

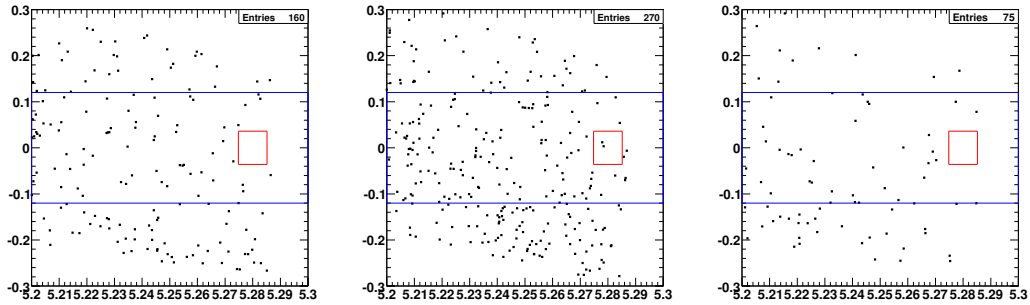


Figure D.5. Simulated $c\bar{c}$ events (L_{eq}=271 fb^{-1}), $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

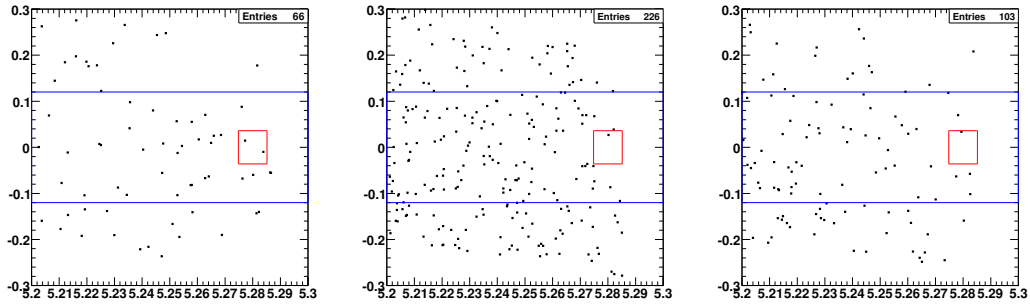


Figure D.6. Simulated $u\bar{u}$,$d\bar{d}$,$s\bar{s}$ events (L_{eq}=241 fb^{-1}), $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

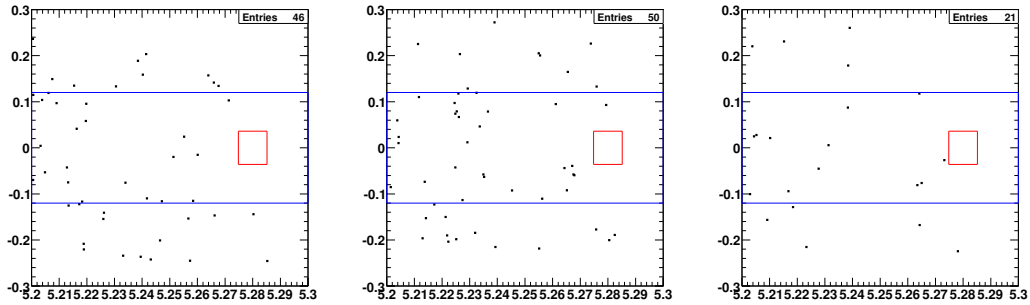


Figure D.7. Simulated $c\overline{c}$ events (L_{eq}=271 fb^{-1}), $B^0 \to D_s^{*-}K^+$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

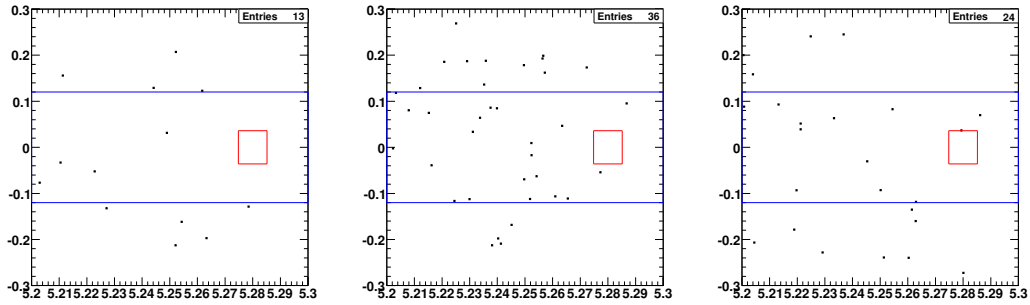


Figure D.8. Simulated $u\bar{u}$,$d\bar{d}$,$s\bar{s}$ events (L_{eq}=241 fb^{-1}), $B^0 \to D_s^{*-}K^+$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

D.2 Generic $B\bar{B}$ Background Studies

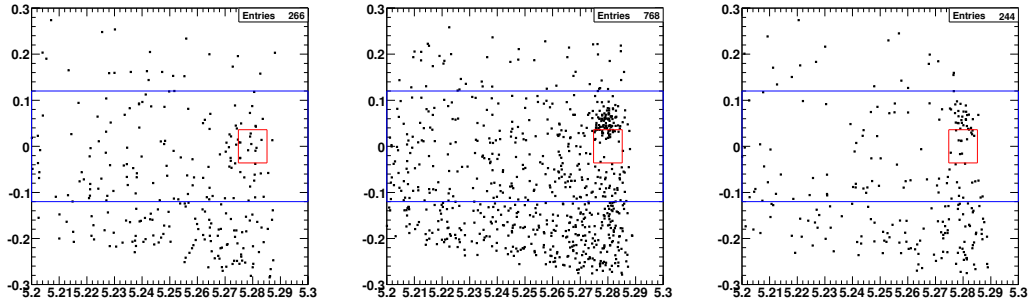


Figure D.9. Simulated $B^0\overline{B}^0$ events (L_{eq}=744 fb^{-1}), $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

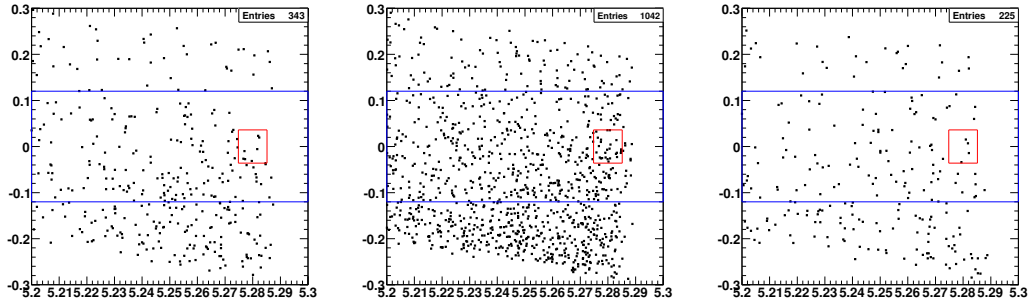


Figure D.10. Simulated B^+B^- events (L_{eq}=744 fb^{-1}), $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

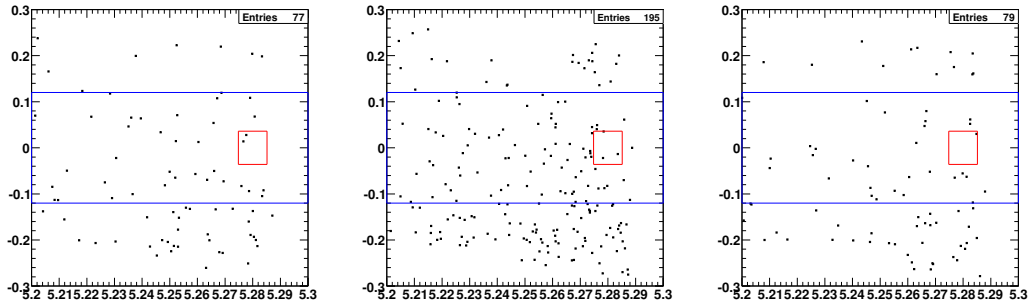


Figure D.11. Simulated $B^0\overline{B}^0$ events (L_{eq}=744 fb^{-1}), $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

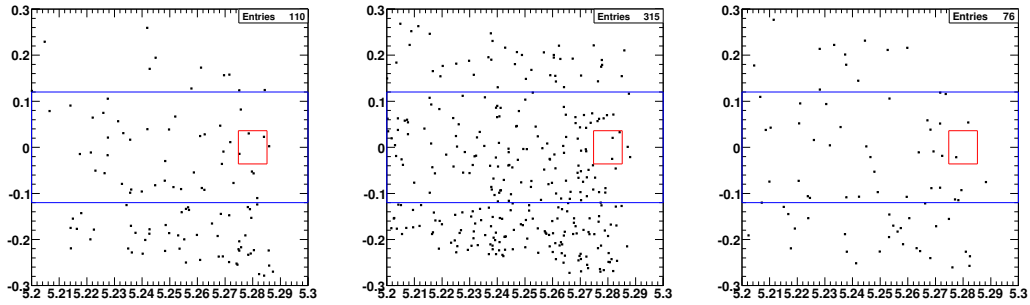


Figure D.12. Simulated B^+B^- events (L_{eq}=744 fb^{-1}), $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

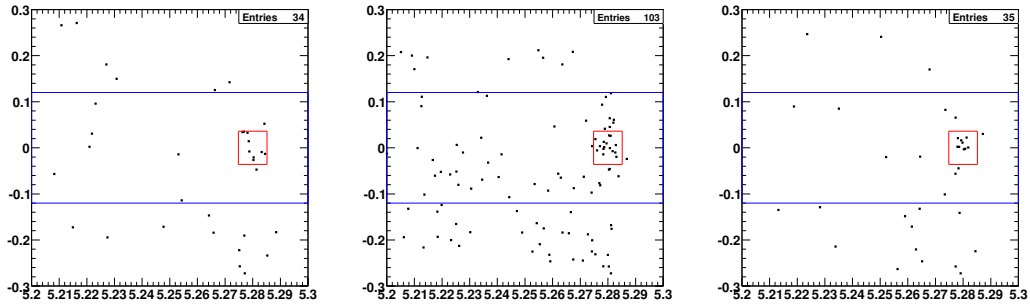


Figure D.13. Simulated $B^0\overline{B}^0$ events (L_{eq}=744 fb^{-1}), $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

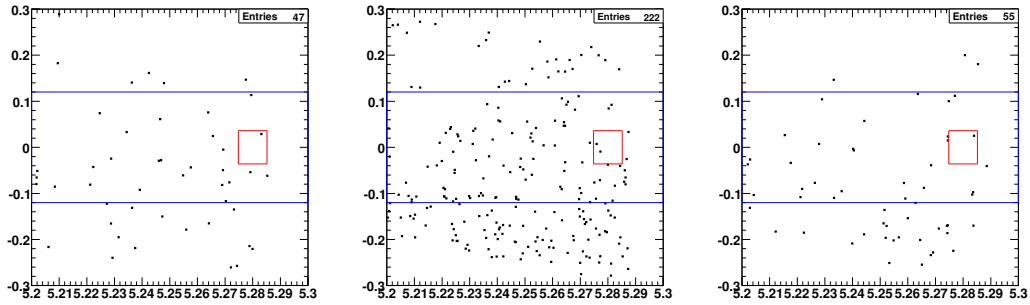


Figure D.14. Simulated B^+B^- events (L_{eq}=744 fb^{-1}), $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

