Search for Rare Quark-Annihilation Decays, Charged B Mesons Decaying to Charged D(S) Mesons And Phi Mesons

J. Adam M. Cunha

SLAC-R-880

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.

UNIVERSITY OF CALIFORNIA
Santa Barbara

Search for Rare Quark-Annihilation Decays, $B^{\pm} \to D_s^{(*)\pm}\phi$

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

Doctor of Philosophy

in

Physics

by

J. Adam M. Cunha

Committee in Charge:

Professor Claudio Campagnari, Chair

Professor David Stuart

Professor Douglas Eardley

September 2006

The Dissertation of
J. Adam M. Cunha is approved:

Professor David Stuart

Professor Douglas Eardley

Professor Claudio Campagnari, Committee Chairperson

August 2006

Search for Rare Quark-Annihilation Decays, $B^{\pm} \to D_s^{(*)\pm}\phi$

by

J. Adam M. Cunha

To Sky:

Thanks for everything.

Compared to planning a wedding, this was a piece of cake.

Acknowledgements

I would like to thank my advisor, Prof. Claudio Campagnari, for all his patient guidance and endless support. I would also like to thank Prof. Jeff Richman for his always thoughtful answers to any of my profound (or more likely, otherwise) questions. And of course my thanks to my close friend and colleague, Tae Min Hong, for too many things to mention.

Cheers to my UCSB graduate student partners in crime, Steve Levy, Michael Mazur, and Bryan Dahmes. And, of course, to the postdocs who put up with us, Wouter Verkerke, Owen Long, and Jeff Berryhill – I owe you all at least one beer.

Finally, I would like to thank my parents, my grandparents, and the rest of my family who have always encouraged me to work hard and reach for the stars (Or more recently, reminded me that there is more to life than research!).

I cannot ever forget my love and soon to be wife, Sky. Words cannot say how much you mean to me and how much your support over the past four and a half years has helped in making my graduate career successful.

Curriculum Vitæ

J. Adam M. Cunha

Education

2004	M.S. U.C. Santa Barbara, California
2001	B.S. Rutgers University, New Brunswick, NJ

Reserach

2002 – 2006	SLAC, Member of the BaBar Experiment. Advisor Claudio Campagnari
1997 – 2000	FNAL, Member of the KTeV Experiment. Advisor Sunil V. Somalwar

Selected Publications

"Search for rare quark-annihilation decays, $B^{\pm} \to D_s^{(*)\pm}\phi$", B. Aubert, et al. (The BaBar Collaboration), Phys. Rev. D 73, 011103 (R) (2006)

"BaBar Silicon Vertex Tracker: Status and Prospects", A.Cunha (For the Babar SVT Group). Accpeted for publicaion in Nuclear Instrumentation and Methods A

"Radiative Decay Width Measurements of Neutral Kaon Excitations Using the Primakoff Effect", A. Alavi-Harati, et al. (The KTeV Collaboration), Phys. Rev. Lett. 89, 072001 (2002)

Abstract

Search for Rare Quark-Annihilation Decays, $B^{\pm} \to D_s^{(*)\pm}\phi$

J. Adam M. Cunha

We report on a search for the decay $B^{\pm} \to D_s^{(*)\pm}\phi$ using 212.2 fb^{-1} of data collected with the *BABAR* detector at the PEP-II B Factory at the Stanford Linear Accelerator Center between 1999 and 2004. This sample of 234×10^6 $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ events yields no significant signal. We report the Bayesian upper limits $\mathcal{B}(B^{\pm} \to D_s^{\pm}\phi)\times\mathcal{B}(D_s^{\pm} \to \phi\pi^{\pm}) < 8.6\times10^{-8}$ and $\mathcal{B}(B^{\pm} \to D_s^{*\pm}\phi)\times\mathcal{B}(D_s^{\pm} \to \phi\pi^{\pm}) < 5.4 \times 10^{-7}$ at the 90% C.L. Using the latest measurement of $\mathcal{B}(D_s^{\pm} \to \phi\pi^{\pm})$ [1], we report: $\mathcal{B}(B^{\pm} \to D_s^{\pm}\phi) < 1.8\times10^{-6}$ and $\mathcal{B}(B^{\pm} \to D_s^{*\pm}\phi) < 1.1\times10^{-5}$ at the 90% C.L.

Professor Claudio Campagnari
Dissertation Committee Chair

Contents

List of Figures

List of Tables

Chapter 1

Document Roadmap

In order to aid in the reading of this document, we provide here a brief roadmap.

In Chapter 2 we discuss the Standard Model expectations for $\mathcal{B}(B^{\pm} \to D_s^{(*)\pm}\phi)$ and why quark annihilation diagrams are interesting. We then follow with a discussion of some of the beyond-the-Standard-Model theories that could lead to enhancements of the rate for $B^{\pm} \to D_s^{(*)\pm}\phi$.

In Chapter 3 we discuss details of the *BABAR* detector and the SLAC facility.

In Chapter 4 we discuss the data samples used in this analysis. Section 4.1 will cover the collection of $e^+e^- \to \Upsilon(4S)$ and $e^+e^- \to q\overline{q}$ (q = u, d, c, or s) data samples while Section 4.2 will cover the generation and use of Monte Carlo simulated data (MC).

Then we get to the meat of this document, the discussion of the $B^{\pm} \to D_s^{(*)\pm}\phi$ analysis. In Chapter 5 we present in detail the roadmap for our analysis. In Chapter 6 we define our selection criteria, discuss the optimization of the selection

criteria to obtain the best possible upper limit on $\mathcal{B}(B^{\pm} \to D_s^{(*)\pm}\phi)$, and explore possible backgrounds. In Chapter 7 we finally take look at the data.

In Chapter 8 we discuss experimental systematic uncertainties. In Chapter 9 we calculate our final results for $\mathcal{B}(B^{\pm} \to D_s^{(*)\pm}\phi)$; and in Chapter 10 we discuss conclusions we can draw from these results.

Finally, in the appendices we provide more detailed discussions of some of the techniques, definitions, etc. used in this analysis. These detailed discussions were moved to the appendices in order to aid in the readability of this document and for ease of reference.

Chapter 2

Theoretical Aspects of the Measurements

2.1 The Standard Model

In the Standard Model (SM), the decay $B^{\pm} \to D_s^{(*)\pm}\phi$ occurs through annihilation of the two quarks in the B meson into a virtual W. The Feynman diagram of this type of process is commonly referred to as an annihilation diagram, see Figure 2.1.

Until this year, no annihilation-type B decays have ever been observed – the Belle Collaboration presented preliminary evidence for the decay, $B^+ \to \tau^+\nu_\tau$ in April (see below). The current 90% confidence level (C.L.) upper limits on the branching fractions, $\mathcal{B}(B^{\pm} \to D_s^{\pm}\phi)$ and $\mathcal{B}(B^{\pm} \to D_s^{*\pm}\phi)$ are 3.2×10^{-4} and 4.0×10^{-4} respectively. These limits come from a CLEO analysis of 1.16 million

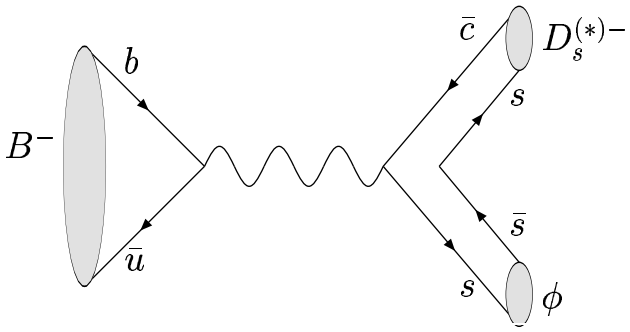

Figure 2.1: The Feynman diagram for process, $B^{\pm} \to D_s^{(*)\pm}\phi$. This is an example of an annihilation diagram

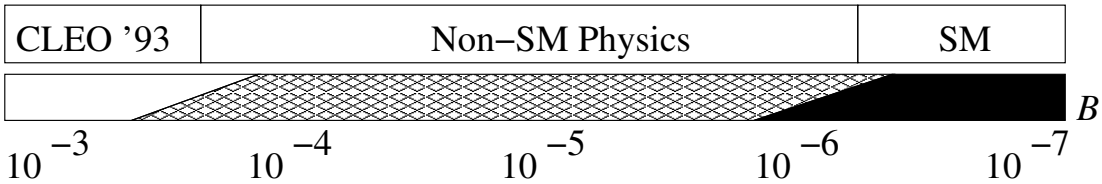

Figure 2.2: The bottom bar in this cartoon depicts the reach we will have on the branching fractions $\mathcal{B}(B^{\pm} \to D_s^{(*)\pm}\phi)$ using the current *BABAR* dataset. The darkened area is out of our reach; the hashed area can be explored with the data available from *BABAR*; the white area has been excluded by CLEO.

$B\overline{B}$ decays [2]. With the present *BABAR* dataset, we should be able to search for these modes with two orders of magnitude more sensitivity (see Figure 2.1).

In the SM, annihilation diagrams of the B meson are highly suppressed due to the smallness of $|V_{ub}|$ and the meson decay constant f_b which describes the overlap of the constituent-quark wavefunctions. Calculations of the $B^{\pm} \to D_s^{(*)\pm}\phi$ branching fractions give predictions of 3×10^{-7} using a perturbative QCD ap-

proach [3], or 7×10^{-7} using QCD-improved factorization [4]. Rescattering[1] could also contribute to $B^{\pm} \to D_s^{(*)\pm}\phi$. Note, however, that there cannot be any rescattering contributions from the dominant $b \to c$ transitions, since starting from a b quark in the initial state we must end up with a $\overline{c}$ quark in the final state, see Figure 2.1. Rescattering can only occur from the suppressed $b \to u\overline{c}s$ transitions where $u\overline{u} \to s\overline{s}$, *e.g.*, $b\overline{u} \to u\overline{c}s\overline{u} \to \overline{c}ss\overline{s}$.

Since the SM predicts a small branching fraction, and since the current experimental limit is about three orders of magnitude higher than the SM expectations, searches for $B^{\pm} \to D_s^{(*)\pm}\phi$ could be sensitive to non-standard-model contributions. For example, in Reference [4] the branching fraction of $B^{\pm} \to D_s^{(*)\pm}\phi$ is estimated to be as high as 8×10^{-6} in a two-Higgs-doublet model and 3×10^{-4} in the MSSM with R-parity violation.

The decay that we are considering, $B^{\pm} \to D_s^{(*)\pm}\phi$, is not the only example of an annihilation decay. From a theoretical point of view, the decay $B^{\pm} \to D_s^{(*)\pm}\phi$ is related simply to a few other decays. Another possibility is $B^{+} \to D^{(*)+}K^{0}$. (Here and throughout the paper the charge-conjugated mode is implied.) The Feynman diagram for this mode can be obtained replacing the $s\overline{s}$ pair with a $d\overline{d}$ pair in Figure 2.1. *BABAR* has completed a search for this decay mode and

[1] Rescattering is a process in which, for example, $u\overline{u} \leftrightarrow s\overline{s}$ in the final state. In general, this process is not well understood and contributions could be large.

placed upper limits on the branching fractions, $\mathcal{B}(B^+ \to D^+ K^0) < 0.5 \times 10^{-5}$ and $\mathcal{B}(B^+ \to D^{*+} K^0) < 0.9 \times 10^{-5}$ [5].

The decay $B^+ \to \tau^+ \nu_\tau$ is also very similar to $B^\pm \to D_s^{(*)\pm}\phi$. The left hand side of the Feynman diagram is the same, but on the right hand side, replace our $D_s^{(*)\pm}$-ϕ system with $\tau\nu_\tau$. A few months ago, the Belle collaboration released a preliminary result claiming evidence for the observation of the $B^+ \to \tau^+ \nu_\tau$ decay [6], and just three weeks ago at the ICHEP 2006 conference, they released an updated, yet still preliminary, result, $\mathcal{B}(B^+ \to \tau^+ \nu_\tau) = (1.79^{+0.56}_{-0.49}(\text{stat})^{+0.39}_{-0.46}(\text{syst})) \times 10^{-4}$ [7].

We note that our experimental signature should be quite clean because of the $D_s^\pm$ and, especially, the ϕ requirement. For example, the *BABAR* searches [8] for $B^0 \to D_s^{(*)\pm}\pi^-$ and $B^0 \to D_s^{(*)\pm}K^-$ resulted in a handful of background events in the signal region. Substituting a ϕ for a pion or a kaon (both of which are ubiquitous at *BABAR*) should greatly improve the rejection of combinatorics, and perhaps even result in a background-free search.

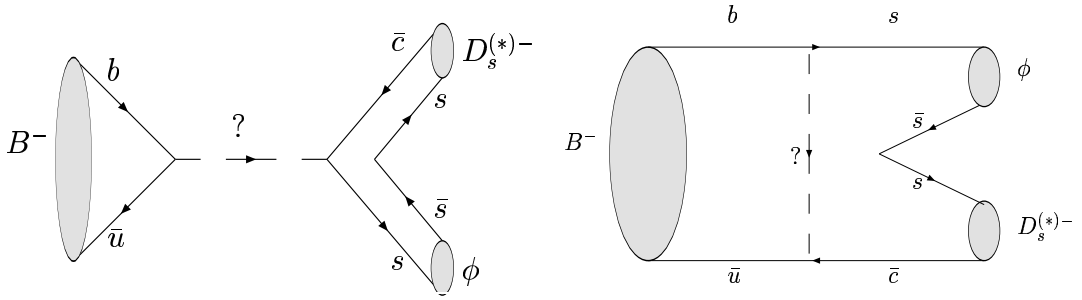

Figure 2.3: Annihilation (left) and exchange (right) tree-level Feynman diagrams for possible NP contributions to $B^{\pm} \to D_s^{(*)\pm}\phi$. Internal gluons (i.e. QCD interactions within each meson) are suppressed for clarity.

2.2 Physics beyond the Standard Model

Non-Standard-Model physics (colloquially refered to as "New Physics", NP) can enhance the $B^{\pm} \to D_s^{(*)\pm}\phi$ rate. Since NP has to compete with a SM tree diagram, it is reasonable to expect that we could only be sensitive to NP that comes in at the tree level. The only tree level diagrams that we can write down are annihilation and exchange diagrams, see Figure 2.3.

The exchange diagram is mediated by some virtual particle that causes a flavor changing neutral current (FCNC) at the tree level, and so it is very unlikely. The annihilation diagram is the same as the SM diagram of Figure 2.1, except here the propagator could be any charged particles that couples to a quark current.

For a NP annihilation diagram to compete with the SM diagram, one or both of the following must be satisfied:

1. The mass of the exchanged particles must be lighter or of the same order as that of the W boson, otherwise the rate, which is inversely related to the mass of the propagator, would be much suppressed relative to the SM contribution. (Of course large couplings could conspire with the large mass to enhance the rate.)

2. The couplings at the $b\overline{u}$ and/or $\overline{c}s$ vertices must be enhanced with respect to the SM.

Because of the results from direct searches, it is quite unlikely that the NP particle in the propagator is lighter than the W. So, in order to have an enhancement of the NP contribution with respect to the SM, we need the NP couplings to be large. Note that the SM couplings are suppressed by the smallness of $|V_{ub}|$. In general NP contributions to $B^{\pm} \to D_s^{(*)\pm}\phi$ will depend on the mass of NP particle(s) and on the couplings to $b\overline{u}$ and $\overline{c}s$.

The two possibilities that have been explored in the literature [4] are R-parity violating supersymmetry (RPV-SUSY) and a two Higgs Doublet Model (2HDM). In RPV-SUSY the exchanged particle is a slepton and the couplings could in principle be quite large. In the 2HDM, the exchanged particle is a charged Higgs, and the couplings can be enhanced by factors of the ratio of vacuum expectation

values for the two Higgs doublets, $\tan\beta$. The remainder of this chapter will be devoted to the discussion of these two models.

R-Parity violating supersymmetry

In the RPV-SUSY model, the Lagrangian includes a term $\lambda'_{ijk}L_iQ_jD_k$ where the λ's are coupling constants, L_i is the SU(2) doublet lepton superfield for the i^{th} generation, Q_j is the SU(2) doublet quark superfield for the j^{th} generation, and D_k is the SU(2) singlet down-type quark superfield for the k^{th} generation. This causes tree level couplings $u-b-slepton$ and $slepton-c-s$ that can then drive the annihilation diagram of Figure 2.3, where the exchanged particle is a slepton.

The RPV-SUSY contribution to $\mathcal{B}(B^{\pm}\to D_s^{(*)\pm}\phi)$ depends on the quantity

$$\frac{\lambda'^2}{M^2}\equiv\sum_i\frac{\lambda'_{i22}\lambda'_{i13}}{M_i^2} \tag{2.1}$$

where M_i is the mass of the i^{th} generation slepton and i runs over the the three matter generations. Mohanta [4] has performed this calculation at tree level and in the context of factorization. Using values of $\lambda'_{i22}\lambda'_{i13}$ and M_i not excluded by other measurements, Mohanta predicts a rate as high as 3×10^{-4}.

Two-Higgs Doublet super-symmetry

In the type II 2HDM model[2], for large $\tan\beta$ the NP contribution to $B^{\pm} \to D_s^{(*)\pm}\phi$ depends on the ratio $\tan\beta/M_{H^{\pm}}$, where $M_{H^{\pm}}$ is the mass of the charged Higgs boson. As for the RPV-SUSY model, Mohanta has calculated the 2HDM $B^{\pm} \to D_s^{(*)\pm}\phi$ contribution at tree level and in the context of factorization and predicts a rate as high as 8×10^{-6}. The type II 2HDM model is quite popular since the minimal super-symmetric standard model (MSSM) is a 2HDM model, so it is worthwhile to spend some time trying to put our result in context.

First of all, searches at LEP resulted in a limit $M_{H^{\pm}} > 78.6$ GeV (at high $\tan\beta$) [9], [10]. In the 2HDM model, the most stringent indirect limits on $M_{H^{\pm}}$ come from $b \to s\gamma$: $M_{H^{\pm}} > 320$ GeV (95% CL) [9], [11]. Information on $B \to D\tau\nu$ yields $\tan\beta/M_{H^{\pm}} < 0.46/\,\text{GeV}$ [12]. The decay $B \to \tau\nu$ also probes the same NP contribution, since it also proceeds via annihilation of the quarks in the $B^{\pm}$ mesons into a $W^{\pm}$ or $H^{\pm}$. From the recently published Belle result, $\mathcal{B}(B^+ \to \tau^+\nu_\tau) = (1.79^{+0.56}_{-0.49}(\text{stat})^{+0.39}_{-0.46}(\text{syst})) \times 10^{-4}$ [6] we can extract the tree-level limit $\tan\beta/M_{H^{\pm}} < 0.31/\,\text{GeV}$.

Note that at tree-level the 2HDM and MSSM models are the same. However, once higher order corrections are included, the predictions can be quite different

[2]In a 2HDM of type II, one Higgs couples to the up-type fermions and the other Higgs couples to the down-type fermions. This is in contrast to a 2HDM of type I in which one Higgs couples to all fermions and the other Higgs couples to nothing.

and model dependent, particularly at high $\tan\beta$, see for example Reference [13] for $B \to \tau\nu$.

Chapter 3

The PEP-II B Factory and the *BABAR* detector

A schematic drawing of the SLAC linear accelerator and the PEP-II storage rings is shown in Figure 3.1. The PEP-II B Factory [15] is an electron-positron collider which operates at center-of-mass energies around 10.56 GeV. This energy was chosen since it is the rest mass of the $\Upsilon(4S)$ resonance, a $b\overline{b}$ bound state which decays almost exclusively to a pair of $B^0\overline{B}^0$ or B^+B^- mesons. The rate of production at peak performance is approximately ten $B\overline{B}$ pairs per second. Figure 3.2 shows a scan across the $\Upsilon(4S)$ resonance. The $\Upsilon(4S)$ resonance has a peak cross-section of 1.1 nb, but it sits on a $e^+e^- \to q\overline{q}$ ($q = u, d, c, s$) background of almost 3 nb. This continuum background will prove to be the largest background in this analysis and its suppression will be critical.

The *BABAR* detector was designed as a general purpose detector. Various publications are available that discuss the detector design and the performance of

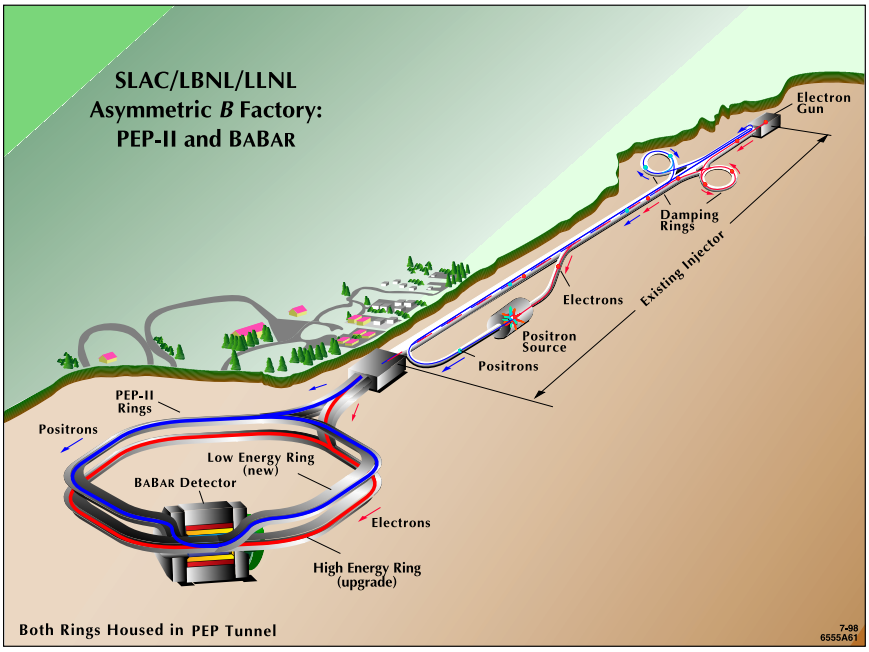

Figure 3.1: The SLAC linear accelerator and the PEP-II storage rings.

the various sub-detectors. An overview, which the reader may use to familiarize himself with the details of detector design, construction, and performance, can be found in [14]. Here we will discuss briefly the design and performance pertinent to this analysis of the sub-detectors that make up *BABAR*.

Figures 3.3 and 3.4 shows a schematic of the *BABAR* detector. It is comprised of five sub-detectors concentrically placed around the e^+-e^- interaction point. From the inside out the detector consists of a silicon vertex tracker (SVT), a drift chamber (DCH), a Cherenkov radiation detector (DIRC), and a CsI calorimeter (EMC). This is followed by a superconducting solenoid which produces a 1.5 T

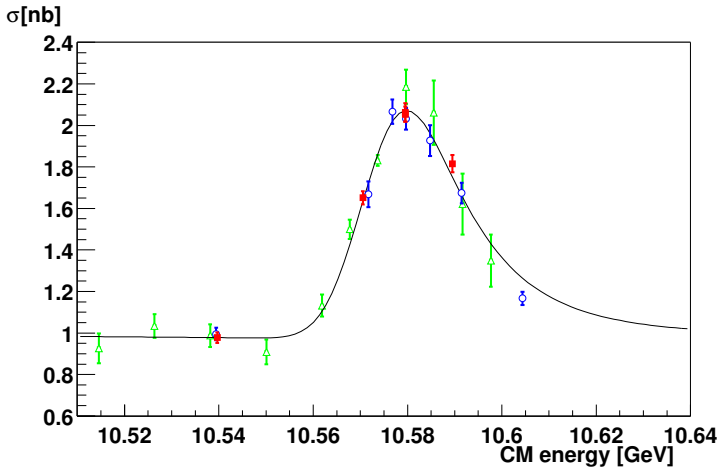

Figure 3.2: e^+e^- scan over the $\Upsilon(4S)$ peak. Note the sizable continuum background upon which the peak sits. The continuum cross-section has been suppressed using event-shape information. (Event shape selection criteria will be discussed in Chapter 6.) Eliminating this will be key to our success in this analysis.

magnetic field oriented along the beam axis. Finally, the magnet's flux return is instrumented for muon and neutral hadron detection.

The coordinate system of *BABAR* is such that the positive z-axis is along the direction of the incoming electron beam. The x-axis is oriented away from the center of the PEP-II rings and the y-axis oriented vertically upward. In reference to Figure 3.3, the z-axis goes from left to right, the x-axis points into the page, and the y-axis points to the top of the page. The azimuthal angle, ϕ, and the polar angle, θ, are the usual for a cylindrical coordinate system.

In the following sections we detail each of the *BABAR* sub-detectors. We will not discuss the instrumented flux return since its primary purpose is the detection

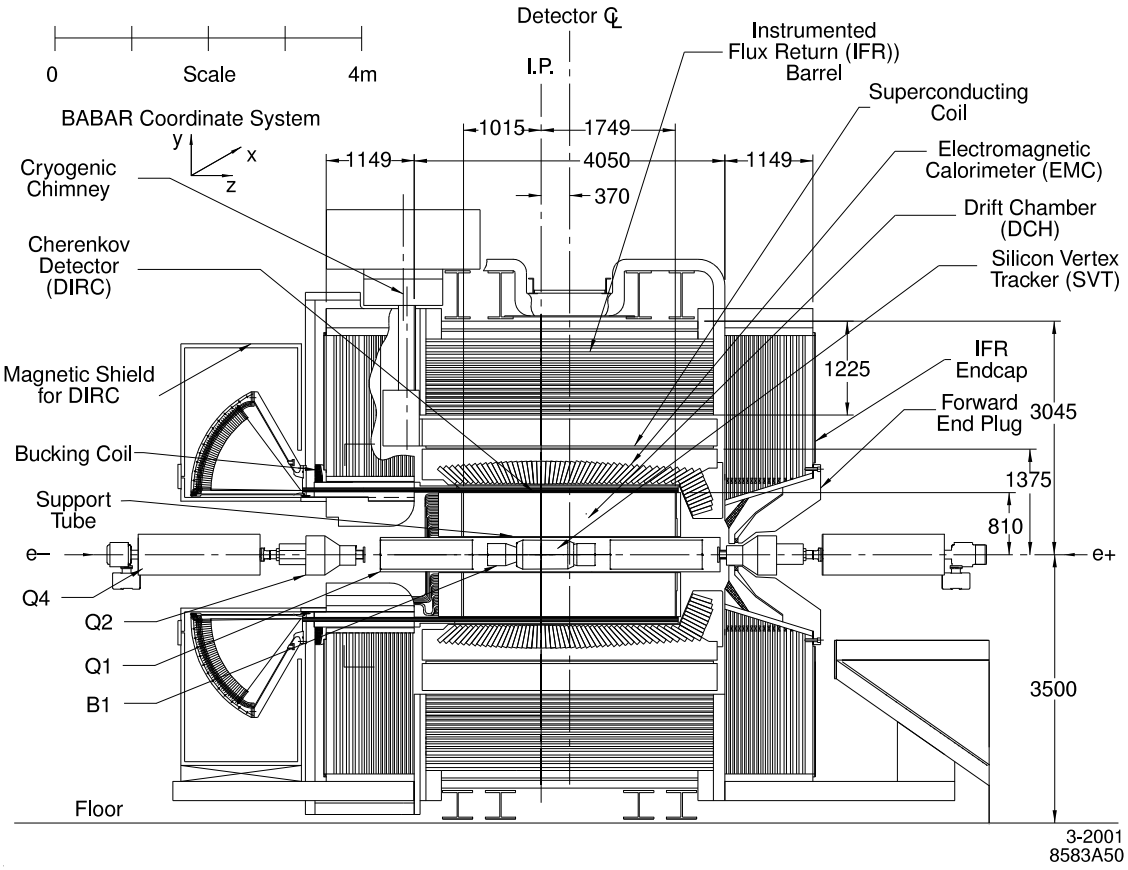

Figure 3.3: A schematic of the longitudinal cross-section of the *BABAR* detector.

of muons and long lived neutral hadrons (K_L^0 mesons) of which we have none in this analysis.

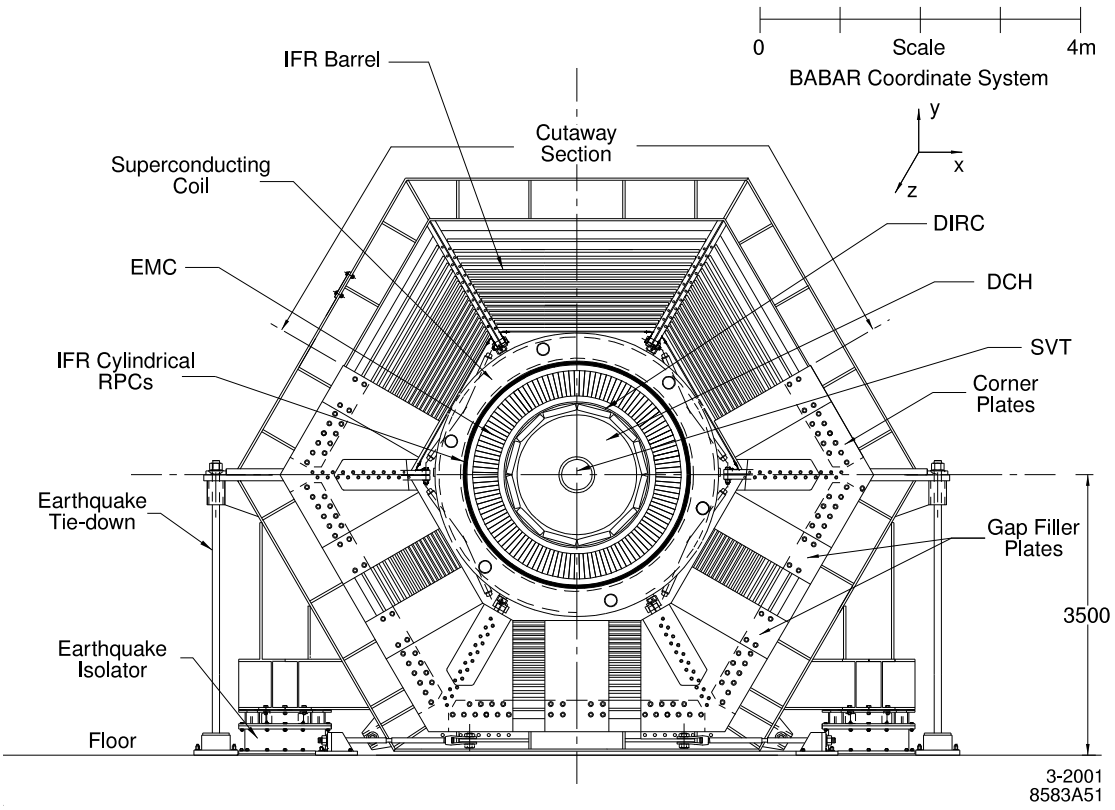

Figure 3.4: A schematic of the end-view cross-section of the *BABAR* detector.

3.1 Silicon vertex tracker

The SVT and DCH comprise the *BABAR* tracking system which is responsible for measuring the position and momentum of charged particles. It sits fully inside the superconducting solenoid and thus inside a uniform magnetic field parallel to the beam axis.

The SVT is a silicon microstrip detector. It consists of five concentric layers of silicon wafers centered about the beam pipe. Figure 3.5 shows the longitudinal cross section of the SVT. Each wafer is instrumented on both sides to detect the passage of charged particles through the wafer. One side of the silicon has strips running perpendicular to the beam direction to record the z coordinate of a passing particle while the other has strips running parallel to the beam direction to record the ϕ coordinate. Thus, along with the radial component which is given by the layer's radial location, we obtain 5 3D points for each particle which traverses the SVT[1].

The SVT was designed to measure the angles and position of charged particles as close to the beam pipe as possible and to provide stand-alone track reconstruc-

[1] This of course assumes that the particle has enough transverse momentum to exit the SVT and enter the DCH, which, if not true, can lead to greater than or less than five position measurements per charged particle.

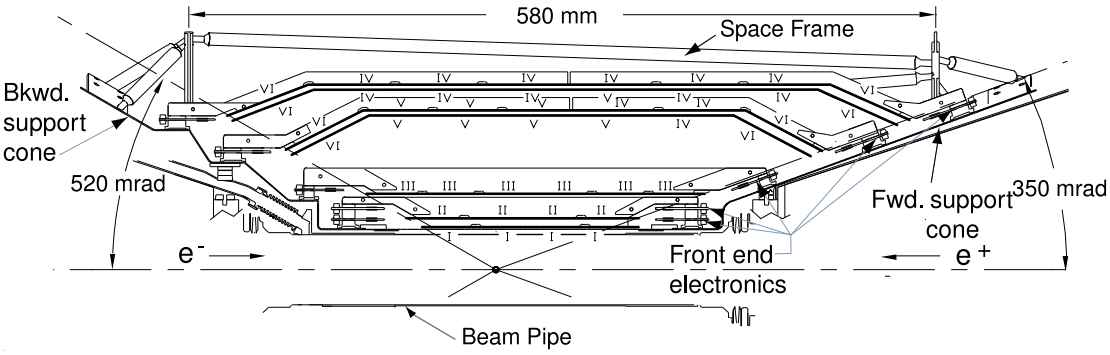

Figure 3.5: This is a longitudinal cross-section of the *BABAR* silicon vertex tracker.

tion for particles with not enough transverse momentum to reach the DCH. The mean three-dimensional vertex resolution[2] for the SVT is around 100 μm.

The SVT can also measure a particle's energy loss while traversing the detector, $\mathrm{d}E/\mathrm{d}x$. Each particle species (pion, kaon, etc.) has a distinctive $\mathrm{d}E/\mathrm{d}x$ distribution as a function of momentum, this can be used to aid in identifying the type of particle traversing the detector. The $\mathrm{d}E/\mathrm{d}x$ resolution of the SVT is about 14%, and charged kaons can be distinguished from pions up to a momentum of 500 MeV/c.

[2]The vertex is the point at which a particle decays. The vertex resolution is the distance apart two vertices must be for our detector to be able to distinguish them.

3.2 Drift chamber

The Drift Chamber consists of 40 cylindrical layers of drift cells centered around the beam pipe, a cross section is shown in Figure 3.6. Each cell is a hexagonal configuration of six field-shaping wires at the corners and one sense wire at the center. Some of the wires are aligned with an angle to the z-axis in order to provide z-coordinate information about a charged particle's track. The DCH volume is filled with an 80:20 helium:isobutane mixture. This mixture keeps the average density of the gas at a minimum to minimize multiple scattering. Tracks are detected by ionization of the gas in the chamber as a charged particle passes through. These ionized atoms and electrons then drift under the electric field present to the wires and the charge is read off as an electrical signal.

As the other sub-detector that makes up the heart of the *BABAR* tracking system along with the SVT, the DCH complements the SVT. Longer lived particles like the K_S^0 will not always decay inside the SVT volume, thus the DCH provides much of the tracking needed to reconstruct the $K_S^0 \to \pi^+\pi^-$ decays used in this analysis. The tracking efficiency in the DCH is around 98% for $p_T > 200$ MeV/c^2 and polar angle greater than 500 mrad, and a $\mathrm{d}E/\mathrm{d}x$ resolution of 7.5% allows for $K^\pm$-$\pi^\pm$ separation up to transverse momenta of 700 MeV/c (see Figure 3.7).

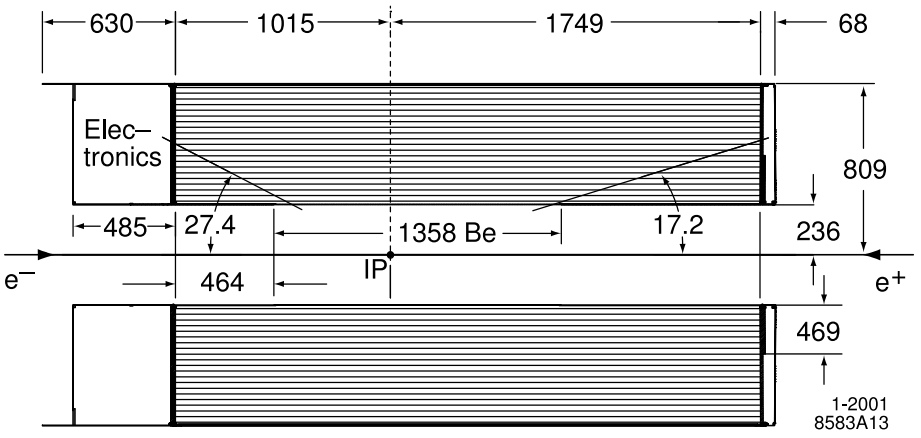

Figure 3.6: Longitudinal cross-section of the DCH.

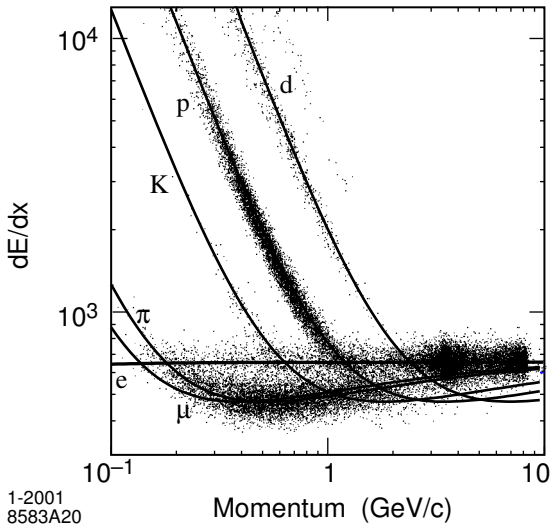

Figure 3.7: Here we show the dE/dx resolution for the DCH (arbitrary units). Note that there is good $K^{\pm}$-$\pi^{\pm}$ separation up to 700 MeV/c.

3.3 Cherenkov radiation detector

The DIRC, an acronym for Detector of Internally Reflected Cherenkov (light), is a novel Cherenkov-based detector devoted to particle identification (PID). Cherenkov radiation is produced when a charged particle traversing a medium is moving faster than the speed of light in that medium. A cone of radiation is emitted around the trajectory of the particle. The apex angle, known as the Cherenkov angle, θ_{C}, is given by,

$$\cos\theta_{\mathrm{C}} = \frac{1}{n\beta} = \frac{c}{nv}. \tag{3.1}$$

The DIRC is made of fused silica which has an index of refraction of n=1.4723.

Figure 3.8 shows a longitudinal cross-section of the DIRC. It consists of 144 silica bars placed cylindrically around the DCH. When a charged track traverses the bar, the Cherenkov light propagates through internal reflection down the quartz bar and out the back end of the *BABAR* detector where it is incident on photo multiplier tubes (PMTs). The geometric placement of the PMTs outside the detector allows for minimal material inside the detector. As the quartz bars are in front of the electromagnetic calorimeter, this geometry is key in maintaining a good energy resolution in the calorimeter by minimizing material in front it.

Figure 3.3 shows the particle identification performance of the DIRC. The DIRC can achieve 4σ or better K/π separation for most particles of momenta

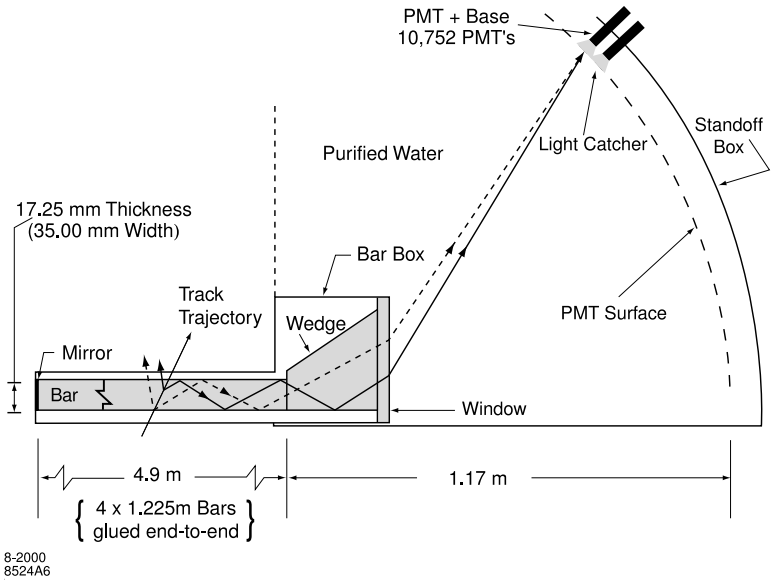

Figure 3.8: Longitudinal cross-section of the DIRC sub-detector.

higher than the threshold for the production of Cherenkov light ($\approx$ 460 MeV/c for kaons) and up to 4.5 GeV/c. This is important in this analysis for the reconstruction of the kaons from the $D_s^{\pm}$ whose momentum peaks at around 1 GeV/c.

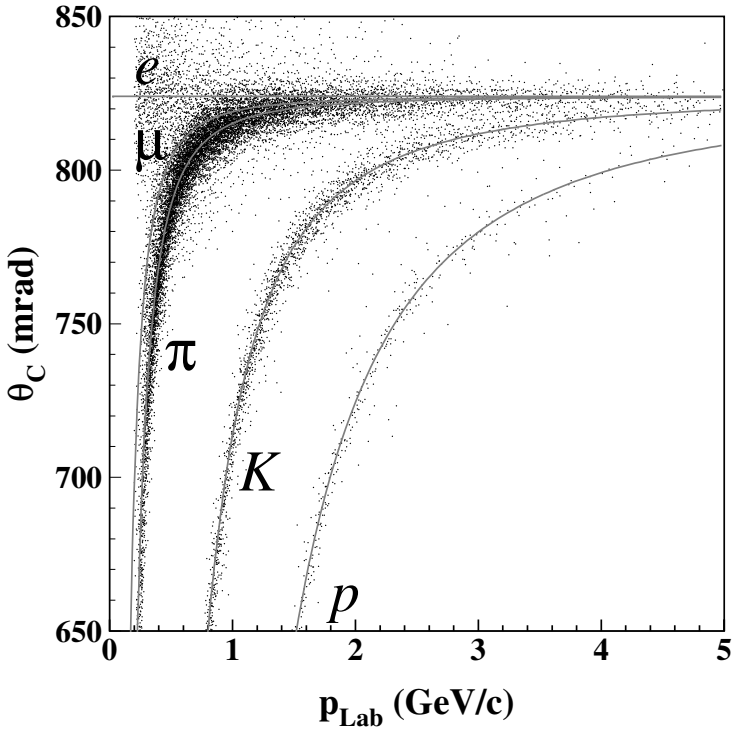

Figure 3.9: Particle identification performance in the DIRC. Note the excellent $K^{\pm}$-$\pi^{\pm}$ separation out to 2 GeV/c.

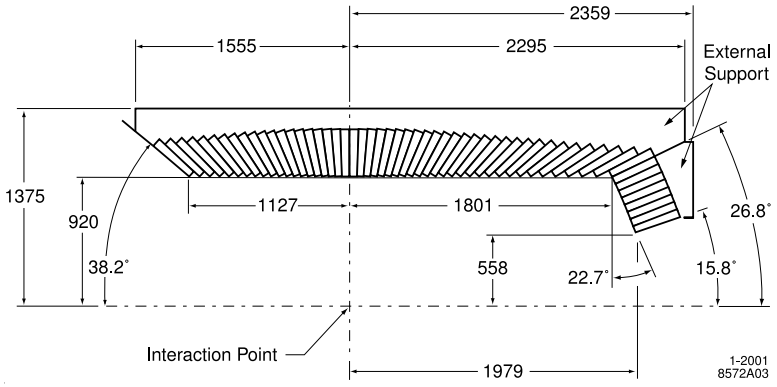

Figure 3.10: Longitudinal cross-section of the electromagnetic calorimeter.

3.4 Electromagnetic calorimeter

The purpose of the EMC is to measure precisely the energy of electrons and photons by causing them to interact electromagnetically with the detector medium and deposit energy. Only electrons and photons deposit most of their energy in the EMC; this allows for discrimination between electrons and other charged tracks and the identification of photons which is important for reconstructing the $D_s^{*\pm} \to D_s^{\pm}\gamma$ in this analysis. The EMC is effective for all energies above 0.03.

A cross-section is shown in Figure 3.10. The EMC consists of an array of finely segmented crystals made of thallium-doped cesium iodide. These crystals have a high light yield and a short radiation length to provide the best detector efficiency and electromagnetic shower containment possible. The single-photon

energy resolution of the EMC can be expressed as

$$\frac{\sigma_{\mathrm{E}}}{\mathrm{E}} = \frac{2.3\%}{\sqrt[4]{\mathrm{E(\,GeV)}}} \oplus 1.9\%. \tag{3.2}$$

The first term is due to fluctuations in photon statistics, electronics noise, and low energy beam-generated backgrounds. The second term arises from non-uniformity in the light collection, leakage, or absorption in front of the crystals.

Chapter 4

Data Samples

4.1 *BABAR* data

In this section we document the data sets used in this analysis. Our results are based on 234×10^6 $\Upsilon(4S) \rightarrow B\overline{B}$ decays, corresponding to an integrated luminosity of 212 fb^{-1}, collected between 1999 and 2004 with the *BABAR* detector at the PEP-II B Factory at the Stanford Linear Accelerator Center. A 12 fb^{-1} data sample taken with a center-of-mass (CM) energy 40 MeV below the $\Upsilon(4S)$ resonance peak (Recall Figure 3.2.) is used to study continuum events, $e^+e^- \rightarrow q\overline{q}$ ($q = u, d, s$, or c)[1]. The number of B mesons in our data sample is two orders of magnitude larger than in the previously-published $B^{\pm} \rightarrow D_s^{(*)\pm}\phi$ search [2]. The event sample used for this analysis was taken during four separate data-taking periods referred to as Runs 1, 2, 3, and 4. This analysis was originally to use

[1] This data sample is know as the off-resonance data sample. The data sample collected at the $\Upsilon(4S)$ mass is referred to as the on-resonance data sample.

data from Runs 1-3 and then extended to include data from Run 4, thus, at times will refer to a distinction between these two groups of data. Also, portions of this analysis, selection criteria optimization for example, were done on the Runs 1-3 dataset. As there is no significant difference between the Runs 1-3 and the Run 4 datasets we chose not to repeat these studies for the Run 4 dataset. It is noted were only Run 1-3 data was used.