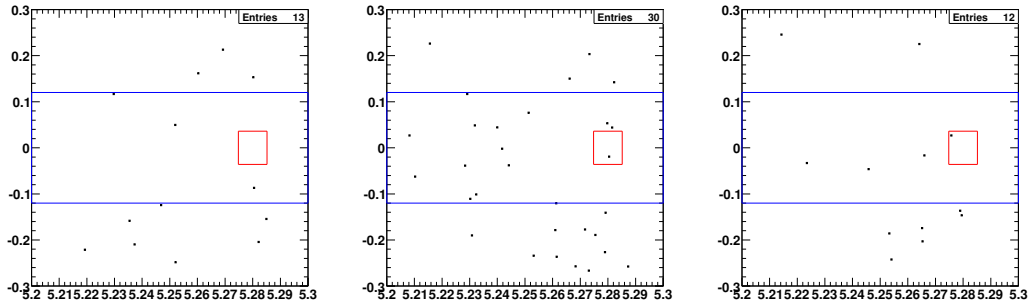


Figure D.15. Simulated $B^0\overline{B}^0$ events (L_{eq}=744 fb^{-1}), $B^0 \to D_s^{*-}K^+$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

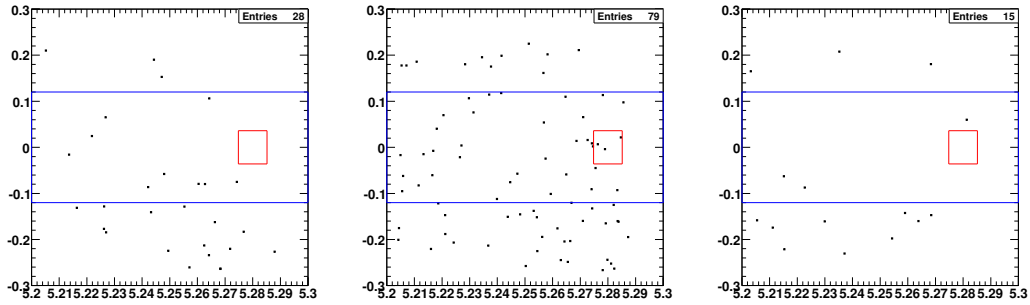


Figure D.16. Simulated B^+B^- events (L_{eq}=744 fb^{-1}), $B^0 \to D_s^{*-}K^+$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

D.3 Exclusive Monte Carlo Background Studies

D.3.1 Peaking Background Studies for $B^0 \to D_s^+\pi^-$

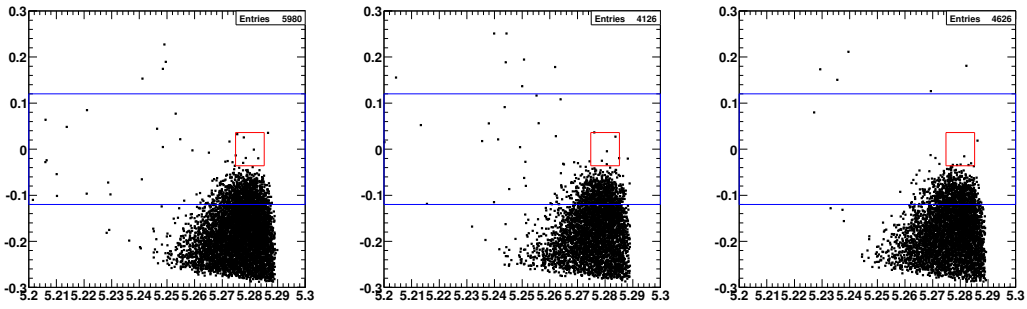


Figure D.17. Simulated $B^0 \to D_s^{*+}\pi^-$ events, $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

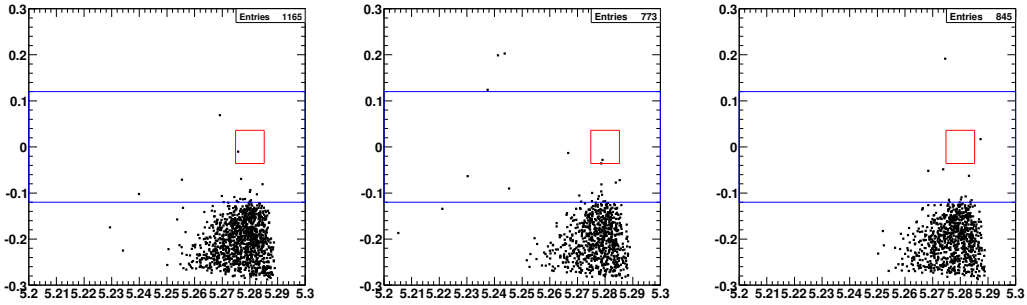


Figure D.18. Simulated $B^0 \to D_s^+\rho^-$ events, $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

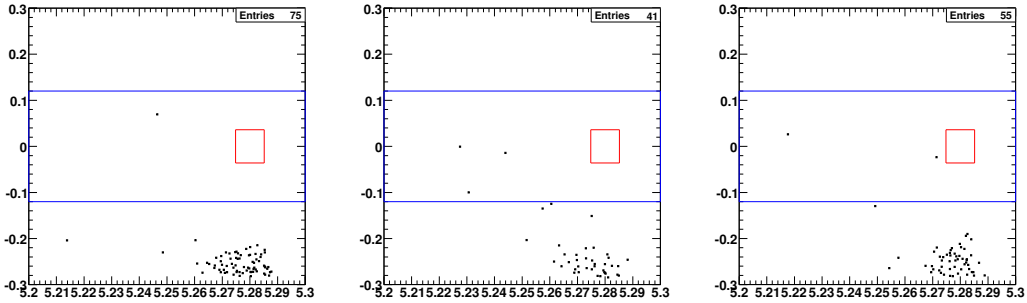


Figure D.19. Simulated $B^0 \to D_s^{*+}\rho^-$ events (longitudinal polarization), $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

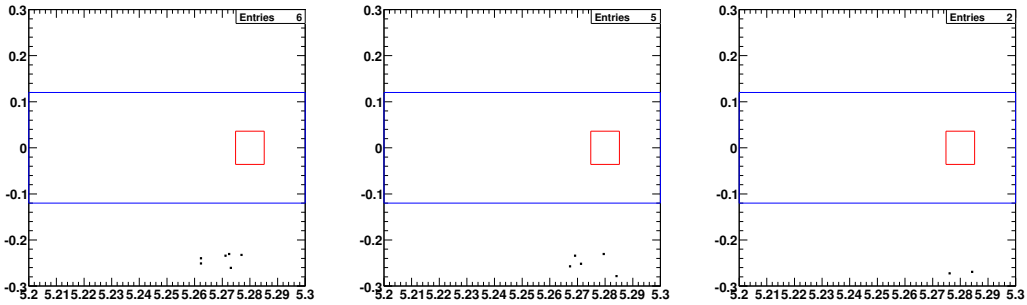


Figure D.20. Simulated $B^0 \to D_s^{*+}\rho^-$ events (transverse polarization), $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

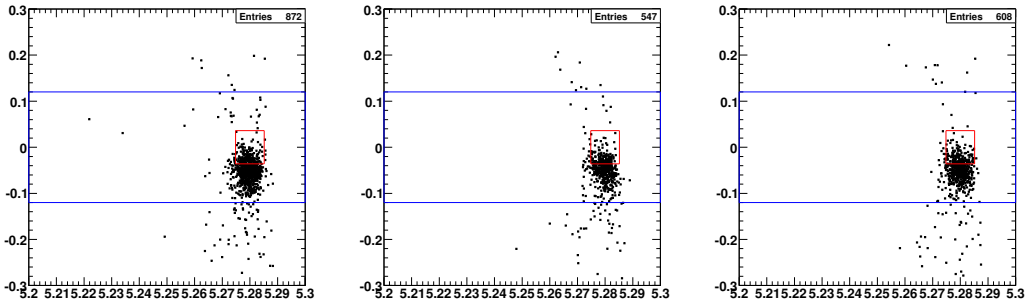


Figure D.21. Simulated $B^0 \to D_s^-K^+$ events, $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

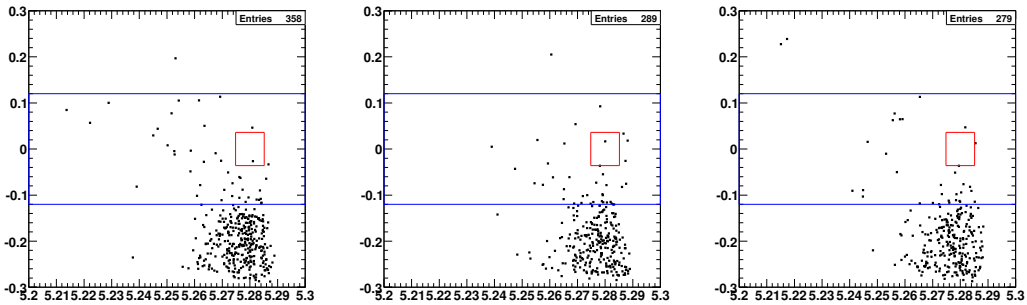


Figure D.22. Simulated $B^0 \to D_s^{*-}K^+$ events, $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

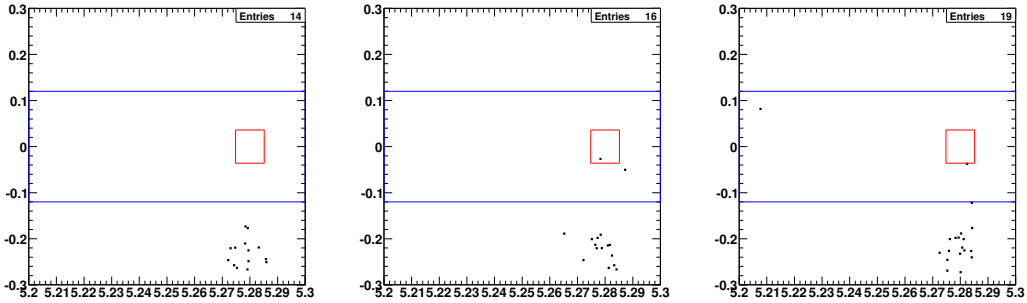


Figure D.23. Simulated $B^0 \to D_s^- K^{*+}$ events, $B^0 \to D_s^+ \pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

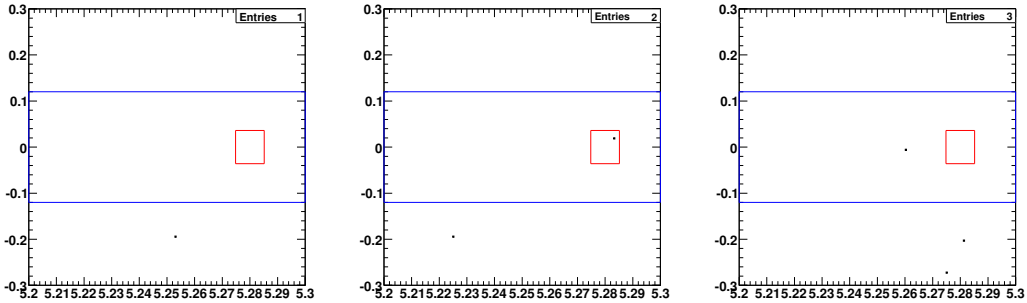


Figure D.24. Simulated $B^0 \to D_s^{*+} K^{*-}$ events (longitudinal polarization), $B^0 \to D_s^+ \pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

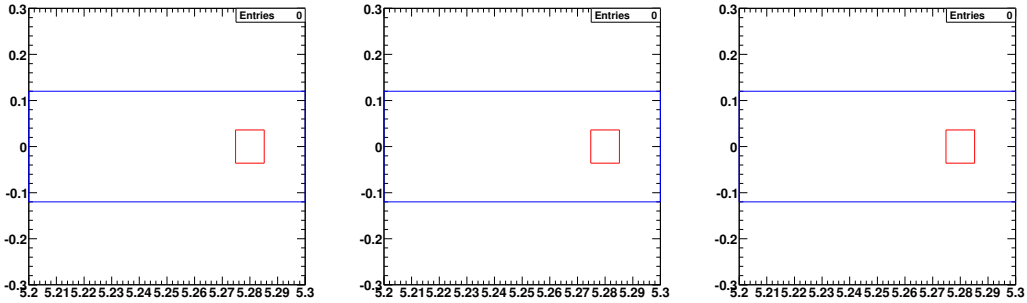


Figure D.25. Simulated $B^0 \to D_s^{*+} K^{*-}$ events (transverse polarization), $B^0 \to D_s^+ \pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

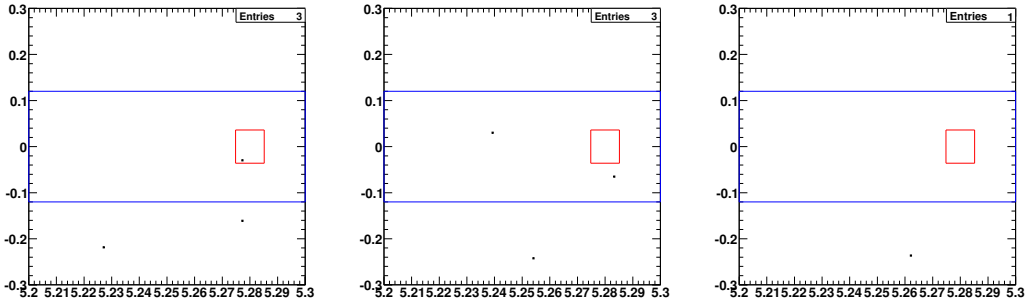


Figure D.26. Simulated $B^+ \to D_s^+\pi^0$ events, $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

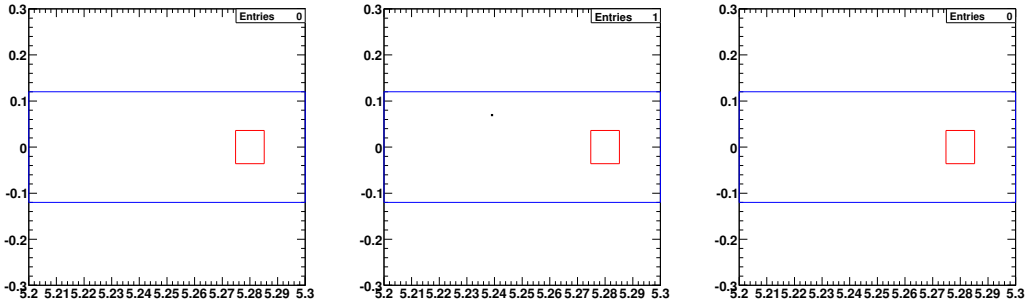


Figure D.27. Simulated $B^+ \to \bar{D}^{0*}\pi^+$ events, $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

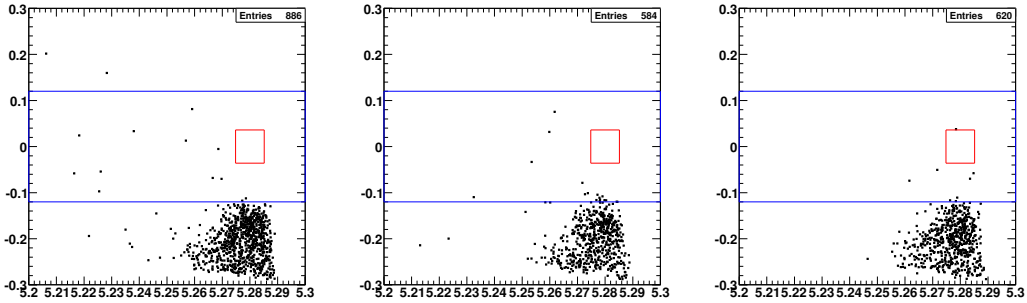


Figure D.28. Simulated $B^+ \to D_s^+\rho^0$ events, $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

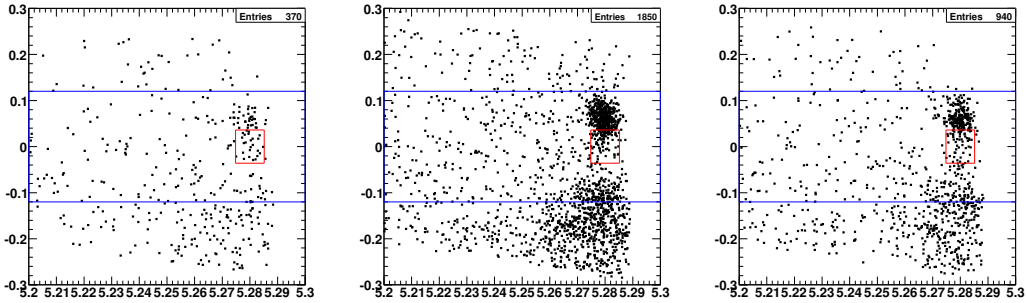


Figure D.29. Simulated $B^0 \to D^{(*)-}X$ "cocktail" events, $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

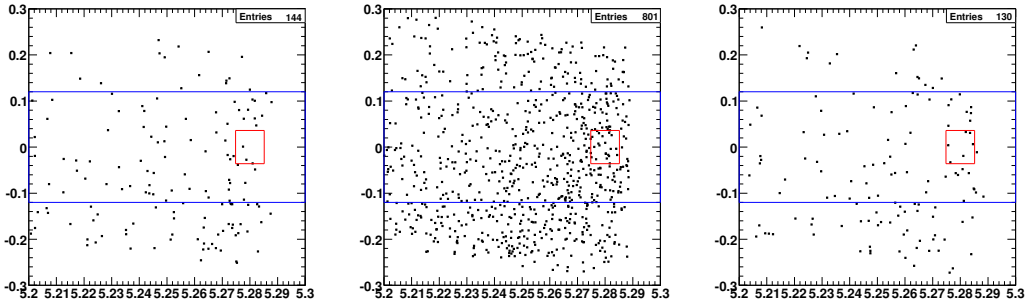


Figure D.30. Simulated $B^+ \to D^{(*)0}\pi^+$ "cocktail" events, $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

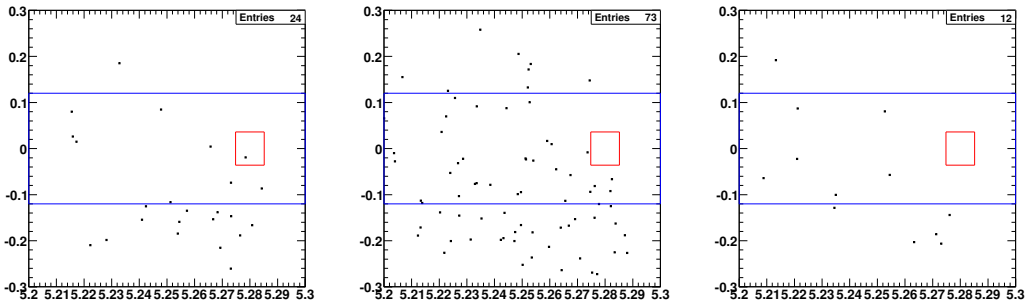


Figure D.31. Simulated $B^+ \to D^{(*)0}X$ "cocktail" events, $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

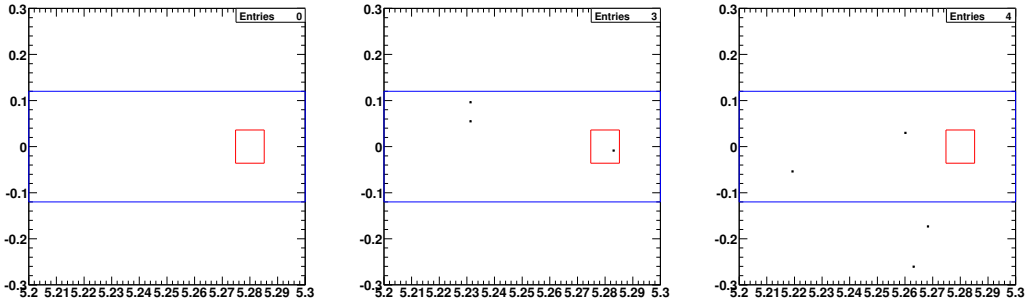


Figure D.32. Simulated $B^0 \to D^- K^+$, $D^- \to K^+\pi^-\pi^-$ events, $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

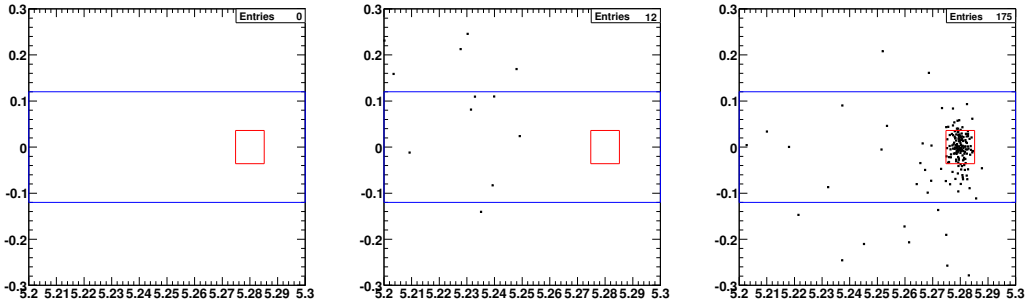


Figure D.33. Simulated $B^0 \to D^- K^+$, $D^- \to K^0\pi^-$ events, $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