In addition, we utilize large samples of Monte Carlo simulated data (MC) events to understand continuum ($q\overline{q}$ MC) and B-meson ($B^0\overline{B}^0$ and B^+B^- MC) background events as well as to understand the detector response to our signal events ($B^\pm \to D_s^{(*)\pm}\phi$ MC). The data sets used are summarized in Table 4.1.

4.2 Monte Carlo simulated data

Monte Carlo simulated data is an essential tool in this analysis. The *BABAR* collaboration has on hand simulations of many $\Upsilon(4S) \to B\overline{B}$ and non-resonant interactions which can be used to study the detector response to various signals. The *BABAR* MC used the GEANT4[2] software package along with detailed detector-geometry and detector-materials models for event simulation. The flow chart depicted in Figure 4.1 shows the steps involved in producing the MC. The main steps to MC production discussed below are:

[2]http://geant4.web.cern.ch/geant4

Table 4.1: Summary of the data sets used in this analysis. The signal Monte Carlo samples, $B^{\pm} \to D_s^{(*)\pm}\phi$, are generated using the appropriate fractions of the decay modes of the $D_s^{(*)\pm}$

Data Type	Number of events (Equivalent luminosity)	Run
On resonance data	124 $\times 10^6$ $B\overline{B}$ events (112.4 fb^{-1})	1-3
On resonance data	110 $\times 10^6$ $B\overline{B}$ events (99.8 fb^{-1})	4
Off resonance data	12 fb^{-1}	1-3
$c\bar{c}$ Monte Carlo	150 fb^{-1}	1-3
uds Monte Carlo	155 fb^{-1}	1-3
$B^0\overline{B}^0$ Monte Carlo	252 $\times 10^6$ events	1-3
B^+B^- Monte Carlo	254 $\times 10^6$ events	1-3
$B^{\pm} \to D_s^{\pm}\phi$ Monte Carlo	60K events	1-3
$B^{\pm} \to D_s^{*\pm}\phi$ Monte Carlo	60K events, longitudinally polarized	1-3
$B^{\pm} \to D_s^{*\pm}\phi$ Monte Carlo	60K events, transversely polarized	1-3
$B^{\pm} \to D_s^{\pm}\phi$ Monte Carlo	57K events	4
$B^{\pm} \to D_s^{*\pm}\phi$ Monte Carlo	57K events, longitudinally polarized	4
$B^{\pm} \to D_s^{*\pm}\phi$ Monte Carlo	57K events, transversely polarized	4

- Generate the physics event
- Propagate the particles through a detector model
- Record positions and idealized energy deposits of the particles
- Simulate detector response
- Add background events
- Reconstruct tracks and clusters as in data

GENERATING THE PHYSICS EVENT

The physics of the e^+e^- collision is simulated by software called an event generator. The initial positron and electron collision is simulated. Then, based on production and decay rules given to the generator, an $e^+e^- \rightarrow X$ event is generated. The production and decay rules are given by the physicist to specify what X should be, e.g. $e^+e^- \rightarrow q\overline{q}$ or $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B^+B^- \rightarrow Y$. The generator output is a set of four-vectors which represents the final state of the collision near the e^+e^- interaction point.

Propagate particles; record idealized information

The four-vectors from the generator stage are transported through a GEANT4 simulation of the detector where energy loss, production of secondary particles, multiple scattering, and decay can occur. As these particles pass through sensitive regions of the detector model, energy, charge, and angle information are used to calculate positions and idealized energy deposits in the detector. These quantities are stored for use in calculating the detector response.

Simulate detector response; add background events

At this stage the idealized position and energy information are transformed into realistic signals which mimic those collected from the detector electronics. In order to better mimic real data, random snapshots are taken of the detector occupancy. That is, at regular (but uncorrelated with the collisions of the two

beams) intervals, the status of all the detector elements is read out and saved as an underlying-background event. These real background events are mixed in with the simulated event to more closely reproduce signal data. The final output from this stage is stored as data for the reconstruction phase.

Reconstruct tracks and clusters

The simulated detector signals are combined into candidate events consisting of particle tracks, energy clusters, probable particle identifications, etc. From this stage on the MC and data are treated the same.

Note that there are two types of signal Monte Carlo data for the $B^{\pm} \to D_s^{*\pm}\phi$ modes, longitudinally polarized and transversely polarized. The $B^{\pm} \to D_s^{*\pm}\phi$ decay is a decay of a scalar meson to two vector mesons. Because of this the spin of the $D_s^{*\pm}$ and the ϕ can have two orientations. The transversely polarized particles have helicity ± 1 while the longitudinally polarized particles have helicity 0.

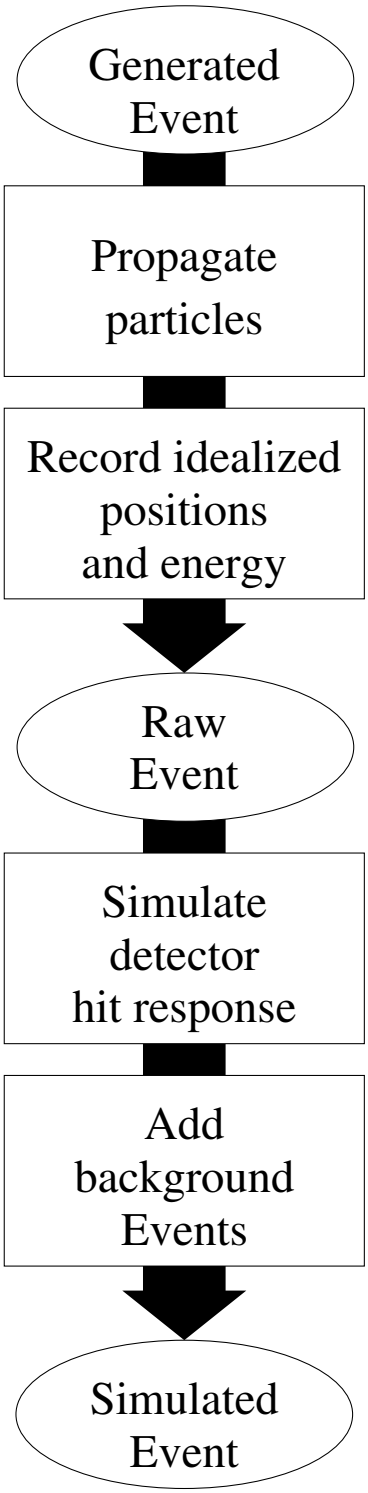

Figure 4.1: Stages of development for Monte Carlo simulated data.

Chapter 5

Roadmap to the analysis

This analysis is a search for a process that in the standard model has a cross section too small for any significant number of events to be seen in our data sample, see Figure 2.1 and Chapter 2. Thus, we tailor our analysis to place the best possible upper limit on the number of signal events observed and therefore the best upper limit on the branching fraction, $\mathcal{B}(B^{\pm} \to D_s^{(*)\pm}\phi)$.

We will perform this analysis blind. That is, we will not look at the on-resonance data sample in the final signal region prior to completing all studies of selection criteria, optimization, etc. In this way, we do not directly see the effects of any analysis choice on the data signal yield. Thus, we do not bias ourselves toward tailoring our analysis to get the largest (or the smallest) signal yield.

The analysis strategy is the following. First, in Section 6.1 we define the full decay chain of the B that we will use. Then we impose loose requirements on the various observables available in the analysis to reduce the size of the data

sample by removing events which are clearly background. We call this "level one selection"; it is described in Sections 6.2 and 6.3. The dataset we have at this point consists of a sample of $B^{\pm} \to D_s^{(*)\pm}\phi$ candidate events that is clean enough to make some observations about our selection efficiency and yields. This is done is Sections 6.4 and 6.5. At this stage in the analysis the main background is from continuum events (i.e. $e^+e^- \to q\overline{q}$), so we study the kinematic distributions and event rates to see how the continuum MC compares to signal MC and off-resonance data in Section 6.6. And in Section 6.7, we compare MC event rates and distributions to the on-resonance datasample in the sideband of our final fit variable.

There are many variables that distinguish between signal and background. Instead of making many cuts (one for each variable), we construct a single variable to cut on. We use a simple likelihood rather than a more sophisticated multivariate (e.g. neural net) because we do not have enough background Monte Carlo events to train a multivariate properly. The likelihood is described in Section 6.8. We optimize this likelihood requirement to get the best expected branching fraction upper limit (Section 6.9).

Now that the selection criteria are determined, in Section 6.10 we take a look at the on-resonance data sample. Note that our final result will be a fit to the number of events in a pre-defined signal region. At this stage we will not look at

the data in this region to keep our analysis blind. The final step before unblinding the signal region is to investigate possible backgrounds that could have a peaking structure in our signal region. We study these backgrounds in Section 6.11 by doing our analysis in a sideband of the D_s^+ mass.

Finally, we are ready to look at the on-resonance data sample and unblind our signal box. This is done in Chapter 7 and covers the results for both the $B^\pm \to D_s^\pm \phi$ and the $B^\pm \to D_s^{*\pm} \phi$ analyses using Run 1-4 data. The presence of a signal is determined by a fit to a signal-plus-background distribution. (Recall, the fit variables will be described in Section 6.2.) Systematic uncertainties due to normalizing PDG branching fractions, particle identification, track reconstruction, signal Monte Carlo statistics, cut efficiency, and determining the number of B mesons in our data sample are detailed in Chapter 8.

Finally in Chapter 9 we discuss the determination of the branching fractions of the $B^\pm \to D_s^{(*)\pm} \phi$ decays. The branching fraction is given by

$$\mathcal{B} = \frac{N_{\rm obs}}{N_{B\overline{B}} \sum_i \epsilon_i \cdot \mathcal{B}_i} \tag{5.1}$$

where $N_{\rm obs}$ is the observed number of events (or an upper limit of the number of events observed in which case the equality becomes an upper limit), $N_{B\overline{B}}$ is the

number of B mesons in our data sample, ϵ_i are the efficiencies for each $D_s^{\pm}$ decay mode, and $\mathcal{B}_i$ are the total branching fractions of each $D_s^{\pm}$ decay mode[1].

[1]This, of course, includes the $\mathcal{B}$ for the decay modes of the $D_s^{\pm}$, e.g. the $\mathcal{B}(K^{*0} \to K^+\pi^-)$; we call these the sub-decay modes.

Chapter 6

Event selection

6.1 Reconstructing the $B^{\pm} \to D_s^{(*)\pm}\phi$ decay

The decay modes used in the search for $B^{\pm} \to D_s^{(*)\pm}\phi$ and their branching fractions are given in Table 6.1. The total branching fraction of each $D_s^{(*)\pm}$ mode is summarized in Table 6.2. Note that the latest *BABAR* result for the branching fraction of $D_s^{\pm} \to \phi\pi^{\pm}$ has not yet been included in the world average published by the Particle Data Group (PDG) [16]. This branching fraction comes up in every part of the analysis since the branching fractions for the other $D_s^{\pm}$ decay modes are normalized to the $D_s^{\pm} \to \phi\pi^{\pm}$ branching fraction. Thus, we use this latest result, $\mathcal{B}(D_s^{\pm} \to \phi\pi^{\pm}) = (4.8 \pm 0.6)\%$ to calculate the total branching fractions for the secondary and tertiary decays of the B meson[1].

[1] To be complete, we redo all calculations using the 2004 PDG [16] world average for $\mathcal{B}(D_s^- \to \phi\pi^-) = 3.6 \pm 0.9$ in Appendix F.

Table 6.1: World-average and most recent branching fractions for all decays used in this analysis. The PDG values were used in the generation of the MC datasets. The most recent $D_s^{\pm}$ branching fractions are from *BABAR* [1].

Decay	Value (2004 PDG)	Value (latest)
$\phi \to K^+K^-$	0.491 ± 0.006	—
$D_s^{*\pm} \to D_s^{\pm}\gamma$	0.942 ± 0.025	—
$D_s^- \to \phi\pi^-$	0.036 ± 0.009	0.048 ± 0.006
$D_s^- \to K^-K_S^0$	0.036 ± 0.011	0.049 ± 0.010
$D_s^- \to K^{*0}K^-$	0.033 ± 0.009	0.044 ± 0.007
$K_S^0 \to \pi^+\pi^-$	0.6895 ± 0.0014	—
$K^0 \to K_S^0$	$1/2$	—
$K^{*0} \to K^+\pi^-$	$2/3$	—

Table 6.2: The total branching fractions for the secondary and tertiary decays in the modes considered in this analysis. The uncertainties listed are the correlated and uncorrelated, respectively.

B decay mode	$D_s^{*\pm}$ mode	$D_s^{\pm}$ mode	$D_s^{\pm}/D_s^{*\pm}$ branching fraction
$D_s^{\pm}\phi$	-	$\phi\pi^-$	$(1.16 \pm 0.15 \pm 0.02)\%$
$D_s^{\pm}\phi$	-	$K^{\pm}K_S^0$	$(0.82 \pm 0.10 \pm 0.13)\%$
$D_s^{\pm}\phi$	-	$\overline{K}^{*0}K^{\pm}$	$(1.45 \pm 0.18 \pm 0.14)\%$
$D_s^{*\pm}\phi$	$D_s^{\pm}\gamma$	$\phi\pi^-$	$(1.09 \pm 0.14 \pm 0.04)\%$
$D_s^{*\pm}\phi$	$D_s^{\pm}\gamma$	$K^{\pm}K_S^0$	$(0.77 \pm 0.10 \pm 0.13)\%$
$D_s^{*\pm}\phi$	$D_s^{\pm}\gamma$	$\overline{K}^{*0}K^{\pm}$	$(1.36 \pm 0.17 \pm 0.14)\%$

6.2 Kinematic variables for B-meson selection

In an experiment using electron-positron collisions the environment is very clean. At hadron colliders, the accelerated particle is not fundamental and therefore the energies of its constituents are unknown; but at e^+e^- colliders, we know precisely the energies of the two initial particles involved. Thus, we can use the beam energies to constrain the kinematics of the $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ event. The way to use this information is not unique and the most optimal will vary depending on the type of analysis performed. We use the energy substituted mass, $m_{\rm ES}$, and the difference in energy between the reconstructed B meson and the center of mass single-beam energy, ΔE:

$$\Delta E \equiv E_B^* - \frac{1}{2}\sqrt{s} = E_B^* - E_{\rm beam}^*, \tag{6.1}$$

$$m_{\rm ES} \equiv \sqrt{(E_{\rm beam}^*)^2 - (\vec{p}_B^*)^2} = \sqrt{(\frac{1}{2}s + \vec{p}_0 \cdot \vec{p}_B)^2/E_0^2 - \vec{p}_B^2}. \tag{6.2}$$

The asterisk denotes the center-of-momentum reference frame and the superscripts 0 and B refer to the initial $\Upsilon(4S)$ and B candidate, respectively. The $m_{\rm ES}$ and ΔE variables are nearly uncorrelated and will serve as our final fit variables that we use to determine our $B^\pm \to D_s^{(*)\pm}\phi$ yield. Note that the $\Upsilon(4S)$ has two main decay modes, $\Upsilon(4S) \to B^+B^-$ and $\Upsilon(4S) \to B^0\overline{B}^0$. For this analysis we do not retain the neutral B events; however, we do assume that $\mathcal{B}(\Upsilon(4S) \to B^+B^-) = \mathcal{B}(\Upsilon(4S) \to B^0\overline{B}^0)$.

6.3 Level one selection criteria

The reconstruction process consists of combining tracks and energy clusters in the *BABAR* detector into $D_s^{\pm}$ candidates, ϕ candidates, and, in the case of $B^{\pm} \to D_s^{*\pm}\phi$ ($D_s^{*\pm} \to D_s^{\pm}\gamma$), photon candidates. These daughter particle candidates are then combined into B-meson candidates. At each stage in the reconstruction, the measurement of the momentum vector of an intermediate particle is improved by refitting the momenta of the decay products with kinematic constraints. These constraints are based on the known mass [16] of the intermediate particle and on the fact that the decay products must originate from a common point in space.

The primary purpose of the level one selection is to reduce dramatically the size of the data set by weeding out events in which there is no chance of a signal candidate. The level one selection consists of a set of basic requirements (otherwise referred to as cuts) that are imposed at the very beginning of the analysis. These requirements are quite minimal, they will be tightened at a later stage.

In order to be considered for this analysis an event much satisfy the following level one requirements[2]

[2]Note that *BABAR* has a set of standard requirements for particle identification and track reconstruction. This is described in more detail in Appendix C and D. Here we simply mention that the categories for distinguishing kaons from pions to increasing degrees of confidence are `NotPion, Loose, Tight,` and `VeryTight`; and the categories for track reconstruction of increasing quality are `ChargedTracks, GoodTracksVeryLoose,` and `GoodTracksLoose.`

- All charged kaons must satisfy the particle identification requirements of `NotPion`[3] or better.

- Tracks from the $K^0_S \to \pi^+\pi^-$ decay must satisfy at least the track-quality requirements of the *BABAR* track reconstruction category, `ChargedTracks`[4].

- All other charged tracks must satisfy all the requirements necessary to make it into the `GoodTracksVeryLoose` particle identificaion category.[5]

- $M(K^0_S)$ within ± 9 MeV/c^2 ($\approx 3\sigma$) of the Particle Data Group (PDG) world average

- Three dimensional flight distance of the K^0_S must be $> 3\,\mathrm{mm}$ [6]

- $M(K^{*0})$ within ± 75 MeV/c^2 ($\approx 1.5\Gamma$) of the PDG world average

- $M(\phi)$ within ± 10 MeV/c^2 ($\approx 2.5\Gamma$) of the PDG world average

- $M(D^\pm_s)$ within ± 15 MeV/c^2 ($\approx 3\sigma$) of the PDG world average

- $\vec{p}^*(D^\pm_s) > 1.3$ GeV/c (Runs 1-3)

[3] See Appendix C for details on particle identification

[4] See Appendix D for details on track-quality definitions at *BABAR*

[5] *ibid.*

[6] The K^0_S has a relatively long lifetime and therefore does not always decay instantaneously. Thus, for any K^0_S with non-zero momentum, its production and decay do not occur at the same point. The distance between the production and decay of the K^0_S is known as its flight length. We can reduce drastically the backgrounds from the random combination of two pion tracks by requiring a flight length of at least few millimeters. The characteristic flight length of a K^0_S is 3 cm.

- $\vec{p}^*(D_s^{\pm}) > 1.6$ GeV/c (Run 4)[7]
- $m_{\rm ES} > 5.20$ GeV/c^2 [8]
- $|\Delta E| < 200$ MeV

For the $B^{\pm} \to D_s^{*\pm}\phi$ analysis the following are also required:

- The reconstructed mass difference, $\Delta m \equiv m(D_s^*) - m(D_s)$, is required to be $130 < \Delta m < 156\,\mathrm{MeV}/c^2$.
- Photons considered for the $D_s^{*\pm} \to D_s^{\pm}\gamma$ decay are reconstructed starting from an energy cluster in the EMC with the following added criteria:
  - Minimum energy: 0.03 GeV,
  - Maximum lateral moment: 0.8[9].

The PDG world averages for various particle masses are given in Table 6.3.

After requiring that the candidate events pass these criteria, there is a possibility that one event will contain more than one valid B-meson candidate. At this

[7] The difference between Runs 1-3 and Run 4 has no effect on the analysis since $D_s^{\pm}$ candidates for this analysis all have $\vec{p}^*(D_s^{\pm}) > 1.6\,\mathrm{GeV}/c$; between Run 3 and Run 4 the definition of a $D_s^{\pm}$ simply changed to reflect this.

[8] Note, for the off resonance data sample, the value of $m_{\rm ES}$ for each event was increased by 20 MeV, to keep the kinematic endpoint the same as in the on-resonance data.

[9] The lateral moment contains information about the azimuthal distribution (with respect to the particle's initial direction) of the shower shape in the EMC. The moment is zero for an energy distribution that is isotropic in azimuth. Hadronic showers are much more irregular than EM showers and thus have a much higher lateral moment.

Table 6.3: World-averages masses of particles used in this analysis. We also include the experimental resolutions for the $D_s^{\pm}$, the ϕ, and Δm, which are used in the best-B-candidate arbitration.

Variable	PDG world average Mass (MeV/c^2)	Mode	Experimental resolution (MeV/c^2)
M(K_S^0)	497.6	—	—
M(K^{*0})	896.1	—	—
M($D_s^{*\pm}$)	2112.1	—	—
		$D_s^- \to \phi\pi^-$	5.0
M($D_s^{\pm}$)	1968.5	$D_s^- \to K^- K_S^0$	6.4
		$D_s^- \to K^{*0} K^-$	5.3
M(ϕ)	1019.5	—	5.7
Δm	4.6	All $D_s^{*\pm}$ modes	144

point we choose the best possible candidate to pass on to the next stage of the analysis. Arbitration between multiple candidates is performed using a χ^2 test in which the best B candidate is defined as that candidate with the smallest value of

$$\chi^2_{D_s} = \frac{(m(D_s) - m_{\rm PDG})^2}{\sigma^2_{\rm exp}} + \frac{(m(\phi) - m_{\rm PDG})^2}{\sigma^2_{\rm exp}}. \tag{6.3}$$

Here, m(D_s) and m(ϕ) are the reconstructed $D_s^{\pm}$ and bachelor ϕ[10] masses, respectively, $m_{\rm PDG}$ represents the Particle Data Group world average for that value, and $\sigma_{\rm exp}$ is the experimental resolution.

Arbitration between multiple candidates is again performed using a χ^2 as in the $B^{\pm} \to D_s^{\pm}\phi$ channel, but with the addition of a term for the Δm mass difference,

[10] To avoid confusion with the ϕ in the $D_s^- \to \phi\pi^-$ decay, we will refer to the ϕ from the $B^{\pm} \to D_s^{(*)\pm}\phi$ decay as the "bachelor ϕ."

$$\chi^2_{D_s^*} = \chi^2_{D_s} + \frac{(\Delta m - \Delta m_{\rm PDG})^2}{\sigma^2_{\rm exp}}, \tag{6.4}$$

where $\chi^2_{D_s}$ is the χ^2 for the $B^\pm \to D_s^\pm \phi$ analysis defined above.

6.4 Level one selection efficiency

The selection efficiency can be studied using signal MC data. The analysis performed in this and the remaining sections of this chapter were done on the Runs 1-3 dataset. These numbers are used in the process of understanding the MC and its relation to the data, but never directly used in the calculation of the final result.

In the final stages of this analysis, the $m_{\rm ES}$ and ΔE selection criteria will be tightened. Thus it makes sense to anticipate these additional requirements and impose a tighter selection on these variables for this study so that we may get a better feel for MC and data that will fall within our final signal region. The level one efficiencies on signal MC we report here include the tighter requirements (otherwise referred to as the "signal box") $m_{\rm ES} > 5.27$ GeV/c^2 and $|\Delta E| < 30$ MeV ($\approx 3\sigma$)[11]. These efficiencies, summarized in Table 6.4, are calculated as the ratio of the number of reconstructed events passing all requirements to the number

[11] These tighter restrictions are very close to what will later be our final cuts in $m_{\rm ES}$ and ΔE. They carve out a box in the $m_{\rm ES}$-ΔE plane and we therefore refer to it as the signal-box.

Table 6.4: Signal Monte Carlo data efficiencies for the level one selection requirements, including the tighter requirements $m_{\rm ES} > 5.27$ GeV/c^2 and $|\Delta E| < 30$ MeV. PID and tracking corrections have not been applied. Uncertainties are statistical only. Each $D_s^{*\pm}$ polarization is listed.

B decay mode	D_s decay mode	MC efficiency
$B^- \to D_s^- \phi$	$D_s^- \to \phi\pi^-$	26.1 ± 0.4 %
$B^- \to D_s^- \phi$	$D_s^- \to K^- K_S^0$	24.4 ± 0.4 %
$B^- \to D_s^- \phi$	$D_s^- \to K^{*0} K^-$	21.6 ± 0.3 %
$B^- \to D_s^{*-} \phi$ (longitudinal)	$D_s^- \to \phi\pi^-$	14.4 ± 0.3 %
$B^- \to D_s^{*-} \phi$ (longitudinal)	$D_s^- \to K^- K_S^0$	13.4 ± 0.3 %
$B^- \to D_s^{*-} \phi$ (longitudinal)	$D_s^- \to K^{*0} K^-$	11.6 ± 0.2 %
$B^- \to D_s^{*-} \phi$ (transverse)	$D_s^- \to \phi\pi^-$	14.8 ± 0.3 %
$B^- \to D_s^{*-} \phi$ (transverse)	$D_s^- \to K^- K_S^0$	14.5 ± 0.3 %
$B^- \to D_s^{*-} \phi$ (transverse)	$D_s^- \to K^{*0} K^-$	12.4 ± 0.2 %

of generated events. These efficiencies do not yet include particle identification (PID) and tracking corrections.

6.5 Level one selection yields

We would like to examine the yields of events passing the level one selection requirements in continuum and $B\overline{B}$ MC. We compare these yields in two categories: one, imposing just the level one selection criteria; and two, also applying the tighter restrictions, $m_{\rm ES} > 5.27$ GeV/c^2 and $|\Delta E| < 30$ MeV. We can also use these numbers to get a preview of whether we will encounter peaking back-

grounds[12]. If we see no enhancement of events in the signal region versus the larger $m_{\rm ES}$ vs. ΔE region allowed by the level one selection we can expect to have small peaking backgrounds.

The yields are summarized in Tables 6.5 and 6.6. There are some immediate, not too surprising, conclusions that can be drawn:

- The main backgrounds are from continuum, mostly $e^+e^- \to c\overline{c}$.
- There is no sign of peaking backgrounds in the $B\overline{B}$ Monte Carlo datasets.
- The $D_s^- \to K^{*0}K^-$ mode is the least pure mode.

In Table 6.7 we compare the continuum MC yields to the actual off-resonance data yields. (The MC yields are the same as those listed in the last column of Tables 6.5 and 6.6.) We note that that the MC overestimates the background levels, especially in the cleaner $D_s^\pm$ modes ($D_s^- \to \phi\pi^-$ and, to some extent, $D_s^- \to K^-K_S^0$).

[12]Peaking background events are events that come from particular real B mesons. Not all B backgrounds are considered peaking, however some decays of the B could show a peaking structure in the $m_{\rm ES}$ and ΔE signal regions. This is in contrast to other backgrounds (non-peaking B decays and continuum) that tend to have a flat distribution in the signal region. Peaking backgrounds will be discussed in Section 6.11

Table 6.5: $B^{\pm} \to D_s^{\pm}\phi$. Number of events passing various levels of selection requirements in Runs 1-3 MC: 1) The level one selection requirements and 2) the level one selection requirements plus $m_{\rm ES} > 5.27$ GeV/c^2 and $|\Delta E| < 30$ MeV (the "signal box"). The generated number of events in the MC samples are listed in Table 4.1. "On res" and "Off res" indicate the on-resonance and off-resonance datasets.

B mode	D_s mode	MC Dataset	L1 Selection MC events	L1 Selection expected on res.	Signal Box MC events	Signal Box expected on res.	L1 Selection expected off res.
$D_s^{\pm}\ \phi$	$\phi\pi^-$	$c\bar{c}$	273	205±12	5	4 ± 2	22 ± 1
$D_s^{\pm}\ \phi$	$K^-K_S^0$	$c\bar{c}$	252	190±12	5	4 ± 2	20 ± 1
$D_s^{\pm}\ \phi$	$K^{*0}K^-$	$c\bar{c}$	741	555±20	22	17 ± 4	59 ± 2
$D_s^{\pm}\ \phi$	$\phi\pi^-$	uds	33	24±4	1	1 ± 1	3 ± 1
$D_s^{\pm}\ \phi$	$K^-K_S^0$	uds	80	58±6	1	1 ± 1	6 ± 1
$D_s^{\pm}\ \phi$	$K^{*0}K^-$	uds	247	179±11	6	4 ± 2	18 ± 1
$D_s^{\pm}\ \phi$	$\phi\pi^-$	$B^0\overline{B}^0$	13	6±2	0	< 1.1 90% CL	N.A.
$D_s^{\pm}\ \phi$	$K^-K_S^0$	$B^0\overline{B}^0$	31	15±3	0	< 1.1 90% CL	N.A.
$D_s^{\pm}\ \phi$	$K^{*0}K^-$	$B^0\overline{B}^0$	70	34±4	2	1 ± 1	N.A.
$D_s^{\pm}\ \phi$	$\phi\pi^-$	$B^+\ B^-$	21	10±2	1	0.5 ± 0.5	N.A.
$D_s^{\pm}\ \phi$	$K^-K_S^0$	$B^+\ B^-$	27	13±2	1	0.5 ± 0.5	N.A.
$D_s^{\pm}\ \phi$	$K^{*0}K^-$	$B^+\ B^-$	113	55±5	4	2 ± 1	N.A.

Table 6.6: $B^{\pm} \to D_s^{*\pm}\phi$. Number of events passing various levels of selection requirements in Runs 1-3 MC: 1) The level one selection requirements and 2) the level one selection requirements plus $m_{\rm ES} > 5.27$ GeV/c^2 and $|\Delta E| < 30$ MeV (the "signal box". The generated number of events in the MC samples are listed in Table 4.1. "On res" and "Off res" indicate the on-resonance and off-resonance datasets.

B mode	D_s mode	MC Dataset	L1 Selection MC events	L1 Selection expected on res.	Signal Box MC events	Signal Box expected on res.	L1 Selection expected off res.
$D_s^{*\pm}\,\phi$	$\phi\pi^-$	$c\bar{c}$	170	128±10	6	5 ± 2	14 ± 1
$D_s^{*\pm}\,\phi$	$K^-K_S^0$	$c\bar{c}$	141	106±9	3	2 ± 1	11 ± 1
$D_s^{*\pm}\,\phi$	$K^{*0}K^-$	$c\bar{c}$	436	328±16	12	9 ± 3	35 ± 2
$D_s^{*\pm}\,\phi$	$\phi\pi^-$	uds	21	15±3	0	< 1.7 90% CL	1.6 ± 0.3
$D_s^{*\pm}\,\phi$	$K^-K_S^0$	uds	49	36±5	2	1.5 ± 1.1	3.8 ± 0.5
$D_s^{*\pm}\,\phi$	$K^{*0}K^-$	uds	121	89±8	1	1 ± 1	9 ± 1
$D_s^{*\pm}\,\phi$	$\phi\pi^-$	$B^0\overline{B}^0$	20	9±2	1	0.5 ± 0.5	N.A.
$D_s^{*\pm}\,\phi$	$K^-K_S^0$	$B^0\overline{B}^0$	27	13±3	0	< 1.1 90% CL	N.A.
$D_s^{*\pm}\,\phi$	$K^{*0}K^-$	$B^0\overline{B}^0$	57	27±4	1	0.5 ± 0.5	N.A.
$D_s^{*\pm}\,\phi$	$\phi\pi^-$	$B^+\,B^-$	19	9±2	0	< 1.1 90% CL	N.A.
$D_s^{*\pm}\,\phi$	$K^-K_S^0$	$B^+\,B^-$	23	11±2	1	0.5 ± 0.5	N.A.
$D_s^{*\pm}\,\phi$	$K^{*0}K^-$	$B^+\,B^-$	106	49±5	2	0.9 ± 0.6	N.A.

Table 6.7: Comparison of the number of events passing the level one selection in the continuum MC (see Table 6.5) and off-resonance data.

B mode	D_s mode	$uds+c\overline{c}$ MC prediction	Off-resonance data
$D_s^{\pm}\phi$	$\phi\pi^-$	25 ± 2	9
$D_s^{\pm}\phi$	$K^-K_S^0$	26 ± 2	14
$D_s^{\pm}\phi$	$K^{*0}K^-$	77 ± 2	64
$D_s^{*\pm}\phi$	$\phi\pi^-$	16 ± 1	6
$D_s^{*\pm}\phi$	$K^-K_S^0$	15 ± 1	7
$D_s^{*\pm}\phi$	$K^{*0}K^-$	44 ± 3	38

6.6 Kinematics of events that pass Level One selection

We now compare various kinematic distributions of signal, background, and off-resonance Monte Carlo data for events that pass the level one selection. As seen in Section 6.5, Tables 6.5 and 6.6, our biggest backgrounds will be from continuum events, as opposed to backgrounds from B decays other than $B^{\pm} \to D_s^{(*)\pm}\phi$. We now introduce two discriminating variables that are based on the distribution of charged tracks and neutral EMC clusters in the event that are very efficient at reducing backgrounds from continuum events

At *BABAR*, the $\Upsilon(4S)$ is created at rest in the center of mass frame; and the mass of the two daughter B mesons is almost equal to the rest mass of the $\Upsilon(4S)$. Because of this the B mesons are moving very slowly in the center of mass frame. Thus when the B mesons decay they do so isotropically in the center of mass frame and the event as a whole has a spherical distribution of reconstructed charged tracks and neutral EMC clusters.

The tracks-and-clusters distribution for a continuum event is very different. In the case of an $e^+e^- \to q\overline{q}$ continuum event, the two quarks in final state system must be moving away from each other rather quickly in order to conserve momentum. Then, as each quark hadronizes the tracks and clusters from these

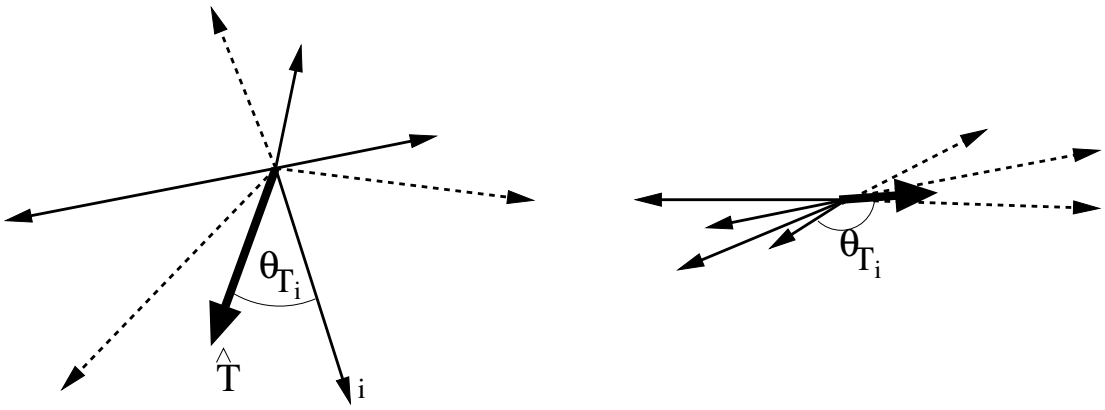

Figure 6.1: The thrust angle, $\theta_{\rm T}$ is represented by the thick solid arrow. The dotted arrows represent the tracks and neutral EMC clusters in the center of mass frame that are used to make up the B we reconstruct as $B^{\pm} \to D_s^{(*)\pm}\phi$. The solid arrows are the other tracks and clusters in the event. The left figure is an example of a $B\overline{B}$ event. Notice that the event has an overall circular shape. The figure on the right is an example of a $e^+e^- \to q\overline{q}$, continuum, event. Notice that the continuum event has much less circular shape.

hadrons tend to be grouped in a cone around the momentum vector of the parent quark. In this case, we say that the tracks and clusters have a jet-like distribution.

In Figure 6.6 we have drawn an example of the spherical and jet-like distribution of tracks and clusters. On the left is a spherical, $\Upsilon(4S) \to B\overline{B}$, event, while on the right is a jet-like $e^+e^- \to q\overline{q}$ event. In order to take advantage of the difference in event shape, we define two event shape variables, the B thrust angle and the Legendre Fisher.

The Legendre Fisher (or simply the Fisher) is a commonly used tool at *BABAR*. A Fisher Discriminant [17] is used to maximize the discrimination power between

two or more variables that are correlated to any degree. The two variables we are trying to distinguish are the momentum-weighted Legendre moments, L_0 and L_2. Hence the name the Legendre Fisher. It is given by[13],

$$\mathcal{F} = C_0 L_0 + C_2 L_2, \tag{6.5}$$

where

$$L_0 = \sum_i^{\text{r.o.e.}} p_i \tag{6.6}$$

$$L_2 = \sum_i^{\text{r.o.e.}} p_i \cos^2 \theta_{Tj}. \tag{6.7}$$

The term "r.o.e." stands for the rest of the event. The sum is over all the tracks and neutral clusters in each event that are not associated with the B that we are reconstructing. The B thrust angle, θ_T, is the angle that the track or neutral cluster makes with the thrust vector. The thrust vector is defined as the vector that satisfies $max(|\Sigma_j \hat{T} \cdot \vec{p}_j|)/(\Sigma_j |\vec{p}_j|)$ for any group of tracks and clusters, j. The thrust vector is defined over all the tracks and clusters in the event. Since we want to maximize the discrimination between continuum events and signal events and the main component of the continuum background is from $e^+e^- \to c\overline{c}$ events, we determined the constants C_0 and C_2 from signal and $c\overline{c}$ MC.

Figures 6.2-6.12 for the $B^\pm \to D_s^\pm \phi$ mode and Figures 6.13-6.23 for the $B^\pm \to D_s^{*\pm} \phi$ mode show the distributions for the variables that we will use in

[13]See Appendix A for more details

this analysis. Some observations about these kinematic distributions are summarized below[14]

- Background MC and off-resonance data show a hint of a $D_s^{\pm}$ peak (Figures 6.2 and 6.13), but a large fraction of the reconstructed $D_s^{\pm}$ are actually just combinatoric.

- The invariant mass distributions for the bachelor ϕ (Figures 6.3 and 6.14) show that many bachelor ϕ candidates in background events are real. Thus, we do not expect to gain much background rejection by cutting tighter on the bachelor ϕ mass or by requiring more stringent kaon particle identification requirements on the bachelor ϕ daughters.

- It appears that the background bachelor ϕ mesons are largely unpolarized (Figures 6.4 and 6.15). The decay (helicity) angle[15] of the ϕ candidate is then a good discriminant against backgrounds in the $B^{\pm} \to D_s^{\pm}\phi$ mode, where the ϕ has a known polarization.

- The $\cos\theta_B$ distributions. The angle θ_B is the polar angle of the B candidate in the CM frame[16]. It should be roughly flat for combinatoric backgrounds

[14] For all these plots, the background MC normalized to the on-resonance data luminosity for Runs 1-3 only, however since Run 4 almost exactly doubled our luminosity, one can easily make the conversion. The shapes of the distributions do not change between Runs 1-3 and Run 4.

[15] See Appendix B for a definition of the decay angle.

[16] See Appendix B for a definition of the angle θ_B.

and $1 - \cos^2\theta_B$ for real $B\overline{B}$ events. Figures 6.5 and 6.16 show that this is indeed what we see.

- The continuum MC distributions of $\cos(\theta_{\rm T})$ peak at $|\cos(\theta_{\rm T})| = 1$ while the signal MC is relatively flat. (Figures 6.6, and 6.17). This is the expected behavior of continuum and signal events.

- Figures 6.7 and 6.18 show that we get good discrimination between the signal and background with the Legendre Fisher variable.

- The distributions of decay (helicity) angle for the ϕ from $D_s^- \to \phi\pi^-$ show that backgrounds are largely uniform in the cosine of this angle (Figures 6.9 and 6.20).

- It seems that most of the K_S^0 candidates in $D_s^- \to K^- K_S^0$ are real K_S^0 (Figures 6.10 and 6.21) Thus, there is no need to tighten the requirements on the K_S^0 further.

- Most of the K^{*0} candidates in $D_s^- \to K^{*0} K^-$ appear to be combinatoric (Figures 6.11 and 6.22). Further study is necessary to determine whether tighter kaon particle identification requirements should be used.

- The decay (helicity) angle of the K^{*0} candidate (Figures 6.12 and 6.23) appears to be a useful variable for background rejection.

- In the $B^{\pm} \to D_s^{*\pm}\phi$ mode, the background MC and the off resonance data do not show a peak in Δm distribution, see Figure 6.24. Thus, it seems that most $D_s^{*\pm}$ candidates are combinatoric.

- In the $B^{\pm} \to D_s^{*\pm}\phi$ mode, the energy of the photon from $D_s^{*\pm} \to D_s^{\pm}\gamma$ in background MC and off resonance data is on average lower than expected for signal (Figure 6.25). This is also what one would expect for combinatorics.

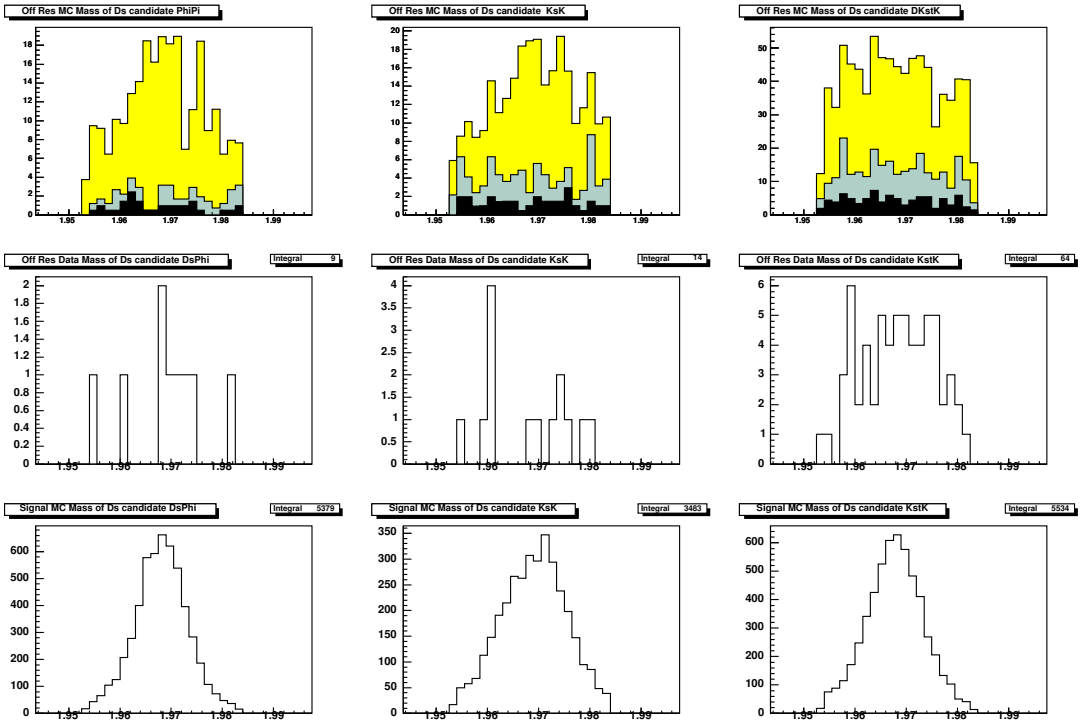

Figure 6.2: $B^{\pm} \to D_s^{\pm}\phi$. Mass of the $D_s^{\pm}$ candidate for all events that pass the level one selection criteria. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, $D_s^- \to K^{*0}K^-$.
Top row: Background MC, normalized to on-resonance data luminosity, black=$B\overline{B}$, blue=uds, yellow=$c\overline{c}$.
Middle row: Off-resonance data.
Bottom row: Signal MC, arbitrary normalization.

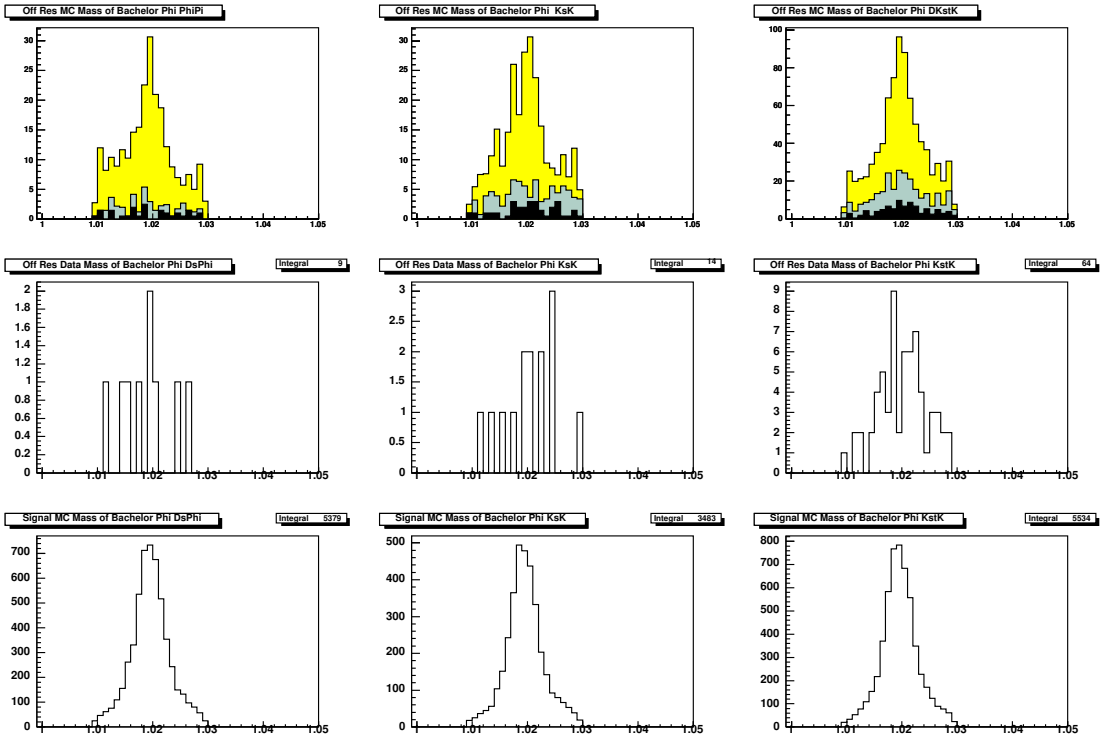

Figure 6.3: $B^{\pm} \to D_s^{\pm}\phi$. Mass of the bachelor ϕ candidate for all events that pass the level one selection criteria. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, $D_s^- \to K^{*0}K^-$.
Top row: Background MC, normalized to on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\bar{c}$.
Middle row: Off-resonance data.
Bottom row: Signal MC, arbitrary normalization.

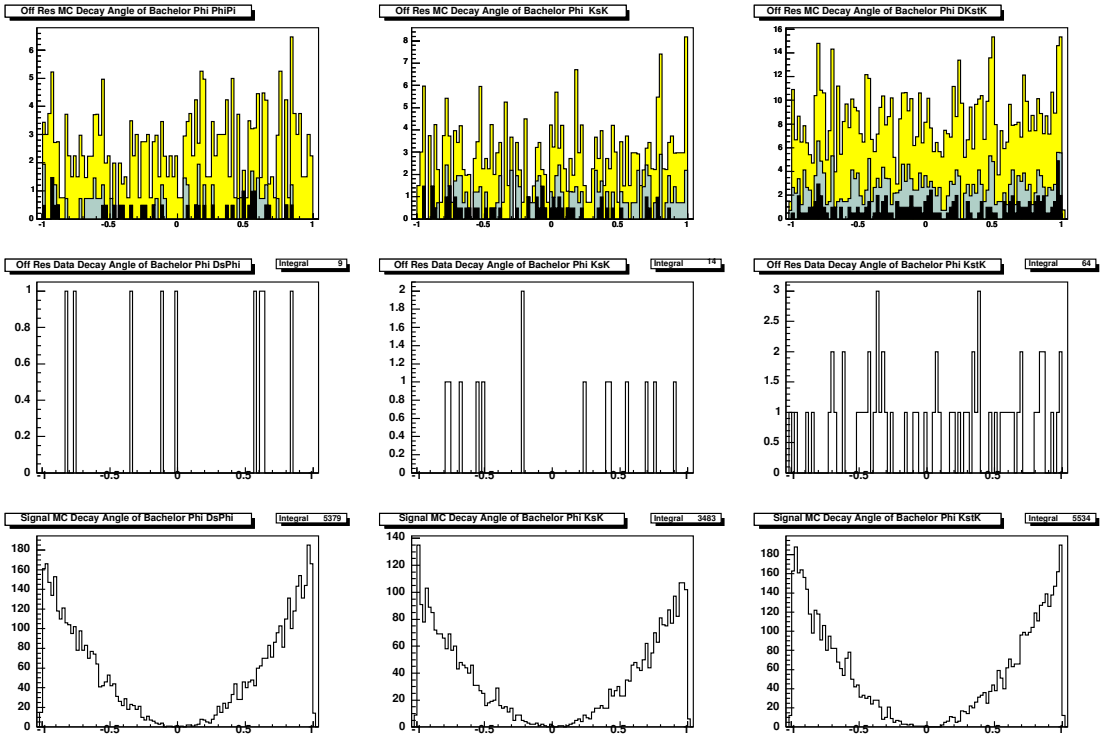

Figure 6.4: $B^{\pm} \to D_s^{\pm}\phi$. Decay (helicity) angle of the bachelor ϕ candidate for all events that pass the level one selection criteria. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^- K_S^0$, $D_s^- \to K^{*0} K^-$.
Top row: Background MC, normalized to on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\bar{c}$.
Middle row: Off-resonance data.
Bottom row: Signal MC, arbitrary normalization.

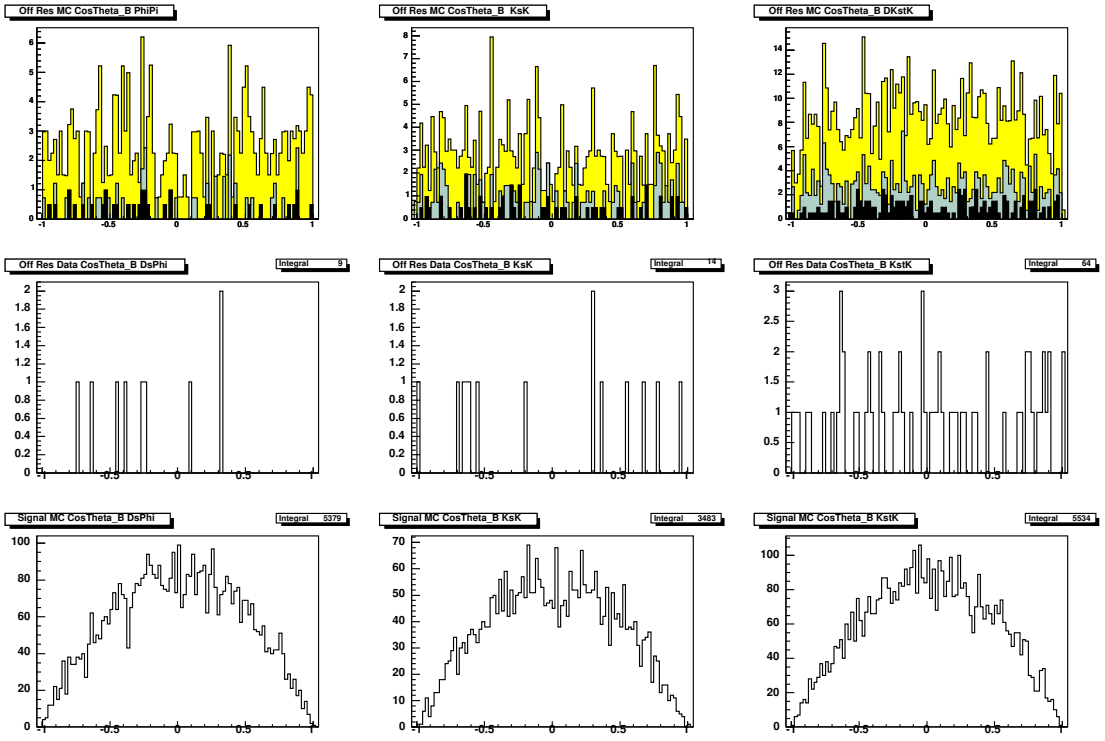

Figure 6.5: $B^{\pm} \to D_s^{\pm}\phi$. Cosine of the CM polar angle of the B candidate (See Appendix B) for all events that pass the level one selection criteria. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^- K_S^0$, $D_s^- \to K^{*0} K^-$.
Top row: Background MC, normalized to on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\bar{c}$.
Middle row: Off-resonance data.
Bottom row: Signal MC, arbitrary normalization.

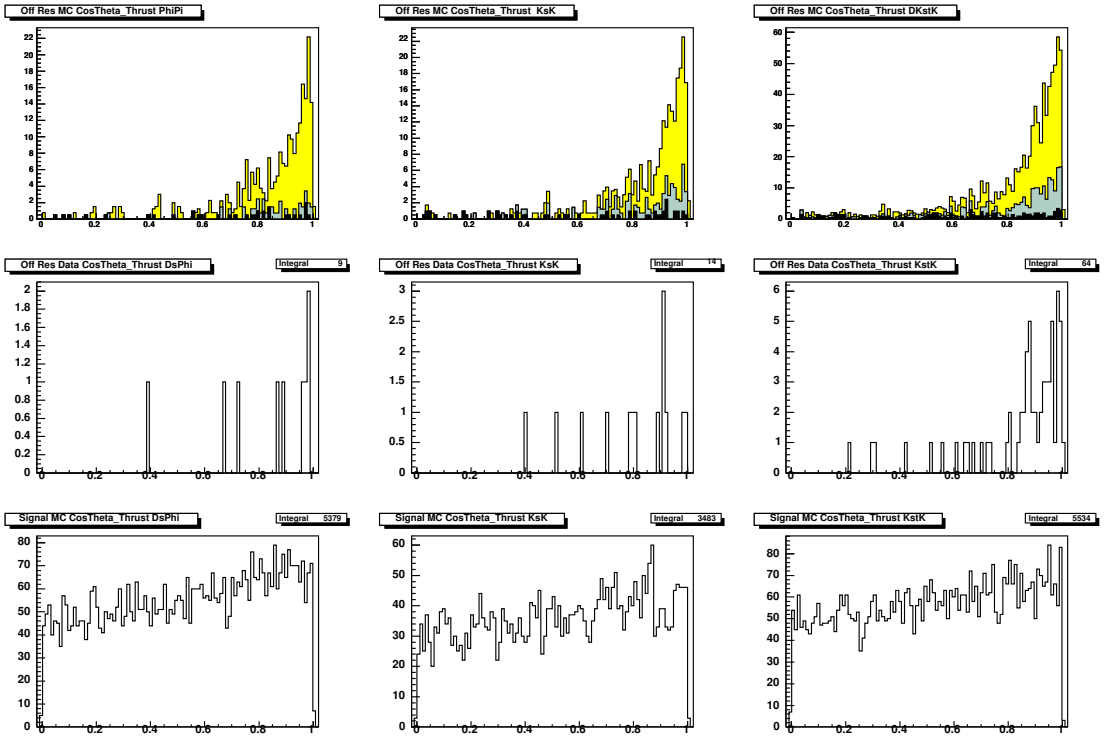

Figure 6.6: $B^{\pm} \to D_s^{\pm}\phi$. Cosine of the angle between the thrusts of the B candidate and of the rest-of-the-event for all events that pass the level one selection criteria. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^- K_S^0$, $D_s^- \to K^{*0} K^-$.
Top row: Background MC, normalized to on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\overline{c}$.
Middle row: Off-resonance data.
Bottom row: Signal MC, arbitrary normalization.

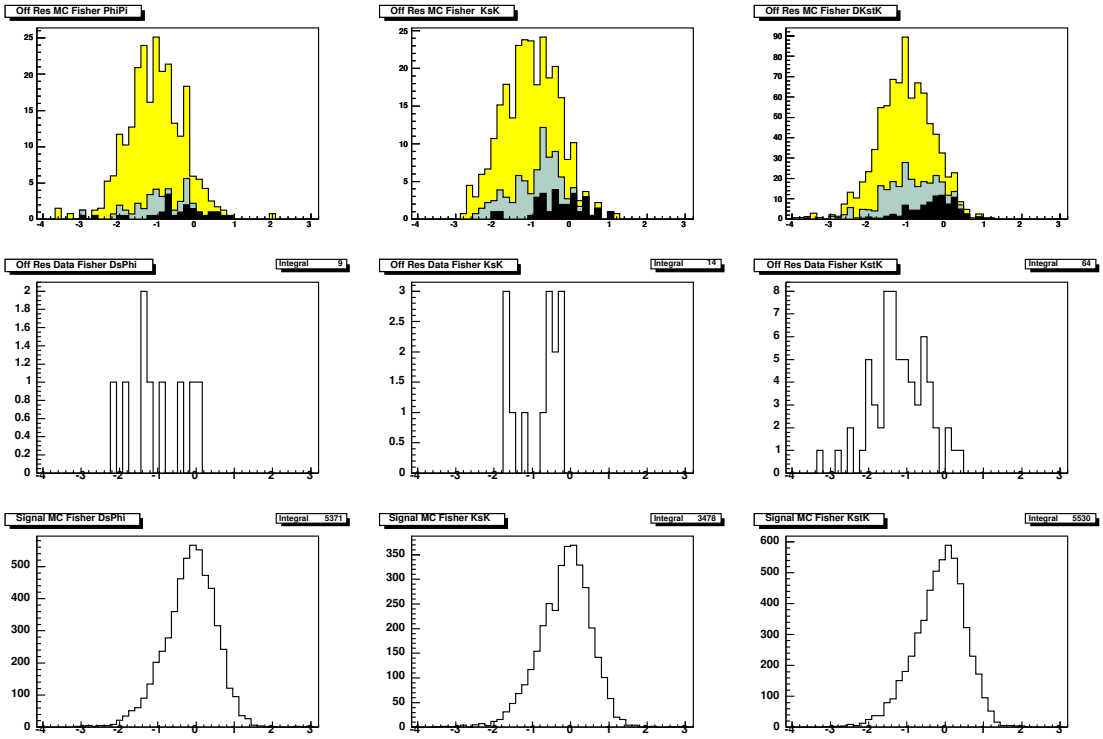

Figure 6.7: $B^{\pm} \to D_s^{\pm}\phi$. Legendre FisherA for all events that pass the level one selection criteria. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, $D_s^- \to K^{*0}K^-$.
Top row: Background MC, normalized to on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\overline{c}$.
Middle row: Off-resonance data.
Bottom row: Signal MC, arbitrary normalization.

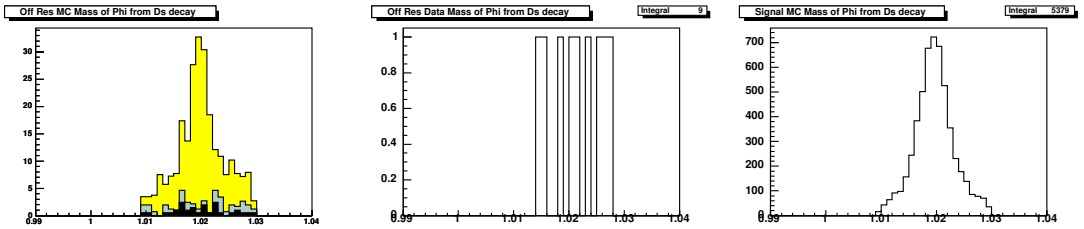

Figure 6.8: $B^{\pm} \to D_s^{\pm}\phi$. Mass of the ϕ from $D_s^- \to \phi\pi^-$ for all events that pass the level one selection criteria.
Left: Background MC, normalized to the on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\overline{c}$.
Middle: Off-resonance data.
Right: Signal MC, arbitrary normalization.

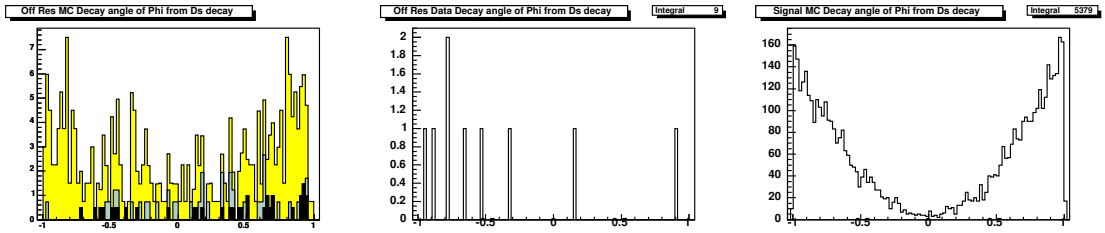

Figure 6.9: $B^{\pm} \to D_s^{\pm}\phi$. Decay (helicity) angle of the ϕ from $D_s^- \to \phi\pi^-$ for all events that pass the level one selection criteria.
Left: Background MC, normalized to the on-resonance luminosity, black=$B\overline{B}$, blue=uds, yellow=$c\overline{c}$.
Middle: Off-resonance data.
Right: Signal MC, arbitrary normalization.

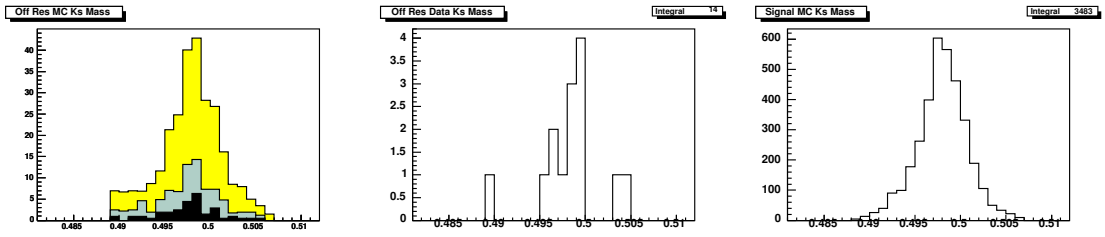

Figure 6.10: $B^{\pm} \to D_s^{\pm}\phi$. Mass of the K_S^0 from $D_s^- \to K^- K_S^0$ for all events that pass the level one selection criteria.
Left: Background MC, normalized to the on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\overline{c}$.
Middle: Off-resonance data.
Right: Signal MC, arbitrary normalization.

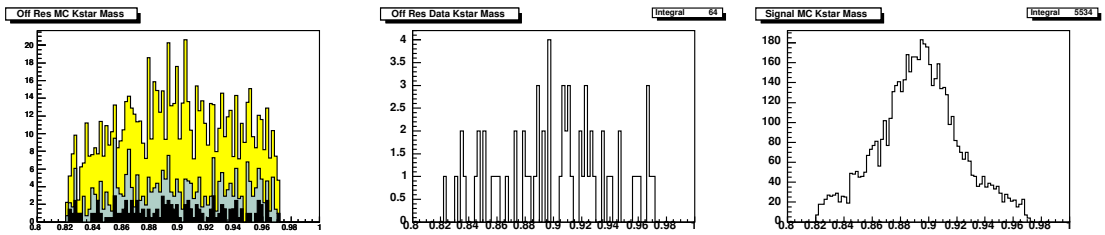

Figure 6.11: $B^{\pm} \to D_s^{\pm}\phi$. Mass of the K^{*0} candidate from $D_s^- \to K^{*0} K^-$ for all events that pass the level one selection criteria.
Left: Background MC, normalized to the on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\overline{c}$.
Middle: Off-resonance data.
Right: Signal MC, arbitrary normalization.

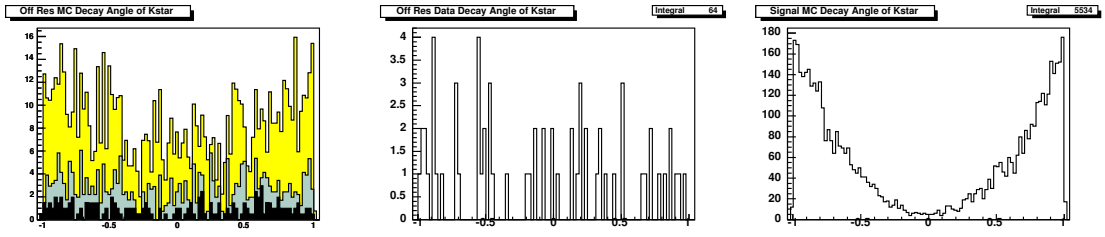

Figure 6.12: $B^{\pm} \to D_s^{\pm}\phi$. Decay (helicity) angle of the K^{*0} from $D_s^- \to K^{*0}K^-$ for all events that pass the level one selection criteria.
Left: Background MC, normalized to the on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\overline{c}$.
Middle: Off-resonance data.
Right: Signal MC, arbitrary normalization.

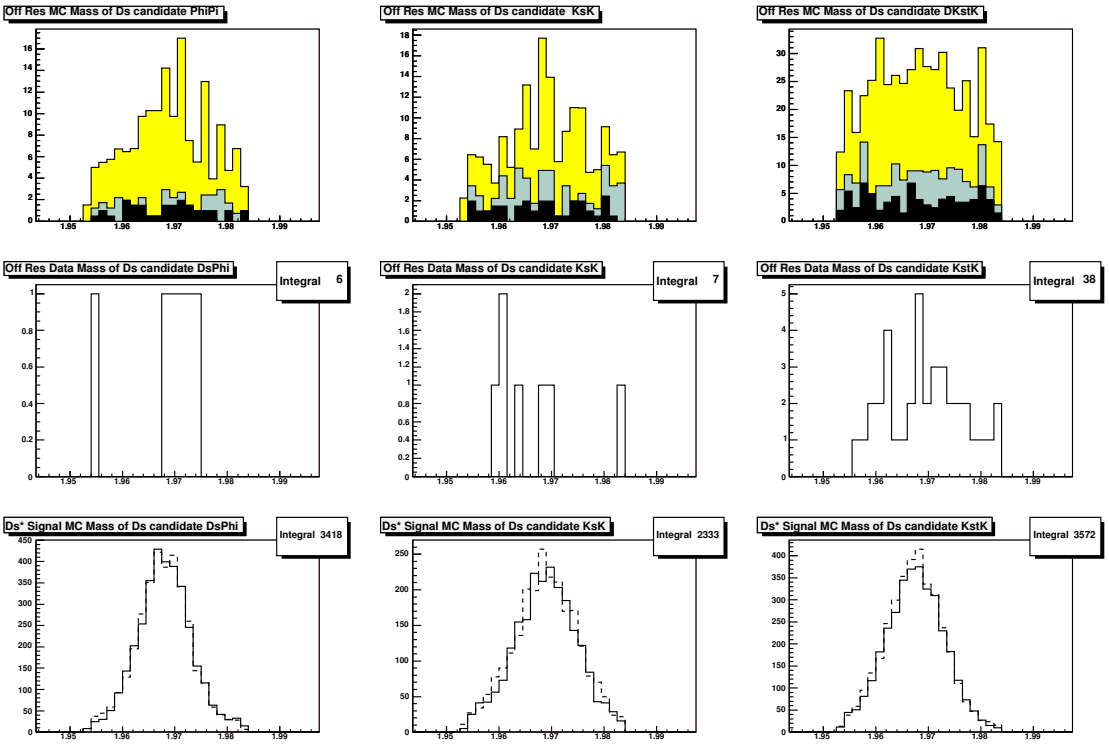

Figure 6.13: $B^{\pm} \to D_s^{*\pm}\phi$. Mass of the $D_s^{\pm}$ candidate for all events that pass the level one selection criteria. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, $D_s^- \to K^{*0}K^-$.
Top row: Background MC, normalized to on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\overline{c}$.
Middle row: Off-resonance data.
Bottom row: Signal MC, arbitrary normalization. Solid line: longitudinal polarization, dashed line: transverse polarization.

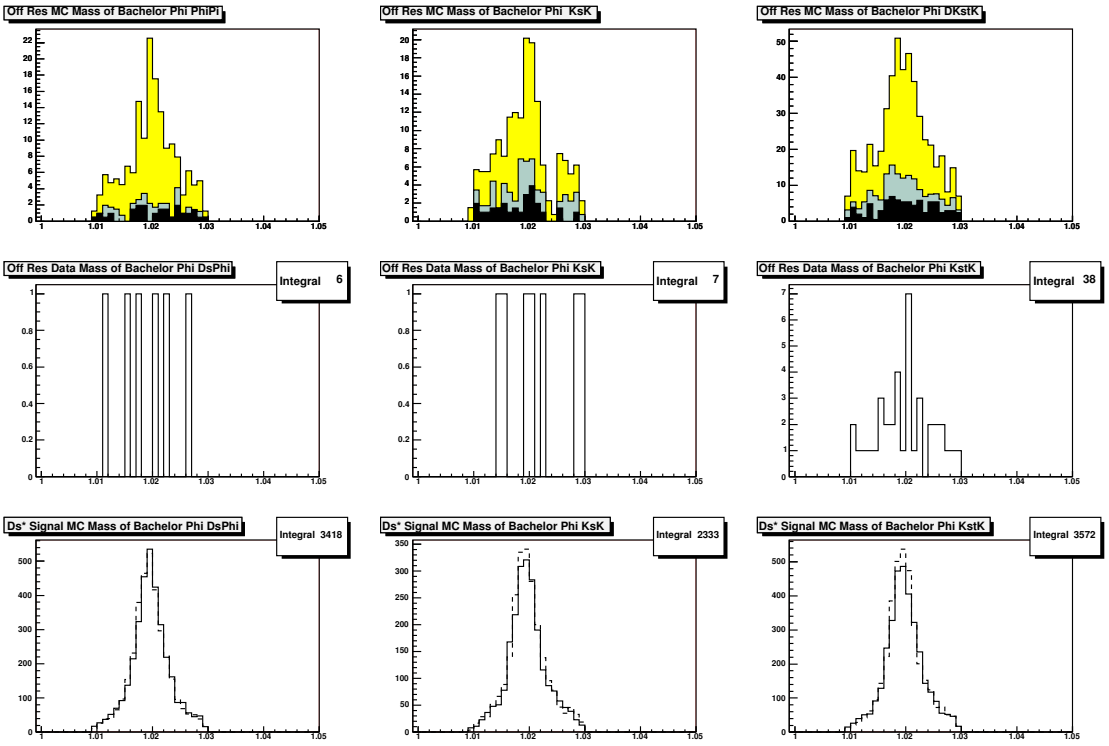

Figure 6.14: $B^{\pm} \to D_s^{*\pm}\phi$. Mass of the bachelor ϕ candidate for all events that pass the level one selection criteria. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^- K_S^0$, $D_s^- \to K^{*0} K^-$.
Top row: Background MC, normalized to on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\bar{c}$.
Middle row: Off-resonance data.
Bottom row: Signal MC, arbitrary normalization. Solid line: longitudinal polarization, dashed line: transverse polarization.

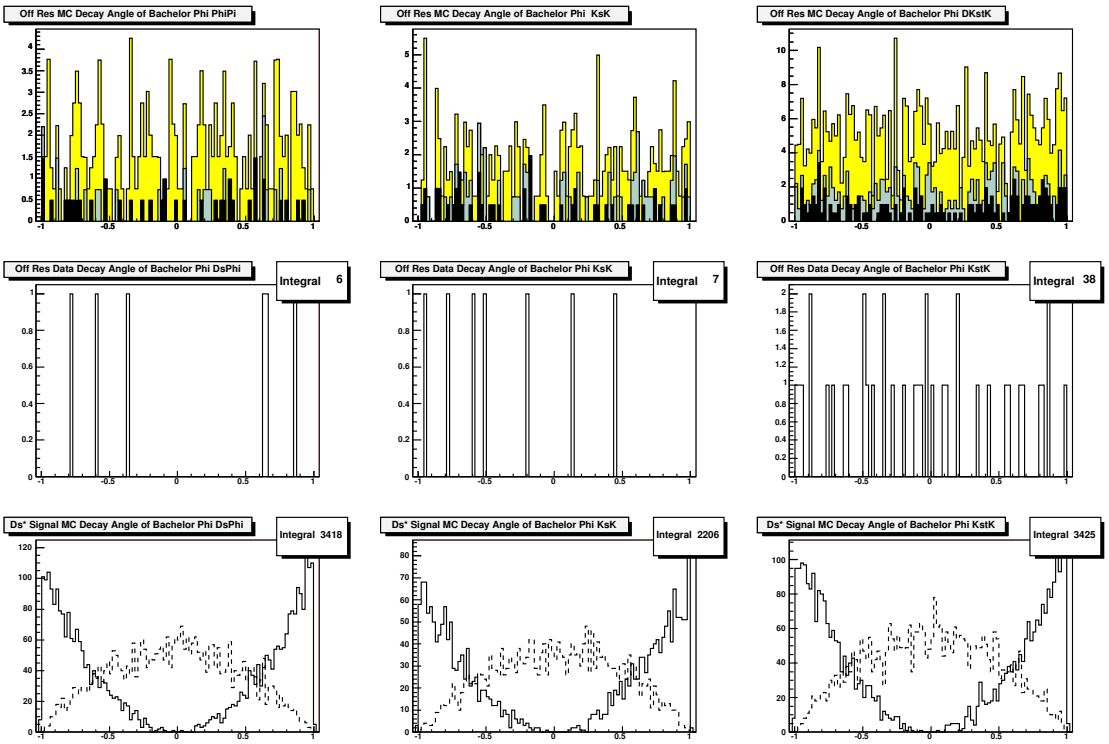

Figure 6.15: $B^{\pm} \to D_s^{*\pm}\phi$. Decay (helicity) angle of the bachelor ϕ candidate for all events that pass the level one selection criteria. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^- K_S^0$, $D_s^- \to K^{*0}K^-$.
Top row: Background MC, normalized to on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\bar{c}$.
Middle row: Off-resonance data.
Bottom row: Signal MC, arbitrary normalization. Solid line: longitudinal polarization, dashed line: transverse polarization.

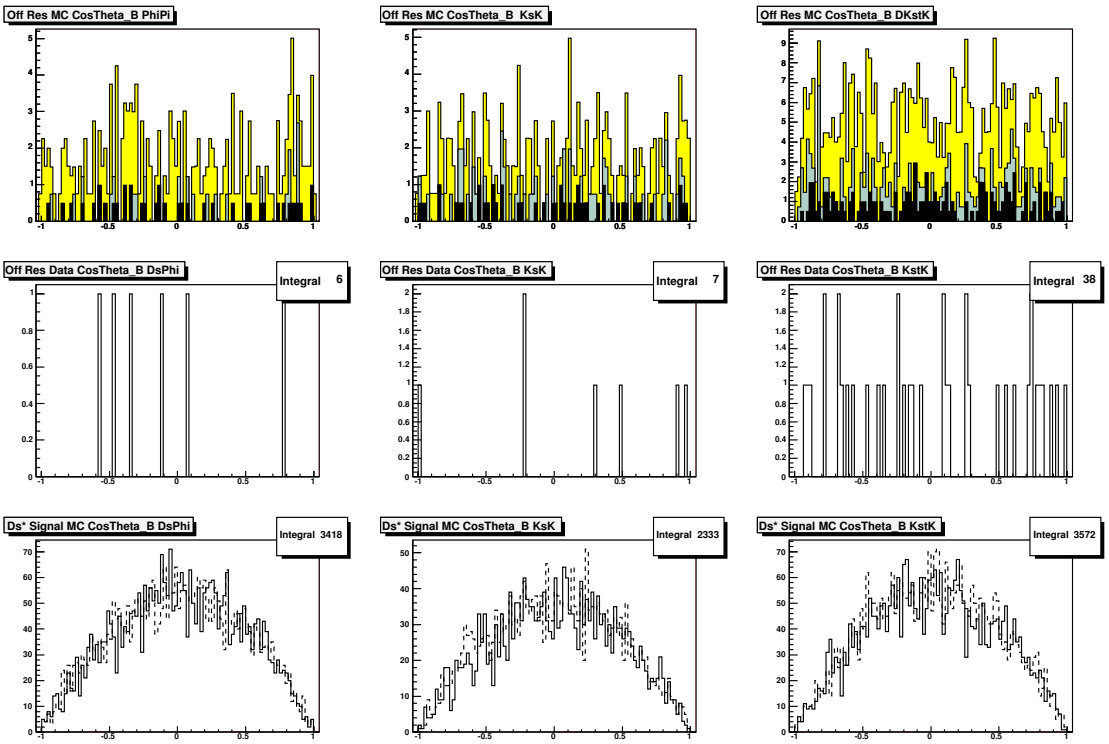

Figure 6.16: $B^{\pm} \to D_s^{*\pm}\phi$. Cosine of the CM polar angle for all events that pass the level one selection criteria. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^- K_S^0$, $D_s^- \to K^{*0} K^-$.
Top row: Background MC, normalized to on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\bar{c}$.
Middle row: Off-resonance data.
Bottom row: Signal MC, arbitrary normalization. Solid line: longitudinal polarization, dashed line: transverse polarization.

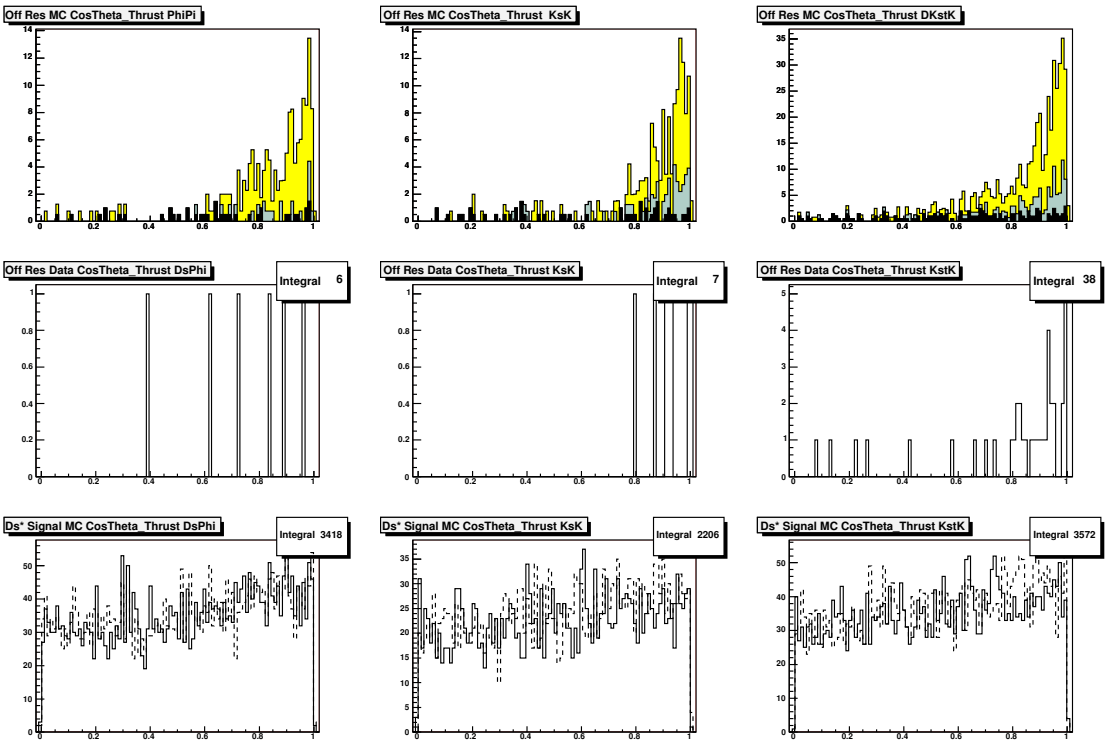

Figure 6.17: $B^{\pm} \to D_s^{*\pm}\phi$. Cosine of the angle between the thrusts of the B candidate and of the rest-of-the-event for all events that pass the level one selection criteria. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^- K_S^0$, $D_s^- \to K^{*0}K^-$.
Top row: Background MC, normalized to on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\bar{c}$.
Middle row: Off-resonance data.
Bottom row: Signal MC, arbitrary normalization. Solid line: longitudinal polarization, dashed line: transverse polarization.

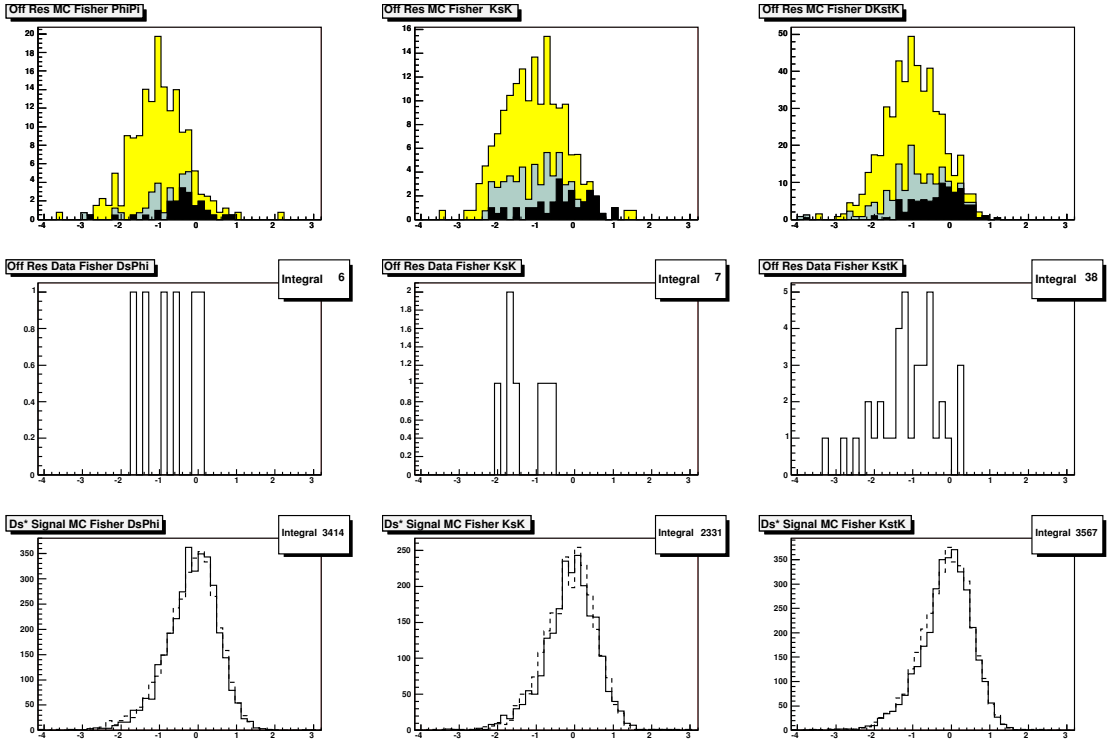

Figure 6.18: $B^{\pm} \to D_s^{*\pm}\phi$. Legendre Fisher for all events that pass the level one selection criteria. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^- K_S^0$, $D_s^- \to K^{*0}K^-$.
Top row: Background MC, normalized to on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\overline{c}$.
Middle row: Off-resonance data.
Bottom row: Signal MC, arbitrary normalization. Solid line: longitudinal polarization, dashed line: transverse polarization.

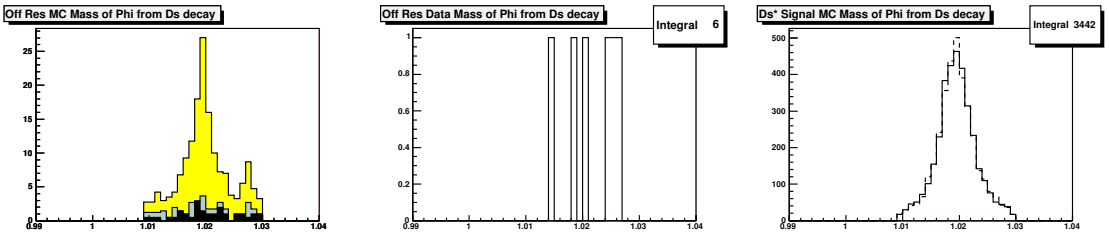

Figure 6.19: $B^{\pm} \to D_s^{*\pm}\phi$. Mass of the ϕ from $D_s^- \to \phi\pi^-$ for all events that pass the level one selection criteria.
Left: Background MC, normalized to the on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\overline{c}$.
Middle: Off-resonance data.
Right: Signal MC, arbitrary normalization. Solid line: longitudinal polarization, dashed line: transverse polarization.

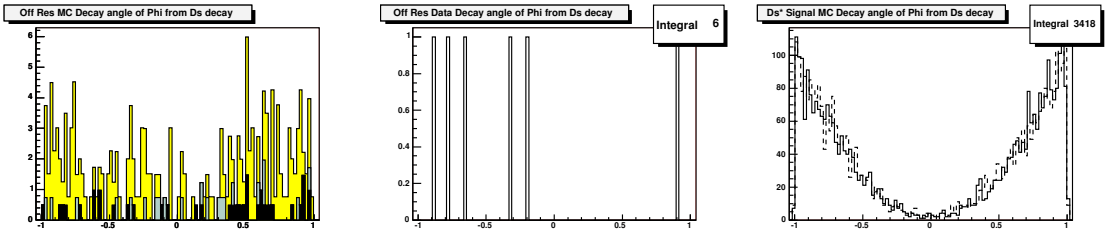

Figure 6.20: $B^{\pm} \to D_s^{*\pm}\phi$. Decay (helicity) angle of the ϕ from $D_s^- \to \phi\pi^-$ for all events that pass the level one selection criteria.
Left: Background MC, normalized to the on-resonance luminosity, black=$B\overline{B}$, blue=uds, yellow=$c\overline{c}$.
Middle: Off-resonance data.
Right: Signal MC, arbitrary normalization. Solid line: longitudinal polarization, dashed line: transverse polarization.

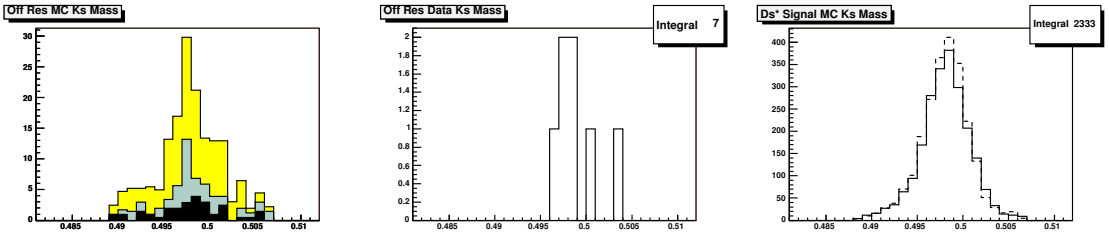

Figure 6.21: $B^{\pm} \to D_s^{*\pm}\phi$. Mass of the K_S^0 from $D_s^- \to K^- K_S^0$ for all events that pass the level one selection criteria.
Left: Background MC, normalized to the on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\overline{c}$.
Middle: Off-resonance data.
Right: Signal MC, arbitrary normalization. Solid line: longitudinal polarization, dashed line: transverse polarization.

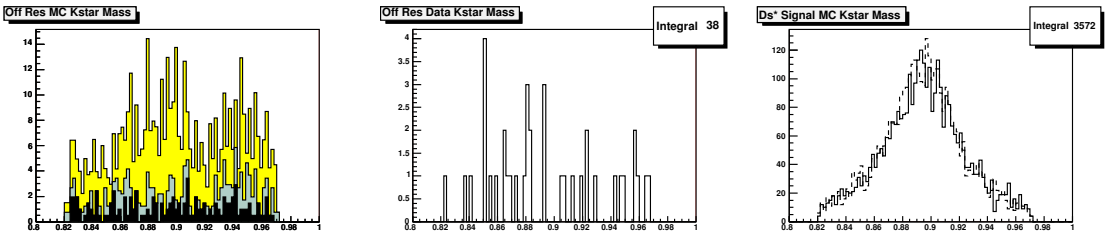

Figure 6.22: $B^{\pm} \to D_s^{*\pm}\phi$. Mass of the K^{*0} candidate from $D_s^- \to K^{*0} K^-$ for all events that pass the level one selection criteria.
Left: Background MC, normalized to the on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\overline{c}$.
Middle: Off-resonance data.
Right: Signal MC, arbitrary normalization. Solid line: longitudinal polarization, dashed line: transverse polarization.

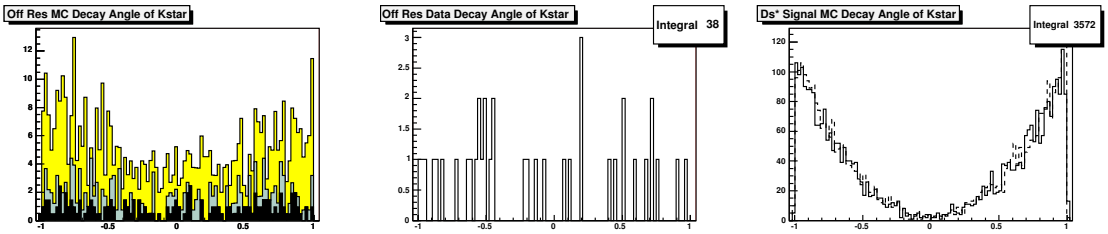

Figure 6.23: $B^{\pm} \to D_s^{*\pm}\phi$. Decay (helicity) angle of the K^{*0} from $D_s^- \to K^{*0}K^-$ for all events that pass the level one selection criteria.
Left: Background MC, normalized to the on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\overline{c}$.
Middle: Off-resonance data.
Right: Signal MC, arbitrary normalization. Solid line: longitudinal polarization, dashed line: transverse polarization.

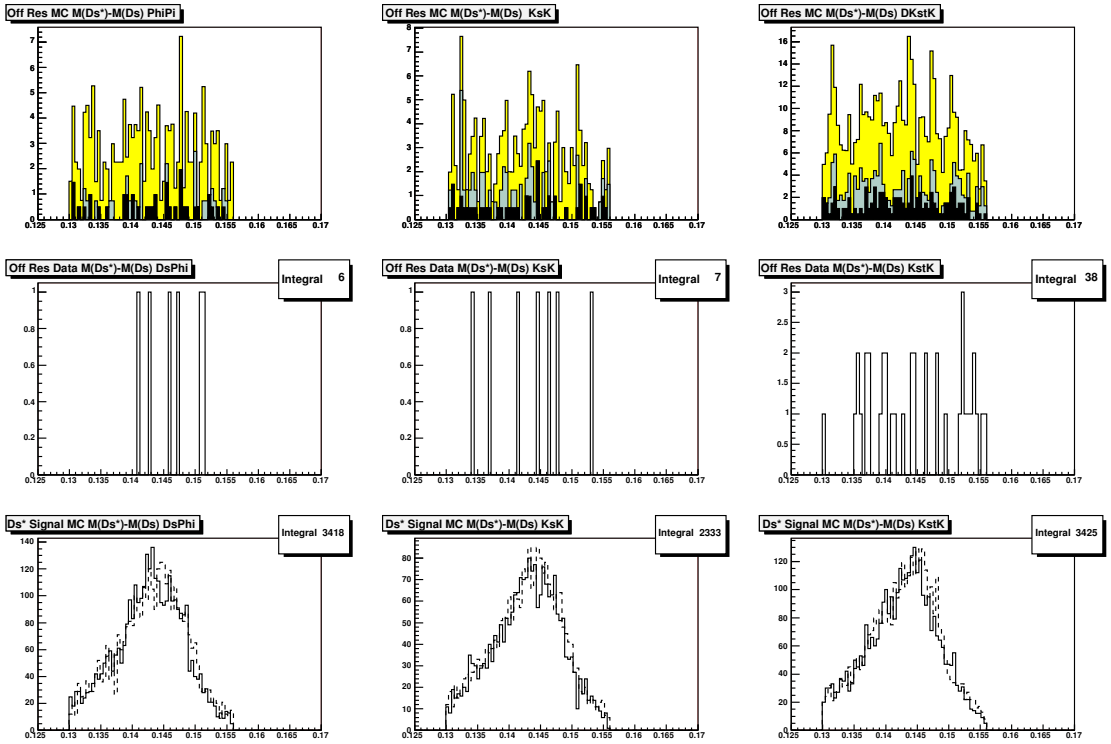

Figure 6.24: $B^{\pm} \to D_s^{*\pm}\phi$. Δm for all events that pass the level one selection criteria.
Left: Background MC, normalized to the on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\overline{c}$.
Middle: Off-resonance data.
Right: Signal MC, arbitrary normalization. Solid line: longitudinal polarization, dashed line: transverse polarization.