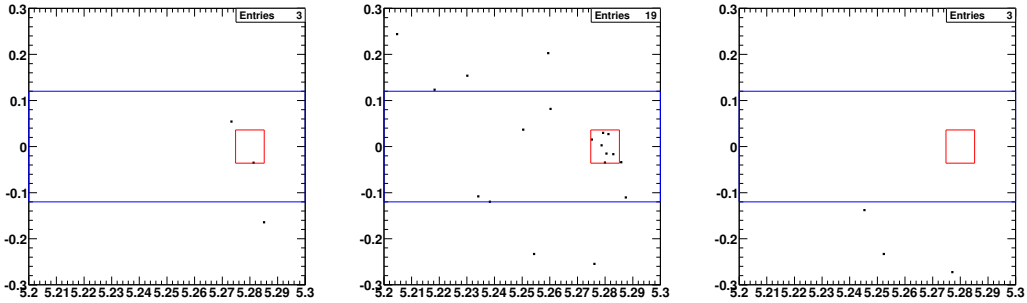


Figure D.34. Simulated $B^0 \to D^- K^+$, $D^- \to K^{*0}\pi^-$ events, $B^0 \to D_s^+\pi^-$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

D.3.2 Peaking Background Studies for $B^0 \to D_s^{*+}\pi^-$

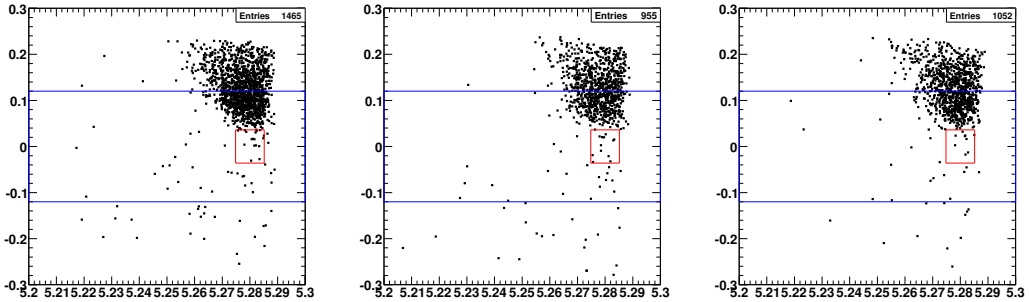


Figure D.35. Simulated $B^0 \to D_s^+\pi^-$ events, $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

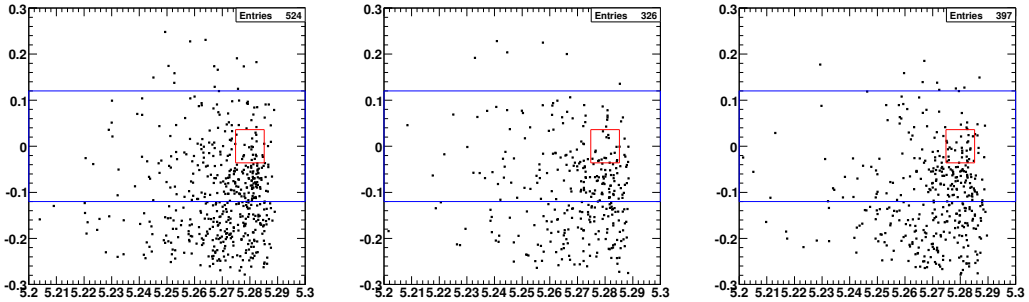


Figure D.36. Simulated $B^0 \to D_s^+\rho^-$ events, $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

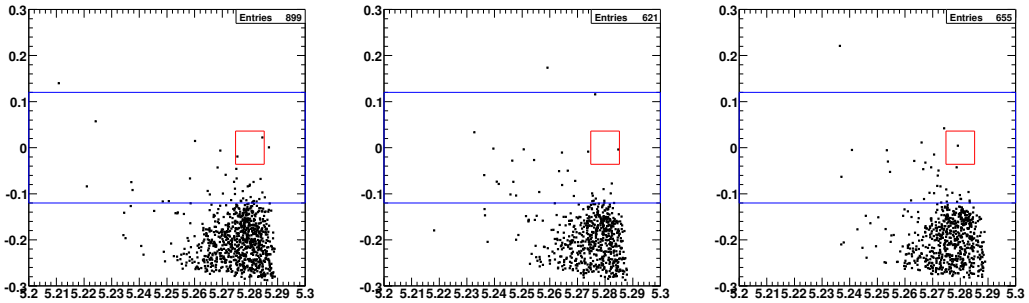


Figure D.37. Simulated $B^0 \to D_s^{*+}\rho^-$ events (longitudinal polarization), $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

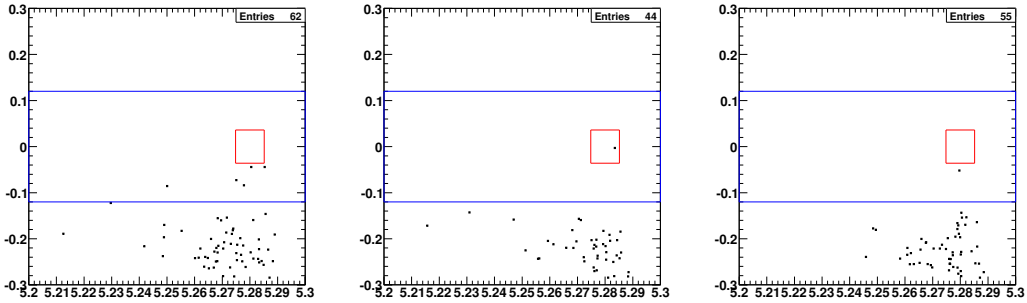


Figure D.38. Simulated $B^0 \to D_s^{*+}\rho^-$ events (transverse polarization), $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

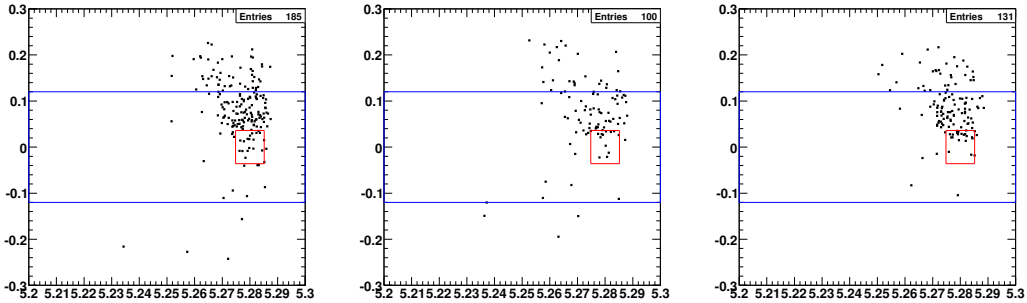


Figure D.39. Simulated $B^0 \to D_s^-K^+$ events, $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

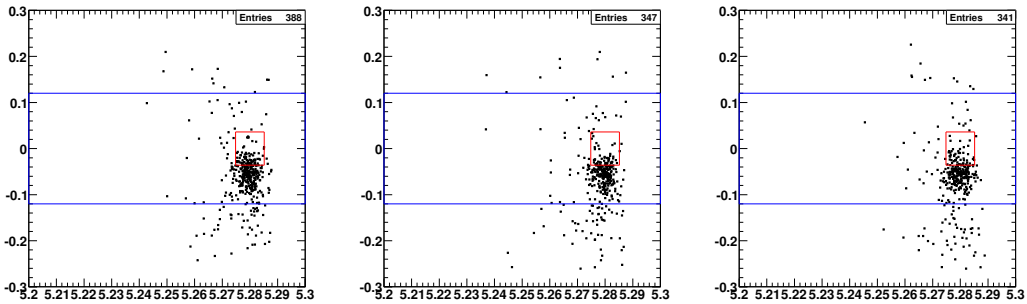


Figure D.40. Simulated $B^0 \to D_s^{*-}K^+$ events, $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

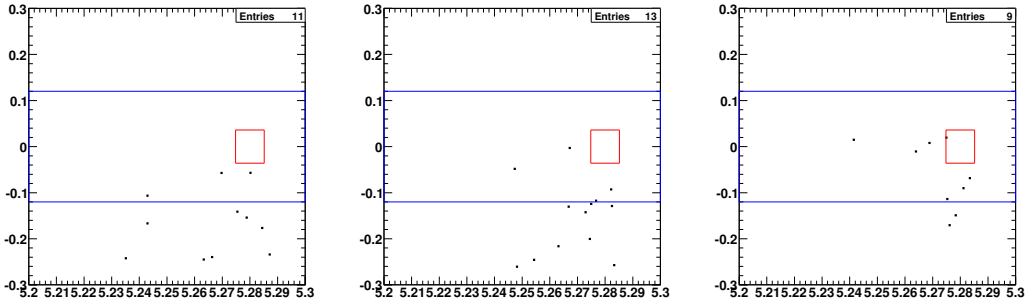


Figure D.41. Simulated $B^0 \to D_s^- K^{*+}$ events, $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

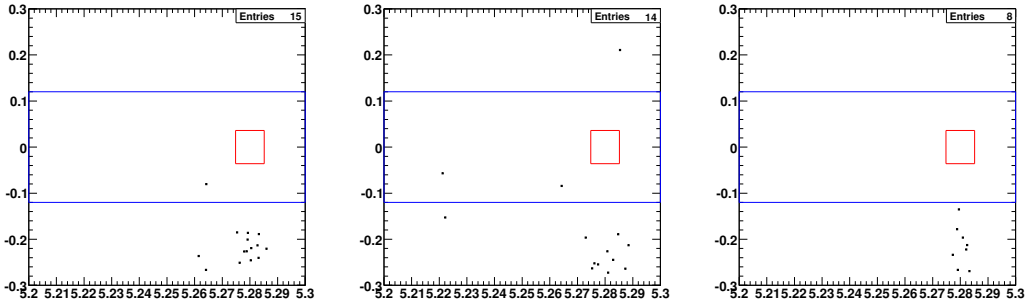


Figure D.42. Simulated $B^0 \to D_s^{*+}K^{*-}$ events (longitudinal polarization), $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

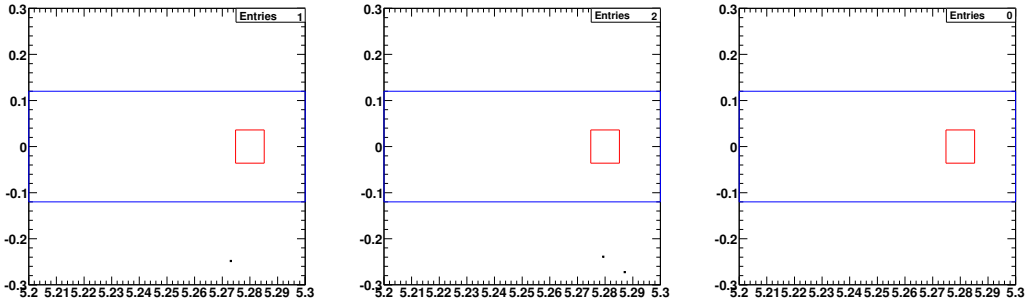


Figure D.43. Simulated $B^0 \to D_s^{*+}K^{*-}$ events (transverse polarization), $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