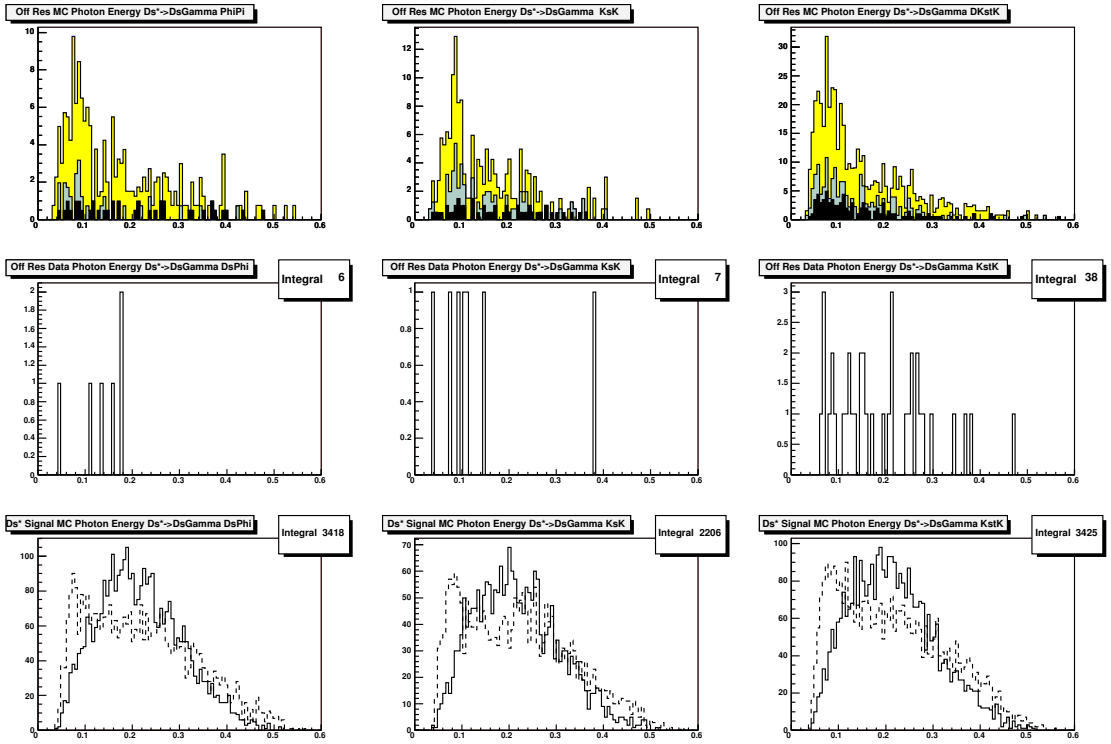

Figure 6.25: $B^{\pm} \to D_s^{*\pm}\phi$. Energy of the photon from $D_s^{*\pm} \to D_s^{\pm}\gamma$ for all events that pass the level one selection criteria.
Left: Background MC, normalized to the on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\overline{c}$.
Middle: Off-resonance data.
Right: Signal MC, arbitrary normalization. Solid line: longitudinal polarization, dashed line: transverse polarization.

6.7 Comparison with on resonance data

The continuum MC predictions for the yields are quite a bit off from what is observed in off-resonance data, see Table 6.7. Also, some of the kinematic distributions, *e.g.* the bachelor ϕ mass, show poor agreement between off resonance data and MC. For these reasons we now examine the larger on-$\Upsilon(4S)$-resonance data sample. We want to determine whether the discrepancy in yields is real or an unlikely statistical fluctuation, and we also want to examine the gross features of the kinematic distributions. In order for the analysis to remain blind, we add the requirement that $m_{\rm ES} < 5.26$ GeV/c^2 to the level one selection criteria.

The comparison of yields between the MC predictions and on-resonance data are listed in Table 6.8. For the $B^{\pm} \to D_s^{\pm}\phi$ modes we see that these results confirm that the continuum MC overestimates the rate of background events, especially in the cleaner $D_s^- \to \phi\pi^-$ and $D_s^- \to K^-K_S^0$ modes. This is also the case for the $B^{\pm} \to D_s^{*\pm}\phi$ modes. Next, we examine the various kinematic distributions for on resonance data. These are displayed in Figures 6.26- 6.33 for $B^{\pm} \to D_s^{\pm}\phi$ and in Figures 6.34- 6.43 for $B^{\pm} \to D_s^{*\pm}\phi$. These distributions all look quite reasonable.

Table 6.8: Number of events passing the level one selection criteria with $m_{\rm ES} < 5.26$ GeV/c^2 in various samples. The numbers in parentheses are normalized to the on-resonance luminosity. Prediction 1 is the prediction for the number of on resonance events from the $c\overline{c}$, uds, and $B\overline{B}$ Monte Carlos. Prediction 2 uses the off resonance data and the $B\overline{B}$ Monte Carlos.

Mode	Off	$c\overline{c}$	uds	$B^0\overline{B}^0$	B^+B^-	On	Pred. 1	Pred. 2
$B \to D_s^- \phi$								
$D_s^- \to \phi\pi^-$	8 (75)	209 (157)	25 (18)	10 (5)	14 (7)	102	187 ± 12	87 ± 26
$D_s^- \to K^- K_S^0$	8 (75)	186 (140)	62 (45)	22 (10)	17 (8)	136	203 ± 12	93 ± 26
$D_s^- \to K^{*0} K^-$	44 (414)	559 (420)	189 (137)	55 (27)	79 (39)	486	623 ± 21	480 ± 62
$B \to D_s^{*-} \phi$								
$D_s^{*-} \to (\phi\pi^-)\gamma$	5 (47)	132 (99)	19 (14)	13 (6)	13 (6)	65	125 ± 9	59 ± 21
$D_s^{*-} \to (K^- K_S^0)\gamma$	7 (66)	108 (81)	38 (28)	17 (8)	14 (7)	77	124 ± 8	81 ± 25
$D_s^{*-} \to (K^{*0} K^-)\gamma$	25 (235)	313 (236)	94 (69)	45 (21)	75 (35)	254	361 ± 16	291 ± 47

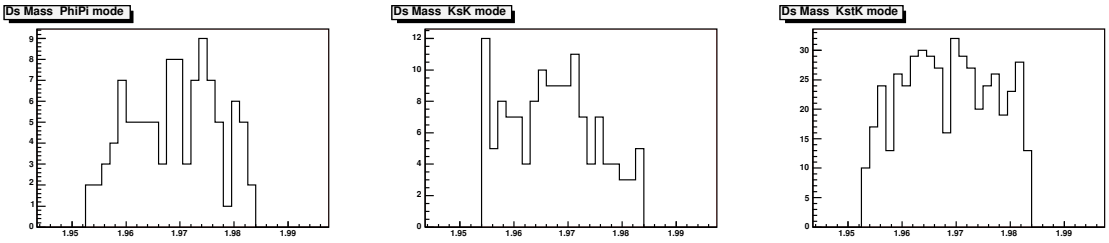

Figure 6.26: $B^{\pm} \to D_s^{\pm}\phi$. Mass of the $D_s^{\pm}$ candidate for on resonance candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, $D_s^- \to K^{*0}K^-$.

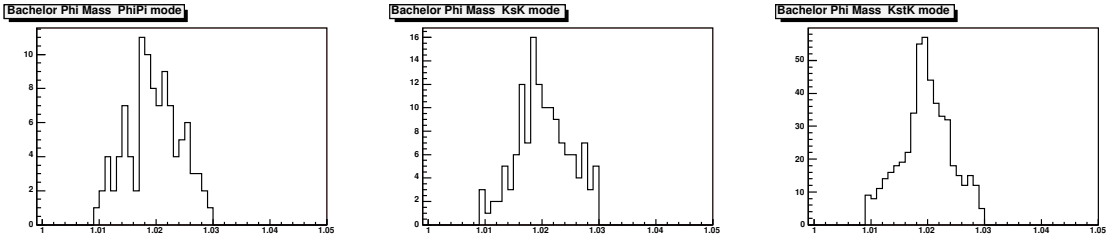

Figure 6.27: $B^{\pm} \to D_s^{\pm}\phi$. Mass of the bachelor ϕ candidate for on resonance candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, $D_s^- \to K^{*0}K^-$.

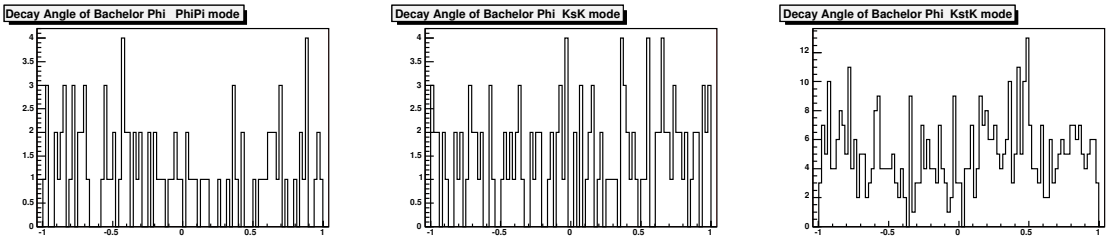

Figure 6.28: $B^{\pm} \to D_s^{\pm}\phi$. Decay (helicity) angle of the bachelor ϕ candidate for on resonance candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, $D_s^- \to K^{*0}K^-$.

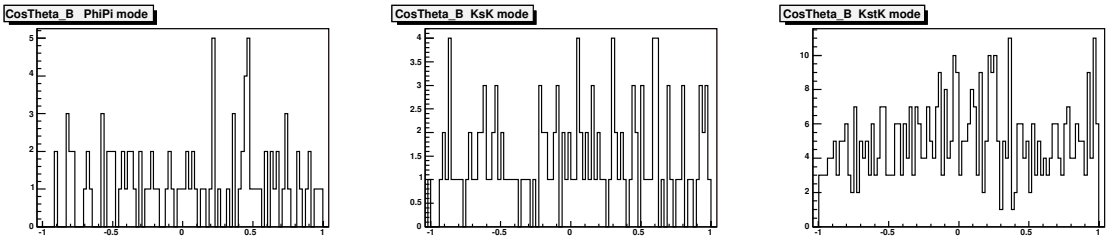

Figure 6.29: $B^{\pm} \to D_s^{\pm}\phi$. Cosine of the CM polar angle for candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^- K_S^0$, $D_s^- \to K^{*0} K^-$.

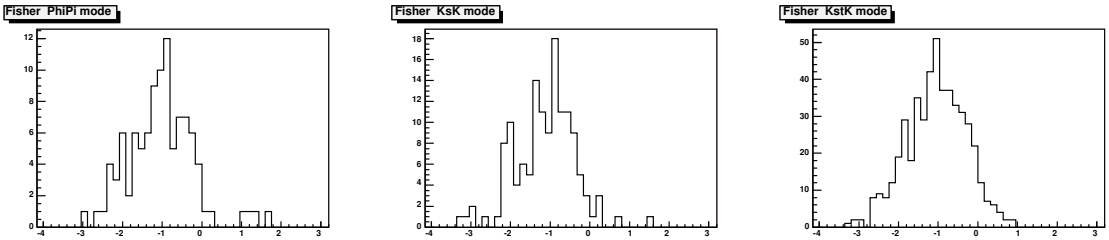

Figure 6.30: $B^{\pm} \to D_s^{\pm}\phi$. Legendre Fisher candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^- K_S^0$, $D_s^- \to K^{*0} K^-$.

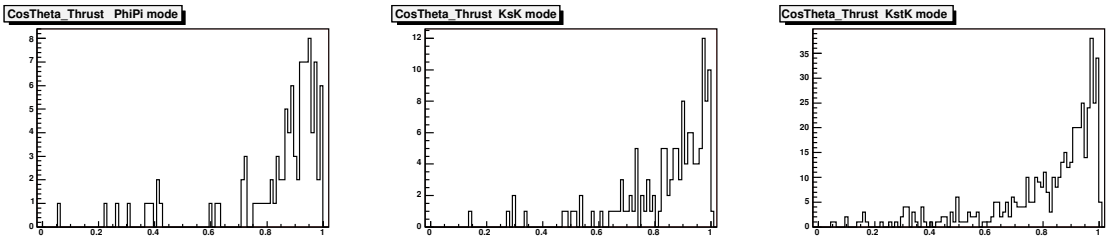

Figure 6.31: $B^{\pm} \to D_s^{\pm}\phi$. $\cos(\theta_{\rm T})$ for candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^- K_S^0$, $D_s^- \to K^{*0} K^-$.

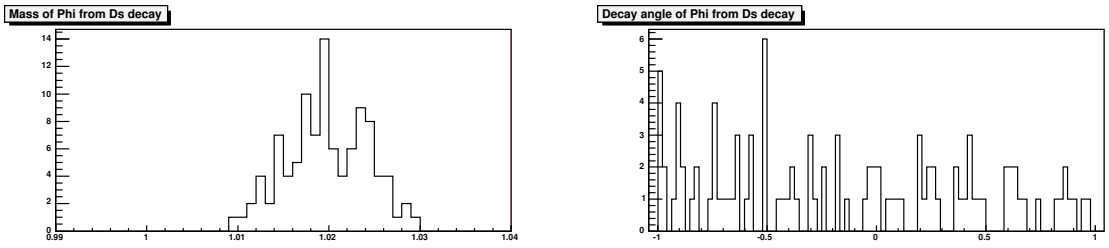

Figure 6.32: $B^{\pm} \to D_s^{\pm}\phi$. Left plot: mass of the ϕ candidate from $D_s^- \to \phi\pi^-$ decay. Right plot: decay (helicity) angle of the ϕ candidate from $D_s^- \to \phi\pi^-$ decay. Distributions for on resonance candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2.

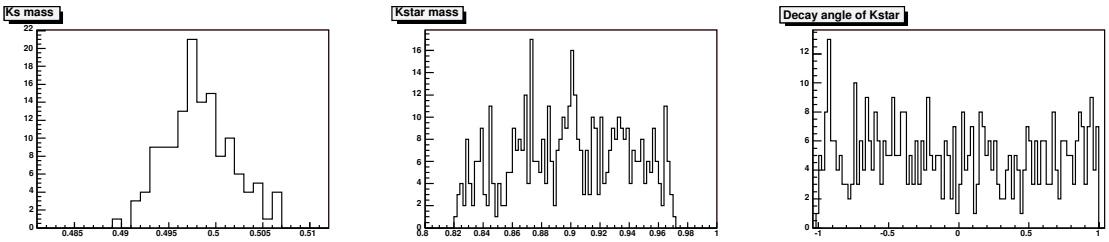

Figure 6.33: $B^{\pm} \to D_s^{\pm}\phi$. Left plot: mass of the K_S^0 candidate from $D_s^- \to K^- K_S^0$ decay. Middle plot: mass of the K^{*0} candidate from $D_s^- \to K^{*0}K^-$ Right plot: decay (helicity) angle of the K^{*0} candidate from $D_s^- \to K^{*0}K^-$ decay. Distributions for on resonance candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2.

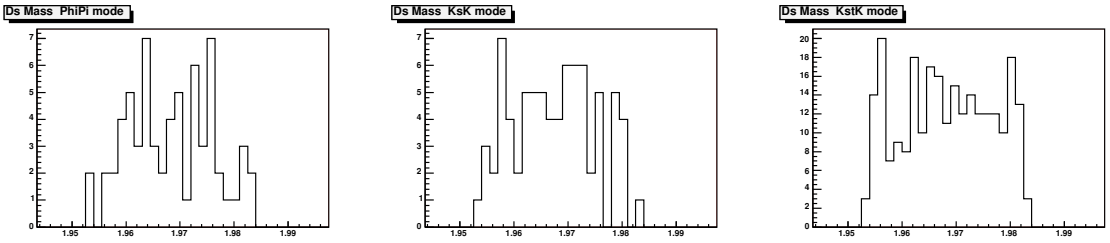

Figure 6.34: $B^{\pm} \to D_s^{*\pm}\phi$. Mass of the $D_s^{\pm}$ candidate for on resonance $B^{\pm} \to D_s^{*\pm}\phi$ candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, $D_s^- \to K^{*0}K^-$.

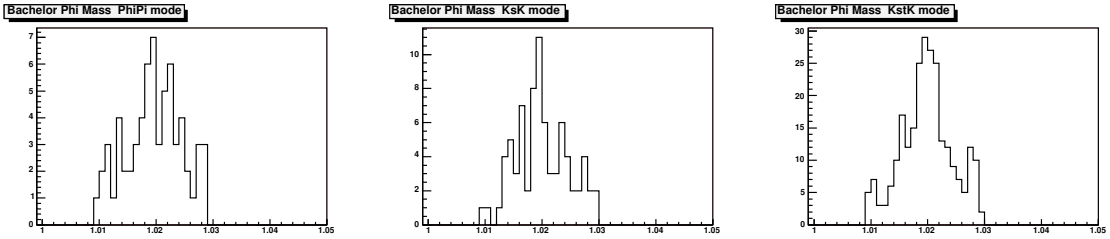

Figure 6.35: $B^{\pm} \to D_s^{*\pm}\phi$. Mass of the bachelor ϕ candidate for on resonance candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, $D_s^- \to K^{*0}K^-$.

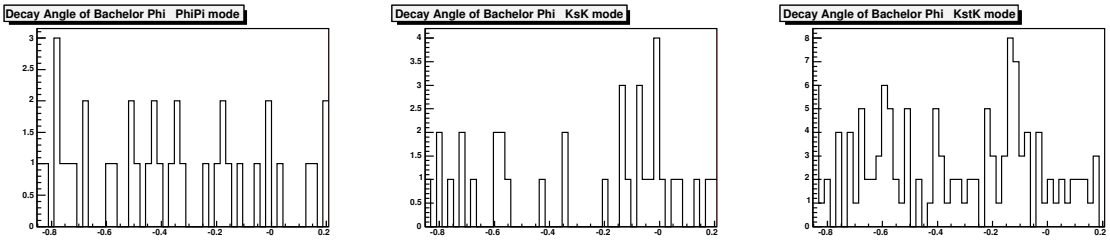

Figure 6.36: $B^{\pm} \to D_s^{*\pm}\phi$. Decay (helicity) angle of the bachelor ϕ candidate for on resonance candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, $D_s^- \to K^{*0}K^-$.

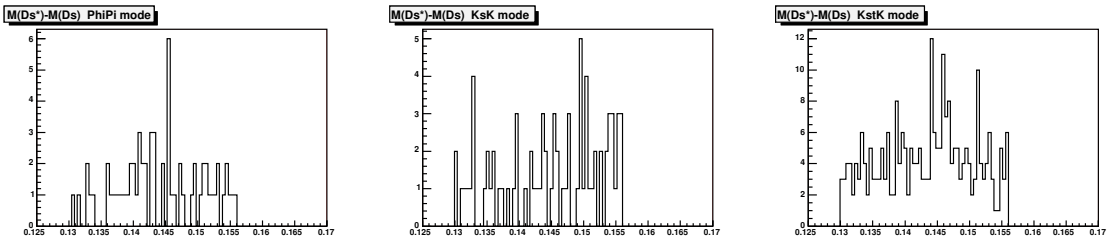

Figure 6.37: $B^{\pm} \to D_s^{*\pm}\phi$. Δm for on resonance candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, $D_s^- \to K^{*0}K^-$.

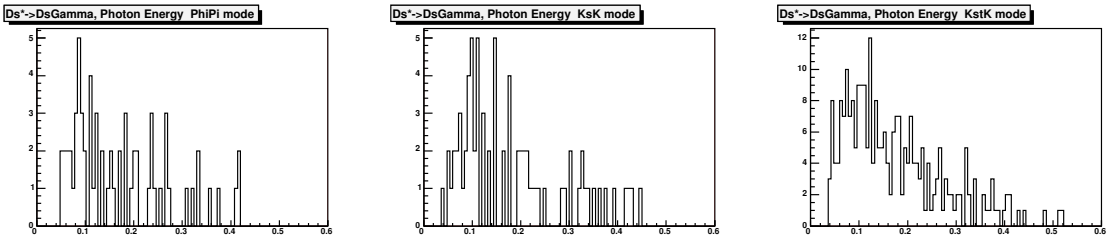

Figure 6.38: $B^{\pm} \to D_s^{*\pm}\phi$. Energy of the photon from the $D_s^{*\pm} \to D_s^{\pm}\gamma$ decay for on resonance candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, $D_s^- \to K^{*0}K^-$.

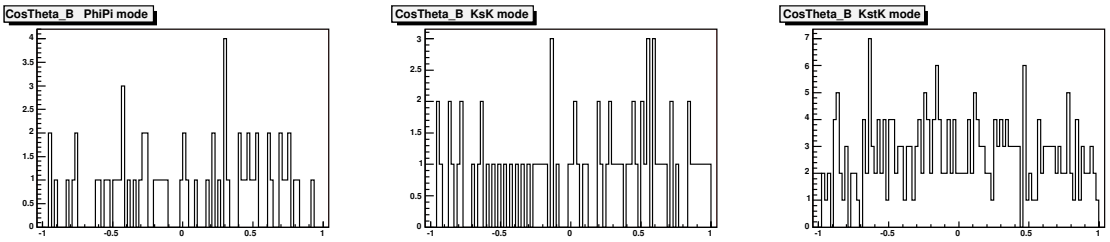

Figure 6.39: $B^{\pm} \to D_s^{*\pm}\phi$. Cosine of the CM polar angle for candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, $D_s^- \to K^{*0}K^-$.

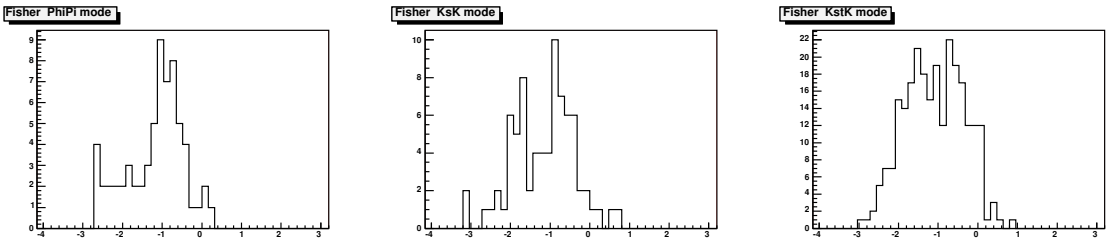

Figure 6.40: $B^{\pm} \to D_s^{*\pm}\phi$. Legendre Fisher candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, $D_s^- \to K^{*0}K^-$.

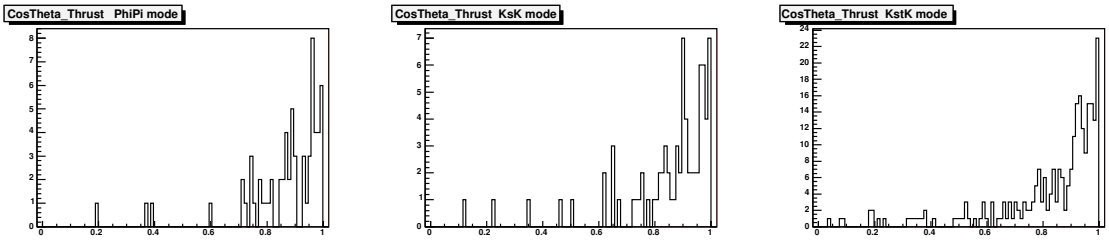

Figure 6.41: $B^{\pm} \to D_s^{*\pm}\phi$. $\cos(\theta_{\rm T})$ for candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, $D_s^- \to K^{*0}K^-$.

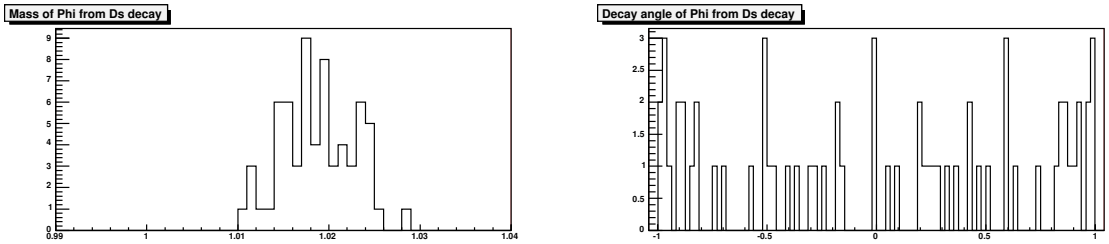

Figure 6.42: $B^{\pm} \to D_s^{*\pm}\phi$. Left plot: mass of the ϕ candidate from $D_s^- \to \phi\pi^-$ decay. Right plot: decay (helicity) angle of the ϕ candidate from $D_s^- \to \phi\pi^-$ decay. Distributions for on resonance candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2.

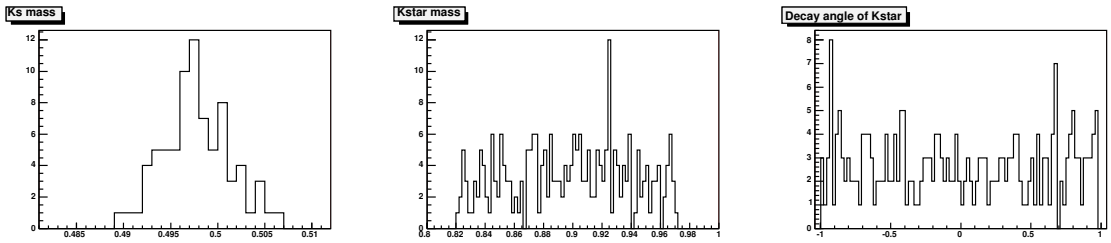

Figure 6.43: $B^{\pm} \to D_s^{*\pm}\phi$. Left plot: mass of the K_S^0 candidate from $D_s^- \to K^- K_S^0$ decay. Middle plot: mass of the K^{*0} candidate from $D_s^- \to K^{*0}K^-$ Right plot: decay (helicity) angle of the K^{*0} candidate from $D_s^- \to K^{*0}K^-$ decay. Distributions for on resonance candidates passing the level one selection criteria and with $m_{\rm ES} < 5.26$ GeV/c^2.

6.8 Constructing a likelihood variable

It is clear that many of the quantities shown in Figures 6.2 to 6.23 can discriminate at some level between signal and background. We have decided to combine some of these quantities into a likelihood, in order to have a single discriminating variable. In order to do this, we define a probability density function (PDF)[17] for each variable. The likelihood is defined as the product of the PDFs over all the quantities. This approach has the advantage that it is relatively simple. Other approaches, *e.g.*, neural networks, are more complicated, but in principle would make optimal use of the available information. However, the statistics of the available background data samples are too limited to train a neural network. In Table 6.9 and 6.10 we list the quantities that were considered in the likelihood, and the forms of the PDFs that we assume. Absent from the likelihood is the thrust angle, $\cos(\theta_{\mathrm{T}})$. This is because $\cos(\theta_{\mathrm{T}})$ is correlated to the Fisher variable, which is included in the likelihood. We will, however, still use the $\cos(\theta_{\mathrm{T}})$ information when we perform our selection criteria optimization in Section 6.9.

Thus our likelihood takes on the form,

$$\mathcal{L} = P_M(M) \times P_\theta(\theta) \times \ldots \tag{6.8}$$

[17] A probability density function is normalized to unity and is defined by: Probability(result lies between x_1 and x_2) = $\int_{x_1}^{x_2} PDF(x)dx$

Table 6.9: $B^{\pm} \to D_s^{\pm}\phi$. Variables used in the definition of the likelihood, and the functional form of the PDF for signal and background.

Mode	Quantity	Signal PDF	Background PDF
ALL	$D_s^{\pm}$ mass	Double Gaus.	Flat
ALL	Polar angle of B in CM	$\sin^2\theta$	Flat
ALL	Helicity of ϕ from B decay	$\cos^2\theta$	Flat
ALL	Legendre Fisher	Bifurcated Gaus.	Bifurcated Gaus.
$D_s^- \to \phi\pi^-$	Mass of ϕ from $D_s^{\pm}$ decay	Voigtian	Flat
$D_s^- \to \phi\pi^-$	Helicity of ϕ from $D_s^{\pm}$ decay	$\cos^2\theta$	Flat
$D_s^- \to K^{*0}K^-$	K^{*0} mass	Breit-Wigner	Flat
$D_s^- \to K^{*0}K^-$	Helicity of K^{*0}	$\cos^2\theta$	Flat

where, for example, $P_M(M)$ is the probability obtained from the PDF representing a particle of mass M, and $P_\theta(\theta)$ is the probability obtained from the PDF representing a decay angle of θ. We define this likelihood for both the signal and the background – $\mathcal{L}_S$ and $\mathcal{L}_B$, respectively. We combine these two likelihood distributions into a single discriminating variable by taking the log of their ratio,

$$\mathcal{R} = \log\left(\frac{\mathcal{L}_S}{\mathcal{L}_B}\right) \tag{6.9}$$

For simplicity, many of the background PDFs are taken as flat, even if some of the distributions are not quite flat in the background MC datasets. The distribution that shows the largest departure from the flatness assumption is the distribution of the mass of the ϕ from $D_s^{\pm}$ decay in the $c\bar{c}$ MC, see Figure 6.8. However, the off-resonance data distribution appear to be quite flat in this vari-

Table 6.10: $B^{\pm} \to D_s^{*\pm}\phi$. Variables used in the definition of the likelihood, and the functional form of the PDF for signal and background.

Mode	Quantity	Signal PDF	Background PDF
ALL	$D_s^{\pm}$ mass	Double Gaus.	Flat
ALL	Polar angle of B in CM	$\sin^2\theta$	Flat
ALL	$D_s^{\pm}$ Δm	Double Gaus.	Flat
ALL	Legendre Fisher	Bifurcated Gaus.	Bifurcated Gaus.
$D_s^- \to \phi\pi^-$	Mass of ϕ from $D_s^{\pm}$ decay	Voigtian	Flat
$D_s^- \to \phi\pi^-$	Helicity of ϕ from $D_s^{\pm}$ decay	$\cos^2\theta$	Flat
$D_s^{\pm} \to \overline{K}^{*0}K^{\pm}$	K^{*0} mass	Breit-Wigner	Flat
$D_s^{\pm} \to \overline{K}^{*0}K^{\pm}$	Helicity angle of K^{*0}	$\cos^2\theta$	Flat

able (again, see Figure 6.8)[18]. The on-resonance distribution (Figure 6.32), looks a little more peaked, but not very sharply.

The parameters of the non-trivial PDFs that enter the likelihood variable, *i.e.,* Gaussians, Voigtian[19], Breit-Wigner, are listed in Table 6.11. Examples of the fitted distributions are shown in Figures 6.44, 6.45, 6.46, and 6.47.

[18]To verify this, we looked at this distribution in off-resonance data opening up the mass cuts around the ϕ and we found no hint of a ϕ peak.

[19]A Voigtian is a convolution of a Breit-Wigner lineshape with a Gaussian, see Appendix A

Table 6.11: Summary of parameters used in the likelihood PDFs. Of course, the Δm is only for the $B^{\pm} \to D_s^{*\pm}\phi$ mode. All means, width, and sigmas are in GeV/c^2, except for Fisher, of course. BW represents a Breit-Wigner lineshape; Bifur. Gaus represents a Bifurcated Gaussian lineshape. Keep in mind that these parameters are from a fit to signal MC, they simply are our best estimate for the shape of these distributions. Thus, for example, one should not be concerned that the parameters for the M($D_s^{\pm}$) fits vary over mode.

Quantity	PDF	Mode	Parameters		
M($D_s^{\pm}$)	Dbl. Gaus	$D_s^- \to \phi\pi^-$	$\mu_1 = 1.9681$	$\sigma_1 = 0.0047$	$\text{Frac}_1 = 0.795$
			$\mu_2 = 1.9667$	$\sigma_2 = 0.0178$	
M($D_s^{\pm}$)	Dbl. Gaus	$D_s^- \to K^- K_S^0$	$\mu_1 = 1.9691$	$\sigma_1 = 0.0065$	$\text{Frac}_1 = 0.832$
			$\mu_2 = 1.9646$	$\sigma_2 = 0.0176$	
M($D_s^{\pm}$)	Dbl. Gaus	$D_s^- \to K^{*0} K^-$	$\mu_1 = 1.9679$	$\sigma_1 = 0.0051$	$\text{Frac}_1 = 0.771$
			$\mu_2 = 1.9645$	$\sigma_2 = 0.0174$	
$D_s^{*\pm}$ Δm	Dbl. Gaus	Signal	$\mu_1 = 0.1438$	$\sigma_1 = 0.0044$	$\text{Frac}_1 = 0.702$
			$\mu_2 = 0.1322$	$\sigma_2 = 0.0140$	
Fisher	Bifur. Gaus	Signal	$\mu = 0.081$	$\sigma_{\text{low}} = 0.82$	$\sigma_{\text{high}} = 0.52$
Fisher	Bifur. Gaus	Continuum	$\mu = -1.00$	$\sigma_{\text{low}} = 0.77$	$\sigma_{\text{high}} = 0.62$
M(ϕ)	Voigtian	$D_s^- \to \phi\pi^-$	$\mu = 1.0195$	$\Gamma = 0.0046$	$\sigma = 0.0011$
M(K^{*0})	BW	$D_s^- \to K^{*0} K^-$	$\mu = 0.8961$	$\Gamma = 0.0512$	

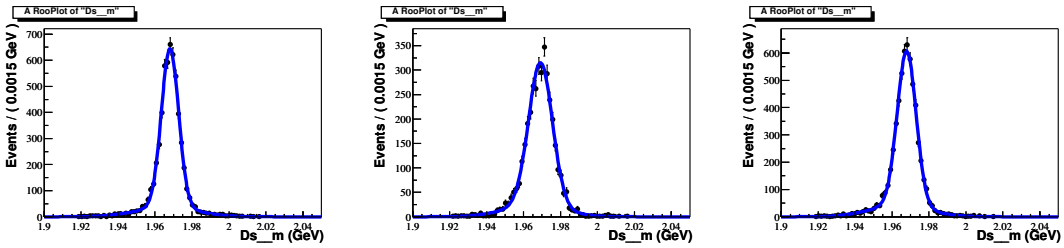

Figure 6.44: Mass of the $D_s^{\pm}$ candidate for events passing the level one selection in the signal MC. Double Gaussian fits are superimposed. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^- K_S^0$, $D_s^- \to K^{*0} K^-$.

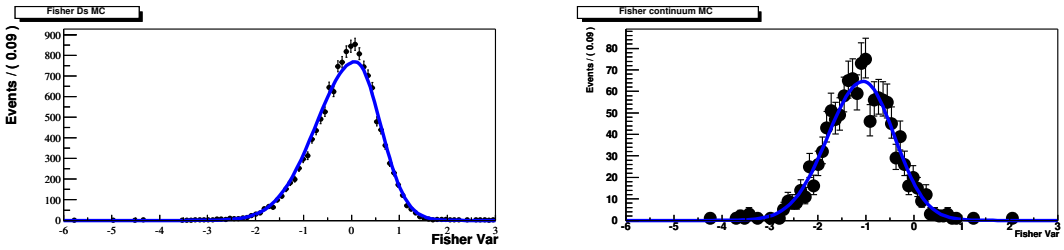

Figure 6.45: Fisher distributions for events passing the level one selection in signal (left plot) and continuum (right plot) MC. Bifurcated Gaussian fits are superimposed.

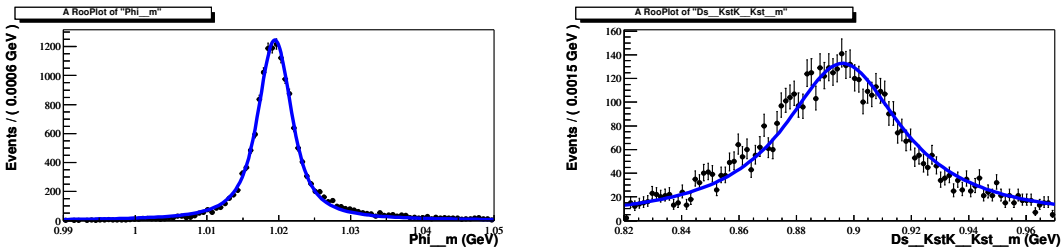

Figure 6.46: Invariant mass of the ϕ in $D_s^- \to \phi\pi^-$ (left plot) and the K^{*0} in $D_s^- \to K^{*0}K^-$ in signal MC with superimposed Voigtian and Breit-Wigner fits, respectively.

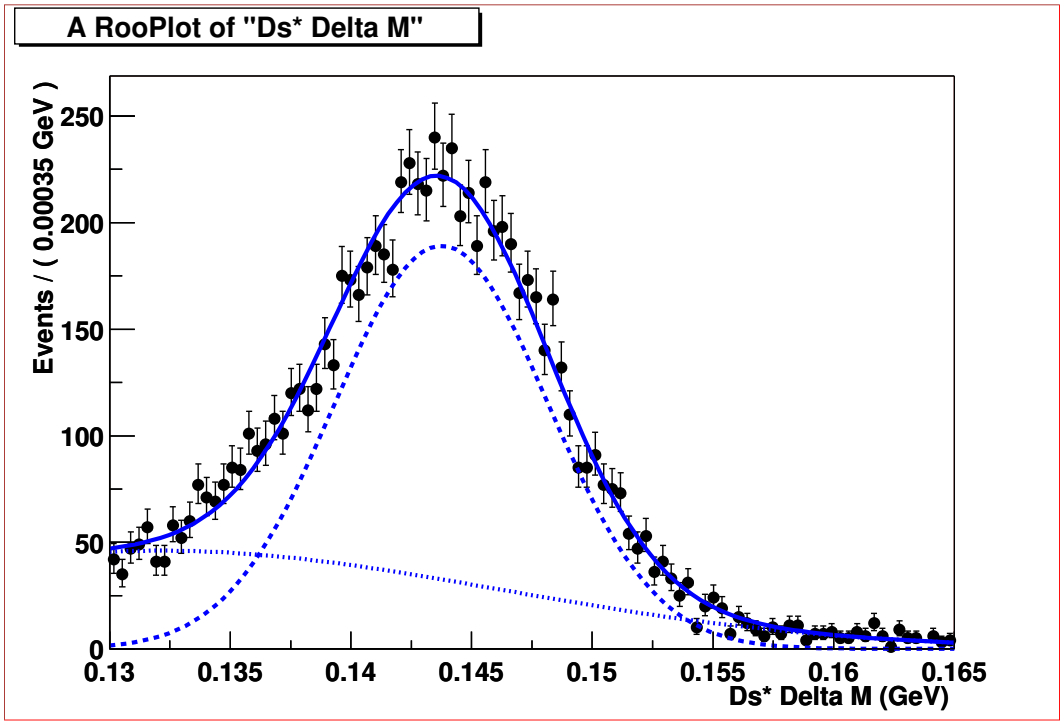

Figure 6.47: Δm for candidates where the photon from the $D_s^{*\pm} \to D_s^{\pm}\gamma$ decay is verified to be the photon created by the signal-MC-generating software. Superimposed are the two Gaussians that compose our double Gaussian fit. Note that the choice of two Gaussians is purely phenomenological in that it simply gives a function with very nice agreement to the MC distribution.

In Figure 6.48(6.49) we show the log of the likelihood function (n.b. we will refer to this as the log-likelihood or simply $\log(\mathcal{L})$) distributions for MC and off-resonance events passing the level one selection in the $B^{\pm} \to D_s^{\pm}\phi$ ($B^{\pm} \to D_s^{*\pm}\phi$) sample.

To get a better visual feeling for the kind of background rejection that can be achieved, we show the effect of cutting on $\log(\mathcal{L})$ for both signal and the dominant background ($e^+e^- \to c\overline{c}$) in figure 6.50 (for the $B^{\pm} \to D_s^{\pm}\phi$ mode). In this figure we show the efficiency in signal MC vs. the efficiency in $c\overline{c}$ MC as the $\log(\mathcal{L})$ cut is moved. We also show, in the same figure, the effect of cutting only on the Legendre Fisher, which is the most effective discriminant between signal and background.

Figure 6.50 shows the improvement in background rejection we get when including a cut on the likelihood over what we see when including only a cut on the Fisher variable. It is possible to achieve a factor of order 5 (10) rejection against continuum background with about 85% (65%) efficiency.

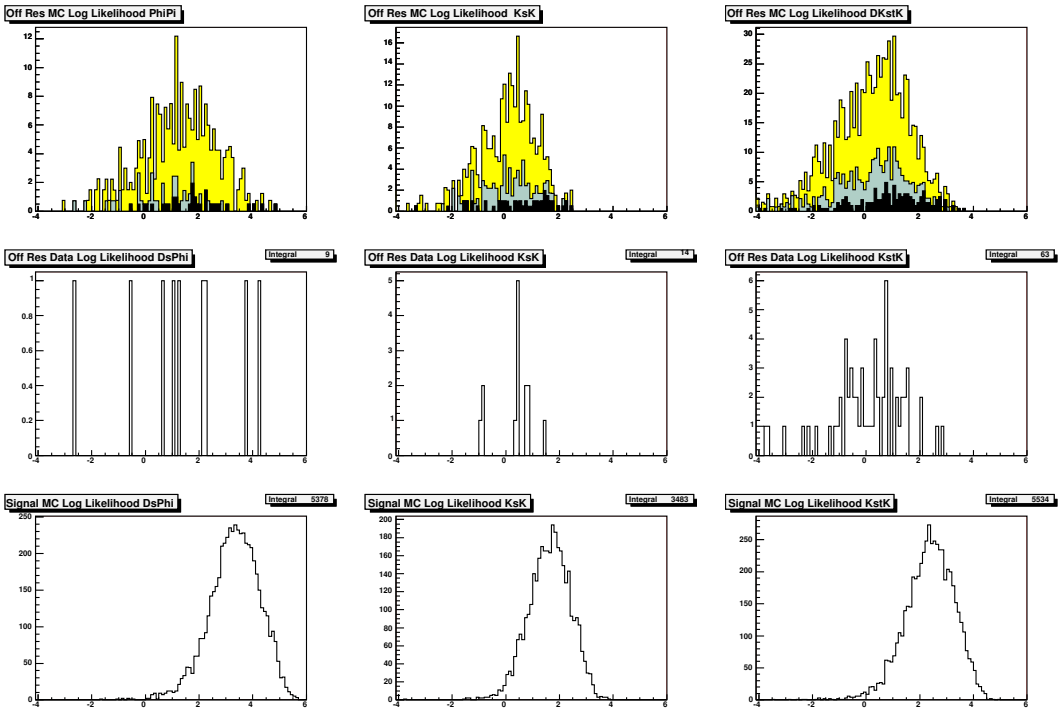

Figure 6.48: $B^{\pm} \to D_s^{\pm}\phi$. $\log(\mathcal{L})$ for candidates passing the level one selection criteria. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, $D_s^- \to K^{*0}K^-$.
Top row: background MC, normalized to on-resonance data, black=$B\bar{B}$, blue=uds, yellow=$c\bar{c}$.
Middle row: Off resonance data.
Bottom row: Signal MC, arbitrary normalization.

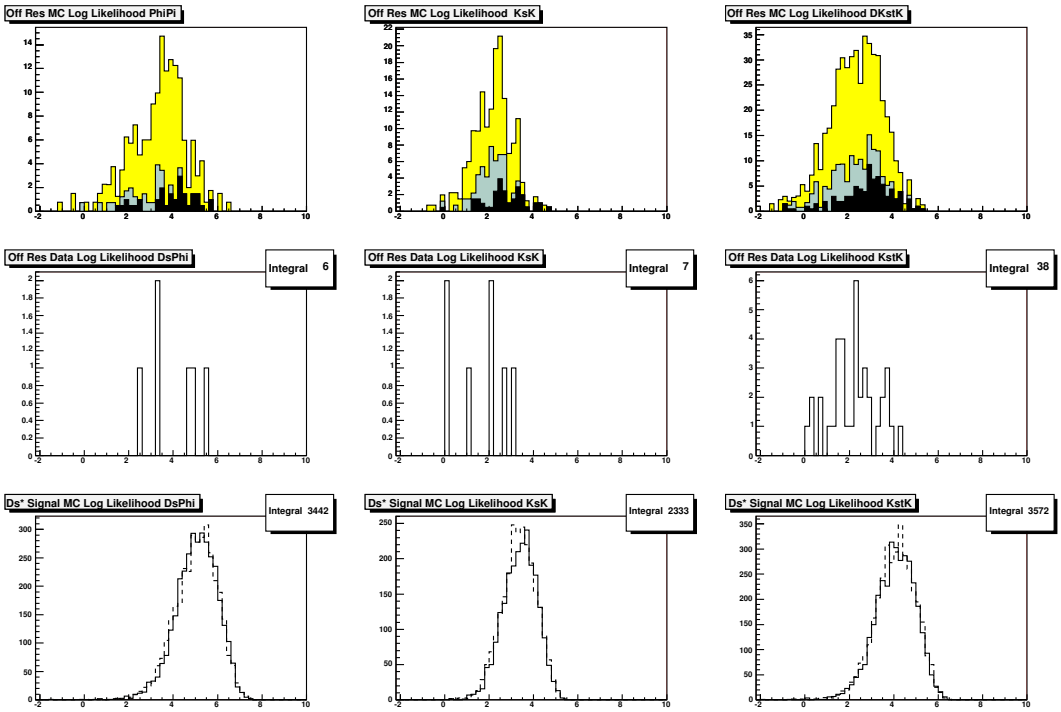

Figure 6.49: $B^{\pm} \to D_s^{*\pm}\phi$. $\log(\mathcal{L})$ for candidates passing the level one selection. From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^- K_S^0$, $D_s^- \to K^{*0} K^-$.
Top row: background MC, normalized to on-resonance data, black=$B\overline{B}$, blue=uds, yellow=$c\bar{c}$.
Middle row: Off resonance data.
Bottom row: Signal MC, arbitrary normalization.