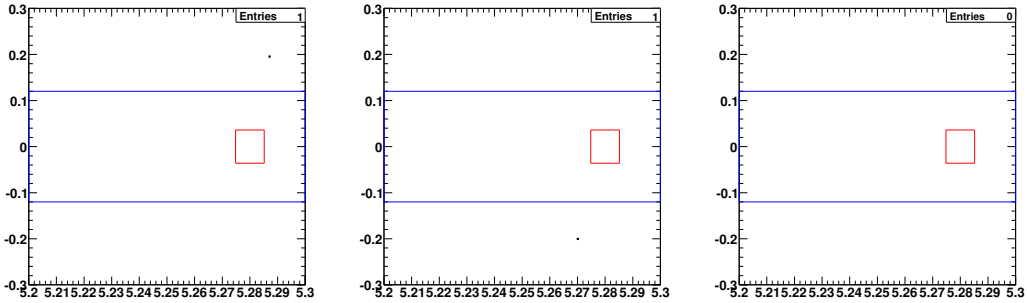


Figure D.44. Simulated $B^+ \to D_s^+\pi^0$ events, $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

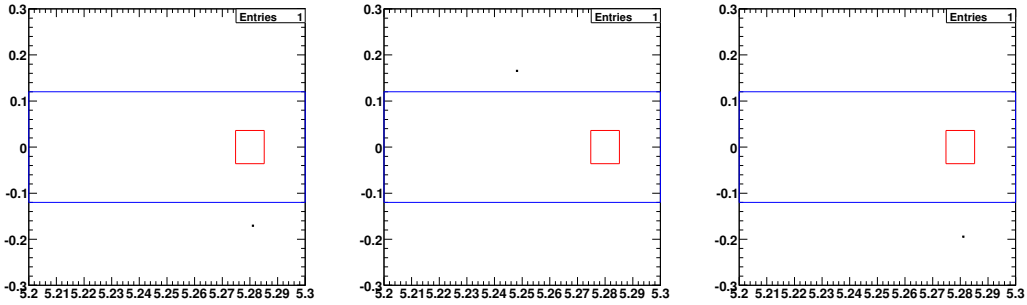


Figure D.45. Simulated $B^+ \to \bar{D}^{0*}\pi^+$ events, $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

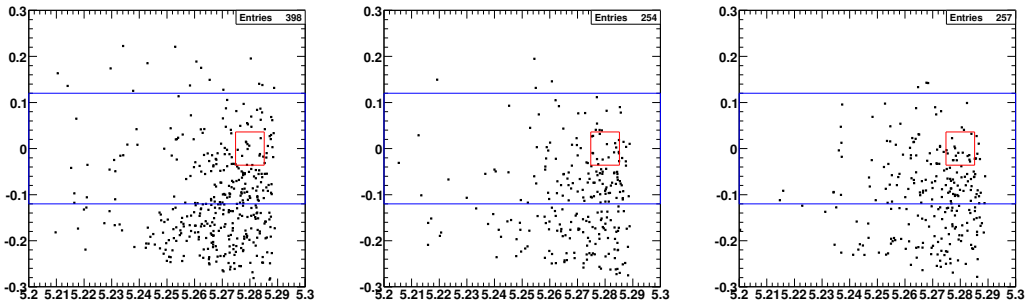


Figure D.46. Simulated $B^+ \to D_s^+\rho^0$ events, $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

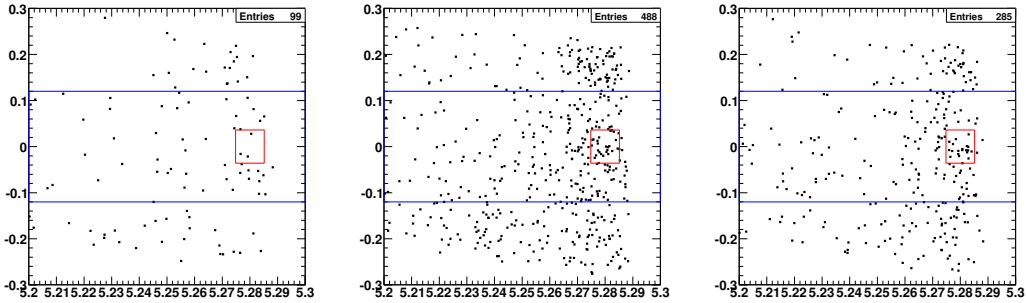


Figure D.47. Simulated $B^0 \to D^{(*)-}X$ "cocktail" events, $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

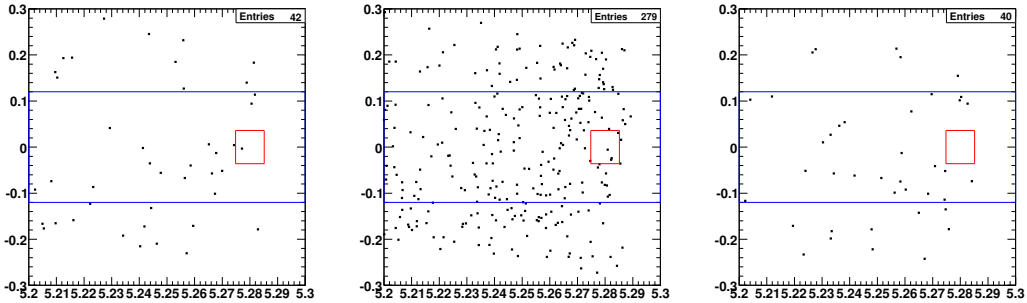


Figure D.48. Simulated $B^+ \to D^{(*)0}\pi^+$ "cocktail" events, $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

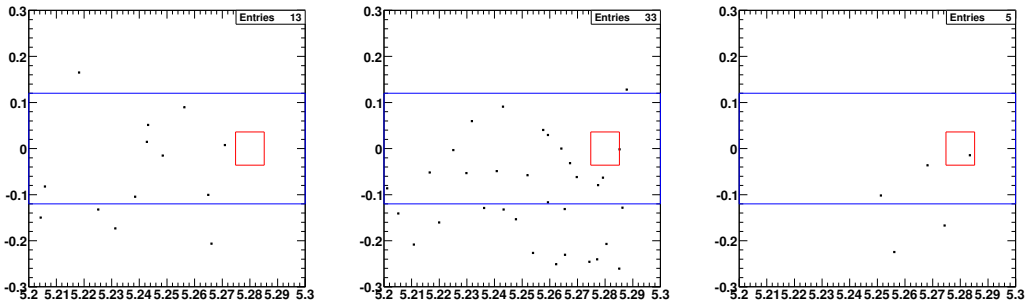


Figure D.49. Simulated $B^+ \to D^{(*)0}X$ "cocktail" events, $B^0 \to D_s^{*+}\pi^-$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

D.3.3 Peaking Background Studies for $B^0 \to D_s^- K^+$

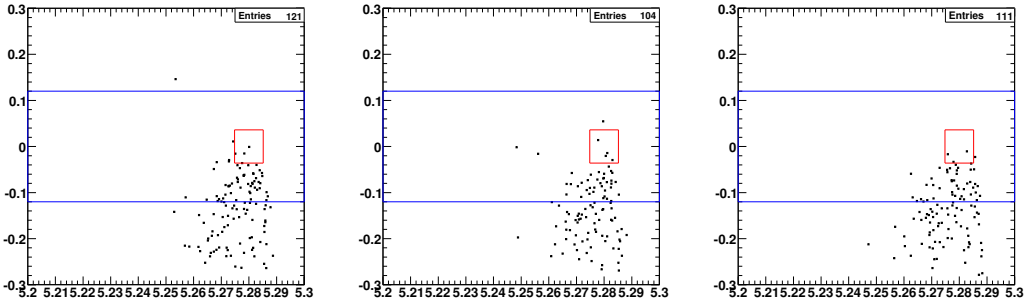


Figure D.50. Simulated $B^0 \to D_s^{*+}\pi^-$ events, $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

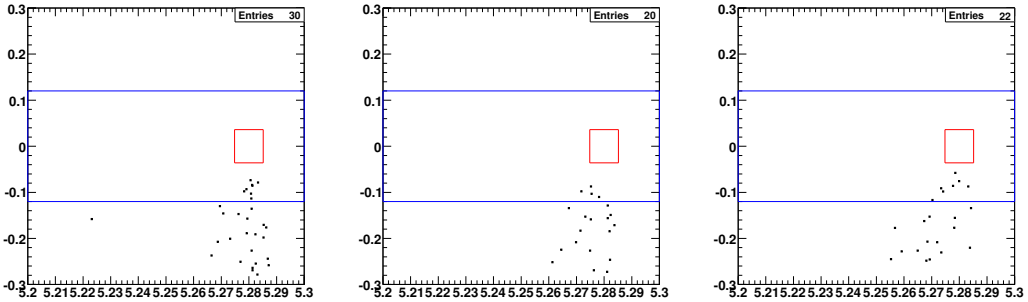


Figure D.51. Simulated $B^0 \to D_s^+\rho^-$ events, $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

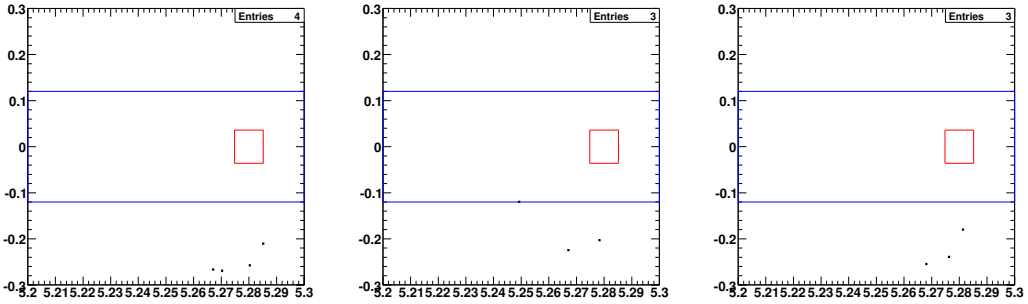


Figure D.52. Simulated $B^0 \to D_s^{*+}\rho^-$ events (longitudinal polarization), $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

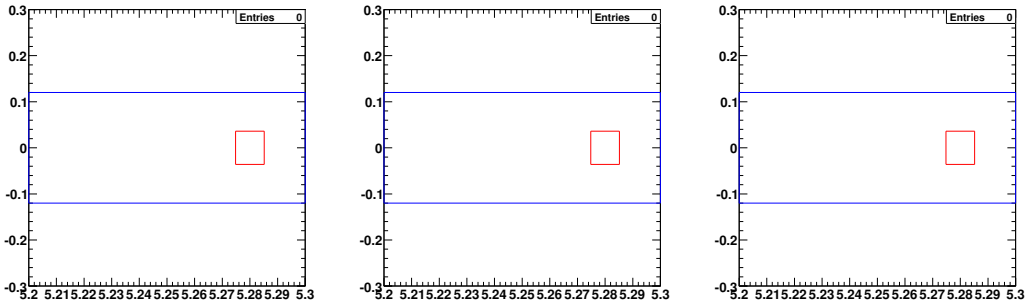


Figure D.53. Simulated $B^0 \to D_s^{*+}\rho^-$ events (transverse polarization), $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