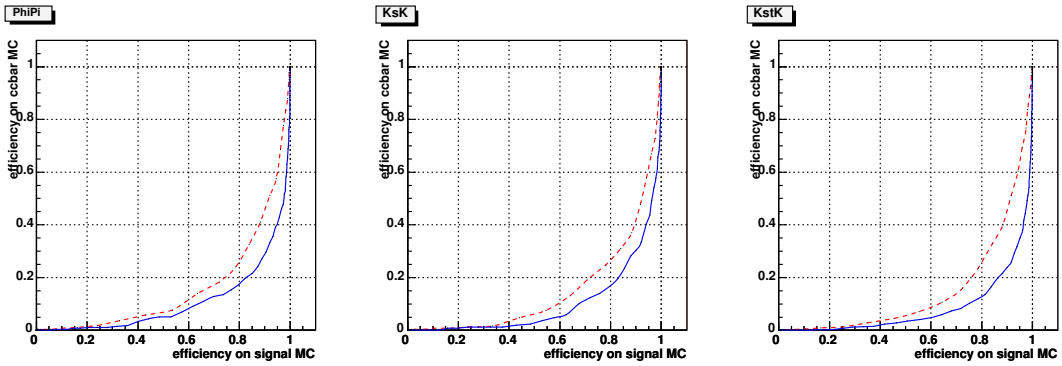

Figure 6.50: $B^{\pm} \to D_s^{\pm}\phi$. Solid blue line: efficiency for a cut on $\log(\mathcal{L})$ in signal MC vs. $c\overline{c}$ MC for events passing the level one selection. Dashed red line: same but for a cut on Fisher only.
From left to right: $D_s^- \to \phi\pi^-$, $D_s^- \to K^- K_S^0$, $D_s^- \to K^{*0} K^-$.

6.9 Selection criteria optimization

In this section we describe the procedure that we used to optimize the selection criteria for the best possible upper limit on the $B^{\pm} \to D_s^{(*)\pm}\phi$ branching fraction[20]. The selection criteria we optimize are

- The likelihood described in Section 6.8.
- The kaon identification category (`NotPion`, `Loose`, `Tight`, and `VeryTight`) of all kaons except the kaons from the decay of the bachelor ϕ.
- Cosine of the thrust angle, $\cos(\theta_{\rm T})$.
- The energy of the photon in $D_s^{*\pm} \to D_s^{\pm}\gamma$ decays, $E(\gamma)$. This applies to the $B^{\pm} \to D_s^{*\pm}\phi$ mode only.

We don't run the optimization for the kaons from the bachelor ϕ because, as demonstrated in Figure 6.3, most of the bachelor ϕ candidates passing the level one selection requirements are real $\phi \to K^+K^-$ decays. Tightening the kaon identification criteria on these kaons beyond the loosest requirements would not provide an improvement in our background rejection power. Note also that the $\cos(\theta_{\rm T})$ variable was not included in the likelihood due to its high degree of correlation with the Fisher variable.

[20] Recall, the optimization was performed only on data from Runs 1-3. The Run 4 data shows no change in the shapes of the discriminating variables and thus the optimization results would be negligible.

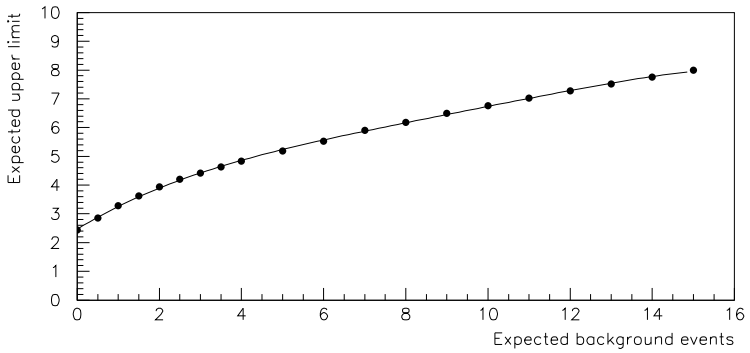

Figure 6.51: Graphical representation of the contents of Table XII from Reference [18]. The horizontal axis is the number of expected background events in a counting experiment. The vertical axis is the corresponding expected 90% confidence level upper limit on the number of signal events. The curve is a 4th order polynomial fit.

The optimization is performed using continuum, generic $B\overline{B}$, and signal MC samples. Since we expect no signal in the Standard Model at our luminosity, the quantity that we optimize is the expected sensitivity. This is defined as the branching fraction upper limit that we should obtain, on average, when performing this experiment many times in the absence of a signal.

Although the signal and/or the upper limit will be extracted from a fit to the $m_{\rm ES}$ distribution, we approximate our experiment as a simple counting experiment. A counting experiment is an experiment where we count the number of events in the signal box and compare it with background expectations. Table XII of the paper by Feldman and Cousins [18] shows the expected upper limit on the number of events for a counting experiment as a function of the expected

background ignoring systematic uncertainties (see Figure 6.51). Our procedure is then the following:

1. For a given set of requirements, calculate N_{BG}, the expected number of background events with $m_{\rm ES} > 5.27$ GeV/c^2 and with $|\Delta E| < 30$ MeV.

2. Given N_{BG}, calculate the expected 90% C.L. upper limit on the number of events (N_{UL}). This calculation is based on the compilation of Table XII from Reference [18] (Figure 6.51).

3. For the same set of requirements, calculate the efficiency on signal MC (ϵ).

4. Finally, calculate the expected sensitivity on the branching fraction using Equation 5.1, $\mathcal{B} = \frac{N_{UL}}{\epsilon_i \cdot \mathcal{B}_i \cdot N_{B\overline{B}}}$.

5. Change the set of requirements and iterate until we converge on the set of requirements that minimizes $\mathcal{B}$.

The procedure outlined above is repeated three times, once for each $D_s^{\pm}$ decay mode, however, there are a couple of subtleties associated with the calculation of N_{BG}. The first is that ideally one would like to simply count the number of events found in the signal box ($m_{\rm ES} > 5.27$ GeV/c^2 and $|\Delta E| < 30$ MeV) in the background MC datasets ($c\overline{c}$, uds, and $B\overline{B}$) for different selection requirements. However, the statistics are too poor for this to be meaningful. Thus, we estimate

N_{BG} starting from the much looser level one selection criteria, $m_{\rm ES} > 5.2$ GeV/c^2 and $|\Delta E| < 200$ MeV, and then rescaling the result by a factor of 0.025. This factor of 0.025 is an empirical factor which relates the numbers of events for the two selections, see Table 6.5[21].

The second subtlety is the following. As discussed in Section 6.5, the continuum MC overestimates the number of background events in the data, see Tables 6.5 and 6.8. On the other hand (within reason given of the lack of statistics) the kinematic distributions show no deviation from the data distributions, see Figures 6.2-6.23. Thus we scale down the predictions of the uds and $c\bar{c}$ MCs to bring them in agreement with the data. We choose the scaling factors by considering the information in Table 6.8. In the $B^{\pm} \to D_s^{\pm}\phi$ mode we choose 1.9 for $D_s^- \to \phi\pi^-$, 1.6 for $D_s^- \to K^-K_S^0$, and 1.3 for $D_s^- \to K^{*0}K^-$; while in the $B^{\pm} \to D_s^{*\pm}\phi$ mode we choose 1.7 for $D_s^- \to \phi\pi^-$, 1.4 for $D_s^- \to K^-K_S^0$, and 1.6 for $D_s^- \to K^{*0}K^-$.

As a graphical example of the optimization procedure, see Figures 6.52, 6.53, and 6.54. In these figures we show the optimization of the likelihood and kaon identification category for three different kaons in the various $D_s^{\pm}$ modes. To interpret these plots we refer the reader to the optimization procedure enumerated

[21] This empirical factor is calculated from the background MC. It is approximately the ratio of the number of events reconstructed in the $B^{\pm} \to D_s^{\pm}\phi$ modes and found in the $m_{\rm ES}$-ΔE signal box to all events reconstructed in the $B^{\pm} \to D_s^{\pm}\phi$ modes.

above. Step one – the calculation of the number of background events as a function of the likelihood variable[22] – is shown in the upper left plot. Step two – using the number of expected background events from step one, calculate the expected 90% C.L. upper limit on the number of events in the signal region, N_{UL} – is shown in the upper right plot. Note that there is a one-to-one correspondence between the upper left and the upper right plots. Step three – calculate the signal efficiency as a function of the log-likelihood – is shown in the bottom left plot. And finally, step four – calculate the expected sensitivity on the branching fraction – is shown in the lower right plot. The minima of the lower right plot is the value of the log-likelihood that our optimization prefers for this step.

[22]These plots show the optimization as a function of the log-likelihood, in general the abscissa can be (and was) any of the optimized variables.

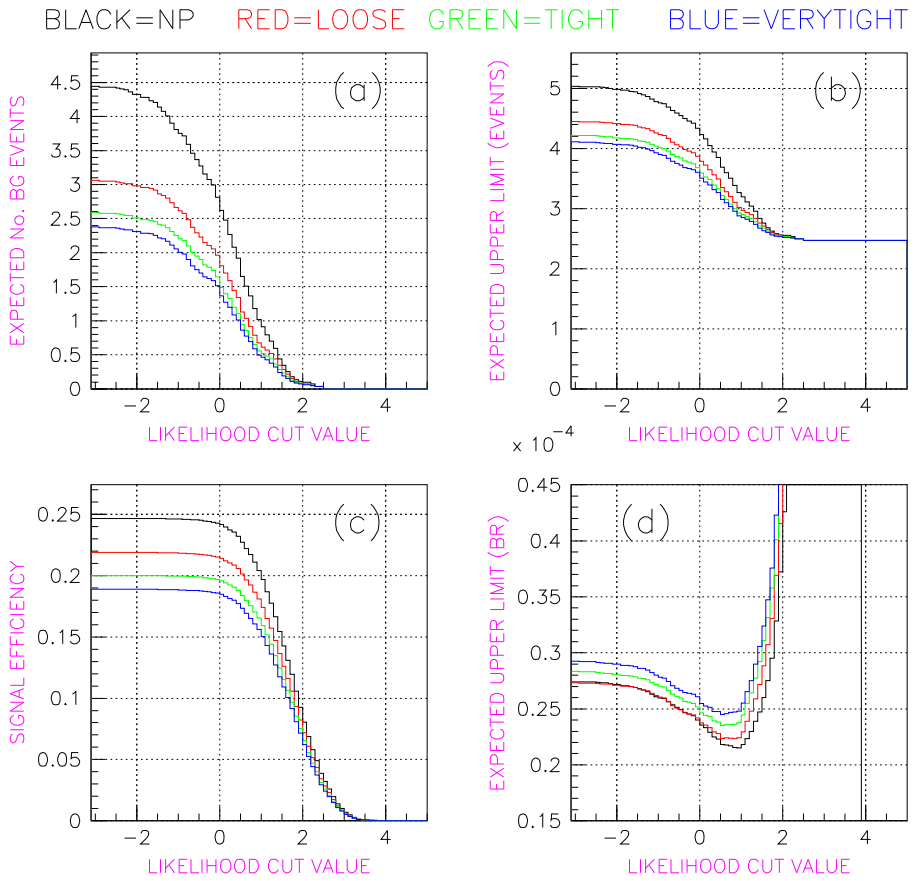

Figure 6.52: Optimization of the log($\mathcal{L}$) cut for $B^{\pm} \to D_s^{\pm}\phi$, $D_s^- \to K^- K_S^0$. As a function of log-likelihood cut value, we show (a) N_{BG} (b) N_{UL} (c) ϵ, and (d) $\mathcal{B}$for four choices of the PID requirement on the K^-, `NotPion` (black), `Loose` (red), `Tight` (green), `VeryTight` (blue). (In case the colors are not clear: For a), b), and c), `NotPion` is the top curve and the others follow in order with `VeryTight` being the lowest curve. For d), this order is reversed.

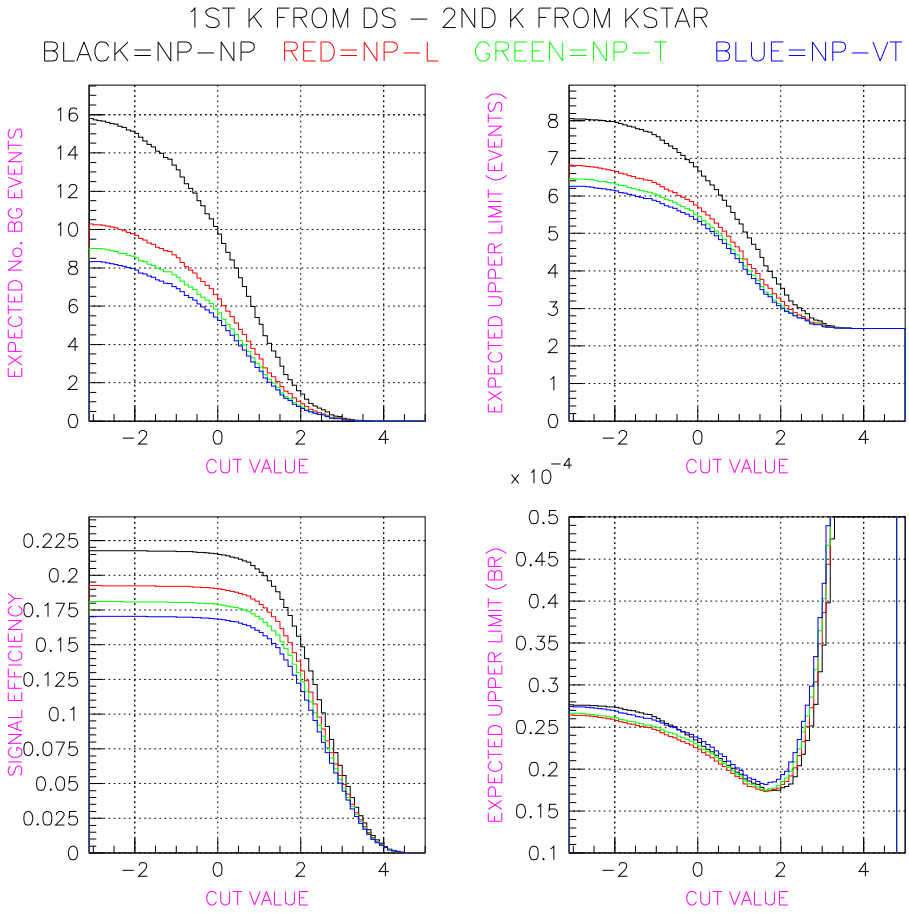

Figure 6.53: Optimization of the $\log(\mathcal{L})$ cut for $B^{\pm} \to D_s^{\pm}\phi$, $D_s^- \to K^{*0}K^-$. As a function of log-likelihood cut value, we show N_{BG} (top left plot), N_{UL} (top right plot) ϵ (bottom left), $\mathcal{B}$(bottom right plot) for four choices of the PID requirement on the kaon from K^{*0} decay., `NotPion` (black), `Loose` (red), `Tight` (green), `VeryTight` (blue). The kaon from $D_s^{\pm}$ decay must pass a `NotPion` requirement.

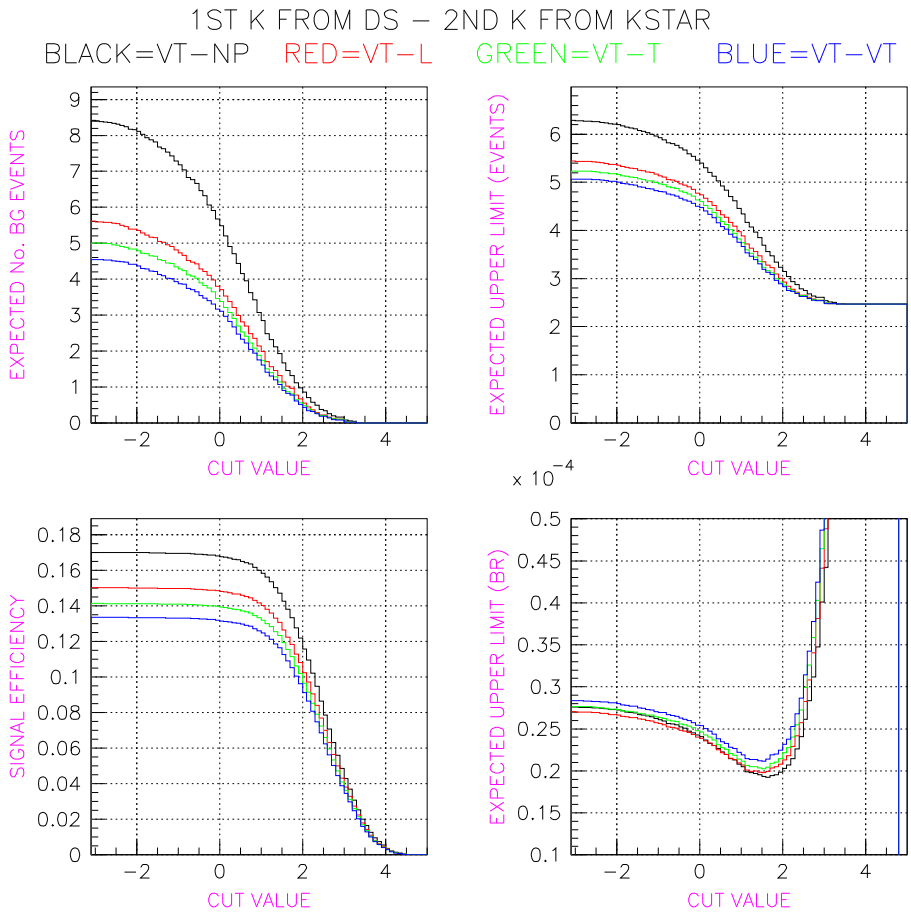

Figure 6.54: Optimization of the log($\mathcal{L}$) cut for $B^{\pm} \to D_s^{\pm}\phi$, $D_s^- \to K^{*0}K^-$. As a function of log-likelihood cut value, we show N_{BG} (top left plot), N_{UL} (top right plot) ϵ (bottom left), $\mathcal{B}$(bottom right plot) for four choices of the PID requirement on the kaon from $D_s^- \to K^{*0}K^-$ decay. `NotPion` (black), `Loose` (red), `Tight` (green), `VeryTight` (blue). The kaon from $D_s^{\pm}$ decay must pass a `VeryTight` requirement.

This procedure was carried out on all variables listed in the beginning of this section. Upon convergence of the optimization we come to the following conclusions.

In general we find that it is not advantageous to tighten any of the kaon PID requirements beyond the least restrictive category, `NotPion`. (Note, since `NotPion` is the least restrictive and is the PID category used in the level one selection, all plots and numbers presented thus far in this paper remain valid.)

Recall we did not include the $\cos(\theta_{\mathrm{T}})$ cut in the likelihood variable because it is highly correlated with the Fisher variable (see Section 6.8). A posteriori, we find that this decision was reasonable: after the $\log(\mathcal{L})$ requirement, the effect of an additional cut on $\cos(\theta_{\mathrm{T}})$ is rather minimal. The efficiency for signal MC is rather flat, especially in the region $0.9 < \cos(\theta_{\mathrm{T}}) < 1.0$. One expects the same upper limit on $\mathcal{B}$ for $\cos(\theta_{\mathrm{T}}) < 0.9$ as for no cut on $\cos(\theta_{\mathrm{T}})$. We have chosen to apply the requirement $\cos(\theta_{\mathrm{T}}) < 0.9$ because we prefer to err on the side of too much background rejection rather than too little.

In performing the optimization procedure we found that the minima in the curves of $\mathcal{B}$ are reasonably broad for all optimized variables. (See, for example, the

Table 6.12: $B^{\pm} \to D_s^{\pm}\phi$. Summary of requirements obtained from the optimization procedure.

$D_s^{\pm}$ Mode	Kaon ID	$\log(\mathcal{L})$ cut	$\cos(\theta_{\rm T})$ cut
$D_s^- \to \phi\pi^-$	`NotPion` for all	> 2.2	< 0.9
$D_s^- \to K^- K_S^0$	`NotPion` for all	> 0.9	< 0.9
$D_s^- \to K^{*0} K^-$	`NotPion` for all	> 1.8	< 0.9

Table 6.13: $B^{\pm} \to D_s^{\pm}\phi$. Expected number of background events in the signal box (N_{BG}), broken out into the various components ($c\overline{c}$, uds, and $B\overline{B}$), and expected upper limit on the branching fraction ($\mathcal{B}$). The efficiency is that of the $\log(\mathcal{L})$ and $\cos(\theta_{\rm T})$ requirements from Table 6.12 on signal MC. Uncertainties are statistical only.

$D_s^{\pm}$ Mode	$c\overline{c}$	uds	$B\overline{B}$	N_{BG}	$\mathcal{B}$	Efficiency
$D_s^- \to \phi\pi^-$	0.45	0.02	0.13	0.61	1.3×10^{-5}	$80.8 \pm 0.5\%$
$D_s^- \to K^- K_S^0$	0.39	0.11	0.27	0.78	2.2×10^{-5}	$78.4 \pm 0.7\%$
$D_s^- \to K^{*0} K^-$	0.89	0.22	0.52	1.64	1.8×10^{-5}	$71.0 \pm 0.6\%$

bottom right plots of Figures 6.52, 6.53, and 6.54.) We are also rather insensitive to our choices for the rescaling factors of the continuum MC[23].

In Table 6.12 we show the results of the optimization procedure. In Table 6.13 we show an estimate of the number of background events we may see in our signal region along with the efficiencies of the requirements in Table 6.12 on signal MC[24].

[23] Recall in Section 6.5 the continuum MC overestimated the number of background events in the data. The rescaling factors we mention here are those discussed just above that were used to bring the MC yields in line with the data.

[24] The efficiencies quoted here are relative to the number of events in our sample after applying the level one cuts, i.e. if we had 10 events after the Level One selection, and we have three after the cuts in Table 6.12, than the efficiency quoted in Table 6.13 would be 30%

At the beginning of this analysis, we made the claim that we would have a good shot at producing a background-free signal region. Note that after applying the requirements of Table 6.12 to the events which pass the level one selection requirements (Section 6.3), the expected number of events in the signal box is of order one for each $D_s^{\pm}$ mode, see Table 6.13. This is, however, not our best estimate of the background level, which will eventually be extracted from a fit to the $m_{\rm ES}$ distribution, but it is a good indication of where we stand.

It is also interesting to note that, within a factor of order two, the expected sensitivities ($\mathcal{B}$) shown in Table 6.13 are the same for all three $D_s^{\pm}$ decay modes. The efficiencies on signal MC range between 70% and 80%, depending on mode with, not surprisingly, the lowest efficiency being $D_s^- \to K^{*0}K^-$ mode – this mode has the highest background and therefore the optimization chooses tighter cuts relative to the other two modes in order to get the best value of $\mathcal{B}$.

In Figure 6.55 we superimpose the $\log(\mathcal{L})$ expectations for signal and background, and we also indicate, graphically, the value of the $\log(\mathcal{L})$ requirement from Table 6.12.

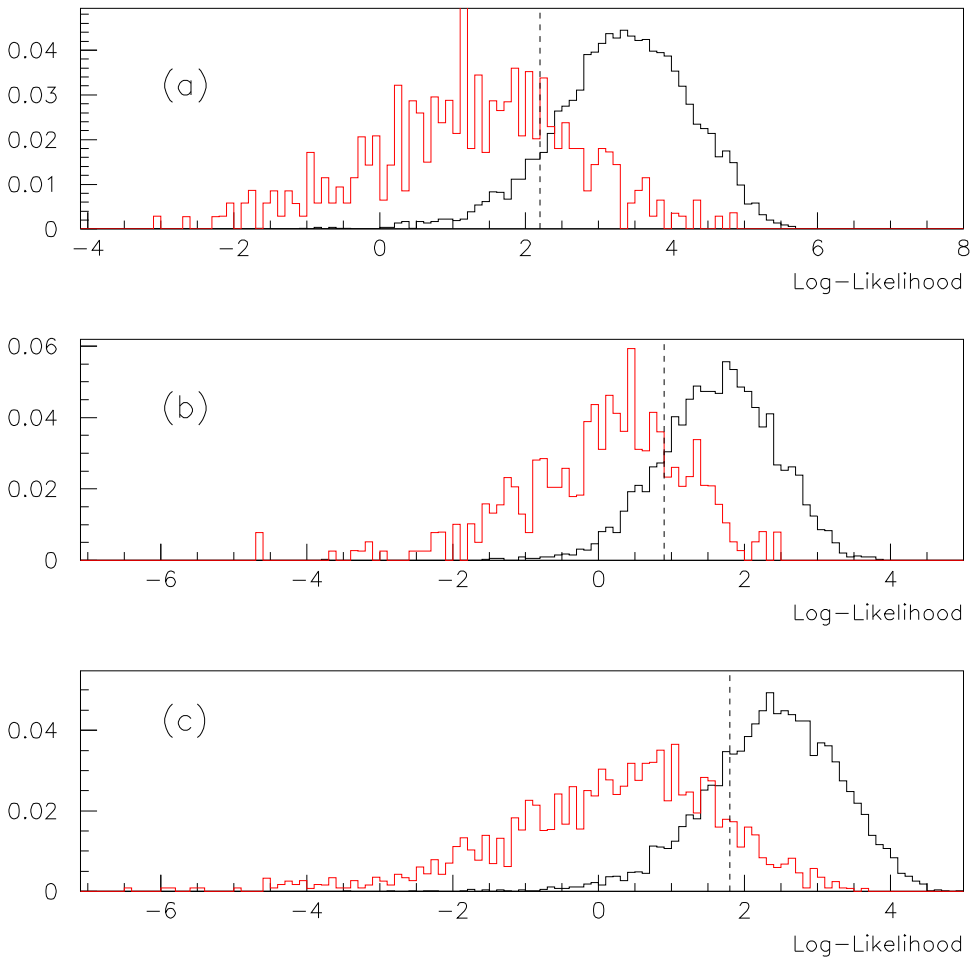

Figure 6.55: $B^{\pm} \to D_s^{\pm}\phi$. $\log(\mathcal{L})$ distributions for signal and background. All distributions are normalized to unity. Black histogram: signal MC. Red histogram: background MC. (a) $B^{\pm} \to D_s^{\pm}\phi$, $D_s^- \to \phi\pi^-$, (b) $B^{\pm} \to D_s^{\pm}\phi$, $D_s^- \to K^- K_S^0$, and (c) $B^{\pm} \to D_s^{\pm}\phi$, $D_s^- \to K^{*0}K^-$. The vertical dashed lines indicate the value of the $\log(\mathcal{L})$ cuts obtained from the optimization procedure, see Table 6.12.

Table 6.14: $B^{\pm} \to D_s^{*\pm}\phi$. Summary of requirements obtained from the optimization procedure.

Mode	Kaon ID	$\log(\mathcal{L})$ cut	$\cos(\theta_{\rm T})$ cut	$E(\gamma)$ cut
$D_s^- \to \phi\pi^-$	`NotPion` for all	> 3.7	< 0.9	> 60 MeV
$D_s^- \to K^-K_S^0$	`NotPion` for all	> 2.5	< 0.9	> 60 MeV
$D_s^- \to K^{*0}K^-$	`NotPion` for all	> 3.3	< 0.9	> 60 MeV

Table 6.15: $B^{\pm} \to D_s^{*\pm}\phi$. Expected number of background events in the signal region (N_{BG}), broken out into the various components ($c\bar{c}$, uds, and $B\overline{B}$), and expected upper limit on the branching fraction for longitudinal ($\mathcal{B}_1$) and transverse ($\mathcal{B}_2$) polarization.

$D_s^{\pm}$ Mode	$c\bar{c}$	uds	$B\overline{B}$	N_{BG}	$\mathcal{B}_1$	$\mathcal{B}_2$
$D_s^- \to \phi\pi^-$	0.38	0.01	0.1	0.49	1.7×10^{-5}	1.7×10^{-5}
$D_s^- \to K^-K_S^0$	0.33	0.04	0.05	0.42	2.6×10^{-5}	2.6×10^{-5}
$D_s^- \to K^{*0}K^-$	0.56	0.08	0.31	0.95	2.1×10^{-5}	2.0×10^{-5}

The optimization of the requirements for the $B^{\pm} \to D_s^{*\pm}\phi$ mode is performed in the same way as for $B^{\pm} \to D_s^{*\pm}\phi$. The two helicity states were optimized separately, but the results are independent of polarization assumptions. In Table 6.14 we show the results of the optimization procedure. After applying the requirements of Table 6.14 to the events which pass the level one selection criteria, the expected number of events in the signal box and the sensitivity of the search are given in Table 6.15. The efficiencies of the requirements of Table 6.14 on signal MC are given in Table 6.16. They range between 74% and 83%, depending on mode.

Table 6.16: Efficiency of the $\log(\mathcal{L})$, $\cos(\theta_{\rm T})$, and $\mathrm{E}(\gamma)$ requirements from Table 6.14 on $B^{\pm} \to D_s^{*\pm}\phi$ signal MC. Uncertainties are statistical only.

$D_s^{\pm}$ Mode	Polarization	Efficiency
$D_s^- \to \phi\pi^-$	Long.	83.1 ± 0.7 %
$D_s^- \to K^- K_S^0$	Long.	82.5 ± 0.9 %
$D_s^- \to K^{*0} K^-$	Long.	76.2 ± 0.8 %
$D_s^- \to \phi\pi^-$	Trans.	81.3 ± 0.7 %
$D_s^- \to K^- K_S^0$	Trans.	77.6 ± 0.9 %
$D_s^- \to K^{*0} K^-$	Trans.	74.8 ± 0.8 %

As we did for the $B^{\pm} \to D_s^{\pm}\phi$ mode, in Figure 6.56 we superimpose the $\log(\mathcal{L})$ expectations for signal and background, and we also indicate, graphically, the value of the $\log(\mathcal{L})$ requirement from Table 6.14. We can see that the optimization produces reasonable results.

6.10 Signal-box excluded look at the on-resonance data

We now take a preview look at the on-resonance data. We will continue to keep our final fit region blinded so as to not bias ourselves in any way. We start with the events passing the level one selection, and we apply the additional requirements on $\log(\mathcal{L})$ and $\cos\theta_T$ listed in Table 6.12. We keep an enlarged signal region $m_{\rm ES} > 5.27$ GeV/c^2 and $|\Delta E| < 40$ MeV blind. As we will discuss in Section 7,

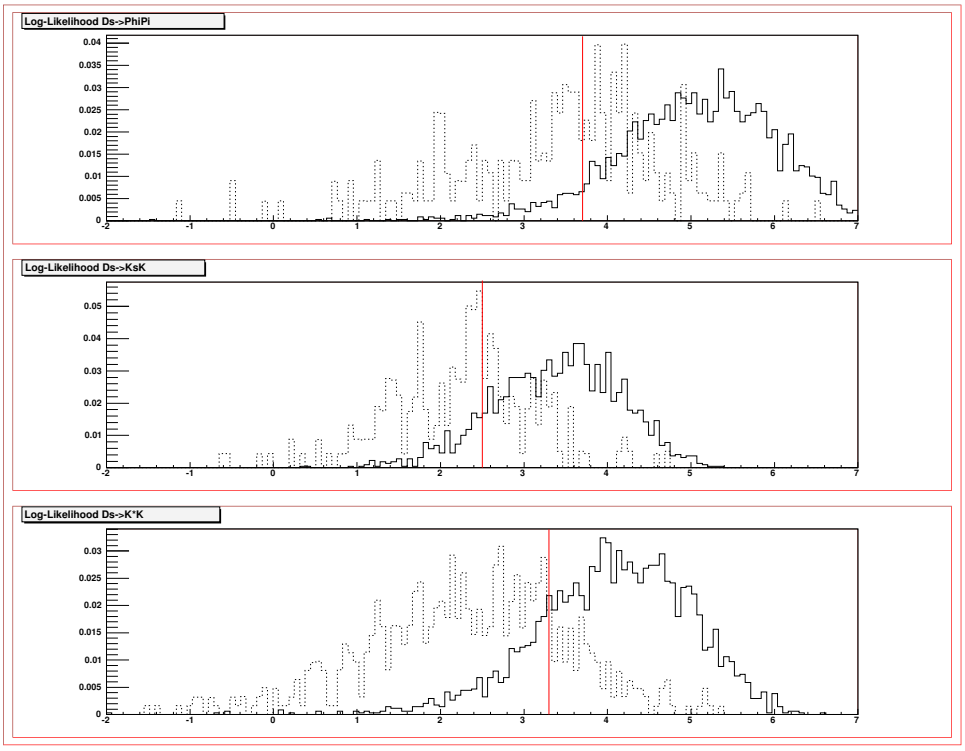

Figure 6.56: $B^{\pm} \to D_s^{*\pm}\phi$. $\log(\mathcal{L})$ distributions for signal and background. All distributions are normalized to unity. Here the longitudinal polarization is shown. Solid histogram: signal MC. Dashed histogram: background MC. (a) $D_s^- \to \phi\pi^-$, (b) $D_s^- \to K^- K_S^0$, and (c) $D_s^- \to K^{*0} K^-$. The red vertical lines indicate the value of the $\log(\mathcal{L})$ cuts obtained from the optimization procedure, see Table 6.14.

our ΔE signal region will be ± 30 MeV around the mean ΔE from signal MC. The mean ΔE in signal MC is shifted by a few MeV from zero, thus, for now we keep a larger ΔE window (± 40 MeV) blind.

The $m_{\rm ES}$ vs. ΔE scatter plot for the surviving $B^{\pm} \to D_s^{\pm}\phi$ candidate events is shown in Figure 6.57. The number of events for each mode is summarized in

Table 6.17: $B^{\pm} \to D_s^{\pm}\phi$. Comparison of the number of data events in Figure 6.57 with the prediction from the MC. An explanation of the method used to obtain the MC prediction is given in the text.

Mode	Data	Prediction
$B^{\pm} \to D_s^{\pm}\phi$, $D_s^- \to \phi\pi^-$	26	24
$B^{\pm} \to D_s^{\pm}\phi$, $D_s^- \to K^- K_S^0$	33	30
$B^{\pm} \to D_s^{\pm}\phi$, $D_s^- \to K^{*0} K^-$	56	63

Table 6.17, together with an expectation from MC. The agreement between data and expectations is excellent. We see nothing out of the ordinary.

The MC expectation in Table 6.17 is derived from the uds, $c\overline{c}$, and generic $B\overline{B}$ MC, after rescaling the $c\overline{c}$ and uds MC predictions by the scaling factors described in Section 6.9. These scaling factors were chosen to make the MC predictions agree with the on-resonance data with $m_{\rm ES} < 5.26$ GeV/c^2 after the level one selection (see Table 6.8). Therefore the comparison between data and expectation in Table 6.17 is essentially a test of the ability of the MC to reproduce the efficiency of the $\log(\mathcal{L})$ and $\cos\theta_T$ requirements on backgrounds. The MC does quite well.

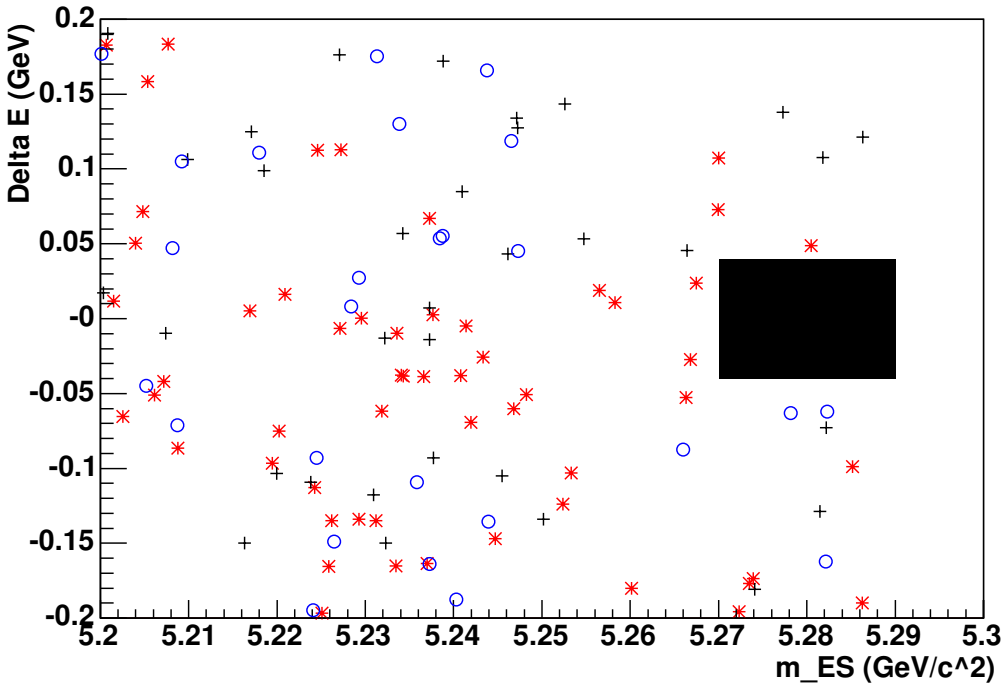

Figure 6.57: $B^{\pm} \to D_s^{\pm}\phi$. Scatter plot of $m_{\rm ES}$ vs. ΔE for events in the on-resonance data sample with all cuts applied. The region $m_{\rm ES} > 5.27$ GeV/c^2 and $|\Delta E| < 40$ MeV is kept blind. Blue circles: $D_s^- \to \phi\pi^-$. Black crosses: $D_s^- \to K^- K_S^0$. Red asterisks: $D_s^- \to K^{*0} K^-$.

Table 6.18: $B^{\pm} \to D_s^{*\pm}\phi$. Comparison of the number of data events in Figure 6.58 with the prediction from the MC.

Mode	Data	Prediction
$B^{\pm} \to D_s^{*\pm}\phi$, $D_s^- \to \phi\pi^-$	20	22
$B^{\pm} \to D_s^{*\pm}\phi$, $D_s^- \to K^- K_S^0$	21	20
$B^{\pm} \to D_s^{*\pm}\phi$, $D_s^- \to K^{*0} K^-$	41	40

The $m_{\rm ES}$ vs. ΔE scatter plot for the surviving $B^{\pm} \to D_s^{*\pm}\phi$ candidate events is shown in Figure 6.58. The number of events for each mode is summarized in Table 6.18, together with an expectation from MC. The agreement between data and expectations is excellent. We see nothing out of the ordinary. The MC expectation in Table 6.18 is derived in the same was as that for Table 6.17.

6.11 Peaking background

In this section we study possible backgrounds that peak in both $m_{\rm ES}$ and ΔE. These could arise, for example, from $B \to \phi\phi\pi$, $B \to K_S^0 K\pi$, $B \to K^* K\pi$. The branching fractions for these modes are not well measured or not measured at all. We choose to estimate these backgrounds directly from the data.

The modes that could contribute to peaking backgrounds do not have a real $D_s^{\pm}$. Thus, we estimate the peaking backgrounds by repeating the analysis in

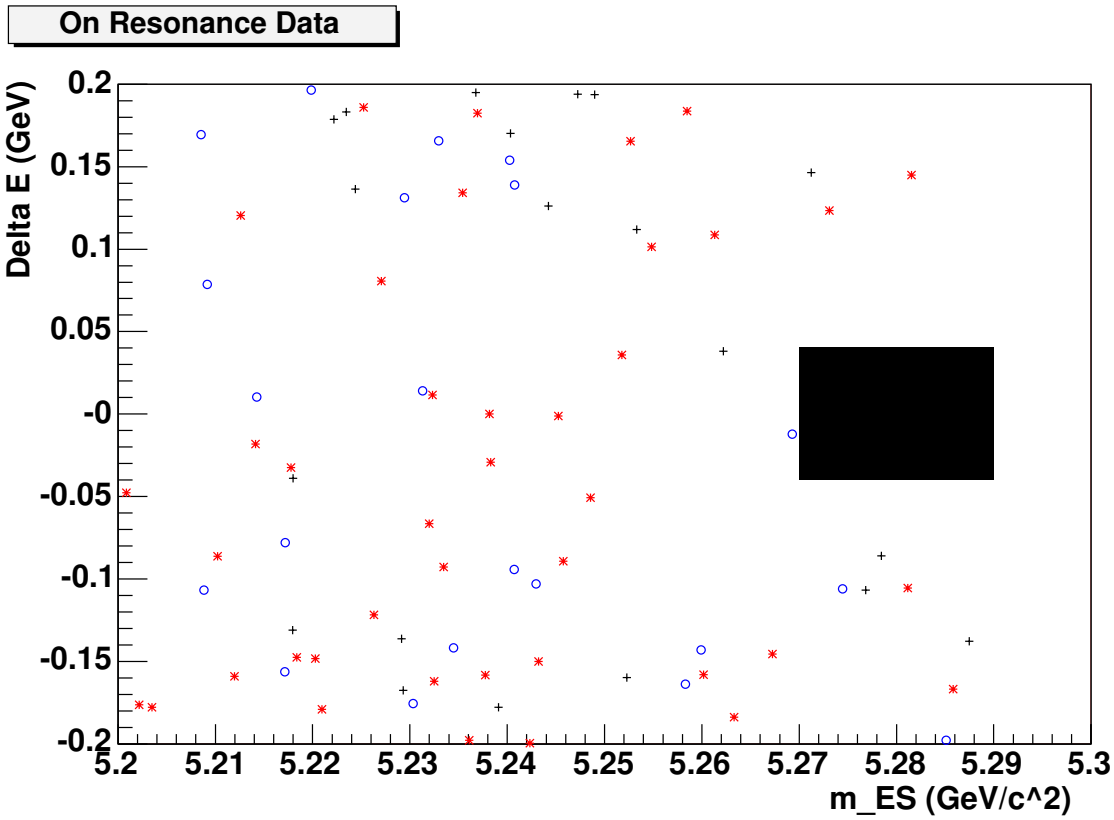

Figure 6.58: $B^{\pm} \to D_s^{*\pm}\phi$. Scatter plot of $m_{\rm ES}$ vs. ΔE for events in the on-resonance data sample with all cuts applied. The region $m_{\rm ES} > 5.27$ GeV/c^2 and $|\Delta E| < 40$ MeV is kept blind. Blue circles: $D_s^- \to \phi\pi^-$. Black crosses: $D_s^- \to K^- K_S^0$. Red asterisks: $D_s^- \to K^{*0} K^-$.

the sideband of the $D_s^{\pm}$. The only assumption that needs to be made is the very reasonable assumption that these backgrounds are linear in the $D_s^{\pm}$ mass region.

The $D_s^{\pm}$ signal region is defined as ± 15 MeV/c^2 around the PDG value of the $D_s^{\pm}$ mass. We define two sideband regions, each of them 15 MeV wide, and well separated (by 15 MeV) from the $D_s^{\pm}$ signal region, see Figure 6.59. Note that the

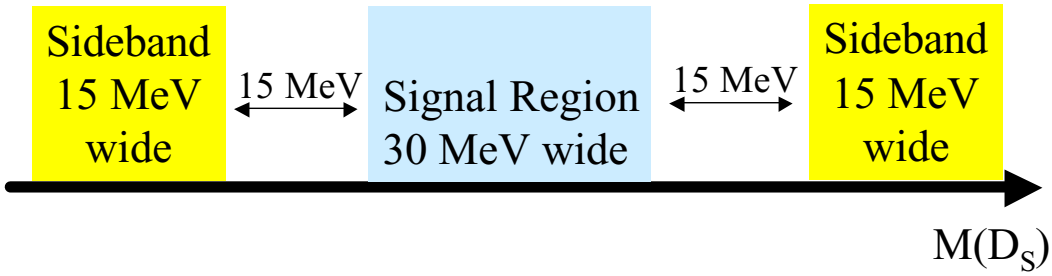

Figure 6.59: A graphical view of the $D_s^{\pm}$ signal and sideband definitions.

total width of the $D_s^{\pm}$ mass ranges covered by the signal and sideband selections are the same.

There are a few subtleties when repeating the analysis in the $D_s^{\pm}$ sideband:

- We remove the selection criterion on the $D_s^{\pm}$ which had been imposed in the level one selection.

- We modify the arbitration procedure, which uses the PDG $D_s^{\pm}$ mass (M_{PDG}), in the obvious way, *i.e.*, we replace M_{PDG} with $M_{PDG} \pm \Delta$, where $\Delta = 30$ MeV/c^2, see Figure 6.59.