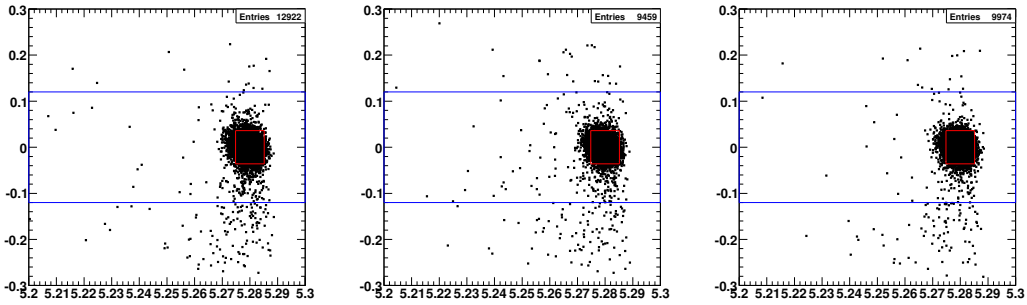


Figure D.54. Simulated $B^0 \to D_s^- K^+$ events, $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

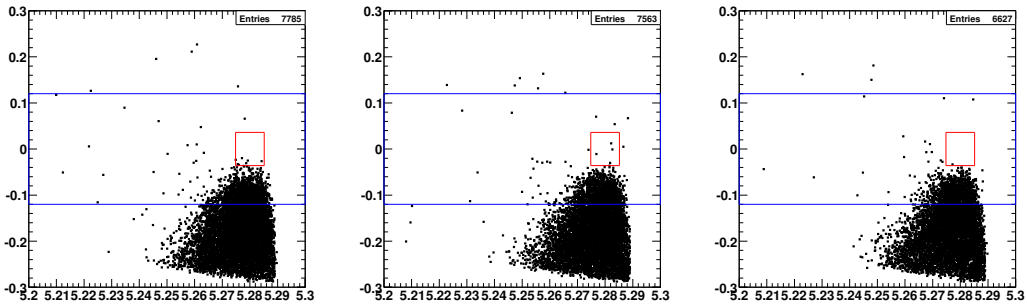


Figure D.55. Simulated $B^0 \to D_s^{*-} K^+$ events, $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

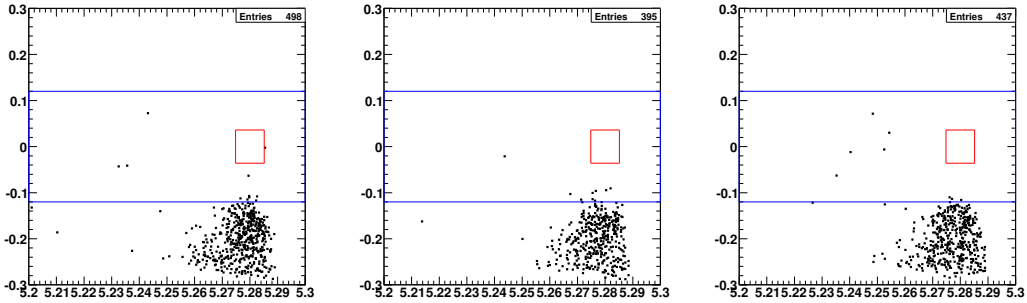


Figure D.56. Simulated $B^0 \to D_s^- K^{*+}$ events, $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

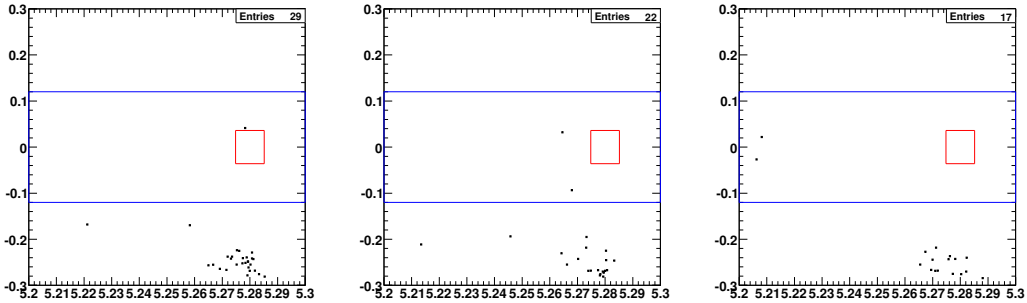


Figure D.57. Simulated $B^0 \to D_s^{*+} K^{*-}$ events (longitudinal polarization), $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

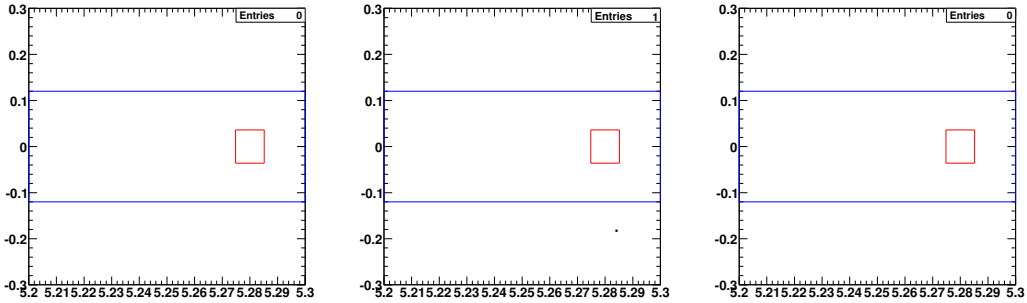


Figure D.58. Simulated $B^0 \to D_s^{*+} K^{*-}$ events (transverse polarization), $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

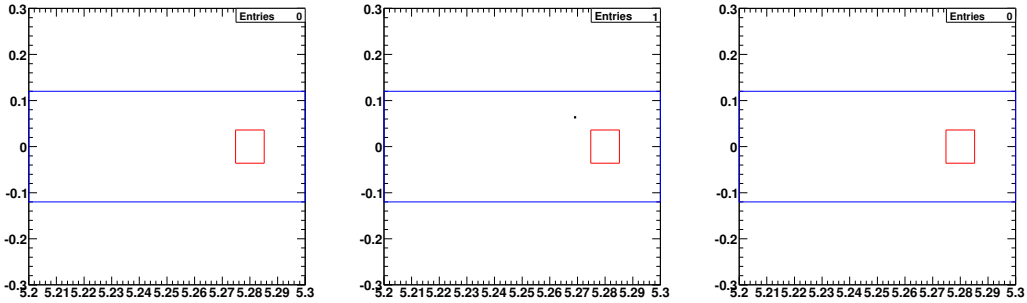


Figure D.59. Simulated $B^+ \to D_s^+ \pi^0$ events, $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

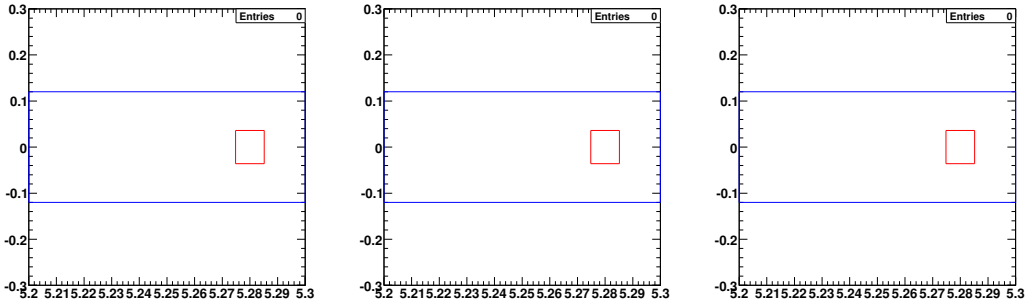


Figure D.60. Simulated $B^+ \to \bar{D}^{0*} \pi^+$ events, $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

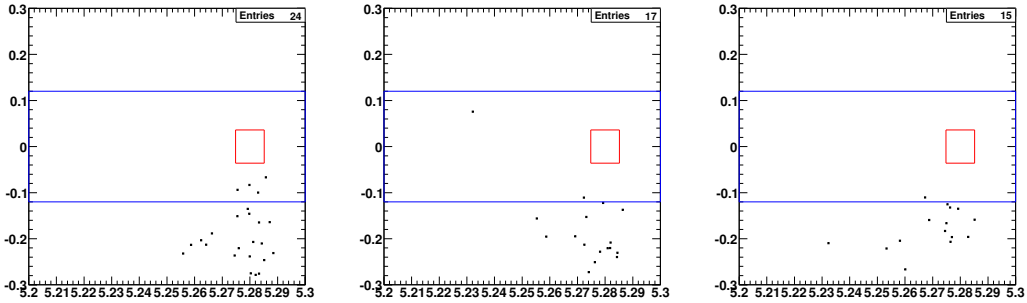


Figure D.61. Simulated $B^+ \to D_s^+ \rho^0$ events, $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

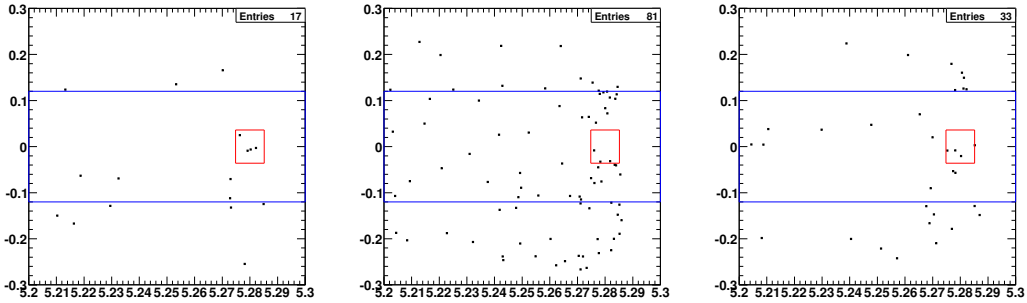


Figure D.62. Simulated $B^0 \to D^{(*)-}X$ "cocktail" events, $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

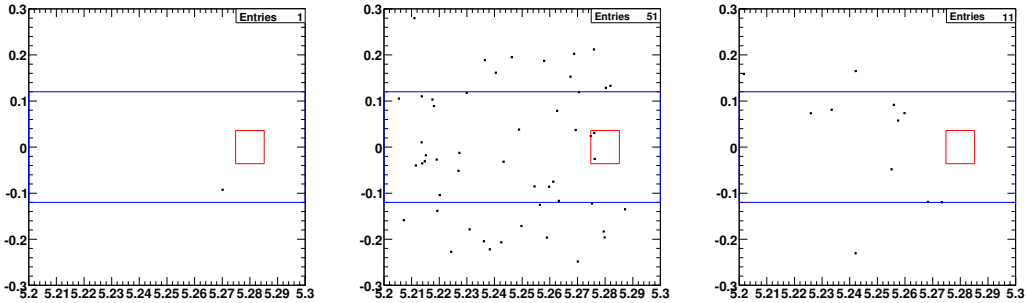


Figure D.63. Simulated $B^+ \to D^{(*)0}\pi^+$ "cocktail" events, $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