- Since the $D_s^{\pm}$ mass is used in the calculation of the likelihood (see Section 6.8), we redefine the likelihood for sideband events to make it consistent (including the normalization) with the likelihood defined for $D_s^{\pm}$ signal region events. To visualize this, say the $D_s^{\pm}$ mass is represented by a single Gaussian. In the normal (i.e. non-sideband) analysis, this is centered

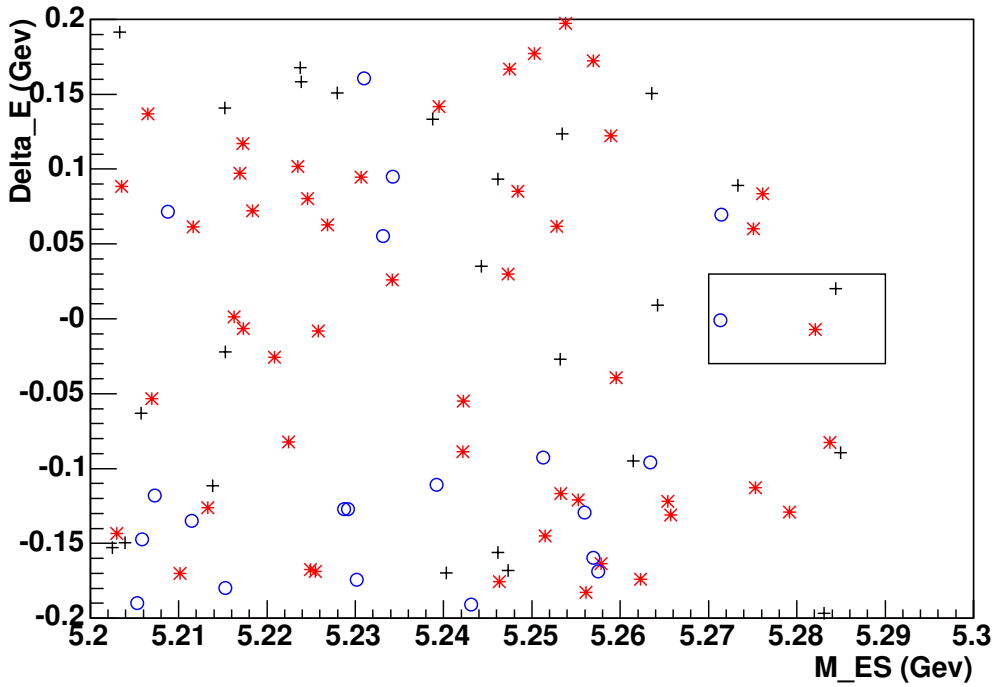

Figure 6.60: $B^{\pm} \to D_s^{\pm}\phi$. Scatter plot of $m_{\rm ES}$ vs. ΔE for events in the on-resonance data sample with all cuts applied but selecting events from the $D_s^{\pm}$ sideband. The box delineates the region $m_{\rm ES} > 5.27$ GeV/c^2 and $|\Delta E| < 30$ MeV. Blue circles: $D_s^- \to \phi\pi^-$. Black crosses: $D_s^- \to K^- K_S^0$. Red asterisks: $D_s^- \to K^{*0} K^-$.

at the $D_s^{\pm}$ mass. Draw a Gaussian centered at the $D_s^{\pm}$ mass in the figure above. For the sideband analysis, we cut this Gaussian in half (draw half the Gaussian in the upper sideband box above and half in the lower sideband box), and define the likelihood in this way. We do this without changing the normalization or the width parameters of the Gaussian. This preserves all attributes of the Gaussian including the normalization.

The resulting scatter plots of $m_{\rm ES}$ vs. ΔE are shown in Figure 6.60 for $B^{\pm} \to D_s^{\pm}\phi$ and Figure 6.61 for $B^{\pm} \to D_s^{*\pm}\phi$. If there were no real $D_s^{\pm}$ mesons in our background, the number of events in Figure 6.60(6.61) should be the same, within statistics, as the number of events in the equivalent plot for the $D_s^{\pm}$ signal region, Figure 6.57 (6.58). (See also Table 6.17)[25].

For the $B^{\pm} \to D_s^{\pm}\phi$ sample, we find 95 events in Figure 6.60. Of these, 21 are $D_s^- \to \phi\pi^-$, 25 are $D_s^- \to K^-K_S^0$, and 49 are $D_s^- \to K^{*0}K^-$. These numbers are about 20% lower than the number of events in the $D_s^{\pm}$ signal region. This is in qualitative agreement with the $D_s^{\pm}$ content of our background that can be inferred from examining Figures 6.2 and 6.26.

There are three events with $m_{\rm ES} > 5.27$ GeV/c^2 and $|\Delta E| < 30$ MeV in Figure 6.60. This is in good agreement with the rough background expectation of $0.025 \times 95 = 2.4$, where 95 is the total number of events in Figure 6.60 and 0.025 is the empirical scaling factor in extrapolating from the larger $m_{\rm ES}$-ΔE region into the signal box[26]. Thus, we see no evidence for peaking backgrounds in $B^{\pm} \to D_s^{\pm}\phi$.

For the $B^{\pm} \to D_s^{*\pm}\phi$ sample, we find 49 events in Figure 6.61. Of these, 12 are $D_s^- \to \phi\pi^-$, 16 are $D_s^- \to K^-K_S^0$, and 21 are $D_s^- \to K^{*0}K^-$. While there

[25]We neglect the fact that a small region of the plots in Figures 6.57 and 6.58 has been excluded by the blinding procedure.

[26]This factor of 0.025 is derived from Table 6.5 and was used in the optimization procedure described in Section 6.9.

are zero events with $m_{\rm ES} > 5.27$ GeV/c^2 and $|\Delta E| < 30$ MeV in Figure 6.61, this is still in agreement with the rough background expectation of $0.025 \times 49 = 1.2$. Thus, as for the $B^{\pm} \to D_s^{\pm}\phi$ modes, we see no evidence for peaking backgrounds in $B^{\pm} \to D_s^{*\pm}\phi$.

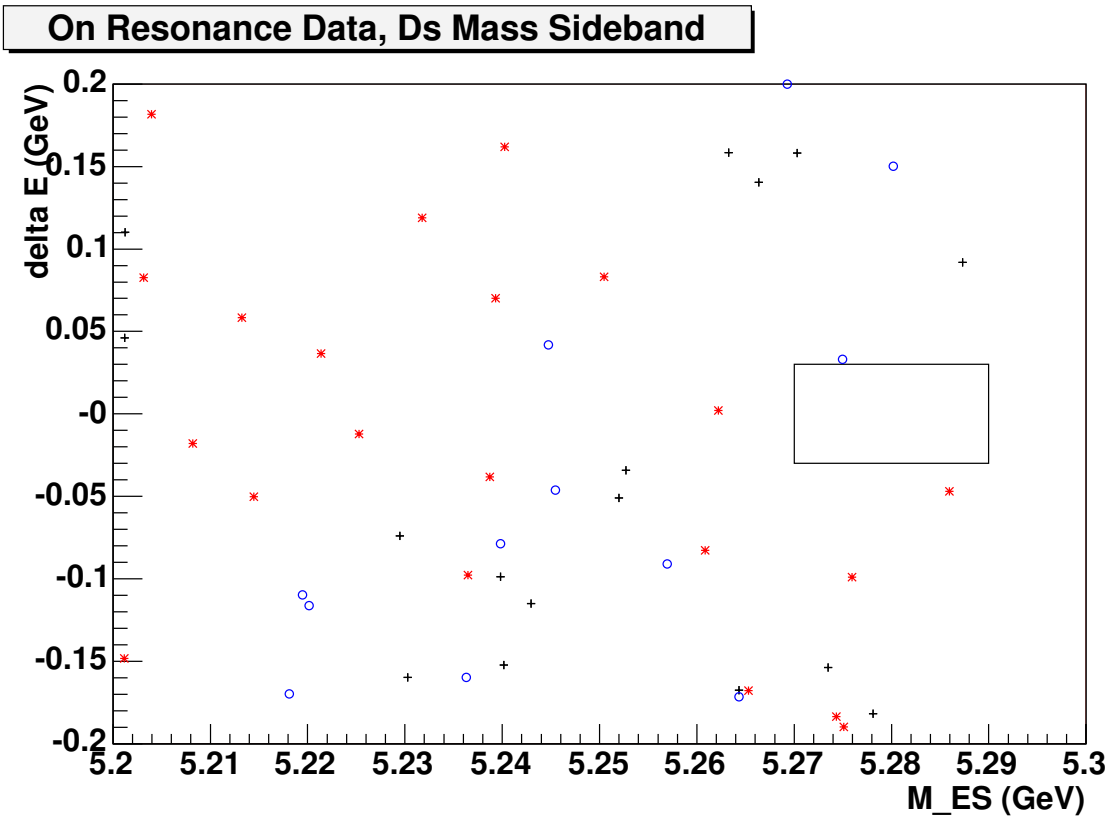

Figure 6.61: $B^{\pm} \to D_s^{*\pm}\phi$. Scatter plot of $m_{\rm ES}$ vs. ΔE for events in the on-resonance data sample with all cuts applied but selecting events from the $D_s^{\pm}$ sideband. The box delineates the region $m_{\rm ES} > 5.27$ GeV/c^2 and $|\Delta E| < 30$ MeV. Blue circles: $D_s^- \to \phi\pi^-$. Black crosses: $D_s^- \to K^- K_S^0$. Red asterisks: $D_s^- \to K^{*0} K^-$.

Chapter 7

Data unblinding and signal extraction

7.1 $B^{\pm} \to D_s^{\pm}\phi$

In this section we discuss the final unblinding of the data. That is, we open up the $m_{\rm ES}$-ΔE signal box to see what is there. Then we will fit the signal region to a signal-plus-background shape and determine our signal yield. We define the ΔE signal region as ± 30 MeV ($\approx 3\sigma$) around the means of the ΔE distributions in signal MC: $< \Delta E >$ = -3 MeV, 0 MeV, and -1 MeV for $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, and $D_s^- \to K^{*0}K^-$, respectively (see Figure 7.1). The signal yield will then be extracted from a fit to the $m_{\rm ES}$ distribution of events passing the ΔE selection.

The unblinded $m_{\rm ES}$ vs. ΔE scatter plot for the $B^{\pm} \to D_s^{\pm}\phi$ mode is shown in Figure 7.2. We see no enhancement of events in the region where signal is expected. There are two events with $m_{\rm ES} > 5.27$ GeV/c^2 in the ΔE signal region

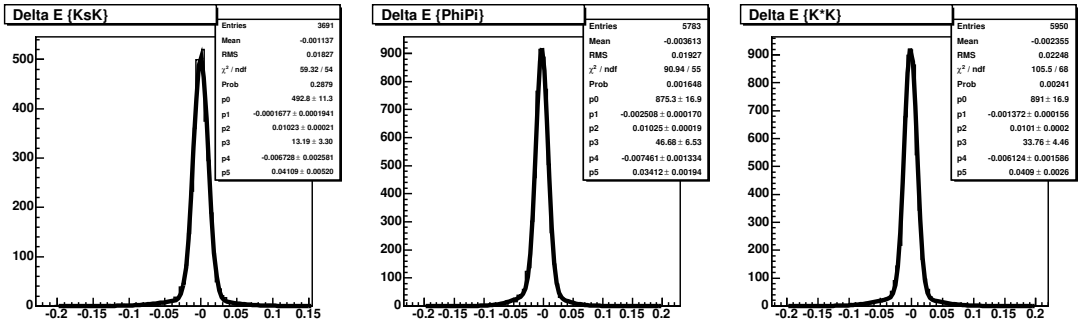

Figure 7.1: ΔE distributions for the three signal MC samples. From left to right: $D_s^- \to K^- K_S^0$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$.

Table 7.1: $B^\pm \to D_s^\pm \phi$. Summary of the $m_{\rm ES}$ and ΔE values for the five events in the on-resonance data sample with $m_{\rm ES}$ and ΔE near the signal region.

$D_s^\pm$ Mode	$m_{\rm ES}$ (GeV/c^2)	ΔE (MeV)	Passes ΔE?
$D_s^- \to K^- K_S^0$	5.271	26.4	YES
$D_s^- \to K^- K_S^0$	5.277	30.2	NO
$D_s^- \to K^{*0} K^-$	5.273	3.0	YES
$D_s^- \to K^{*0} K^-$	5.279	29.8	NO
$D_s^- \to K^{*0} K^-$	5.273	30.5	NO

and three events just outside it. The $m_{\rm ES}$ and ΔE values of these five events are listed in Table 7.1. Note that the background predictions from the optimization procedure was 5.7 events (using the numbers from Table 6.13 scaled up to include Run 4 luminosity).

Next, we perform an unbinned extended maximum likelihood fit of the data $m_{\rm ES}$ distribution. The fit method consists of fitting simultaneously the $m_{\rm ES}$ distributions for the ΔE signal region and the ΔE sideband region ($|\Delta E| < 200$

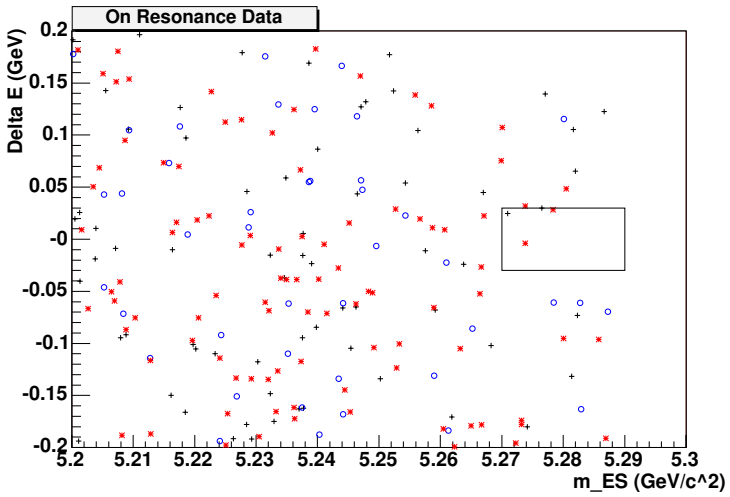

Figure 7.2: $B^{\pm} \to D_s^{\pm}\phi$. Scatter plot of $m_{\rm ES}$ vs. ΔE for events in the on-resonance data sample with all cuts applied. This is the same as Figure 6.57 but with the signal box revealed and including Run 4. There are 206 entries in this plot; if one assumes a flat background, one can get an idea as to the number of background events in our signal box by multiplying this by 0.025 (206∗0.025 = 5.2) which is an empirical factor which relates the numbers of events in this whole plot to the number in the signal region. The box delimits the region $m_{\rm ES} > 5.27$ GeV/c^2 and $|\Delta E| < 30$ MeV. Blue circles: $D_s^- \to \phi\pi^-$. Black crosses: $D_s^- \to K^- K_S^0$. Red asterisks: $D_s^- \to K^{*0} K^-$.

MeV, excluding the signal region). The ΔE signal region is fit to the sum of a background distribution and a signal distribution. The signal distribution is a Gaussian with mean and width fixed at 5.2799 GeV/c^2 and 2.7 MeV/c^2, respectively. For the background distribution we use a threshold function first used by the ARGUS experiment [19] (See Appendix A). Its endpoint is determined by the total CM energy of the electron and positron beams and is set to 5.291 GeV/c^2. The ΔE sideband region is fit to an Argus function only, and the slope parame-

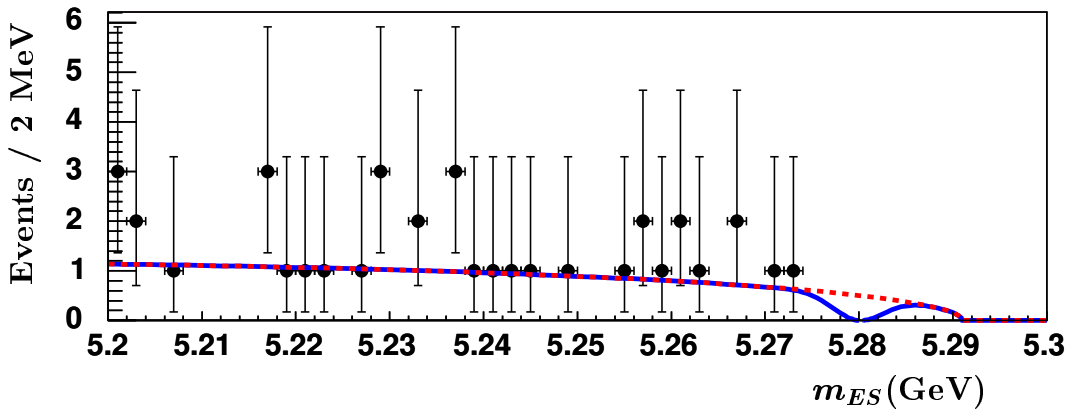

Figure 7.3: $B^{\pm} \to D_s^{\pm}\phi$. The $m_{\rm ES}$ distribution of the events in the ΔE signal region, with the superimposed Argus + Gaussian fit.

ter of the Argus is taken to be the same in the ΔE signal and sideband regions. This allows us to use the events in the ΔE sideband region to help pin down the background in the signal region. Note, we make the reasonable assumption that the slope of the Argus is independent of ΔE.

In the fit, we allow the Gaussian yield to fluctuate negative, but we constrain the sum of the Argus and Gaussian function to remain positive. We need to do this because otherwise the fit is highly unstable and tends to converge to an infinitely negative yield. This can happen in low statistics maximum likelihood fits when there are no events in the signal region [20].

The data and the corresponding fit are shown in Figures 7.3 and 7.4. The result of the fit for the number of signal events is $N_{sig} = -1.6^{+0.7}_{-0.0}$, where the

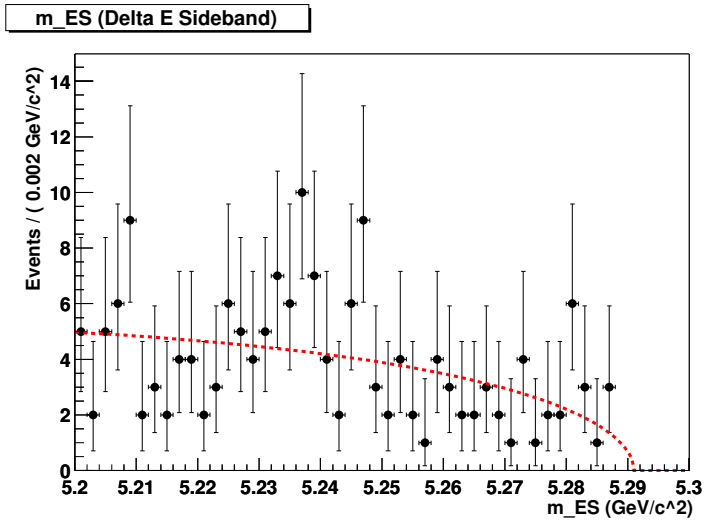

Figure 7.4: $B^{\pm} \to D_s^{\pm}\phi$. The $m_{\rm ES}$ distribution of the events in the ΔE sideband region, with the superimposed Argus fit. This is the result of a simultaneous fit to the signal and sideband ΔE regions.

uncertainty corresponds to a change in 0.5 in $\log(\mathcal{L})$. The value of the Argus slope parameter returned by the fit is $\kappa = -7.5 \pm 7.8$.

In Figure 7.5 we show the likelihood as a function of N_{sig}. The likelihood function is precisely an exponential function, $\exp(-N_{sig})$. This is simply because there are essentially no events in the $m_{\rm ES}$ signal region[1], so the likelihood as a function of N_{sig} is just the Poisson probability for zero events observed, which goes like $\exp(-N_{sig})$.

[1]The closest event to the mean expected signal $m_{\rm ES}$ is at $m_{\rm ES} = 5.273$ GeV/c^2, *i.e.*, over 3σ away from the mean, see Table 7.1

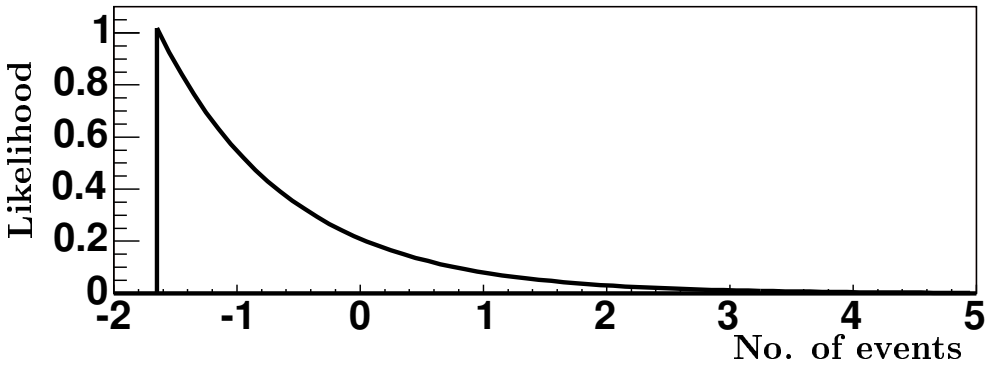

Figure 7.5: $B^{\pm} \to D_s^{\pm}\phi$. The likelihood function, in arbitrary units, from the fit of Figure 7.3 as a function of the number of signal events (N_{sig}). The likelihood function is identical to $\exp(-N_{sig})$. Note that at $N_{sig} = -1.6$ the likelihood function hits a "brick wall", since for $N_{sig} < -1.6$ the total fitted function (Argus + Gaussian) becomes negative and this is not allowed in the fit. The fact that the likelihood function goes to zero at $N_{sig} = -1.6$ is simply an artifact of plotting.

7.2 $B^{\pm} \to D_s^{*\pm}\phi$

Just as for the $B^{\pm} \to D_s^{\pm}\phi$ mode, we define the ΔE signal region as ± 30 MeV ($\approx 3\sigma$) around the means of the ΔE distributions in signal Monte Carlo. In this mode with a $D_s^{*\pm}$ there is a large tail at negative values of ΔE which is due to picking the wrong photon in the $D_s^{*\pm} \to D_s^{\pm}\gamma$ decay, see Figure 7.6. Thus, we use the means of the ΔE distributions for which the photon in the reconstructed $D_s^{*\pm}$ is required to be the photon that was produced in the MC generation for the $D_s^{*\pm} \to D_s^{\pm}\gamma$ decay. These give similar results for both helicity states: $B^{\pm} \to D_s^{\pm}\phi$, $< \Delta E > =$ -3 MeV, 0 MeV, and -2 MeV for $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, and $D_s^- \to K^{*0}K^-$, respectively. The signal yield will then be extracted from a fit to the $m_{\rm ES}$ distribution of events passing the ΔE selection.

The unblinded $m_{\rm ES}$ vs. ΔE scatter plot for the $B^{\pm} \to D_s^{*\pm}\phi$ mode is shown in Figure 7.7. There are seven events with $m_{\rm ES} > 5.27$ GeV/c^2 in the ΔE signal region and one event just outside it. The $m_{\rm ES}$ and ΔE values of these eight events are listed in Table 7.2. The (rough) background predictions from the optimization procedure was 3.6 events in the signal box, (using the numbers from Table 6.15 scaled up to include Run 4 luminosity).

We use the same method as for the $B \to D_s^-\phi$ modes (see Section 7.1), to perform on the data an unbinned extended maximum likelihood fit of $m_{\rm ES}$. The data

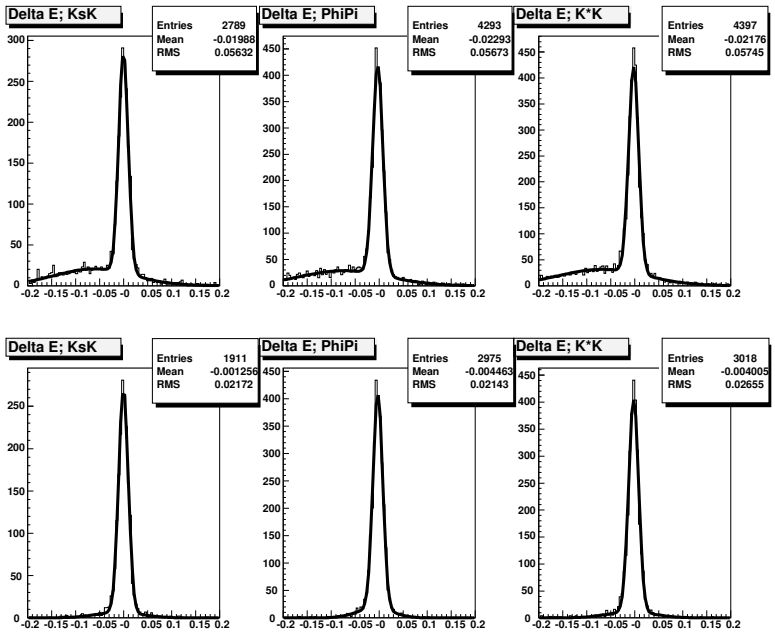

Figure 7.6: $B^{\pm} \to D_s^{*\pm}\phi$. ΔE distributions in $B^{\pm} \to D_s^{*\pm}\phi$ Monte Carlo for the three $D_s^{\pm}$ modes. From left to right: $D_s^- \to K^-K_S^0$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$. The bottom row of plots includes the requirement that the photon from the $D_s^{*\pm} \to D_s^{\pm}\gamma$ decay be the same photon that was generated in the MC. The top row of plots does not include this requirement. Note, the low mass tail is the result of making a $D_s^{*\pm}$ from a $D_s^{\pm}$ and a random photon found in the detector. Since there are a lot of low energy photons we make lots of $D_s^{*\pm}$ mesons (and hence $B^{\pm}$ mesons) with too low of a mass. This leads to a negative ΔE (see 6.1).

and the corresponding fits are shown in Figure 7.8 and Figure 7.9. In Figure 7.10 we show the likelihood as a function of N_{sig}. The result of the fit for the number of signal events is $N_{sig} = 3.4^{+2.8}_{-2.1}$, where the uncertainty corresponds to a change of 0.5 in $\log(\mathcal{L})$. The value of the Argus slope parameter returned by the fit is $\kappa = -30.8 \pm 9.1$.

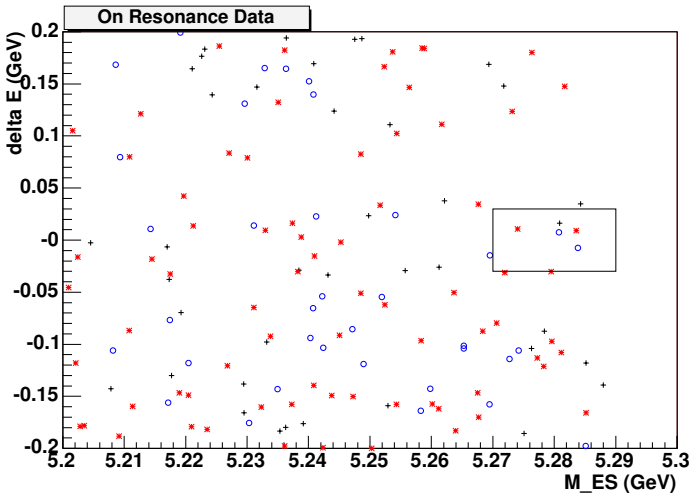

Figure 7.7: $B^{\pm} \to D_s^{*\pm}\phi$. Scatter plot of $m_{\rm ES}$ vs. ΔE for events in the on-resonance data sample with all cuts applied. This is the same as Figure 6.58 but with the signal box revealed and including Run 4 data. There are 151 entries in this plot; if one assumes a flat background, one can get an idea as to the number of background events in our signal box by multiplying this by 0.025 ($151{*}0.025 = 3.8$) which is an empirical factor which relates the numbers of events in this whole plot to the number in the signal region. The box delineates the region $m_{\rm ES} > 5.27$ GeV/c^2 and $|\Delta E| < 30$ MeV. Blue circles: $D_s^- \to \phi\pi^-$. Black crosses: $D_s^- \to K^- K_S^0$. Red asterisks: $D_s^- \to K^{*0} K^-$.

Table 7.2: $B^{\pm} \to D_s^{*\pm}\phi$. Summary of the $m_{\rm ES}$ and ΔE values for the eight events in the on-resonance data sample with $m_{\rm ES}$ and ΔE near the signal region.

Mode	$m_{\rm ES}$ (GeV/c^2)	ΔE (MeV)	Passes ΔE?
$D_s^- \to \phi\pi^-$	5.284	-7.3	YES
$D_s^- \to \phi\pi^-$	5.280	5.2	YES
$D_s^- \to K^- K_S^0$	5.280	16.8	YES
$D_s^- \to K^- K_S^0$	5.284	35.6	NO
$D_s^- \to K^{*0} K^-$	5.284	9.1	YES
$D_s^- \to K^{*0} K^-$	5.274	11.9	YES
$D_s^- \to K^{*0} K^-$	5.280	-29.7	YES
$D_s^- \to K^{*0} K^-$	5.272	-32.0	YES

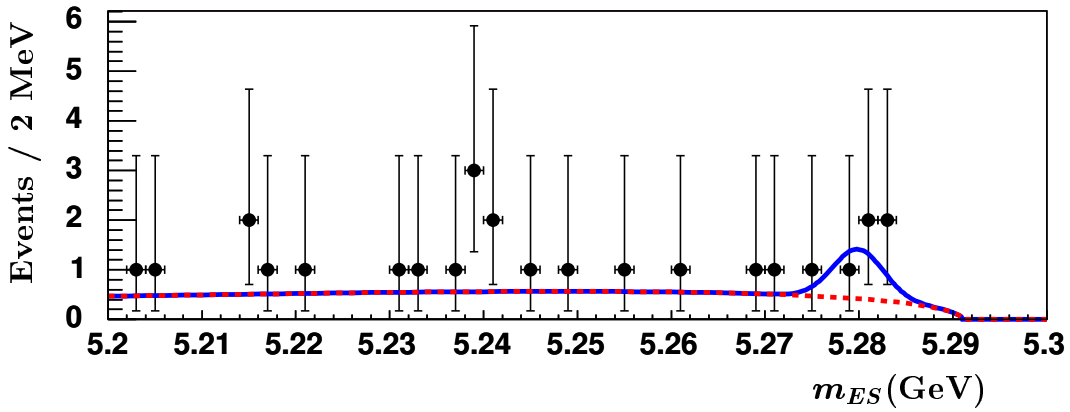

Figure 7.8: $B^{\pm} \to D_s^{*\pm}\phi$. The $m_{\rm ES}$ distribution of events in the ΔE signal region, with the superimposed Argus + Gaussian fit. This is the result of a simultaneous fit to the signal and sideband ΔE regions.

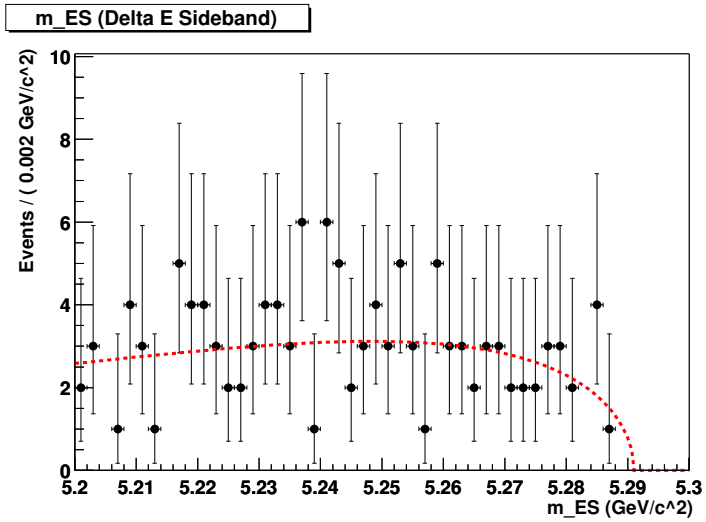

Figure 7.9: $B^{\pm} \to D_s^{*\pm}\phi$. The $m_{\rm ES}$ distribution of events in the ΔE sideband region, with the superimposed Argus fit. This is the result of a simultaneous fit to the signal and sideband ΔE regions.

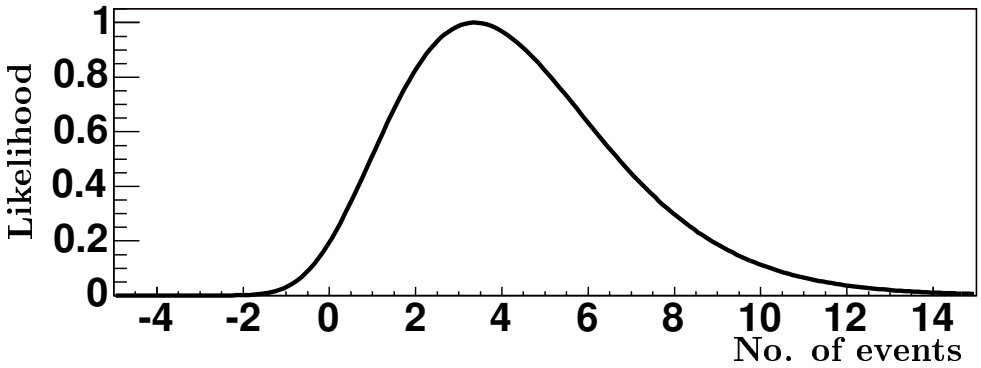

Figure 7.10: $B^{\pm} \to D_s^{*\pm}\phi$. The likelihood function, in arbitrary units, from the fit of Figure 7.8 as a function of the number of signal events (N_{sig}).

7.3 Data unblinding: summary

Now that we have unblinded the signal region, we would like to make a few comments about what we see. It is interesting to note that while we see no signal in either the $B^{\pm} \to D_s^{\pm}\phi$ or the $B^{\pm} \to D_s^{*\pm}\phi$ modes, in the former, the background fluctuated low, while in the latter, the background fluctuated high. If we had not done this analysis in the $B^{\pm} \to D_s^{\pm}\phi$ mode, after seeing Figure 7.7 one might have jumped to the conclusion that we are beginning to see a signal in the $B^{\pm} \to D_s^{*\pm}\phi$ mode. However, the fit yield is less than 2σ above zero and this is certainly not the least bit a decisive signal. As we can see from the $B^{\pm} \to D_s^{\pm}\phi$ there is just as good of a chance for the us to see not only zero signal events, but zero background events.

This is a good example of how much background events can actually fluctuate in any type of statistical analysis such as this. It should be kept in mind when one is trying to claim discovery how easily it is for backgrounds to fool us into seeing something that is not really there (especially if we really want to see something).

Chapter 8

Evaluation of Systematic Uncertainties

Now that we have fit for our signal yield we would like to calculate our limits on $\mathcal{B}(B^{\pm} \to D_s^{(*)\pm}\phi)$. Prior to doing this we need to evaluate the reliability of the MC to correctly model the signal selection efficiency of the data and determine our systematic uncertainties. Section 8.1 will discuss corrections necessary to bring the MC efficiencies into line with the data.

Systematic uncertainties take various forms but each enters into the calculation of the $B^{\pm} \to D_s^{(*)\pm}\phi$ branching fractions using (Recall Equation 5.1):

$$\mathcal{B} < \frac{N_{UL}}{N_{B\overline{B}} \sum_i \epsilon_i \times \mathcal{B}_i}$$

Each of the quantities in the denominator will have a systematic uncertainty associated with it, but the experimental systematic uncertainties from this analysis all relate to the determination of the efficiency, ϵ.

The uncertainty in ϵ is discussed in Sections 8.2, 8.3 and 8.4. The uncertainty associated with identifying kaons is discussed in Section 8.2, with finding tracks in the detector in Section 8.3, and with our choice of the selection criteria in Section 8.4. Monte Carlo simulated data statistics are covered in Section 8.5.

Remaining in the denominator of Equation 5.1 is the number of B mesons in our sample, $N_{B\bar{B}}$, and the branching fractions of all the daughter particles in the $B^{\pm} \to D_s^{(*)\pm}\phi$ decays, $\mathcal{B}_i$. The systematic uncertainties associated with these quantities are discussed in Sections 8.6 and 8.7.

One final note: Since our final result is the sum of three different decay modes of the $D_s^{\pm}$, the treatment of correlations between uncertainties is a non-trivial calculation. All care was taken to account for correlations between systematic uncertainties, but in presenting the uncertainties here we will present them as averages over all modes so the reader may get a feel for their magnitude.

8.1 Corrections to MC efficiencies

Monte Carlo simulated data is a very useful tool in this type of analysis. In a perfect world, our MC would precisely mimic real data. This, of course, is not the case. Even given the high level of sophistication of the *BABAR* MC, there are still differences between the simulated data and the real thing. One of the

most noticeable places where this difference could occur is in particle identification (PID) and track reconstruction. In both cases, we are dealing with a situation in which the MC might over- or underestimate compared to the data our efficiency for selecting our signal events. In the case of track reconstruction, our ability to correctly identify a charged track in our detector may have a dependency on momentum, or the geometrical location of the track in the detector that is not perfectly modeled by the MC. In the case of PID, our ability to correctly identify whether a track should be a kaon rather than a pion may be imperfect – The following two sections discuss in turn studies performed to determine the extent to which the MC efficiency misrepresents the data.

8.2 PID uncertainties

To study if this is a potential source of error we would like to compare how well we identify kaons in the data to how well we do using MC in a way that does not use any particle identification techniques. To do this we start with a sample of D^{*+} decays ($D^{*+} \to D^0\pi^+$ with $D^0 \to K^-\pi^+$) that are kinematically identified[1]. We obtain this sample for both data and MC. Since the kaons and

[1] The mass difference between the D^{*+} and the D^0 is $145\,\mathrm{MeV}/c^2$. The mass of the π^+ is $140\,\mathrm{MeV}/c^2$. Because most of the energy available in the $D^{*+} \to D^0\pi^+$ decay must go to the masses of the D^0 and the π^+, there is only $5\,\mathrm{MeV}/c^2$ to be shared for the momentum of the D^0 and the π^+. For this reason we call the π^+ a "slow pion". The daughter particles of the D^0 decay have much higher momentum (they share the mass energy of the D^0, $1869\,\mathrm{MeV}/c^2$).

pions are absolutely identified by the kinematics, we can determine how often the tracks that are combined to make the D^{*+} are correctly identified as kaons and pions.

These studies lead us assign a 2% systematic uncertainty per kaon. In addition to this we assign another 2.5% uncertainty due to the fact that our kaons have a slightly different momentum spectrum from the control sample. These are added in quadrature to obtain 3.6% systematic uncertainty per kaon. This results in a systematic uncertainty of 14% for the efficiency of the modes with four charged kaons ($D_s^- \to K^{*0}K^-$, $D_s^- \to \phi\pi^-$), and 11% for the mode with three charged kaons ($D_s^- \to K^-K_S^0$). This may seem high, and indeed it is our largest experimental systematic uncertainty. However, as we shall see in Section 8.7, the branching fractions of the $D_s^\pm$ sub-decay modes have a comparable uncertainty.

In the $B^\pm \to D_s^{*\pm}\phi$ search, there is an additional systematic due to the uncertainty on the efficiency to reconstruct the photon in $D_s^{*\pm} \to D_s^\pm\gamma$. The photon reconstruction efficiency is studied using a sample of π^0 mesons. π^0 meson candidates are reconstructed from the kinematic combination of any two photons in the detector. The photon uncertainty is determined from the efficiency to reconstruct neutral pions. (Neutral pions decay via $\pi^0 \to \gamma\gamma$.) The efficiency for reconstructing $\tau^\pm \to \rho^\pm\nu$ is compared to that for reconstructing $\tau^\pm \to \pi^\pm\nu$. Since

Because of these two reasons, it is entirely possible to identify this D^{*+} decay chain with solely kinematic information.

$\rho^{\pm} \to \pi^{\pm}\pi^0$, this comparison gives us information on the π^0 reconstruction efficiency. The extent to which these two efficiencies agree is taken as a systematic uncertainty in the pion reconstruction efficiency and since $\pi^0 \to \gamma\gamma$ the photon systematic uncertainty is half this value, or 1.8%. The reconstruction efficiencies of the signal MC differ by at most 1% between two polarization states of the $D_s^{*\pm}$, thus we take a 1% uncertainty due to the fact that in the data, the polarization of the $D_s^{*\pm}$ is unknown.

8.3 Track reconstruction uncertainties

Determining the quality of track reconstruction has two parts: correcting the MC efficiency to bring it in line with the data, and assigning a track reconstruction systematic uncertainty. There are many methods employed at *BABAR* to understand our charged track reconstruction efficiency. The main strategy is the following. In order to measure the efficiency for charged track reconstruction one can isolate a sample of events for which the physics requires a definite track multiplicity, N. If one can select such events by criteria placed on N-1 of the tracks, the fraction, f_N, of selection events in which an N^{th} track is found is a measure of efficiency times acceptance for that track. By measuring f_N for data and MC

one can determine any correction that is necessary to bring the MC in line with the data.

This strategy is applied to the process $e^+e^- \to \tau^+\tau^-$. Two common decay modes of the tau are called one-prong and three-prong. In the one-prong decay mode, the tau decays to one charged particle (e.g. a muon, electron, kaon, or pion) and any number of neutral particles. The three-prong decay mode has three charged tracks in the final state and any number of neutral particles; the charged tracks are either kaons or pions. A sample of $e^+e^- \to \tau^+\tau^-$ events where one tau decay is three-prong and one tau decay is one-prong is selected using only criteria based on two of the tracks on the three-prong side and the single track (lepton) of the other tau. One measures the fraction of selected of events which contain a reconstructed fourth track (according to some criteria, such as `GoodTracksLoose`). This fraction, corrected by the detector acceptance, gives the absolute tracking efficiency.

As usual, data and MC efficiencies are then compared. If the data efficiency differs from the MC efficiency, the MC efficiency is multiplied by a correction factor to make it agree with the data. The corrections for this analysis are tracking-category dependent. For `GoodTracksVeryLoose` the MC efficiency is too high and must be reduced by 0.5% per track. For `GoodTracksLoose` the MC efficiency is again too high and must be reduced by 0.8% per track. The systematic uncer-

Table 8.1: The track reconstruction Systematic uncertainty assignment per track. The total systematic uncertainty is a weighted sum of the tracks that fall into each category.

Track-Quality Category	$p_T < 200$MeV/c	$p_T > 200$MeV/c
`GoodTracksVeryLoose`	1.4%	0.9%
`GoodTracksLoose`	1.4%	0.6%

tainties on the tracking efficiency can be obtained by comparing the results from the different three-prong and one-prong decay modes (i.e. where the three prongs are kaons vs. where they are all pions).

Our total tracking uncertainty is calculated by adding linearly the uncertainty from each track. All tracks in our sample are either `GoodTracksVeryLoose` or `GoodTracksLoose`. Table 8.2 shows which category is used for each track in this analysis. The uncertainty associated with the tracking correction of the MC is dependent on whether or not the tracks have transverse momentum, $p_T >$ 200MeV/c. Except for the pion from the K^{*0} decay, more than 99.5% of tracks for each particle have $p_T > 200\,\mathrm{MeV}/c^2$. For the pion from the K^{*0} decay, 90% of the tracks have $p_T > 200\,\mathrm{MeV}/c$. For this pion (`GoodTracksVeryLoose`), we split the pion tracks into two subsets, $p_T < 200$MeV/c and $p_T > 200$MeV/c and assign a systematic uncertainty weighted by the distribution across these two categories. Table 8.1 gives the uncertainty assignment for each set of tracks.

Table 8.2: The track quality categories for all tracks used to reconstruct the $B^{\pm} \to D_s^{(*)\pm}\phi$ decay.

Mode	Track	Track Quality
$B^{\pm} \to D_s^{(*)\pm}\phi$	Charged kaons from the ϕ	One K is `GoodTracksVeryLoose` the other `GoodTracksLoose`
$D_s^- \to \phi\pi^-$	π Charged kaons from the ϕ	`GoodTracksVeryLoose` One K is `GoodTracksVeryLoose` the other `GoodTracksLoose`
$D_s^- \to K^- K_S^0$	$K^{\pm}$	`GoodTracksLoose`

$D_s^- \to K^{*0} K^-$	$K^{\pm}$
	$K^{\pm}$ from the K^{*0}
	$\pi^{\pm}$ from the K^{*0}
		`GoodTracksLoose`
		`GoodTracksLoose`
		`GoodTracksVeryLoose`

For K_S^0 mesons, we include a 1.9% total systematic uncertainty for every K_S^0. The K_S^0 has a non-zero flight length and thus the pion tracks from its decay do not come from the beam line (or more precisely the e^+-e^- interaction point). The 1.9% uncertainty comes from the fact that the *BABAR* track reconstruction assumes tracks come from the interaction point and thus the K_S^0 reconstruction efficiency is affected. In table 8.3 we list the total systematic uncertainty due to tracking for each mode. The results are independent of $D_s^{\pm}$ vs $D_s^{*\pm}$. Note, this is reasonable given that the transverse momentum distribution of the $D_s^{\pm}$ daughter particles in the $B^{\pm} \to D_s^{\pm}\phi$ vs. $B^{\pm} \to D_s^{*\pm}\phi$ modes is such that the ratio (# tracks with $p_T < 200$ MeV/c)/(# of tracks with $p_T > 200$ MeV/c) is the same for $B^{\pm} \to D_s^{\pm}\phi$ and $B^{\pm} \to D_s^{*\pm}\phi$. The average tracking uncertainty is 3.7%.

Table 8.3: Corrections to and systematic uncertainties of the charged track reconstruction efficiency for each mode.

Mode	Total Efficiency Correction for MC		Total Uncertainty	
	Runs 1-3	Run 4	Runs 1-3	Run 4
$D_s^- \to \phi\pi^-$	0.969	0.969	4.0%	4.0%
$D_s^- \to K^- K_S^0$	0.967	0.946	2.9%	3.1%
$D_s^- \to K^{*0} K^-$	0.966	0.966	3.8%	3.8%

8.4 Selection criteria efficiency uncertainties

There is also an uncertainty associated with the efficiency of all the other requirements in the analysis; *e.g.*, the uncertainty associated with the efficiency of the mass requirements in the level one selection criteria, the uncertainty associated with the efficiency of likelihood requirements, etc. These efficiencies have been computed from MC, and we know that the MC is not perfect.

We consider separately the systematic uncertainties on the following components of the selection:

- **The efficiency of the ΔE requirement.** Experience with previous *BABAR* analyses suggests that the central value of ΔE could be off by a few MeV in the MC. We artificially shift the reconstructed ΔE in signal MC by 3 MeV, keeping the ΔE requirements unchanged. When we do this we find a change of 0.3% in the acceptance.