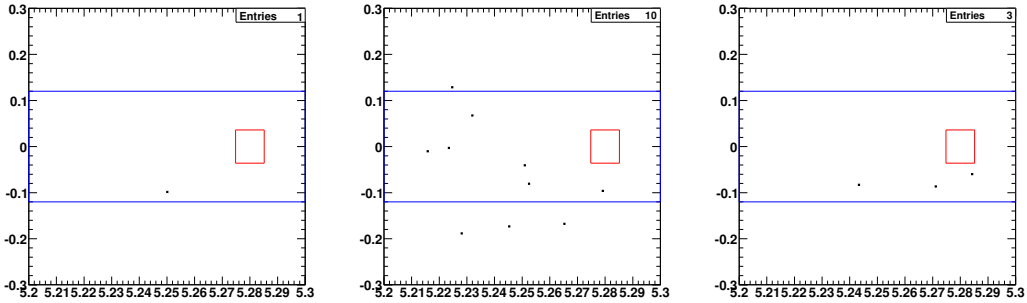


Figure D.64. Simulated $B^+ \to D^{(*)0}X$ "cocktail" events, $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

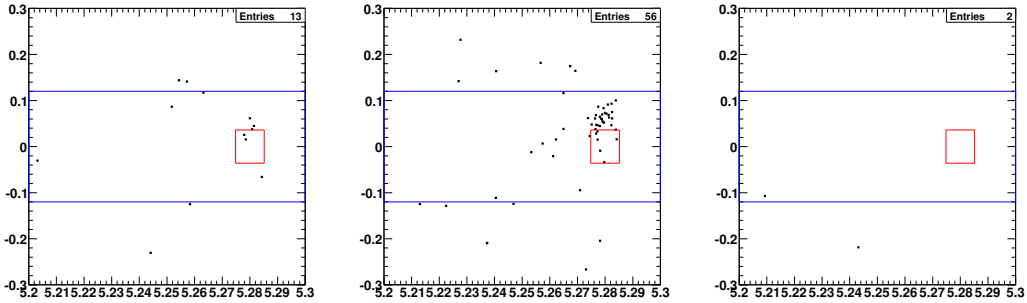


Figure D.65. Simulated $B^0 \to D^- K^+$, $D^- \to K^+\pi^-\pi^-$ events, $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

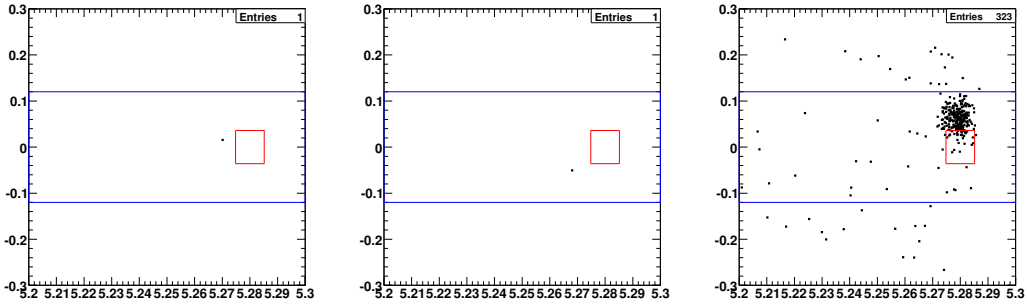


Figure D.66. Simulated $B^0 \to D^- K^+$, $D^- \to K^0\pi^-$ events, $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

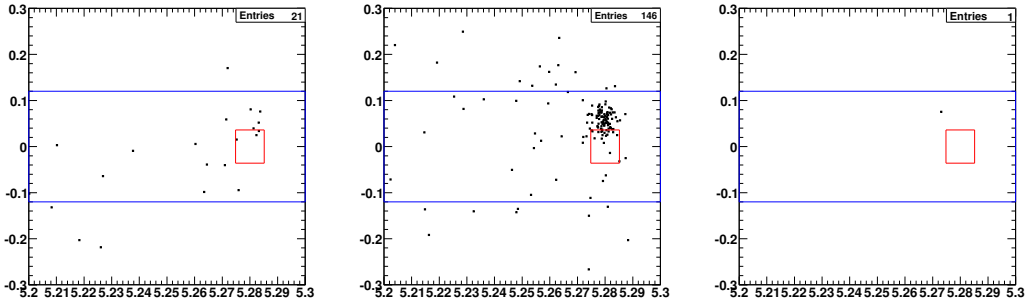


Figure D.67. Simulated $B^0 \to D^- K^+$, $D^- \to K^{*0}\pi^-$ events, $B^0 \to D_s^- K^+$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

D.3.4 Peaking Background Studies for $B^0 \to D_s^{*-} K^+$

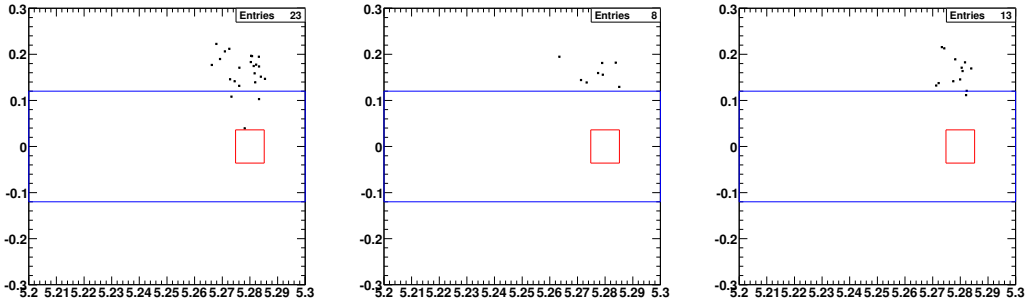


Figure D.68. Simulated $B^0 \to D_s^+ \pi^-$ events, $B^0 \to D_s^{*-} K^+$, $D_s^{*+} \to D_s^+ \gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

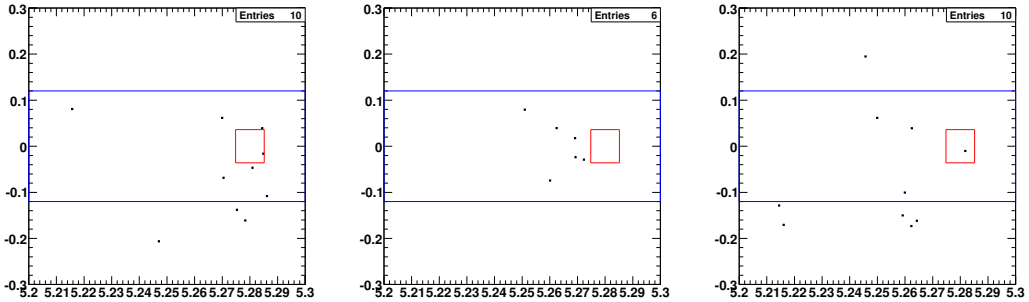


Figure D.69. Simulated $B^0 \to D_s^+ \rho^-$ events, $B^0 \to D_s^{*-} K^+$, $D_s^{*+} \to D_s^+ \gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

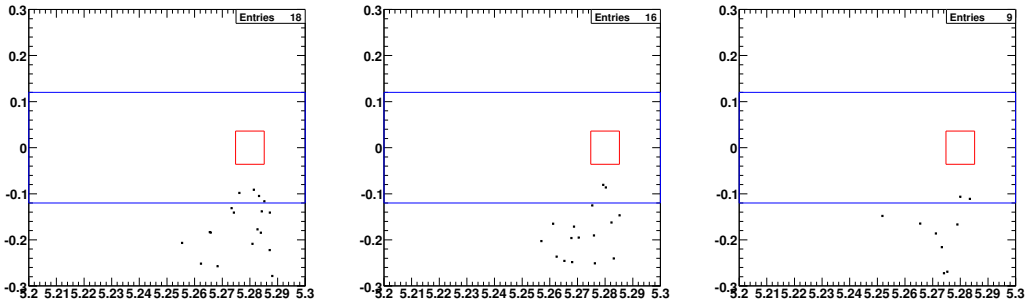


Figure D.70. Simulated $B^0 \to D_s^{*+} \rho^-$ events (longitudinal polarization), $B^0 \to D_s^{*-} K^+$, $D_s^{*+} \to D_s^+ \gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

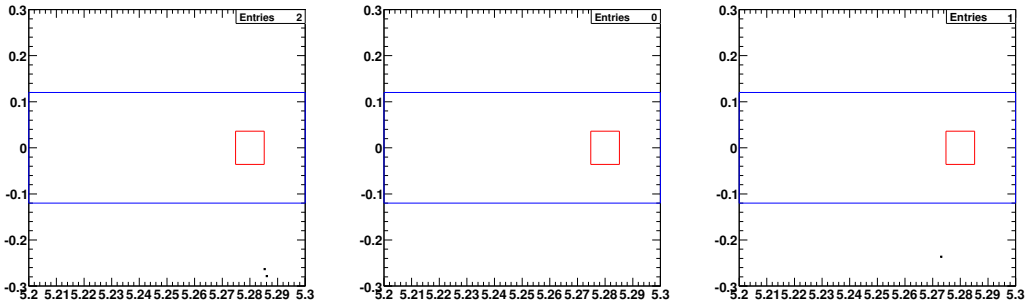


Figure D.71. Simulated $B^0 \to D_s^{*+}\rho^-$ events (transverse polarization), $B^0 \to D_s^{*-}K^+$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

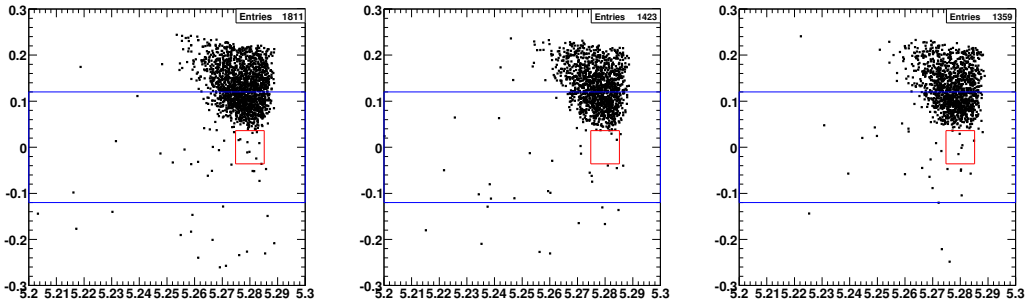


Figure D.72. Simulated $B^0 \to D_s^-K^+$ events, $B^0 \to D_s^{*-}K^+$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

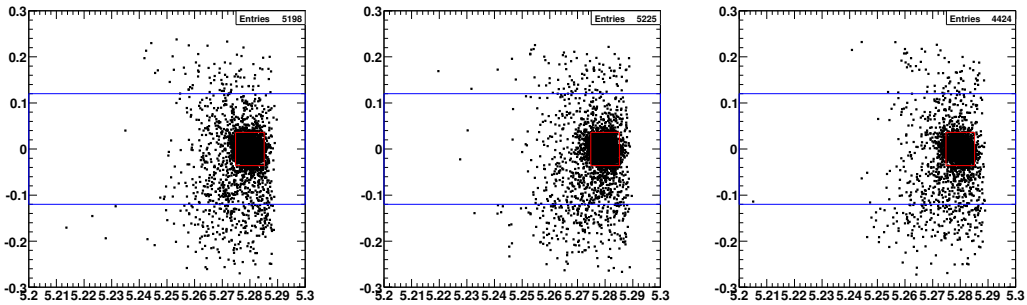


Figure D.73. Simulated $B^0 \to D_s^{*-}K^+$ events, $B^0 \to D_s^{*-}K^+$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