- **Effects of mass shifts in the likelihood.** The likelihood described in Section 6.8 is constructed using several invariant mass variables. Again, experience with previous *BABAR* analyses suggests that the the MC could be wrong by of order 1 MeV in reconstructing mass states. We artificially shift the reconstructed $D_s^{\pm}$ and ϕ masses in signal MC by 1 MeV. We then construct the standard likelihood (using unshifted PDFs), and apply the

likelihood cut. We find a change of 0.2% (for the $D_s^{\pm}$) and 0.1% (for the ϕ) in the efficiency of the likelihood requirement. In the case of $B^{\pm} \to D_s^{*\pm}\phi$, we also move Δm by 1 MeV. We find a change of 0.2% in the acceptance.

- **Effects of mass resolution effects in the likelihood.** We also add an extra smear to the reconstructed $D_s^{\pm}$ mass and Δm in signal MC in such a way that the σ of the reconstructed quantity increases by 1 MeV. The efficiency of the likelihood requirement changes by 1.5% for both the $D_s^{\pm}$ and 1.5% for Δm.

- **Effects of Fisher discriminant simulation.** We use a sample of data $B^+ \to D^0 K^+$ with $D^0 \to K^-\pi^+$ events from another analyses. From this sample we obtain a data Fisher distribution, see Figure 8.1. This can be compared with the Fisher distribution from our signal MC, see Figures 6.45 and 8.2. We replace the reconstructed Fisher discriminant in signal MC with a random variable drawn from the distribution of Figure 8.1. We then use this new variable to calculate the likelihood for a signal event. The efficiency changes by about 3%.

The results of these studies of the selection criteria efficiencies are summarized in Table 8.4. Based on these results, we assign a systematic uncertainty of 5% to the understanding of the cut efficiency.

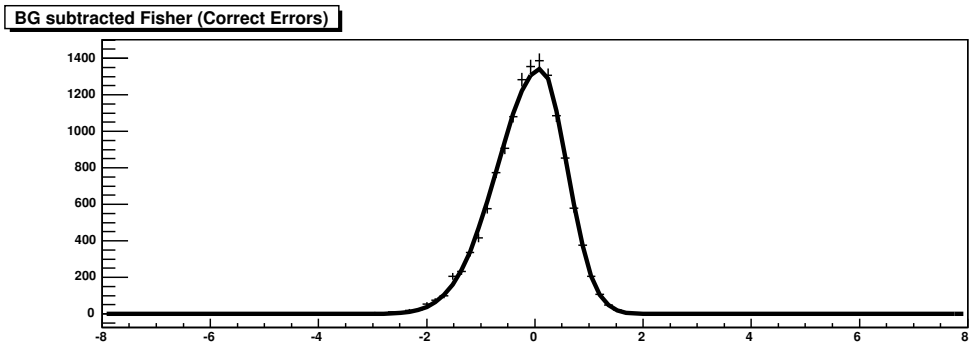

Figure 8.1: Fisher discriminant from a data $B \to D\pi$ sample, with a superimposed bifurcated-Gaussian fit.

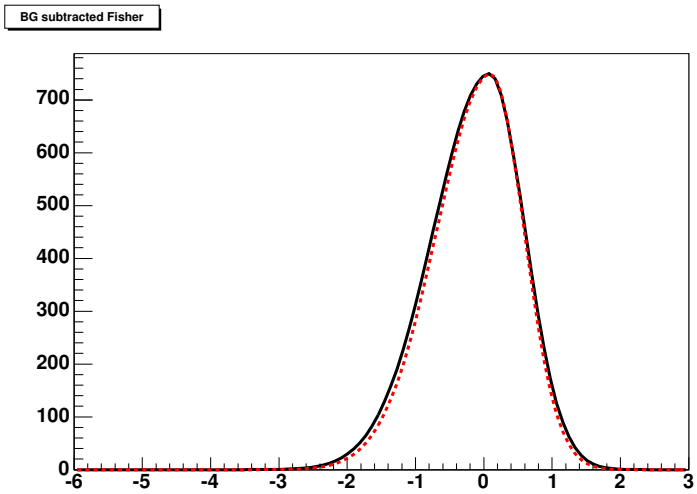

Figure 8.2: Solid black line: bifurcated-Gaussian fit to the $B \to D\pi$ sample of Figure 8.1. Dashed red line: bifurcated-Gaussian fit to the signal MC sample of Figure 6.45.

Table 8.4: Summary of studies of cut systematic uncertainties. Obviously the Δm just applies to the $D_s^{*\pm}$ modes.

Source	What was done	Acceptance shift
ΔE	3 MeV shift	0.3%
$D_s^{\pm}$ mass	1 MeV shift	0.2%
$D_s^{\pm}$ mass	Increase σ by 1 MeV	1.5%
ϕ mass	1 MeV shift	0.1%
ϕ mass	Increase σ by 1 MeV	0.4%
Δm	1 MeV shift	0.2%
Δm	Increase σ by 1 MeV	1.5%
Fisher	Use data $B \to D\pi$,$D \to K\pi$	3 %
Total		$\sim$5 %

8.5 Signal Monte Carlo statistics

We take as systematic the uncertainty associated with the finite statistics of the signal MC sample. We calculate the uncertainty separately for each mode so that we can correctly combine correlated and uncorrelated errors when we combine all modes in our final measurement. The uncertainty is that associated with binomial statistics. Recall in binomial statistics, the variance is simply $\sqrt{Np(1-p)}$, where N is the number of trials and p is the probability of success. We apply these statistics to our MC sample by letting N be the number of events generated in each sample (e.g. $B^{\pm} \to D_s^{\pm}\phi$, $D_s^- \to \phi\pi^-$) and p be the acceptance, ϵ, for that mode:

$$\sigma_{\text{MC stat}} = \sqrt{N_{\text{gen}}\epsilon(1-\epsilon)} \quad . \tag{8.1}$$

The total acceptance and its error are calculated from the weighted average of the MC samples associated with Runs 1-3 and Run 4.

8.6 B-counting uncertainties

In order to calculate a branching fraction we need to know the number of B mesons we started with, otherwise the number we observe is meaningless. The process used to count the number of B mesons we have in our data sample is conceptually simple: we count the number of hadronic events in our sample and

then subtract off the number of hadronic events that are not due to the process $e^+e^- \to \Upsilon(4S) \to B\overline{B}$. We determine the number of events not due to $\Upsilon(4S)$ production by looking at a data sample that was taken with a center-of-mass energy just below the $b\bar{b}$ production threshold (off resonance data). To do this calculation we need to know the luminosities in our data sample and our off resonance data sample; these are calculated by counting the number of $e^+e^- \to \mu^+\mu^-$. The systematic uncertainty in our knowledge of $B\overline{B}$ pairs is 1.1%. In Runs 1-3 we have $N_{B\overline{B}} = (123.8 \pm 1.4) \times 10^6$. For Run 4 we have $N_{B\overline{B}} = (110.3 \pm 1.2) \times 10^6$.

8.7 Branching fraction uncertainties

The uncertainty on the branching fractions of the daughter decays, $\mathcal{B}_i$ contribute the largest uncertainty to the extraction of the limit on the $B^\pm \to D_s^{(*)\pm}\phi$ branching ratio. Using the latest *BABAR* measurement of $\mathcal{B}(D_s^- \to \phi\pi^-)$ [1] and including all correlations between the modes[2], we calculate these uncertainties to be 13%, 20%, and 16% for the $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, and $D_s^- \to K^{*0}K^-$ branching ratios respectively. They are all correlated, since the $D_s^- \to \phi\pi^-$ mode

[2] All branching fractions of the $D_s^\pm$ are calculated in reference to $\mathcal{B}(D_s^- \to \phi\pi^-)$. So a change in the uncertainty in the $D_s^- \to \phi\pi^-$ mode will affect all others.

is used to normalize the branching ratios of the two other modes. (See Table 6.2 for a breakdown.)

We calculate the uncertainty on $\sum_i \epsilon_i \cdot \mathcal{B}_i$ due to the uncertainty on $\mathcal{B}_i$ starting from the values in Table 6.1, and taking account the correlated and uncorrelated uncertainties on the $D_s^{\pm}$ branching ratios. The resulting uncertainty on $\mathcal{B}_i$ is 14%. Uncertainties due to the $K^{*0} \to K^+\pi^-$ and $K_S^0 \to \pi^+\pi^-$ branching ratios are negligible.

8.8 Summary of systematic uncertainties

The systematic uncertainties are summarized in Table 8.5. Note, these are only approximate as all calculations were done using spreadsheet software in order to correctly account for correlations.

Table 8.5: Systematic uncertainties on $\sum_i \epsilon_i \cdot \mathcal{B}_i$. The index i runs over the three $D_s^{\pm}$ modes used in this analysis, ϵ_i are the experimental efficiencies, and $\mathcal{B}_i$ are the branching fractions for the i^{th} mode.

Source	Applies To	$B^{\pm} \to D_s^{\pm}\phi$	$B^{\pm} \to D_s^{*\pm}\phi$
D_s^- branching fraction	$\mathcal{B}_i$	14 %	14 %
D_s^{*-} branching fraction	$\mathcal{B}_i$	–	2.5%
Other branching fractions	$\mathcal{B}_i$	1.5%	1.5%
Charged kaon ID	ϵ_i	13.2%	13.3%
Tracking and K_S^0 efficiency	ϵ_i	3.7%	3.7%
Photon efficiency	ϵ_i	–	1.8%
Final state polarization	ϵ_i	–	1 %
Selection requirements	ϵ_i	5 %	5 %
Simulation statistics	ϵ_i	0.6%	0.6%
B counting	$N_{B\overline{B}}$	1.1%	1.1%
Total		20 %	21 %

Chapter 9

Upper Limit Calculations

We see no evidence of an enhancement in the signal region. Thus, we choose to extract the upper limit on the number of events and therefore the branching fraction. We choose to use a Bayesian approach with a flat prior. That is, we start from the likelihoods $\mathcal{L}(N)$ from Figures 7.5 and 7.10, where N is the number of events. We then define the 90% confidence level upper limit on the number of events, N_{UL}, to be such that

$$\int_0^{N_{UL}} \mathcal{L}(N)\ dN \quad = \quad 0.9 \int_0^{+\infty} \mathcal{L}(N)\ dN \tag{9.1}$$

Then, if we neglect systematic uncertainties, we can simply set the 90% confidence level upper limit on the branching fraction, $\mathcal{B}$, as[1]

$$\mathcal{B} < \frac{N_{UL}}{N_{B\overline{B}} \sum_i \epsilon_i \times \mathcal{B}_i} \tag{9.2}$$

[1] This implicitly assumes that $\mathrm{BR}(\Upsilon(4S) \to B^+B^-) = \mathrm{BR}(\Upsilon(4S) \to B^0\overline{B}^0)$.

where $(N_{BB}) = (233.9 \pm 2.5) \times 10^6$, is the number of $B\bar{B}$ events in Runs 1-4; i is an index that runs through the three $D_s^{\pm}$ decay modes, and ϵ_i is the experimental efficiency in the i^{th} mode[2].

Section 9.1 summarizes and tabulates for reference the values of and uncertainties for the quantity $\sum_i \epsilon_i \cdot \mathcal{B}_i$ which is used in the final calculation of the upper limits on the branching fractions $\mathcal{B}(B^{\pm} \to D_s^{(*)\pm}\phi)$. These calculations are summarized in Section 9.2 for the $B^{\pm} \to D_s^{\pm}\phi$ mode and in Section 9.3 for the $B^{\pm} \to D_s^{*\pm}\phi$ mode.

[2]The efficiencies are the luminosity-weighted average of Runs 1-3 and Run 4.

9.1 Final efficiencies

We tabulate here for succinctness the acceptance efficiencies we use for our limit calculations. For all calculations we use the latest result, $\mathcal{B}(D_s^- \to \phi\pi^-)=$ $(4.8 \pm 0.6)\%$ [1]. As of the time of writing this result has not yet been included in the world average set by the Particle Data Group. Thus, in Appendix F we redo all calculations using the 2004 world average of $\mathcal{B}(D_s^- \to \phi\pi^-)$.

The efficiency of all our selection criteria on signal MC is called the raw efficiency. They are shown in Table 9.1 along with the track reconstruction efficiency correction. The efficiency, ϵ, used in Equation 9.2 is the product of the raw efficiency and the tracking correction.

Table 9.1 shows the raw efficiency, the charged track reconstruction efficiency correction to be applied to the MC to bring it in line with the data, and the final efficiency, ϵ_i. The uncertainties on the raw efficiency are from the limited MC statistics (Section 8.5).

Note that the tracking corrections are identical for Runs 1-3 and Run 4, except for the $D_s^- \to K^-K_S^0$ and $D_s^{*-} \to (K^-K_S^0)\gamma$ modes since the prescription for tracking corrections is the same for all tracks, but differs for the K_S^0 reconstruction.

Table 9.2 summarizes our knowledge of $\epsilon_i \cdot \mathcal{B}_i$. The branching fractions, $\mathcal{B}$, include the branching fractions for all the daughter meson decays, $D_s^{*\pm}$, $D_s^{\pm}$, ϕ,

K_S^0, and K^{*0}. Note that we tabulate these numbers so the reader might get a feel for the level of uncertainty and efficiency that is present in this analysis. The actual calculation of the total systematic uncertainties and the final branching fraction upper limits were performed via spreadsheet software. All correlations between systematic uncertainties were taken into account in the spreadsheet. The final uncertainty for both the $B^{\pm} \to D_s^{\pm}\phi$ and $B^{\pm} \to D_s^{*\pm}\phi$ modes is 20%.

Table 9.3 is similar to Table 9.2, except we average over the $D_s^{(*)\pm}$ decay modes and not over the Runs. The values for Runs 1-3 and Run 4 are shown separately for the $B^{\pm} \to D_s^{\pm}\phi$, the $B^{\pm} \to D_s^{*\pm}\phi$ longitudinal polarization, and the $B^{\pm} \to D_s^{*\pm}\phi$ transverse polarization modes.

In Table 9.4 we list the values of the $\Sigma_i \epsilon_i \cdot \mathcal{B}_i$ for $B^{\pm} \to D_s^{\pm}\phi$ and $B^{\pm} \to D_s^{*\pm}\phi$ averaged by luminosity over all runs and broken down into uncertainties on ϵ and $\mathcal{B}$ respectively. Here we include all systematic uncertainties on ϵ_i and $\mathcal{B}_i$ (recall Table 8.5). The $\Sigma_i \epsilon_i \cdot \mathcal{B}_i$ and its uncertainty are independent of $D_s^{*\pm}$ polarization.

Table 9.1: This table shows the raw MC efficiency, the MC tracking efficiency correction factors, and the final MC efficiency, ϵ_i = (Raw Efficiency)*(Tracking Correction). The uncertainties on the raw efficiency are from the limited MC statistics (Section 8.5).

Runs 1-3			
$D_s^{(*)\pm}$ Decay Mode	Raw Efficiency (%)	Tracking Correction	Final Efficiency, ϵ_i (%)
$D_s^- \to \phi\pi^-$	21.0 ± 0.3	0.969 ± 0.039	20.3 ± 0.9
$D_s^- \to K^- K_S^0$	18.9 ± 0.3	0.967 ± 0.028	18.3 ± 0.6
$D_s^- \to K^{*0} K^-$	15.2 ± 0.2	0.966 ± 0.038	14.7 ± 0.6
Long. $D_s^{*-} \to (\phi\pi^-)\gamma$	11.9 ± 0.2	0.969 ± 0.039	11.5 ± 0.5
Long. $D_s^{*-} \to (K^- K_S^0)\gamma$	10.9 ± 0.3	0.967 ± 0.028	10.5 ± 0.4
Long. $D_s^{*-} \to (K^{*0} K^-)\gamma$	8.8 ± 0.2	0.966 ± 0.037	8.5 ± 0.4
Trans. $D_s^{*-} \to (\phi\pi^-)\gamma$	11.9 ± 0.2	0.969 ± 0.039	11.6 ± 0.5
Trans. $D_s^{*-} \to (K^- K_S^0)\gamma$	11.1 ± 0.3	0.967 ± 0.028	10.8 ± 0.4
Trans. $D_s^{*-} \to (K^{*0} K^-)\gamma$	9.2 ± 0.2	0.966 ± 0.037	8.9 ± 0.4
Run 4			
$D_s^{(*)\pm}$ Decay Mode	Raw Efficiency (%)	Tracking Correction	Final Efficiency, ϵ_i (%)
$D_s^- \to \phi\pi^-$	18.5 ± 0.3	0.969 ± 0.039	18.0 ± 0.8
$D_s^- \to K^- K_S^0$	18.0 ± 0.3	0.946 ± 0.029	17.0 ± 0.6
$D_s^- \to K^{*0} K^-$	13.8 ± 0.2	0.966 ± 0.038	13.3 ± 0.6
Long. $D_s^{*-} \to (\phi\pi^-)\gamma$	10.5 ± 0.2	0.969 ± 0.039	10.1 ± 0.5
Long. $D_s^{*-} \to (K^- K_S^0)\gamma$	9.9 ± 0.3	0.946 ± 0.029	9.4 ± 0.4
Long. $D_s^{*-} \to (K^{*0} K^-)\gamma$	8.3 ± 0.2	0.966 ± 0.037	8.0 ± 0.3
Trans. $D_s^{*-} \to (\phi\pi^-)\gamma$	10.5 ± 0.2	0.969 ± 0.039	10.2 ± 0.5
Trans. $D_s^{*-} \to (K^- K_S^0)\gamma$	9.7 ± 0.3	0.946 ± 0.029	9.2 ± 0.4
Trans. $D_s^{*-} \to (K^{*0} K^-)\gamma$	8.0 ± 0.2	0.966 ± 0.037	7.7 ± 0.3

Table 9.2: Signal MC efficiency (ϵ_i), averaged between Runs 1-3 and Run 4 according to luminosity (see Table 9.1), total branching fractions ($\mathcal{B}_i$) of the reconstructed daughter decay modes, and product of acceptance and branching fraction for each of the three modes used in the $B^{\pm} \to D_s^{*\pm}\phi$ search. When two uncertainties are listed, the first is the uncertainty associated with the ϵ_i (only the uncertainties due to MC statistics and tracking efficiency corrections, i.e. those in Table 9.1, are included) and the second is that associated with $\mathcal{B}_i$.

$D_s^{(*)\pm}$ Decay Mode	Polarization	ϵ_i (%)	$\mathcal{B}_i$ (10^{-3})	$\epsilon_i \cdot \mathcal{B}_i$ (10^{-3})
$D_s^- \to \phi\pi^-$	—	19.2 ± 0.6	11.57 ± 1.46	$2.22 \pm 0.07 \pm 0.28$
$D_s^- \to K^-K_S^0$	—	17.7 ± 0.4	8.21 ± 1.66	$1.45 \pm 0.04 \pm 0.29$
$D_s^- \to K^{*0}K^-$	—	14.0 ± 0.4	14.46 ± 2.30	$2.03 \pm 0.06 \pm 0.32$
$D_s^{*-} \to (\phi\pi^-)\gamma$	Longitudinal	10.9 ± 0.4	10.90 ± 1.41	$1.19 \pm 0.04 \pm 0.15$
$D_s^{*-} \to (K^-K_S^0)\gamma$	Longitudinal	10.0 ± 0.3	7.73 ± 1.58	$0.77 \pm 0.02 \pm 0.16$
$D_s^{*-} \to (K^{*0}K^-)\gamma$	Longitudinal	8.3 ± 0.3	13.62 ± 2.20	$1.13 \pm 0.03 \pm 0.18$
$D_s^{*-} \to (\phi\pi^-)\gamma$	Transverse	10.9 ± 0.4	10.90 ± 1.41	$1.19 \pm 0.04 \pm 0.15$
$D_s^{*-} \to (K^-K_S^0)\gamma$	Transverse	10.0 ± 0.3	7.73 ± 1.58	$0.77 \pm 0.02 \pm 0.16$
$D_s^{*-} \to (K^{*0}K^-)\gamma$	Transverse	8.3 ± 0.3	13.62 ± 2.20	$1.14 \pm 0.03 \pm 0.18$

Table 9.3: The sum over modes of the acceptance times branching fraction. The index i runs over the $D_s^{\pm}$ decay modes. When two uncertainties are listed, the first is that associated with ϵ_i (only the uncertainties due to MC statistics and tracking efficiency corrections, i.e. those in Table 9.1, are included) and the second is that associated with $\mathcal{B}_i$. The row marked "All" is the weighted-by-luminosity average of the Runs 1-3 and Run 4 results. See Table 9.4 for the final result including all systematic uncertainties. Note that these results are independent of $D_s^{*\pm}$ helicity.

Runs	$B^{\pm} \to D_s^{\pm}\phi$ $\sum_i \epsilon_i \cdot \mathcal{B}_i$ (10^{-3})	$B^{\pm} \to D_s^{*\pm}\phi$ Long. $\sum_i \epsilon_i \cdot \mathcal{B}_i$ (10^{-3})	$B^{\pm} \to D_s^{*\pm}\phi$ Trans. $\sum_i \epsilon_i \cdot \mathcal{B}_i$ (10^{-3})
1-3	$5.97 \pm 0.22 \pm 0.81$	$3.23 \pm 0.08 \pm 0.44$	$3.30 \pm 0.08 \pm 0.45$
4	$5.40 \pm 0.21 \pm 0.74$	$2.92 \pm 0.08 \pm 0.39$	$2.87 \pm 0.07 \pm 0.38$
All	$5.70 \pm 0.21 \pm 0.78$	$3.09 \pm 0.08 \pm 0.42$	$3.10 \pm 0.09 \pm 0.42$

Table 9.4: The sum over modes of the acceptance times branching fraction with all systematic uncertainties involving ϵ and $\mathcal{B}_i$ listed in Table 8.5 included except that for the number of $B\overline{B}$ pairs. The first uncertainty is that related to ϵ_i. The second uncertainty is from the uncertainty in $\mathcal{B}_i$.

B decay mode	$\sum_i \epsilon_i \cdot \mathcal{B}_i$ (10^{-3})
$B^{\pm} \to D_s^{\pm}\phi$	$5.70 \pm 0.57 \pm 0.78$
$B^{\pm} \to D_s^{*\pm}\phi$	$3.09 \pm 0.31 \pm 0.42$

9.2 Upper limit on $\mathcal{B}(B^{\pm} \to D_s^{\pm}\phi)$

The likelihood curve for the number of signal events was obtained by scanning over the number of signal events when determining the yield in Section 7.1, it is shown in Figure 9.1. To obtain the Bayesian 90% confidence level upper limit on the number of signal events we first multiply the likelihood curve by our Bayesian prior. Since the number of signal events must a non-negative number, our prior function is simply the step function, $\Theta(0)$. $\Theta(0)$ evaluates to zero for all negative values of N and to unity otherwise. After applying the prior to the likelihood function (see Figure 9.2), we then use Equation 9.1 to calculate, $N_{UL} = 2.3$. If we were to assume no experimental uncertainties, we would calculate directly from Equation 9.2, $\mathcal{B}(B^{\pm} \to D_s^{\pm}\phi) < 1.7 \times 10^{-6}$ at 90% C.L.

In order to account for the systematic uncertainties described in Chapter 8, we use a Monte Carlo technique to obtain a PDF for $\mathcal{B}$ starting from $\mathcal{L}(N)$ and a PDF for $N_{B\overline{B}} \sum_i \epsilon_i \cdot \mathcal{B}_i$. (See Appendix E for justification of this method.) For lack of a better term we will refer to this Monte Carlo technique as smearing. We will then substitute this PDF representing $\mathcal{B}$ for the likelihood, $\mathcal{L}(N)$, in Equation 9.1 to obtain our 90% C.L. upper limit on the branching fractions, $\mathcal{B}(B^{\pm} \to D_s^{(*)\pm}\phi)$. The dotted lines in Figures 9.1-9.4 are the PDF representing $\mathcal{B}$[3]. The PDF for

[3]Note that the abscissa for the dotted line is no longer N and the PDF is normalized to the likelihood PDF but the point here is simply to show the shape that resulted in this smearing

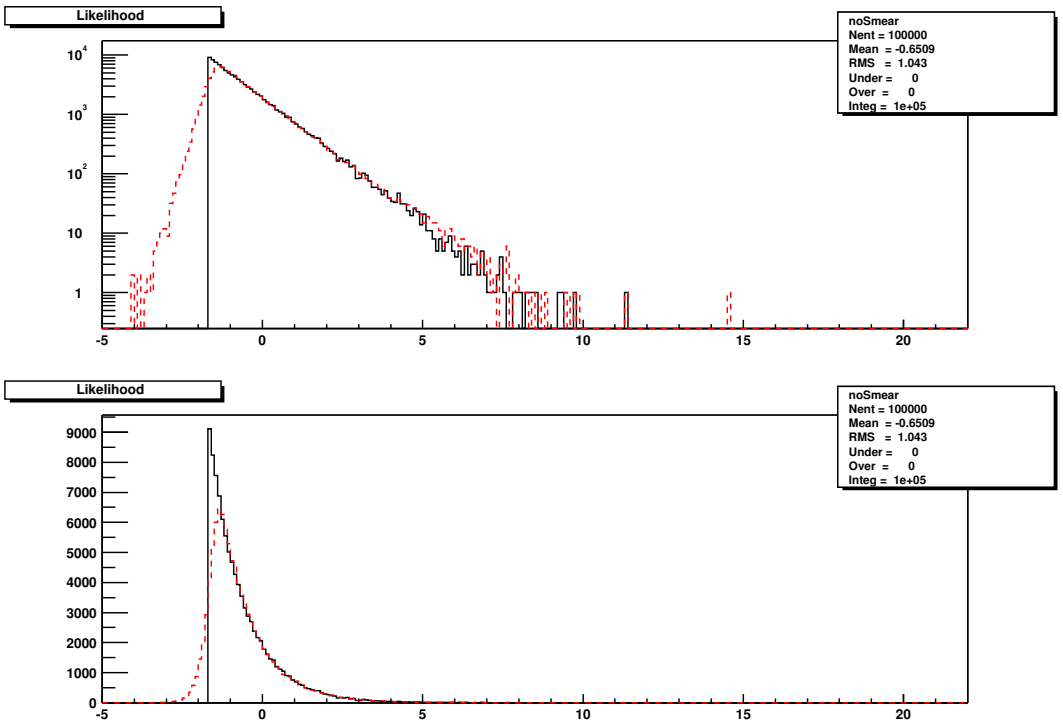

Figure 9.1: $B^{\pm} \to D_s^{\pm}\phi$. The likelihood curve for the number of signal events (black solid) and smeared likelihood (red dashed), in arbitrary units as a function of the number of signal events (N_{sig}). Top: log-scale ordinate. Bottom: linear-scale ordinate.

the denominator in Equation 9.2 is taken to be a Gaussian of mean $\mu = 1.33 \times 10^6$ and $\sigma = 0.20 \times \mu$ [4] The result is

$$\mathcal{B}(B^{\pm} \to D_s^{\pm}\phi) < 1.8 \times 10^{-6} \quad \text{at } 90\% \text{ C.L.} \tag{9.3}$$

[4] The value for μ and σ are value of $N_{B\overline{B}} \sum_i \epsilon_i \cdot \mathcal{B}_i$ and its uncertainty.

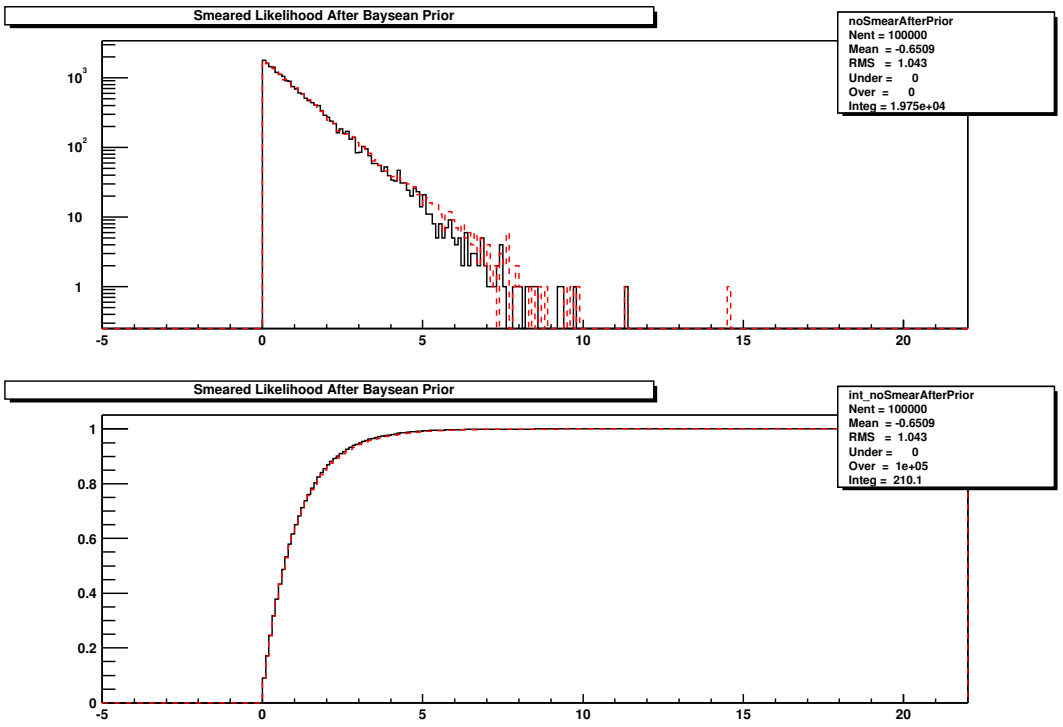

Figure 9.2: $B^{\pm} \to D_s^{\pm}\phi$. Here we show the likelihood of Figure 9.1 after we apply the Bayesian prior. The top plot is the likelihood (black solid) and smeared likelihood (red dashed), as a function of N_{sig} after applying the flat Bayesian prior. The bottom plot shows the integration (Equation 9.1) of the binned likelihood function.

The measurement of $\mathcal{B}(D_s^- \to \phi\pi^-)$ will certainly improve in the coming years, thus we also calculate the limit on $\mathcal{B}(B^{\pm} \to D_s^{\pm}\phi) \times \mathcal{B}(D_s^{\pm} \to \phi\pi^{\pm})$. The result is

$$\mathcal{B}(B^{\pm} \to D_s^{\pm}\phi) \times \mathcal{B}(D_s^- \to \phi\pi^-) < 8.6 \times 10^{-8} \quad \text{at } 90\% \text{ C.L.} \tag{9.4}$$

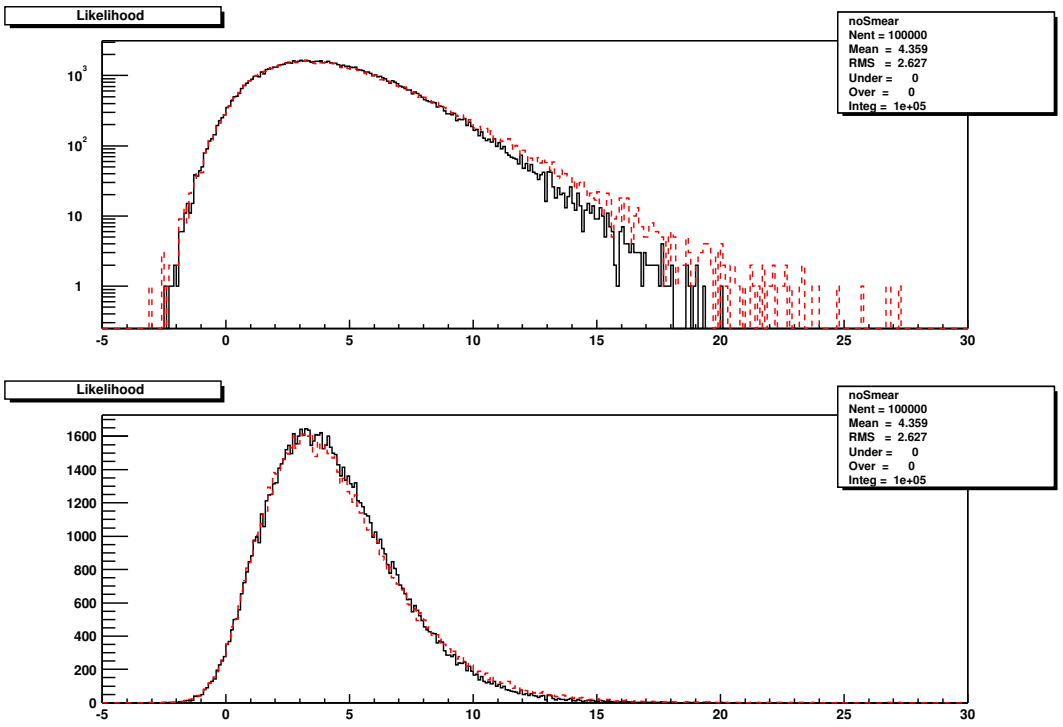

Figure 9.3: $B^{\pm} \to D_s^{*\pm}\phi$. The likelihood curve for the number of signal events (black solid) and smeared likelihood (red dashed), in arbitrary units as a function of the number of signal events (N_{sig}). Top: log-scale ordinate. Bottom: linear-scale ordinate.

9.3 Upper limit on $\mathcal{B}(B^{\pm} \to D_s^{*\pm}\phi)$

The likelihood curve for the $B^{\pm} \to D_s^{*\pm}\phi$ mode is shown in Figure 9.3. As for the $B^{\pm} \to D_s^{\pm}\phi$ mode, we multiply by our Bayesian prior which yields Figure 9.4. Using Equation 9.1, we obtain N_{UL} =7.8. As in the previous section, if we assume no experimental uncertainties, we can calculate $\mathcal{B}(B^{\pm} \to D_s^{*\pm}\phi) < 1.1 \times 10^{-5}$ at 90% C.L.

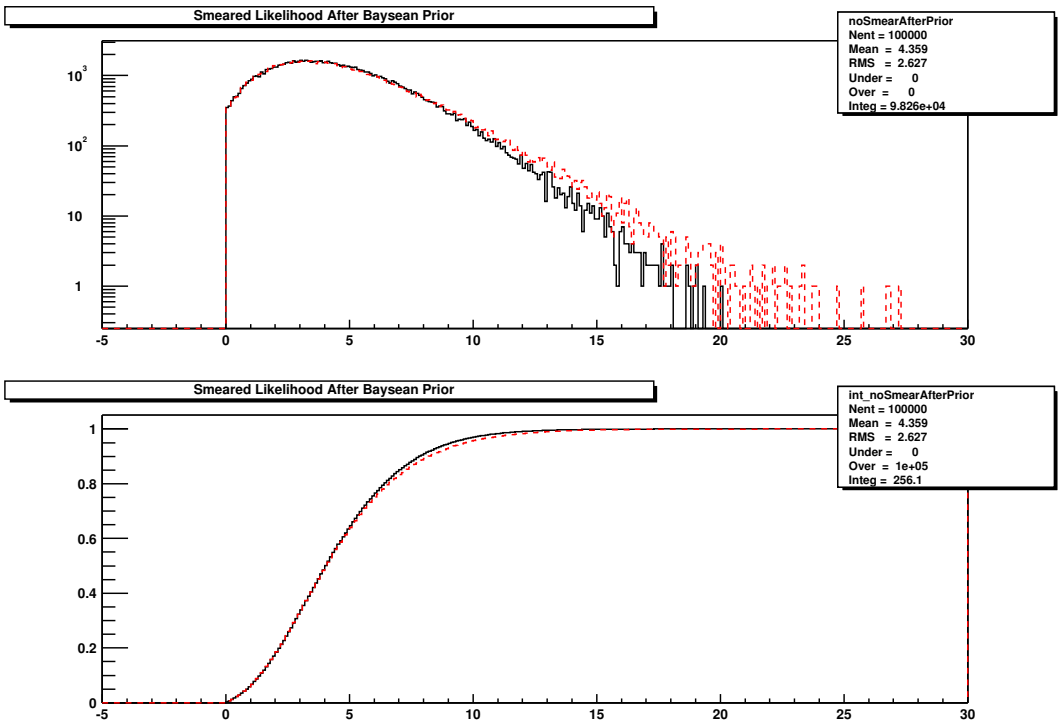

Figure 9.4: $B^{\pm} \to D_s^{*\pm}\phi$. Here we show the likelihood of Figure 9.3 after we apply the Bayesian prior. The top plot is the likelihood (black solid) and smeared likelihood (red dashed), as a function of N_{sig} after applying the flat Bayesian prior. The bottom plot shows the integration (Equation 9.1) of the binned likelihood function.

We account for experimental uncertainties as we did for the $B^{\pm} \to D_s^{\pm}\phi$ modes. The Gaussian PDF for the denominator in Equation 9.2 has $\mu = 7.15 \times 10^6$ and $\sigma = 0.20 \times \mu$. The result is

$$\mathcal{B}(B^{\pm} \to D_s^{*\pm}\phi) < 1.1 \times 10^{-5} \quad \text{at } 90\% \text{ C.L.} \tag{9.5}$$

And,

$$\mathcal{B}(B^{\pm} \to D_s^{*\pm}\phi) \times \mathcal{B}(D_s^{\pm} \to \phi\pi^{\pm}) < 5.3 \times 10^{-7} \quad \text{at } 90\% \text{ C.L.} \tag{9.6}$$

Chapter 10

Conclusions

The search for $B^{\pm} \to D_s^{(*)\pm}\phi$ found no evidence for these decays. Our limits are about two orders of magnitude lower than the previous results, but are still one order of magnitude higher than the Standard Model expectation. However, the limits obtained in Chapter 9 can be used to examine the validity of the theories detailed in Chapter 2.

Using the calculation in [4], and conservatively assuming fully destructive interference between the SM and RPV-SUSY amplitudes, we find that our limit on $B^{\pm} \to D_s^{(*)\pm}\phi$ implies (recall Equation 2.1)

$$|\frac{\lambda^2}{M^2}| < \frac{4 \times 10^{-4}}{(100\,\mathrm{GeV})^2}, \tag{10.1}$$

where M is the slepton mass and λ is quark-quark-slepton coupling constant. Note that this limit does not include any of the hadronic uncertainties associated with the factorization hypothesis.

Using the calculation in [4] and our limit on $\mathcal{B}(B^{\pm} \to D_s^{(*)\pm}\phi)$ we can extract a limit $\tan\beta/M_{H^{\pm}} < 0.37/\,\mathrm{GeV}$. Just as in the RPV-SUSY model, the limit does not include any hadronic uncertainties. Recall, the Belle preliminary result on $\mathcal{B}(B^{+} \to \tau^{+}\nu_{\tau})$ gives $\tan\beta/M_{H^{\pm}} < 0.31/\,\mathrm{GeV}$ while information on $B \to D\tau\nu$ yields $\tan\beta/M_{H^{\pm}} < 0.46/\,\mathrm{GeV}$ [12]. So we have produced a competitive result.

The CDF collaboration has recently published new limits in the $\tan\beta - M_{H^{\pm}}$ plane from a study of the consistency of their $t\bar{t}$ sample in the various decay modes with the SM hypothesis [21]. Their results, which are shown in Figure 10.1, are obtained in the contexts of the MSSM and include higher order corrections for a given choice of MSSM parameters.

We have also obtained from CDF their exclusion regions in the $\tan\beta - M_{H^{\pm}}$ at tree level [22]. This is shown in Figure 10.2, including the tree-level limits from B decays. It seems that at high $\tan\beta$, the indirect limits from B decays are more powerful than those from the TeVatron

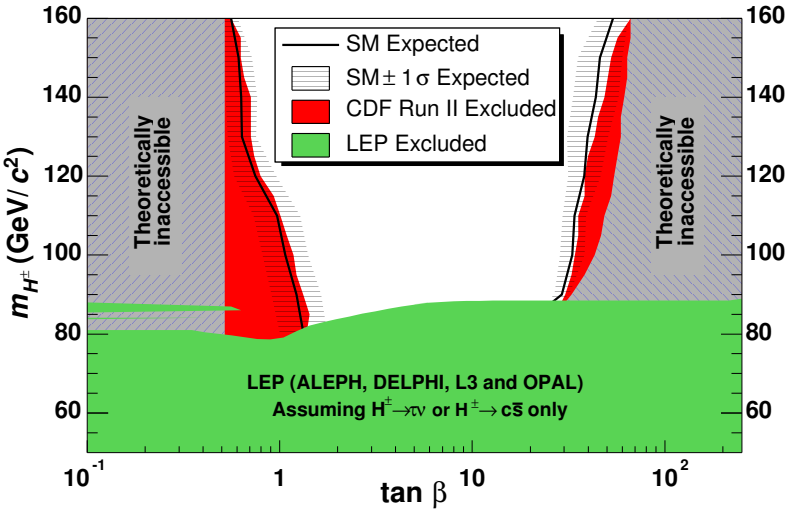

Figure 10.1: Excluded regions (95% CL) in the $\tan\beta - M_{H^\pm}$ plane from an analysis of CDF's $t\bar{t}$ events. These are in the context of the MSSM for a certain choice of parameters including radiative corrections. Note that the regions of very high and very low $\tan\beta$ are not probed, since the tbH couplings become very large. See Reference [21] for details.

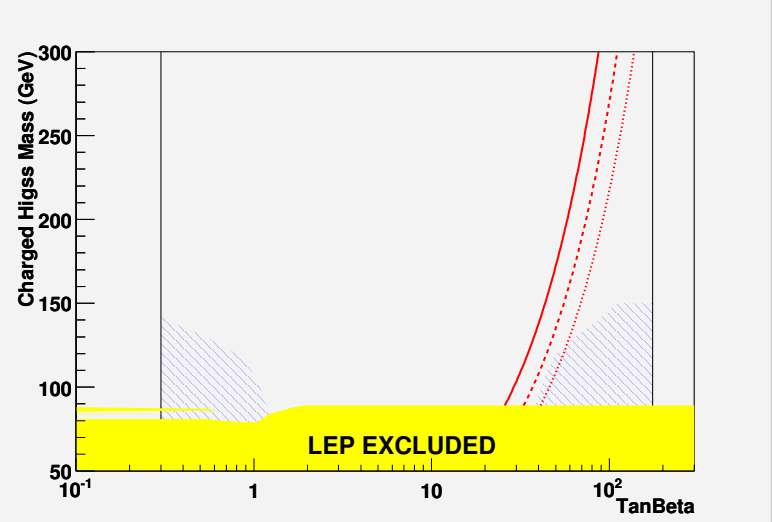

Figure 10.2: Excluded regions (95% CL) in the $\tan\beta - M_{H^\pm}$ plane from CDF's $t\bar{t}$ events. Shaded region: CDF's $t\bar{t}$ events, tree-level only. The solid (dashed, dot-dashed) lines delimit the 90% CL limits from $B \to \tau\nu$ ($B^\pm \to D_s^{(*)\pm}\phi$, $B \to D\tau\nu$) described in the text. The area to the right of the curves is excluded by B decays. Note that the regions of very high and very low $\tan\beta$ correspond to very large tbH couplings.

Bibliography

[1] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. **D71**, 091104(R) (2005).

[2] J. P. Alexander *et al.*, Phys. Lett. **B319**, 365 (1993).

[3] C. D. Lu, Eur. Phys. J. **C24**, 121 (2002).

[4] R. Mohanta, Phys. Lett. **B540**, 241 (2002).

[5] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. **D72**, 011102(R) (2005).

[6] Belle Collaboration, hep-ex/0605018.

[7] Belle Collaboration, "Rare B decays with missing energy at Belle", http://ichep06.jinr.ru/reports/287_8s5_15p18_Browder(1).ppt.

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

[9] PDG, "Searches for Higgs Bosons", W.-M. Yao et al., J. Phys. **G33**, 1 (2006)

[10] A. Heister *et al.*, Phys. Lett. **B543**, 1 (2002).

[11] P. Gambino and M. Misiak, Nucl. Phys. **B611**, 338 (2001).

[12] ALEPH Collaboration, ICHEP96 proceedings, 1997.

[13] Ackeroyd and Recksiegel, J. Phys. **G29** 2311 (2003). (also hep-ph/0306037).

[14] *BABAR* Collaboration, B. Aubert *et al.*, Nucl. Instr. and Methods **A479**, 1 (2002).

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

[16] PDG, "Review of Particle Physics: Mesons", S. Eidelman et al., Phys. Lett. **B592**, 1 (2004).

[17] Fisher, R. A., "The use of multiple measurements in taxonomic problems", Annals of Eugenics, 7, Part II, 179188 (1936).

[18] G. J. Feldman and R. D. Cousins, Phys. Rev. **D57** (1998) 3873-3889.

[19] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. **B192**, 245 (1987).

[20] F. Porter, arxiv: physics/0311092; also, see `http://www.cithep.caltech.edu/~fcp/statistics/maxlikfitnote.ps`.

[21] CDF collaboration, A. Abulencia, *et al.*, Phys. Rev. Lett. 96 (2006) 042003

[22] Ricardo Eusebi, personal communications. See also page 9 of `http://b0urpc.fnal.gov/~Eeusebi/TeV4LHC_nonotes.pdf`.

Appendices

Appendix A

Function Definitions

Here we define the functions used in this analysis..

Gaussian:

$$G(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\exp(\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}) \tag{A.1}$$

Double Gaussian (Sum of two Gaussians):

$$\begin{aligned} DG(x;f,\mu_1,\mu_2,\sigma_1,\sigma_2) = &\frac{1}{\sigma_1\sqrt{2\pi}}(f)\exp(\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}) \\ &+ \frac{1}{\sigma_2\sqrt{2\pi}}(1-f)\exp(\frac{1}{2}\frac{(x-\mu_2)^2}{\sigma_2^2}) \end{aligned} \tag{A.2}$$

Bifurcated Gaussian:

$$BG(x;\mu,\sigma_1,\sigma_2) = \begin{cases} \frac{1}{\sigma_1\sqrt{2\pi}}\exp(\frac{1}{2}\frac{(x-\mu)^2}{\sigma_1^2}), & x < \mu \\ \frac{1}{\sigma_2\sqrt{2\pi}}\exp(\frac{1}{2}\frac{(x-\mu)^2}{\sigma_2^2}), & x \geq \mu \end{cases} \tag{A.3}$$

BREIT-WIGNER:

$$BW(x; N, x_0, \Gamma) = \frac{1}{N} \frac{1}{[(x - x_0)^2 + (\Gamma/2)^2]} \tag{A.4}$$

VOIGTIAN: A Voigtian is a convolution between a Gaussian (centered at zero with unit normalization) and a Breit-Wigner,

$$V(x; N, \sigma, x_0, \Gamma) = \int_{-\infty}^{+\infty} BW(y) \cdot G(y - x; 0, \sigma) dy. \tag{A.5}$$

To understand the what a convolution is doing let the integral be represented by a sum, then we have,

$$V(x; N, \sigma, x_0, \Gamma) = \sum_{\Delta x} BW(x + \Delta x) \cdot G(\Delta x; 0, \sigma). \tag{A.6}$$

The effect of convolving the Breit-Wigner with a Gaussian is to smear out the Breit-Wigner distribution. E.g. let us measure something that has a perfectly precise value, x, with an experimental setup with imperfect resolution. Because our resolution is not perfect, we may actually measure $x' = x \pm \Delta x$. If we make this measurement an infinite number of times we will get a distribution for the measurement which is centered at x but with some finite width.

If we represent BW by

$$BW(x) = \begin{cases} 1, & x = -1 \\ 4, & x = 0 \\ 1, & x = +1 \end{cases}$$

and a Gaussian by

$$G(x) = \begin{cases} 2, & x = -1 \\ 3, & x = 0 \\ 2, & x = +1 \end{cases},$$

then by A.6 (note the Gaussian should be normalized to unity, hence the factor of 1/7 in front),

$$\begin{aligned} V(-2) &= (1/7)\,[BW(-2-1)G(-1) + BW(-2+0)G(0) + BW(-2+1)G(1)] \\ &= (1/7)\,[0*2 + 0*3 + 1*2] = 2/7 \\ V(-1) &= (1/7)\,[BW(-1-1)G(-1) + BW(-1+0)G(0) + BW(-1+1)G(1)] \\ &= (1/7)\,[0*2 + 1*3 + 4*2] = 11/7 \\ V(0) &= (1/7)\,[BW(0-1)G(-1) + BW(0+0)G(0) + BW(0+1)G(1)] \\ &= (1/7)\,[1*2 + 4*3 + 1*2] = 16/7 \\ V(+1) &= (1/7)\,[BW(1-1)G(-1) + BW(1+0)G(0) + BW(1+1)G(1)] \\ &= (1/7)\,[4*2 + 1*3 + 0*2] = 11/7 \\ V(+2) &= (1/7)\,[BW(2-1)G(-1) + BW(2+0)G(0) + BW(2+1)G(1)] \\ &= (1/7)\,[1*2 + 0*3 + 0*2] = 2/7 \end{aligned}$$

Thus we have the new function, V(x)

$$V(x) = \begin{cases} 2/7, & x = -2 \\ 11/7, & x = -1 \\ 16/7, & x = 0 \\ 11/7, & x = +1 \\ 2/7, & x = +2 \end{cases},$$

which has the same normalization as the original BW, but with a wider spread.

LEGENDRE FISHER: The Legendre Fisher is a commonly used tool at *BABAR*. A Fisher Discriminant is used maximize the discrimination power between two or more variables that are correlated to any degree. The two variables we are trying to distinguish are the momentum-weighted Legendre moments, L_0 and L_2. Hence the name the Legendre Fisher. It is given by,

$$\mathcal{F} = C_0 L_0 + C_2 L_2, \tag{A.7}$$

where

$$L_0 = \sum_i^{\text{r.o.e.}} p_i \tag{A.8}$$

$$L_2 = \sum_i^{\text{r.o.e.}} p_i \cos^2 \theta_{Ti}. \tag{A.9}$$

The term "r.o.e." stands for the rest of the event. The sum is over all the tracks and neutral clusters in each event that are not associated with the B that we are reconstructing. θ_{Ti} is the angle that the track or neutral cluster, i, makes with the thrust vector.

The thrust vector is defined as the vector that satisfies $max(|\Sigma_j \hat{T} \cdot \vec{p}_j|)/(\Sigma_j |\vec{p}_j|)$ for any group of tracks and clusters, j. The thrust vector is defined over all the tracks and clusters in the event. See Appendix B for a diagram.

The constants C_0 and C_2 are determined by the following:

$$S_{22}C_2 + S_{20}C_0 = D_2 \tag{A.10}$$

$$S_{20}C_2 + S_{00}C_0 = D_0, \tag{A.11}$$

where

$$D_2 = < L_2^m > - < L_2^n > \tag{A.12}$$

$$D_0 = < L_0^m > - < L_0^n > \tag{A.13}$$

for two different sets of data, m and n. In our case m and n are the $c\overline{c}$ and signal MC and the L are determined from the MC distributions. Finally, Equations A.10 and A.11 can we solved for D_2 and D_0:

$$D_0 = \frac{S_{22}D_2 - S_{22}D_0}{S_{20}^2 - S_{00}S_{22}} \tag{A.14}$$

$$D_2 = \frac{S_{20}D_0 - S_{00}D_2}{S_{20}^2 - S_{00}S_{22}}. \tag{A.15}$$

ARGUS FUNCTION: The Argus function was developed by the ARGUS collaboration [19]. It is threshold function with two parameters: the endpoint, m_0, and the slope, ζ. For our use, ζ is the fit parameter and m_0 is fixed at one-half the average CM energy.

$$f(m_{ES}) \propto m_{ES}\sqrt{1-x^2}\exp[-\zeta(1-x^2)] \tag{A.16}$$

$$x = m_{ES}/m_0. \tag{A.17}$$

Appendix B

Decay Angles Definitions

In this appendix we summarize the various angles used in this analysis.

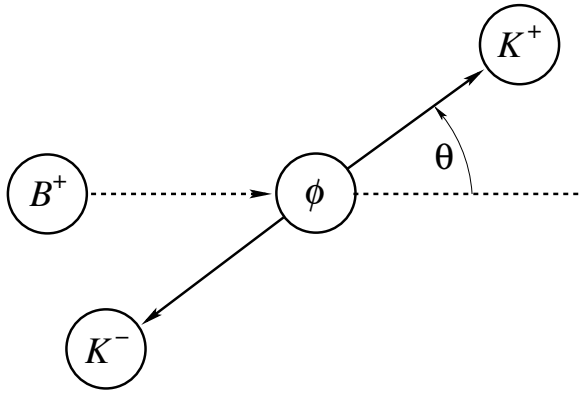

Figure B.1: Decay (helicity) angle for the $B^{\pm} \to D_s^{*\pm}\phi$ decays. This figure shows the decay angle for the bachelor ϕ however, the decay angle for the K^{*0} from the $D_s^{\pm}$ decay is obtained by replacing the B with the $D_s^{\pm}$, the ϕ with a K^{*0} and the K^- with a π^-.

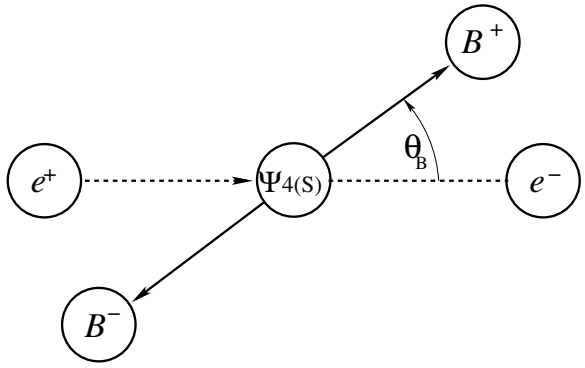

Figure B.2: The production angle of the B meson in the $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ process. Since the B meson is a scalar particle, we expect a $1 - \cos^2(\theta_B)$ distribution.

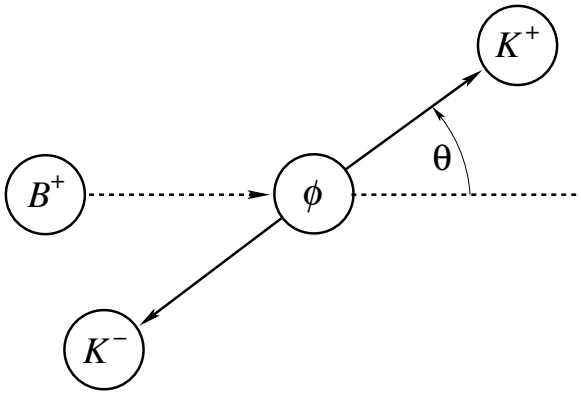

Figure B.3: Decay (helicity) angle for the $B^{\pm} \to D_s^{*\pm}\phi$ decays. This figure shows the decay angle for the bachelor ϕ however, the decay angle for the K^{*0} from the $D_s^{\pm}$ decay is obtained by replacing the B with the $D_s^{\pm}$, the ϕ with a K^{*0} and the K^- with a π^-.

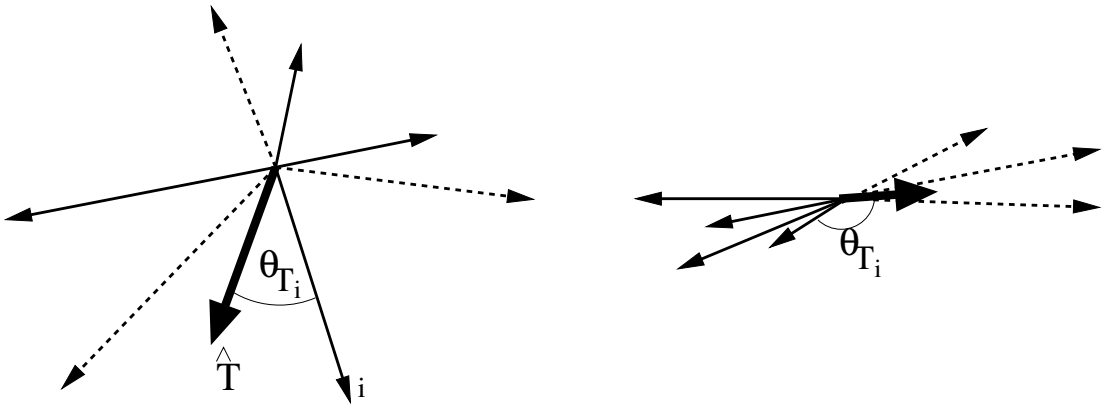

Figure B.4: The thrust angle, $\theta_{\rm T}$ is represented by the thick solid arrow. The dotted arrows represent the tracks and neutral EMC clusters in the center of mass frame that are used to make up the B we reconstruct as $B^{\pm} \to D_s^{(*)\pm}\phi$, the solid arrows are the other tracks and clusters in the event. The left figure is an example of a $B\overline{B}$ event. Notice that the event has an overall circular shape. The figure on the right is an example of a $e^+e^- \to q\overline{q}$, continuum, event. Notice that the continuum event has much less circular shape.

Appendix C

Particle Identification Categories

The kaon selection at *BABAR* is based on the information from the silicon vertex tracker, the drift chamber and the Cherenkov radiation detector. From measurements of a particle's energy loss per unit length dE/dx in the SVT and the DCH we can reliably distinguish between pions and kaons. From Cherenkov angle we can distinguish kaons from pions up to a momentum of $\approx 4\,\mathrm{GeV}/c$. When determining the particle type we start with the assumption that we have a pion and ask the question, "Could this particle be a kaon?"

Each particle is placed in a category depending on our level of confidence as to its species. The categories are defined by how well they discern between koans and other particles in our detector, mainly pions and protons. The categories are, `NotPion, Loose, Tight, and VeryTight`. The `VeryTight` list is optimized for the cleanest sample of kaons, with the rate of mis-identifying a pion as a kaon kept

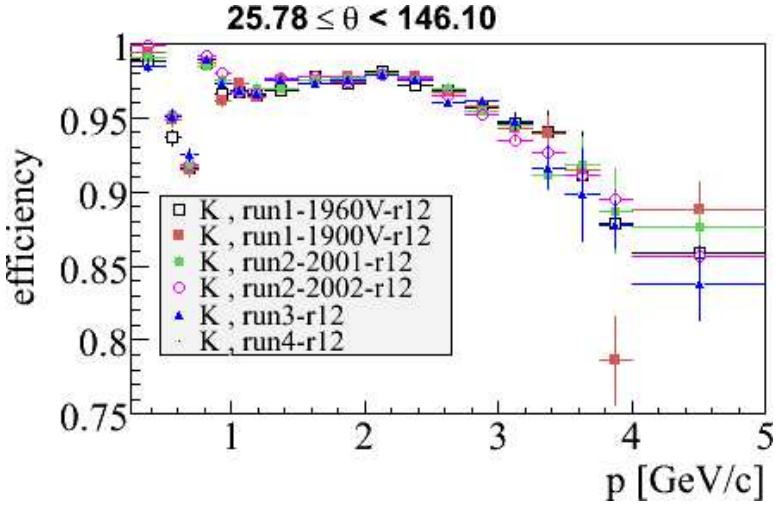

Figure C.1: The kaon identification efficiency for the `NotPion` particle identification category.

below 2%, while the `NotPion` list is optimized to have the highest kaon efficiency possible while still rejecting most pions.

Figures C.1-C.4 show the kaon identification efficiency and the rate for misidentifying a pion as a koan for the `NotPion` and `Loose` categories.

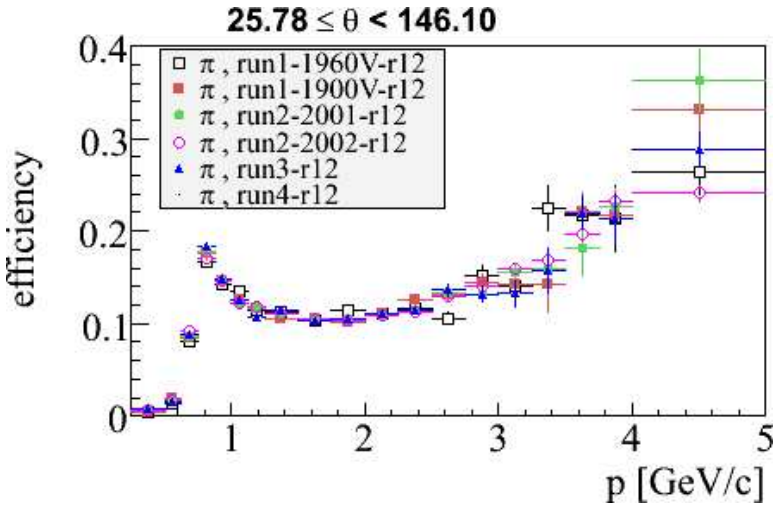

Figure C.2: The rate for mis-identifying a pion as a kaon for the `NotPion` particle identification category.

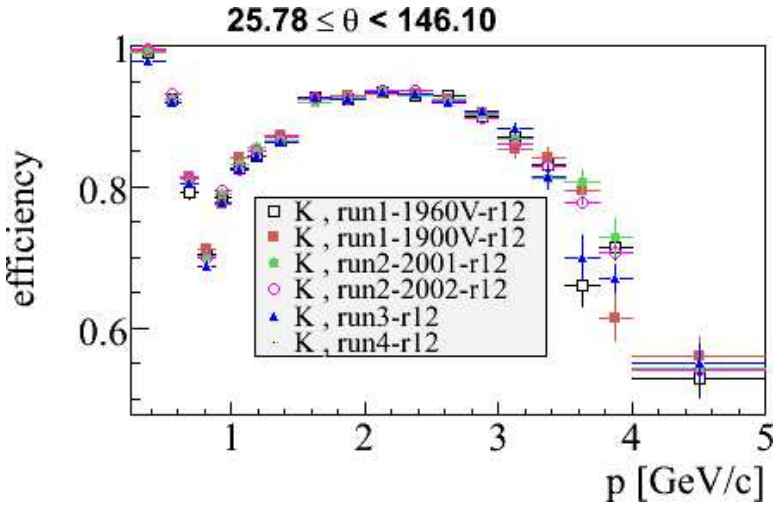

Figure C.3: The kaon identification efficiency for the `Loose` particle identification category.

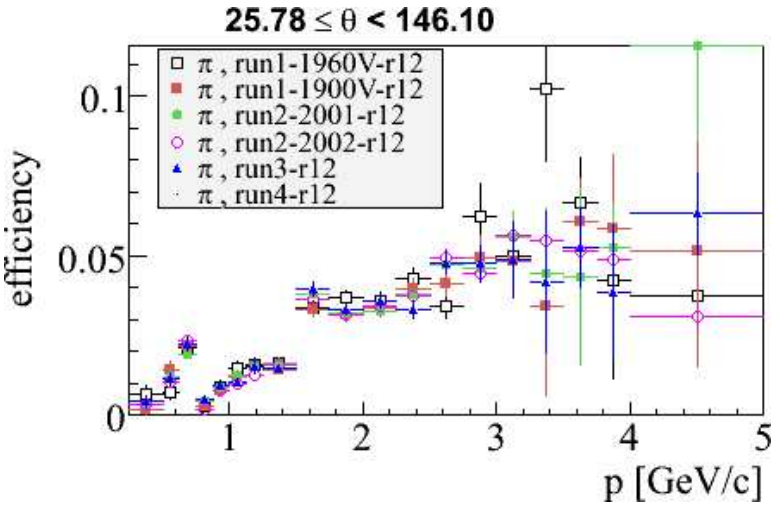

Figure C.4: The rate for mis-identifying a pion as a kaon for the `Loose` particle identification category.

Appendix D

Track Quality Classifications

Here we would like to discuss the concept of tracking and how it is used in *BABAR* and this analysis. The following will be covered:

- Definition of tracking,
- Track-finding theory and its use at *BABAR*, and
- Definition of *BABAR* track-quality categories.

Definition of tracking

Tracking at *BABAR* has two components, track finding and track fitting. A track is defined as any collection of detector hits which can be connected to form a valid trajectory through the detector. Track finding is the process by which we use pattern recognition to collect the many hits in the various sub-detectors of *BABAR* into groups such that it represents a particle's possible path. What is a valid trajectory is determined by track fitting algorithms. These algorithms

involve a series of least-squares fits, called a Kalman fit. In a Kalman fit one starts with a track seed (a set of n detector measurements of the passage of a charged particle) and fits a track to these n measurements. Then all other measurements of a charged particle's position is sequentially added with the fit being re-calculated for $n' = n + 1$ measurements after each addition. Track fitting would be trivial if we didn't have to account for experimental effects like, for example, multiple scattering, energy loss, and a non-uniform magnetic field in our detector.

Track-finding theory and its use at *BABAR*

In the ideal scenario in which the experimental effects like those listed above do not occur, the only force on a charged particle would come from its passage through a magnetic field. As a rough rule of thumb, the following holds true in a uniform magnetic field:

$$p_T = 0.3Br \ ,$$

where p_T is the magnitude of the momentum transverse to the z-axis in GeV/c, B is the magnetic field strength in Tesla, and r is the radius the particle's curved tack in meters. If the particle has a non-zero component of its momentum along the z-axis, the particle's trajectory is stretched out into a helix, but, of course, the above equation still holds for the projection of the track onto the x-y plane.

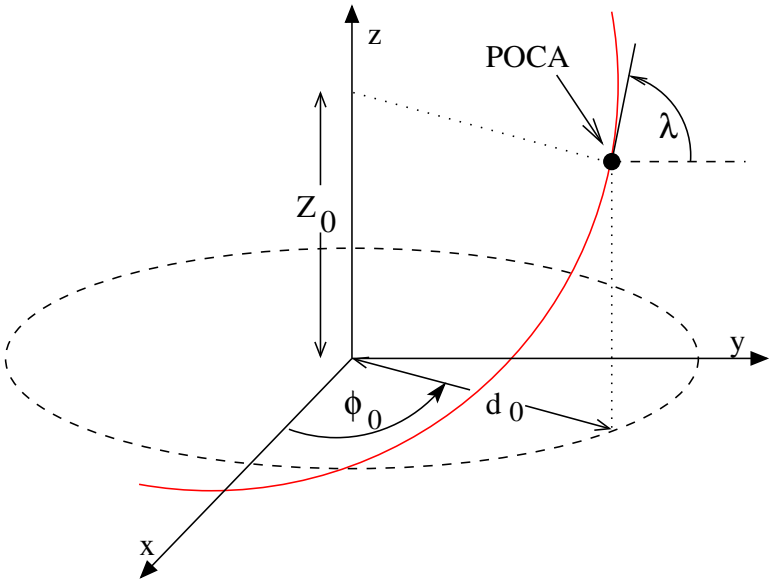

Figure D.1: Track helix parameters.

A helix is defined by five parameters (three parameters describing a circle plus two parameters defining a line). The definitions of the z offset (z_0), the radius (d_0), the azimuthal angle of the POCA (ϕ_0), and the dip angle (λ) are depicted in Figure D.1. The fifth parameter, the geometrical curvature (ω) is related to the radius of curvature and is defined from the transverse momentum,

$$r = \frac{1}{\omega} = \frac{-10^{13} p_T}{q B_z c} = -\frac{333.6 p_T}{q B_z} \ .$$

Definition of *BABAR* track-quality categories

The *BABAR* collaboration has defined track-quality categories. The categories relevant for this analysis are:

- `ChargedTracks`. All the tracks in the detector that correspond to charged particles.

- `GoodTracksVeryLoose`. Take the tracks in the `ChargedTracks` category and impose the following:

  - $p_T \geq 0$ GeV/c,
  - $p < 10$ GeV/c,
  - The χ^2 from the track-fitting algorithm must have a probability ≥ 0,
  - The distance of closest approach in the x-y plane (DOCA XY) must be < 1.5 cm.[1]
  - The z position of the point of closest approach (POCA) must be < 10 cm from the beam spot z position.[2]

- `GoodTracksLoose`. Take the tracks in the `GoodTracksVeryLoose` category and impose the following:

  - $p_T \geq 0.1$ GeV/c
  - Number of hits in the drift chamber ≥ 12

[1] To calculate DOCA XY, project the particle track and the location of the intersection of the two beams (the beam spot) onto the x-y plane. The DOCA XY is the minimum distance in the x-y plane between the projected track projected beam spot.

[2] The point of closest approach is the (x,y,z) position along the track where the (x,y) coordinate is that which satisfies the conditions of the DOCA XY; *ibid.* footnote 1.

Appendix E

Justification: Incorporating Systematic Uncertainties

In this appendix we describe the method used in Chapter 9 to incorporate the systematic uncertainties into the the upper limit calculation. Recall the branching fraction upper limit is obtained using Equation 9.2,

$$\mathcal{B} < \frac{N_{UL}}{N_{B\bar{B}} \sum_i \epsilon_i \cdot \mathcal{B}_i}.$$

In our approach we interpret the likelihood as a PDF for the true number of events (i.e. the number of events one would observe if this experiment was performed many times). The numerator is a PDF. The denominator, which we will call the "B yield," is a number with an uncertainty. To perform this division, we represent the denominator by a unit-normalized Gaussian distribution with mean=1 and width equal to the value and uncertainty of $N_{B\bar{B}} \sum_i \epsilon_i \cdot \mathcal{B}_i$, respectively.

We saw no excess of signal events in the signal box. Now, if the B yield had no uncertainty, we could simply divide the value of N_{UL} obtained by integrating

the likelihood function by the B yield to obtain our upper limit. On the other hand, if we had seen a signal, irregardless of whether there is an uncertainty on the B yield, the numerator and the denominator would each be a number with an uncertainty; from this we could use the normal propagation-of-errors rules for quotients to obtain a number with an uncertainty for $\mathcal{B}$. However, the B yield has an uncertainty and we observed no signal; thus, the numerator is not a number with an uncertainty but rather it is a PDF. This PDF is a function of the number of events expected, N, and represents the likelihood of observing zero events if in the limit of infinite statistics we would have measured N. Since the numerator is a PDF, and the denominator is a number with an uncertainty it does not make sense to divide the former by the latter. To progress, we will perform the division using a numerical method in which we represent the denominator by a PDF.

The denominator is a number with an uncertainty that represents our best estimate of the value of $N_{B\overline{B}} \sum_i \epsilon_i \cdot \mathcal{B}_i$ and its uncertainty. We extract the central value of this quantity and then, since the uncertainty is symmetric around the central value, we assume the uncertainty can be represented by a unit-normalized Gaussian distribution with mean and width equal to the value and uncertainty of $N_{B\overline{B}} \sum_i \epsilon_i \cdot \mathcal{B}_i$, respectively. Thus the B yield is represented by a number multiplied by a Gaussian distribution.

Now that both the numerator and denominator of 9.2 are represented by PDFs the division can be meaningful. For each point along the curve of $\mathcal{L}(N)$:

- Draw a random number from the Gaussian distribution that represents the B yield,
- Divide the value of $\mathcal{L}(N)$ by this random number,
- Plot this new value.

The result is a new curve, call it $\mathcal{B}'$. When $\mathcal{L}(N)$ is replaced by $\mathcal{B}'$ in Equation 9.2 and the integration is performed we get the 90% C.L. upper limit on the branching fraction, $\mathcal{B}$.

Appendix F

Repeat of Section 9.1, using $\mathcal{B}(D_s^- \to \phi\pi^-)=(3.6 \pm 0.9)\%$

In this appendix we tabulate as we did in Section 9.1 the values of and uncertainties for $\sum_i \epsilon_i \cdot \mathcal{B}_i$. This time we use the world average from the 2004 Review of Particle Physics published by the Particle Data Group at Berkeley [16]. Again, these are intended to allow the reader to get a feel for the level of uncertainty involved in this analysis and make a rough calculation of $\mathcal{B}$.

All the branching fractions of all $D_s^{(*)\pm}$ decay modes are normalized to $\mathcal{B}(D_s^- \to \phi\pi^-)$, thus $\mathcal{B}(D_s^- \to K^{*0}K^-)$ and $\mathcal{B}(D_s^- \to K^-K_S^0)$ need to be recalculated. Table F.1 shows the total branching fractions of the three $D_s^\pm$ decay modes, given that $\mathcal{B}(D_s^- \to \phi\pi^-)= (3.6\pm0.9)\%$. Then in Table F.2 we summarize our knowledge of $\epsilon_i \cdot \mathcal{B}_i$. The branching fractions, $\mathcal{B}$, include the branching fractions for all the daughter meson decays, $D_s^{*\pm}$, $D_s^\pm$, ϕ, K_S^0, and K^{*0}.

Table F.1: The total branching fractions for the secondary and tertiary decays in the modes considered in this analysis. The uncertainties listed are the correlated and uncorrelated, respectively.

B decay mode	$D_s^{*\pm}$ mode	$D_s^{\pm}$ mode	$D_s^{\pm}/D_s^{*\pm}$ branching fraction
$D_s^{\pm}\phi$	-	$\phi\pi^-$	$(0.87 \pm 0.22 \pm 0.02)\%$
$D_s^{\pm}\phi$	-	$K^{\pm}K_S^0$	$(0.62 \pm 0.15 \pm 0.10)\%$
$D_s^{\pm}\phi$	-	$\overline{K}^{*0}K^{\pm}$	$(1.08 \pm 0.27 \pm 0.11)\%$
$D_s^{*\pm}\phi$	$D_s^{\pm}\gamma$	$\phi\pi^-$	$(0.82 \pm 0.20 \pm 0.03)\%$
$D_s^{*\pm}\phi$	$D_s^{\pm}\gamma$	$K^{\pm}K_S^0$	$(0.58 \pm 0.15 \pm 0.09)\%$
$D_s^{*\pm}\phi$	$D_s^{\pm}\gamma$	$\overline{K}^{*0}K^{\pm}$	$(1.02 \pm 0.26 \pm 0.10)\%$

According to the 2004 PDG, the uncertainties are 25%, 30%, and 27% for the $D_s^- \to \phi\pi^-$, $D_s^- \to K^-K_S^0$, and $D_s^- \to K^{*0}K^-$ branching ratios respectively. (See Table F.1 for a breakdown.) Accounting for correlations the total uncertainty is on $\mathcal{B}$ is 26%.

Table F.3 is similar to table two, except we average over the $D_s^{(*)\pm}$ decay modes and not over the Runs. The values for Runs 1-3 and Run 4 are shown separately for the $B^{\pm} \to D_s^{\pm}\phi$, the $B^{\pm} \to D_s^{*\pm}\phi$ longitudinal polarization, and the $B^{\pm} \to D_s^{*\pm}\phi$ transverse polarization modes.

In Table F.4 we list the values of the $\sum_i \epsilon_i \cdot \mathcal{B}_i$ for $B^{\pm} \to D_s^{\pm}\phi$ and $B^{\pm} \to D_s^{*\pm}\phi$ averaged by luminosity over all runs. Here we include all systematic uncertainties on ϵ_i and $\mathcal{B}_i$ (recall Table 8.5). Note, the $\sum_i \epsilon_i \cdot \mathcal{B}_i$ and its uncertainty are independent of $D_s^{*\pm}$ polarization.

Table F.2: Signal MC efficiency (ϵ_i), averaged between Runs 1-3 and Run 4 according to luminosity (see Table 9.1), total branching fractions ($\mathcal{B}_i$) of the reconstructed daughter decay modes, and product of acceptance and branching fraction for each of the three modes used in the $B^\pm \to D_s^{*\pm}\phi$ search. When two uncertainties are listed, the first is the uncertainty associated with the ϵ_i (only the uncertainties due to MC statistics and tracking efficiency corrections, i.e. those in Table 9.1, are included) and the second is that associated with $\mathcal{B}_i$.

$D_s^{(*)\pm}$ Decay Mode	Polarization	ϵ_i (%)	$\mathcal{B}_i$ (10^{-3})	$\epsilon_i \cdot \mathcal{B}_i$ (10^{-3})
$D_s^- \to \phi\pi^-$	—	19.2 ± 0.6	8.68 ± 2.17	$1.67 \pm 0.05 \pm 0.42$
$D_s^- \to K^- K_S^0$	—	17.7 ± 0.4	6.15 ± 1.82	$1.09 \pm 0.03 \pm 0.32$
$D_s^- \to K^{*0} K^-$	—	14.0 ± 0.4	10.84 ± 2.91	$1.52 \pm 0.04 \pm 0.41$
$D_s^{*-} \to (\phi\pi^-)\gamma$	Longitudinal	10.9 ± 0.4	8.18 ± 2.06	$0.89 \pm 0.03 \pm 0.23$
$D_s^{*-} \to (K^- K_S^0)\gamma$	Longitudinal	10.0 ± 0.3	5.80 ± 1.72	$0.58 \pm 0.02 \pm 0.18$
$D_s^{*-} \to (K^{*0} K^-)\gamma$	Longitudinal	8.3 ± 0.3	10.21 ± 2.76	$0.84 \pm 0.03 \pm 0.24$
$D_s^{*-} \to (\phi\pi^-)\gamma$	Transverse	10.9 ± 0.4	8.18 ± 2.06	$0.89 \pm 0.03 \pm 0.23$
$D_s^{*-} \to (K^- K_S^0)\gamma$	Transverse	10.0 ± 0.3	5.80 ± 1.72	$0.58 \pm 0.02 \pm 0.18$
$D_s^{*-} \to (K^{*0} K^-)\gamma$	Transverse	8.3 ± 0.3	10.21 ± 2.76	$0.85 \pm 0.03 \pm 0.24$

Using all this but not including systematic uncertainties yet, we can calculate, just as in Chapter 9, the upper limit on $\mathcal{B}(B^\pm \to D_s^\pm\phi)$ using $\mathcal{B}(D_s^\pm \to \phi\pi^\pm)=(3.6 \pm 0.9)\%$, $\mathcal{B}(B^\pm \to D_s^\pm\phi) < 2.3 \times 10^{-6}$ at the 90% confidence level.

Our systematic uncertainties are a bit higher when using $\mathcal{B}(D_s^- \to \phi\pi^-)=(3.6\pm 0.9)\%$. The PDF for the acceptance is taken to be a Gaussian of mean $\mu = 1.33 \times 10^6$ and of $\sigma = 0.28 \times \mu$. The result is

$$\mathcal{B}(B^\pm \to D_s^\pm\phi) < 2.6 \times 10^{-6} \quad \text{at 90\% C.L.}$$

For the $B^\pm \to D_s^{*\pm}\phi$ mode, if we first neglecting systematic uncertainties, $\mathcal{B}(B^\pm \to D_s^{*\pm}\phi) < 1.4 \times 10^{-5}$ at the 90% confidence level. And, if the Gaussian

Table F.3: The sum over modes of the acceptance times branching fraction. The index i runs over the $D_s^\pm$ decay modes. When two uncertainties are listed, the first is that associated with ϵ_i (only the uncertainties due to MC statistics and tracking efficiency corrections, i.e. those in Table 9.1, are included) and the second is that associated with $\mathcal{B}_i$. The row marked "All" is the weighted-by-luminosity average of the Runs 1-3 and Run 4 results. See Table 9.4 for the final result including all systematic uncertainties. Note that these results are independent of $D_s^{*\pm}$ helicity.

Runs	$B^\pm \to D_s^\pm\phi$ $\sum_i \epsilon_i \cdot \mathcal{B}_i$ (10^{-3})	$B^\pm \to D_s^{*\pm}\phi$ Long. $\sum_i \epsilon_i \cdot \mathcal{B}_i$ (10^{-3})	$B^\pm \to D_s^{*\pm}\phi$ Trans. $\sum_i \epsilon_i \cdot \mathcal{B}_i$ (10^{-3})
1-3	$4.48 \pm 0.17 \pm 1.14$	$2.42 \pm 0.09 \pm 0.62$	$2.48 \pm 0.09 \pm 0.63$
4	$4.05 \pm 0.15 \pm 1.04$	$2.19 \pm 0.09 \pm 0.71$	$2.16 \pm 0.09 \pm 0.70$
All	$4.27 \pm 0.16 \pm 1.09$	$2.31 \pm 0.06 \pm 0.59$	$2.32 \pm 0.06 \pm 0.60$

Table F.4: The sum over modes of the acceptance times branching fraction with all systematic uncertainties involving ϵ and $\mathcal{B}_i$ listed in Table 8.5 included except that for the number of $B\overline{B}$ pairs. All care was taken to account for correlations between systematic uncertainties. The first uncertainty is that related to ϵ_i. The second uncertainty is from the uncertainty in $\mathcal{B}_i$.

B decay mode	$\sum_i \epsilon_i \cdot \mathcal{B}_i$ (10^{-3})
$B^\pm \to D_s^\pm\phi$	$5.70 \pm 0.57 \pm 0.78$
$B^\pm \to D_s^{*\pm}\phi$	$3.09 \pm 0.31 \pm 0.42$

PDF for the denominator of Equation 5.1 is taken to have a mean of $\mu = 7.15 \times 10^6$ and of $\sigma = 0.28 \times \mu$. The result is

$$\mathcal{B}(B^\pm \to D_s^{*\pm}\phi) < 1.7 \times 10^{-5} \quad \text{at } 90\% \text{ C.L.}$$

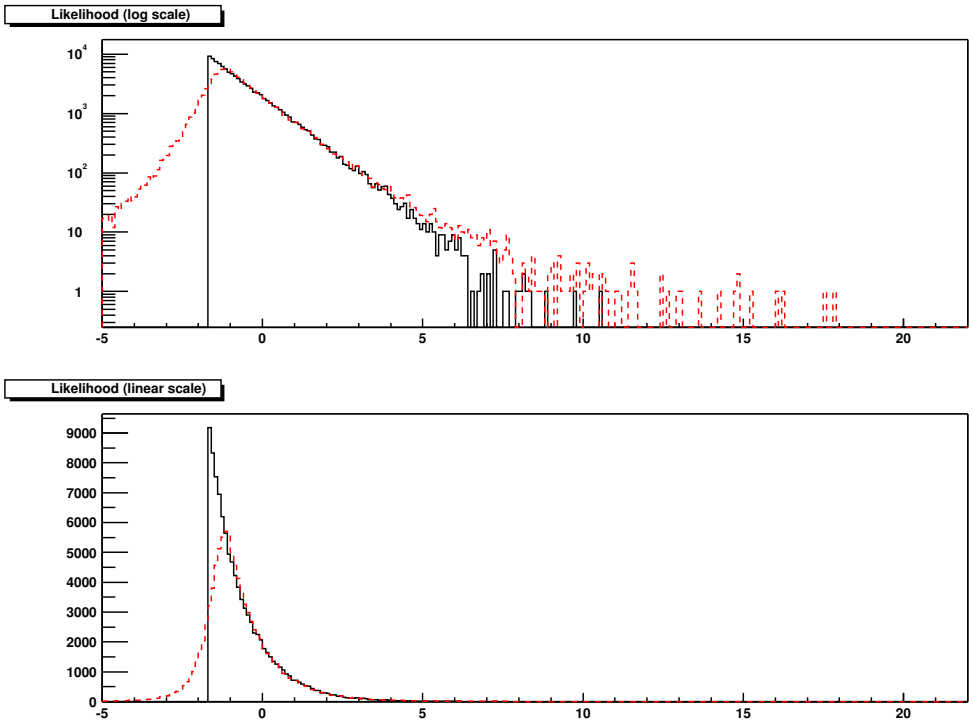

Figure F.1: $B^\pm \to D_s^\pm \phi$, $\mathcal{B}(D_s^\pm \to \phi\pi^\pm)$=$(3.6 \pm 0.9)\%$. The likelihood (black solid) and "smeared" likelihood (red dashed), in arbitrary units as a function of the number of signal events (N_{sig}). Top: log-scale ordinate. Bottom: linear-scale ordinate.

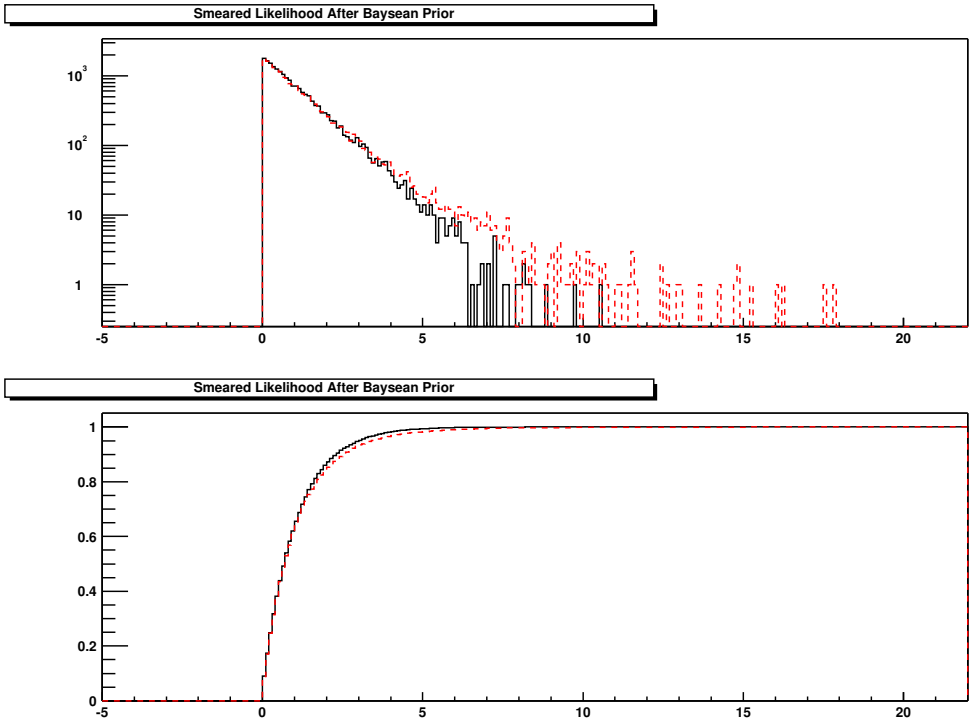

Figure F.2: $B^{\pm} \to D_s^{\pm}\phi$, $\mathcal{B}(D_s^{\pm} \to \phi\pi^{\pm})$=$(3.6 \pm 0.9)\%$. Top: The likelihood (black solid) and "smeared" likelihood (red dashed), as a function of N_{sig} after applying the flat Bayesian prior. Bottom: The integration of the top plot.

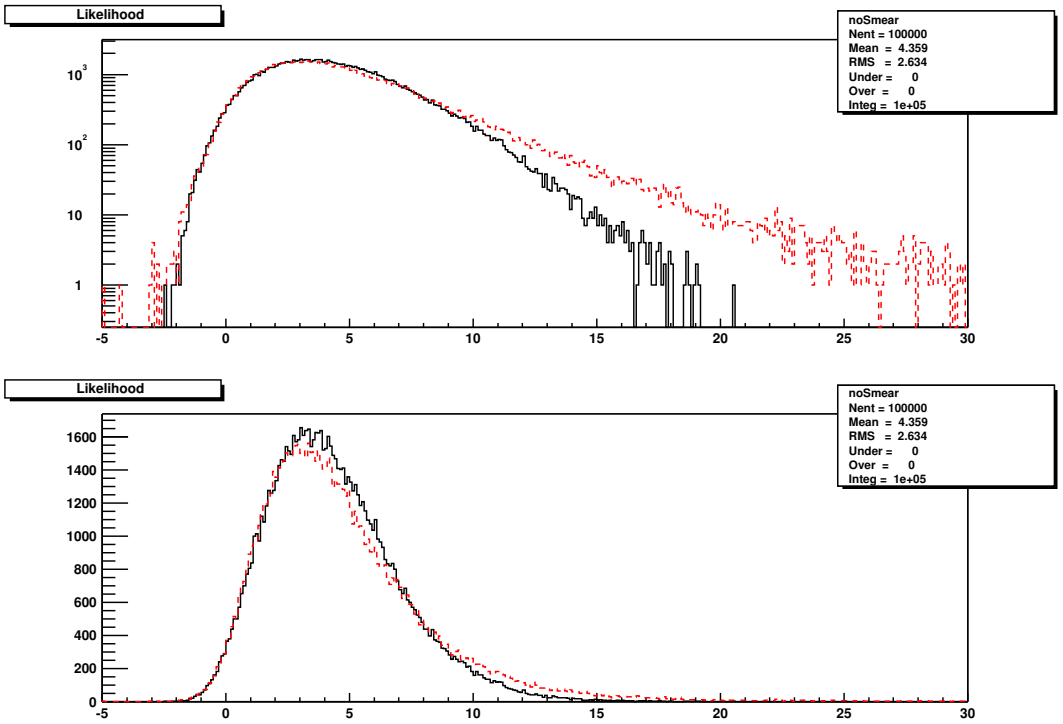

Figure F.3: $B^{\pm} \to D_s^{*\pm}\phi$, $\mathcal{B}(D_s^{\pm} \to \phi\pi^{\pm})$=$(3.6 \pm 0.9)\%$. The likelihood (black solid) and "smeared" likelihood (red dashed), in arbitrary units as a function of the number of signal events (N_{sig}). Top: log-scale ordinate. Bottom: linear-scale ordinate.

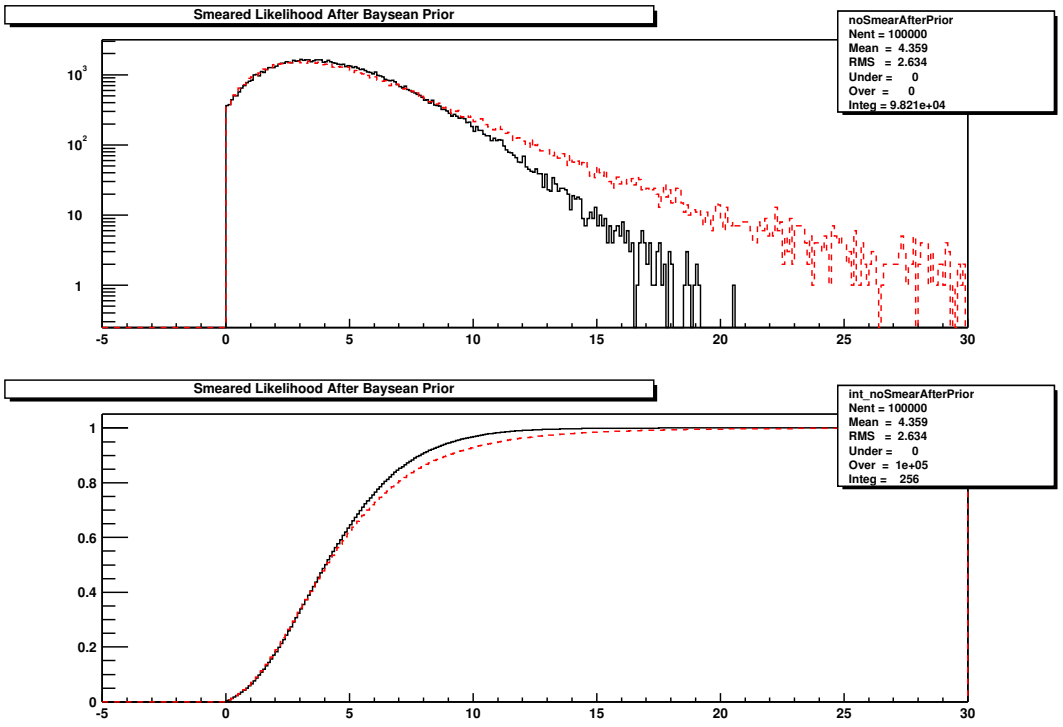

Figure F.4: $B^{\pm} \to D_s^{*\pm}\phi$, $\mathcal{B}(D_s^{\pm} \to \phi\pi^{\pm})$=$(3.6 \pm 0.9)\%$. Top: The likelihood (black solid) and "smeared" likelihood (red dashed), as a function of N_{sig} after applying the flat Bayesian prior. Bottom: The integration of the top plot.