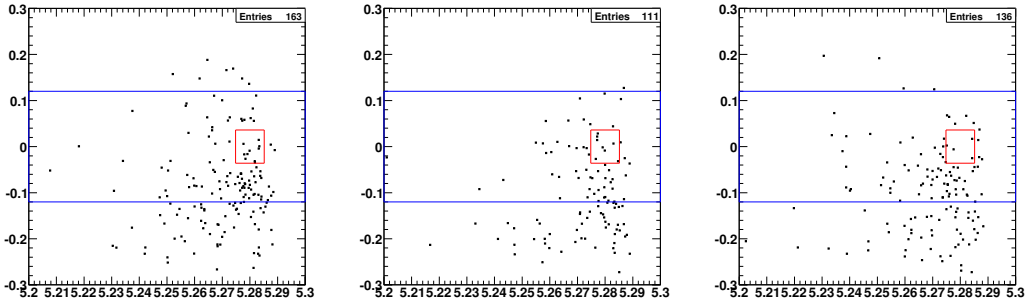


Figure D.74. Simulated $B^0 \to D_s^- K^{*+}$ events, $B^0 \to D_s^{*-} K^+$, $D_s^{*+} \to D_s^+ \gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

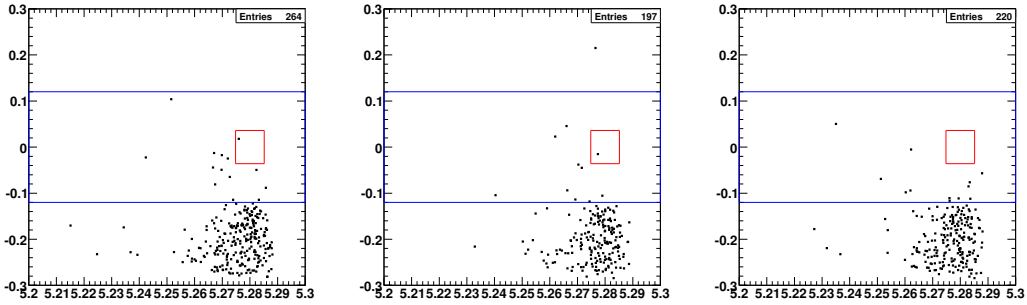


Figure D.75. Simulated $B^0 \to D_s^{*+} K^{*-}$ events (longitudinal polarization), $B^0 \to D_s^{*-} K^+$, $D_s^{*+} \to D_s^+ \gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

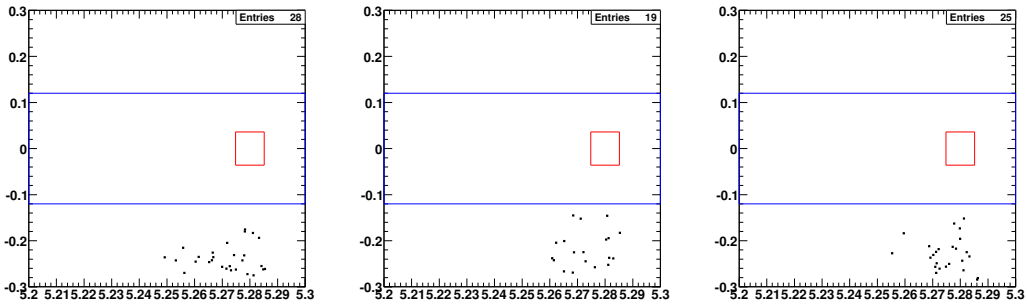


Figure D.76. Simulated $B^0 \to D_s^{*+} K^{*-}$ events (transverse polarization), $B^0 \to D_s^{*-} K^+$, $D_s^{*+} \to D_s^+ \gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*} K^+$(middle) and $D_s^+ \to \bar{K}^0 K^+$(right) selection.

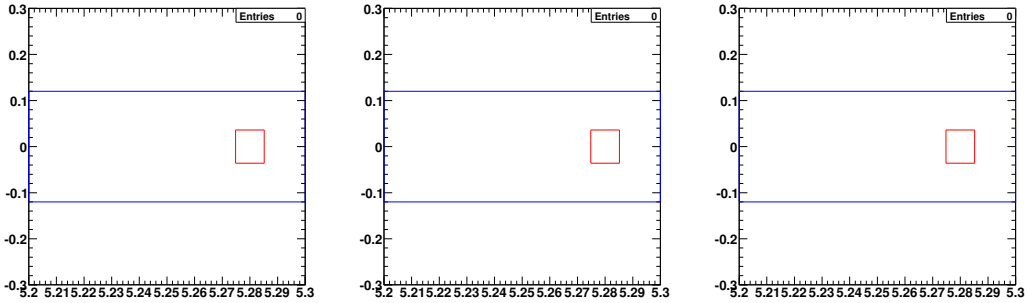


Figure D.77. Simulated $B^+ \to D_s^+\pi^0$ events, $B^0 \to D_s^{*-}K^+$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

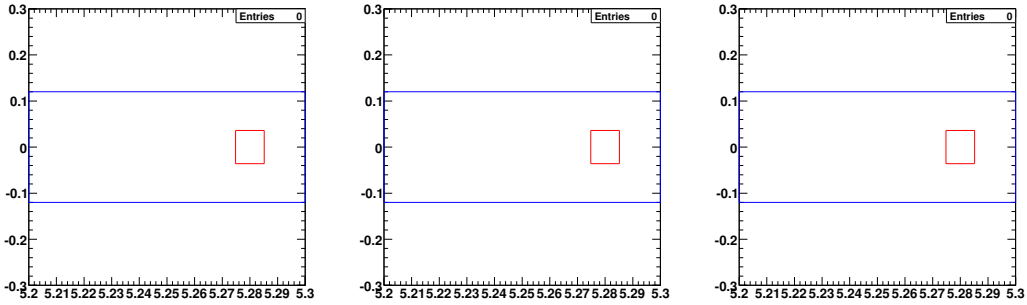


Figure D.78. Simulated $B^+ \to \bar{D}^{0*}\pi^+$ events, $B^0 \to D_s^{*-}K^+$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

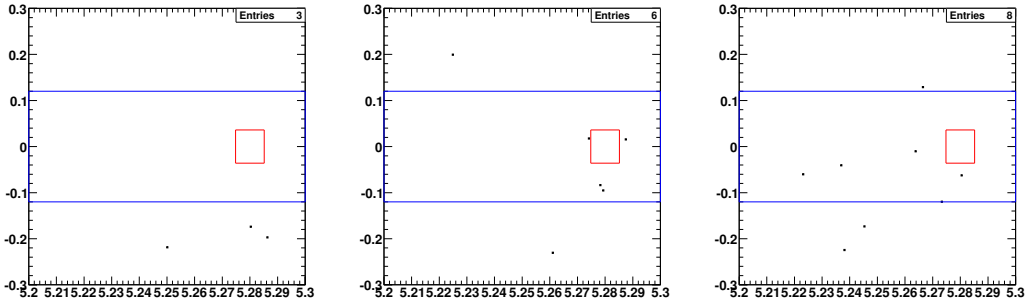


Figure D.79. Simulated $B^+ \to D_s^+\rho^0$ events, $B^0 \to D_s^{*-}K^+$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

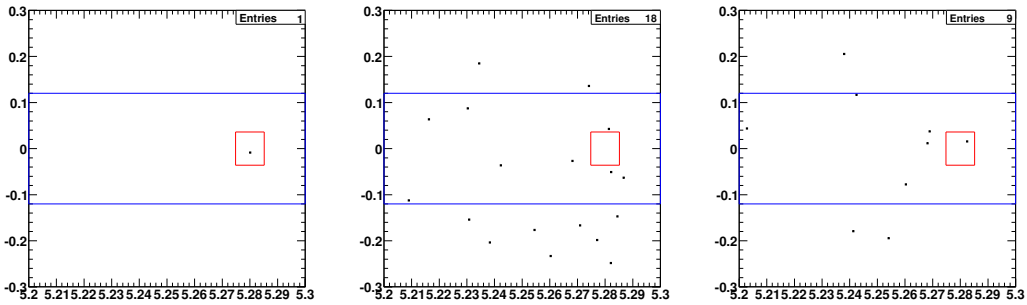


Figure D.80. Simulated $B^0 \to D^{(*)-}X$ "cocktail" events, $B^0 \to D_s^{*-}K^+$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

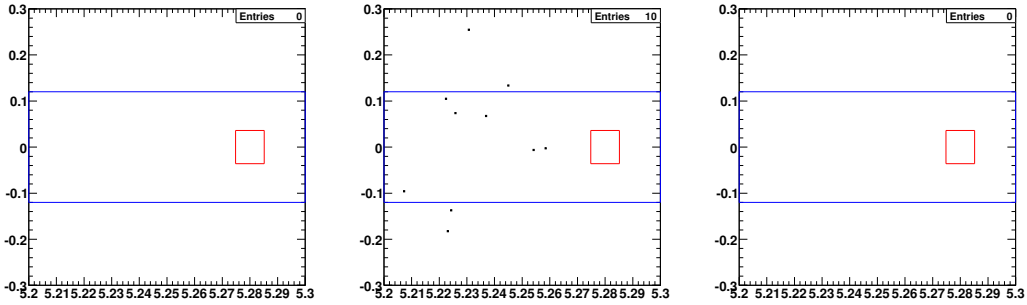


Figure D.81. Simulated $B^+ \to D^{(*)0}\pi^+$ "cocktail" events, $B^0 \to D_s^{*-}K^+$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.

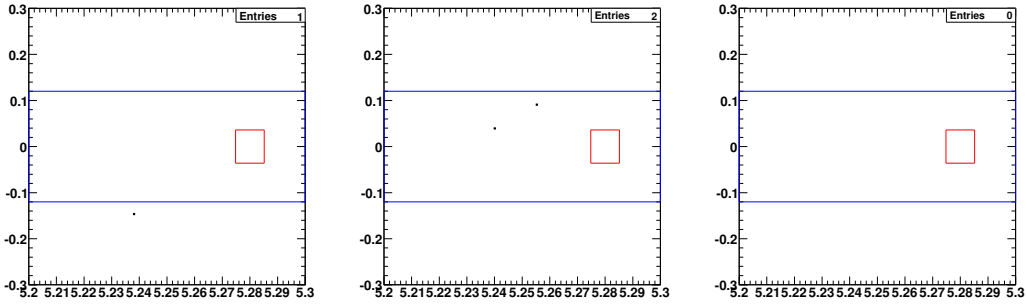


Figure D.82. Simulated $B^+ \to D^{(*)0}X$ "cocktail" events, $B^0 \to D_s^{*-}K^+$, $D_s^{*+} \to D_s^+\gamma$, $D_s^+ \to \phi\pi^+$(left), $D_s^+ \to \bar{K}^{0*}K^+$(middle) and $D_s^+ \to \bar{K}^0K^+$(right) selection.